An Introductory Guide in the Construction of Actuarial Models: A Preparation for the Actuarial Exam C/4

Marcel B. Finan
Arkansas Tech University

Preliminary Draft
Last Updated

November 4, 2017

To my son
Amin

Preface

This is the fifth of a series of lecture notes intended to help individuals to pass actuarial exams. The topics in this manuscript parallel the topics tested on Exam C of the Society of Actuaries exam sequence. As with the previous manuscripts, the main objective of the present manuscript is to increase users' understanding of the topics covered on the exam.

The flow of topics follows very closely that of Klugman et al. *Loss Models: From Data to Decisions*. The lectures cover designated sections from this book as suggested by the 2012 SOA Syllabus.

The recommended approach for using this manuscript is to read each section, work on the embedded examples, and then try ALL the problems given in the text. An answer key is provided by request. Email:mfinan@atu.edu.

Problems taken from previous SOA/CAS exams will be indicated by the symbol ‡.

This manuscript can be used for personal use or class use, but not for commercial purposes. If you find any errors, I would appreciate hearing from you: mfinan@atu.edu

Marcel B. Finan
Russellville, Arkansas
February 15, 2013.

Contents

Preface iii

Actuarial Modeling 1
1 Understanding Actuarial Models 2

A Review of Probability Related Results 5
2 A Brief Review of Probability 6
3 A Review of Random Variables 18
4 Raw and Central Moments 34
5 Empirical Models, Excess and Limited Loss variables 45
6 Median, Mode, Percentiles, and Quantiles 62
7 Sum of Random Variables and the Central Limit Theorem 69
8 Moment Generating Functions and Probability Generating Functions 78

Tail Weight of a Distribution 87
9 Tail Weight Measures: Moments and the Speed of Decay of $S(x)$ 88
10 Tail Weight Measures: Hazard Rate Function and Mean Excess Loss Function 95
11 Equilibrium Distributions and Tail Weight 102

Risk Measures 109
12 Coherent Risk Measurement 110
13 Value-at-Risk 115
14 Tail-Value-at-Risk 119

Characteristics of Actuarial Models 125
15 Parametric and Scale Distributions 126
16 Discrete Mixture Distributions 131
17 Data-dependent Distributions 138

Generating New Distributions **143**
18 Scalar Multiplication of Random Variables 144
19 Powers and Exponentiation of Random Variables 148
20 Continuous Mixing of Distributions 153
21 Frailty (Mixing) Models 160
22 Spliced Distributions 165
23 Limiting Distributions 168
24 The Linear Exponential Family of Distributions 172

Discrete Distributions **177**
25 The Poisson Distribution 178
26 The Negative Binomial Distribution 183
27 The Bernoulli and Binomial Distributions 189
28 The $(a, b, 0)$ Class of Discrete Distributions 195
29 The Class $C(a, b, 1)$ of Discrete Distributions 201
30 The Extended Truncated Negative Binomial Model 206

Modifications of the Loss Random Variable **211**
31 Ordinary Policy Deductibles 212
32 Franchise Policy Deductibles 220
33 The Loss Elimination Ratio and Inflation Effects for Ordinary Deductibles 226
34 Policy Limits 233
35 Combinations of Coinsurance, Deductibles, Limits, and Inflations 237
36 The Impact of Deductibles on the Number of Payments 250

Aggregate Loss Models **257**
37 Individual Risk and Collective Risk Models 258
38 Aggregate Loss Distributions via Convolutions 264
39 Stop Loss Insurance 284
40 Closed Form of Aggregate Distributions 294
41 Distribution of S via the Recursive Method 300
42 Discretization of Continuous Severities 308
43 Individual Policy Modifications Impact on Aggregate Losses . . 314
44 Aggregate Losses for the Individual Risk Model 321
45 Approximating Probabilities in the Individual Risk Model . . . 328

Review of Mathematical Statistics **333**
46 Properties of Point Estimators 334
47 Interval Estimation 341

48 Hypothesis Testing 344

The Empirical Distribution for Complete Data **351**
49 The Empirical Distribution for Individual Data 352
50 Empirical Distribution of Grouped Data 357

Estimation of Incomplete Data **363**
51 The Risk Set of Incomplete Data 364
52 The Kaplan-Meier and Nelson-Åalen Estimators 370
53 Mean and Variance of Empirical Estimators with Complete Data 379
54 Greenwood Estimate for the Variance of the Kaplan-Meier Estimator 387
55 Variance Estimate of the Nelson-Åalen Estimator and Confidence Intervals 393
56 Kernel Density Estimation 400
57 The Kaplan-Meier Approximation for Large Data Sets 409

Methods of parameter Estimation **419**
58 Method of Moments and Matching Percentile 421
59 Maximum Likelihood Estimation for Complete Data 434
60 Maximum Likelihood Estimation for Incomplete Data 445
61 Asymptotic Variance of MLE 455
62 Information Matrix and the Delta Method 463
63 Non-Normal Confidence Intervals for Parameter Estimation . . 471
64 Basics of Bayesian Inference 474
65 Bayesian Parameter Estimation 485
66 Conjugate Prior Distributions 493
67 Estimation of Class $(a, b, 0)$ 497
68 MLE with $(a, b, 1)$ Class 506

Model Selection and Evaluation **513**
69 Assessing Fitted Models Graphically 514
70 Kolmogorov-Smirnov Hypothesis Test of Fitted Models 521
71 Anderson-Darling Hypothesis Test of Fitted Models 527
72 The Chi-Square Goodness of Fit Test 532
73 The Likelihood Ratio Test 540
74 Schwarz Bayesian Criterion 546

Credibility Theory 551
75 Limited Fluctuation Credibility Approach: Full Credibility . . . 552
76 Limited Fluctuation Credibility Approach: Partial Credibility . 561
77 Greatest Accuracy Credibility Approach 567
78 Conditional Distributions and Expectation 571
79 Bayesian Credibility with Discrete Prior 581
80 Bayesian Credibility with Continuous Prior 594
81 Bühlman Credibility Premium 600
82 The Bühlmann Model with Discrete Prior 605
83 The Bühlmann Model with Continuous Prior 618
84 The Bühlmann-Straub Credibility Model 627
85 Exact Credibility 637
86 Non-parametric Empirical Bayes Estimation for the Bühlmann Model 643
87 Non-parametric Empirical Bayes Estimation for the Bühlmann-Straub Model 652
88 Semiparametric Empirical Bayes Credibility Estimation 661

Basics of Stochastic Simulation 667
89 The Inversion Method for Simulating Random Variables 668
90 Applications of Simulation in Actuarial Modeling 675
91 Estimating Risk Measures Using Simulation 684
92 The Bootstrap Method for Estimating Mean Square Error . . . 686

Answer Key 691

Exam C Tables 781

BIBLIOGRAPHY 796

Index 798

Actuarial Modeling

This book is concerned with the construction and evaluation of actuarial models. The purpose of this chapter is to define models in the actuarial setting and suggest a process of building them.

1 Understanding Actuarial Models

Modeling is very common in actuarial applications. For example, life insurance actuaries use models to arrive at the likely mortality rates of their customers; car insurance actuaries use models to work out claim probabilities by rating factors; pension fund actuaries use models to estimate the contributions and investments they will need to meet their future liabilities.

A "**model**" in actuarial applications is a simplified mathematical description of a certain actuarial task. Actuarial models are used by actuaries to form an opinion and recommend a course of action on contingencies relating to uncertain future events.

Commonly used actuarial models are classified into two categories:

(I) **Deterministic Models.** These are models that produce a unique set of outputs for a given set of inputs such as the future value of a deposit in a savings account. In these models, the inputs and outputs don't have associated probability weightings.

(II) **Stochastic or Probabilistic Models.** In contrast to deterministic models, these are models where the outputs or/and some of the inputs are random variables. Examples of stochastic models that we will discuss in this book are the asset model, the claims model, and the frequency-severity model.

The book in [4] explains in enormous detail the advantages and disadvantages of stochastic (versus deterministic) modeling.

Example 1.1
Determine whether each of the model below is deterministic or stochastic.
(a) The monthly payment P on a home or a car loan.
(b) A modification of the model in (a) is $P+\xi$, where ξ is a random variable introduced to account for the possibility of failure of making a payment.

Solution.
(a) In this model, the element of randomness is absent. This model is a deterministic one.
(b) Because of the presence of the random variable ξ, the given model is stochastic ■

In [1], the following process for building an actuarial model is presented.

Phases of a Good Modeling Process

A good modeling requires a thorough understanding of the problem modelled. The following is a helpful checklist of a modeling process which is by no means complete:

Choice of Models. Appropriate models are selected based on the actuary's prior knowledge and experience and the nature of the available data.

Model Calibration. Available data and existing techniques are used to calibrate a model.

Model Validation. Diagnostic tests are used to ensure the model meets its objectives and adequately conforms to the data.

Adequacy of Models. There is a possibility that the models in the previous stage are inadequate in which case one considers other possible models. Also, there is a possibility of having more than one adequate models.

Selection of Models. Based on some preset criteria, the best model will be selected among all valid models.

Modification of Models. The model selected in the previous stage needs to be constantly updated in the light of new future data and other changes.

Practice Problems

Problem 1.1
After an actuary being hired, his or her annual salary progression is modeled according to the formula $S(t) = \$45,000e^{0.06t}$, where t is the number of years of employment.

Determine whether this model is deterministic or stochastic.

Problem 1.2
In the previous model, a random variable ξ is introduced: $S(t) = \$45,000e^{0.06t} + \xi$.

Determine whether this model is deterministic or stochastic.

Problem 1.3
Consider a model that depends on the movement of a stock market such as the pricing of an option with an underlying stock.

Does this model considered a deterministic or stochastic model?

Problem 1.4
Consider a model that involves the life expectancy of a policyholder.

Is this model categorized as stochastic or deterministic?

Problem 1.5
Insurance companies use models to estimate their assets and liabilities.

Are these models considered deterministic or stochastic?

A Review of Probability Related Results

One aspect of insurance is that money is paid to policyholders upon the occurrence of a random event. For example, a claim in an automobile insurance policy will be filed whenever the insured auto is involved in a car wreck. In this chapter a brief outline of the essential material from the theory of probability is given. Almost all of the material presented here should be familiar to the reader. A more thorough discussion can be found in [2] and a listing of important results can be found in [3]. Probability concepts that are not usually covered in an introductory probability course will be introduced and discussed in futher details whenever needed.

2 A Brief Review of Probability

In probability, we consider **experiments** whose results cannot be predicted with certainty. Examples of such experiments include rolling a die, flipping a coin, and choosing a card from a deck of playing cards.

By an **outcome** or **simple event** we mean any result of the experiment. For example, the experiment of rolling a die yields six outcomes, namely, the outcomes 1,2,3,4,5, and 6.

The **sample space** Ω of an experiment is the set of all possible outcomes for the experiment. For example, if you roll a die one time then the experiment is the roll of the die. A sample space for this experiment could be $\Omega = \{1, 2, 3, 4, 5, 6\}$ where each digit represents a face of the die.

An **event** is a subset of the sample space. For example, the event of rolling an odd number with a die consists of three simple events $\{1, 3, 5\}$.

Example 2.1
Consider the random experiment of tossing a coin three times.
(a) Find the sample space of this experiment.
(b) Find the outcomes of the event of obtaining more than one head.

Solution.
We will use T for tail and H for head.
(a) The sample space is composed of eight simple events:

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}.$$

(b) The event of obtaining more than one head is the set

$$\{THH, HTH, HHT, HHH\} \blacksquare$$

The **complement** of an event E, denoted by E^c, is the set of all possible outcomes not in E. The **union** of two events A and B is the event $A \cup B$ whose outcomes are either in A or in B. The **intersection** of two events A and B is the event $A \cap B$ whose outcomes are outcomes of both events A and B.

Two events A and B are said to be **mutually exclusive** if they have no outcomes in common. Clearly, for any event E, the events E and E^c are mutually exclusive.

Remark 2.1
The above definitions of intersection, union, and mutually exclusive can be extended to any number of events.

Probability Axioms
Probability is the measure of occurrence of an event. It is a function $\Pr(\cdot)$ defined on the collection of all (subsets) events of a sample space Ω and which satisfies **Kolmogorov axioms**:

Axiom 1: For any event $E \subseteq \Omega$, $0 \le \Pr(E) \le 1$.
Axiom 2: $\Pr(\Omega) = 1$.
Axiom 3: For any sequence of mutually exclusive events $\{E_n\}_{n \ge 1}$, that is $E_i \cap E_j = \emptyset$ for $i \ne j$, we have

$$\Pr\left(\cup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \Pr(E_n). \text{ (Countable Additivity)}$$

If we let $E_1 = \Omega$, $E_n = \emptyset$ for $n > 1$ then by Axioms 2 and 3 we have $1 = \Pr(\Omega) = \Pr\left(\cup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \Pr(E_n) = \Pr(\Omega) + \sum_{n=2}^{\infty} \Pr(\emptyset)$. This implies that $\Pr(\emptyset) = 0$. Also, if $\{E_1, E_2, \cdots, E_n\}$ is a finite set of mutually exclusive events, then by defining $E_k = \emptyset$ for $k > n$ and Axiom 3 we find

$$\Pr\left(\cup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \Pr(E_k).$$

Any function Pr that satisfies Axioms 1-3 will be called a **probability measure**.

Example 2.2
Consider the sample space $\Omega = \{1, 2, 3\}$. Suppose that $\Pr(\{1, 3\}) = 0.3$ and $\Pr(\{2, 3\}) = 0.8$. Find $\Pr(1), \Pr(2)$, and $\Pr(3)$. Is Pr a valid probability measure?

Solution.
For Pr to be a probability measure we must have $\Pr(1) + \Pr(2) + \Pr(3) = 1$. But $\Pr(\{1, 3\}) = \Pr(1) + \Pr(3) = 0.3$. This implies that $0.3 + \Pr(2) = 1$ or $\Pr(2) = 0.7$. Similarly, $1 = \Pr(\{2, 3\}) + \Pr(1) = 0.8 + \Pr(1)$ and so $\Pr(1) = 0.2$. It follows that $\Pr(3) = 1 - \Pr(1) - \Pr(2) = 1 - 0.2 - 0.7 = 0.1$. It can be easily seen that Pr satisfies Axioms 1-3 and so Pr is a probability measure ■

Probability Trees

For all multistage experiments, the probability of the outcome along any path of a tree diagram is equal to the product of all the probabilities along the path.

Example 2.3

In a city council, 35% of the members are female, and the other 65% are male. 70% of the male favor raising city sales tax, while only 40% of the female favor the increase. If a member of the council is selected at random, what is the probability that he or she favors raising sales tax?

Solution.

Figure 2.1 shows a tree diagram for this problem.

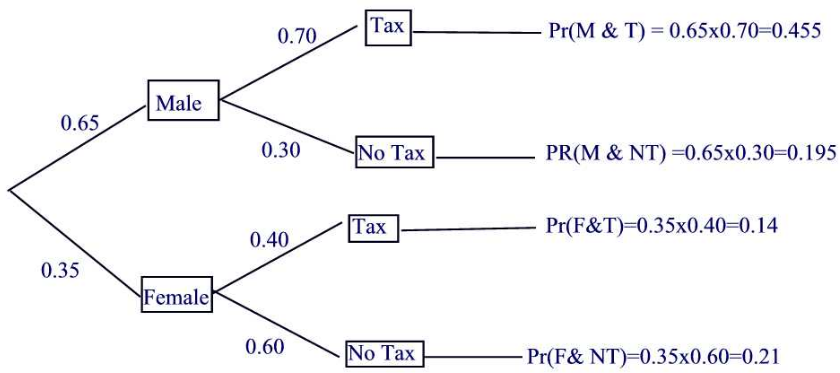


Figure 2.1

The first and third branches correspond to favoring the tax. We add their probabilities.

$$\Pr(\text{tax}) = 0.455 + 0.14 = 0.595 \blacksquare$$

Conditional Probability and Bayes Formula

Consider the question of finding the probability of an event A given that another event B has occurred. Knowing that the event B has occurred causes us to update the probabilities of other events in the sample space.

To illustrate, suppose you roll two dice of different colors; one red, and one green. You roll each die one at time. Our sample space has 36 outcomes. The probability of getting two ones is $\frac{1}{36}$. Now, suppose you were told that the green die shows a one but know nothing about the red die. What would be the probability of getting two ones? In this case, the answer is $\frac{1}{6}$. This shows that the probability of getting two ones changes if you have partial information, and we refer to this (altered) probability as a **conditional probability**.

If the occurrence of the event A depends on the occurrence of B then the conditional probability will be denoted by $P(A|B)$, read as the *probability of A given B*. It is given by

$$\Pr(A|B) = \frac{\text{number of outcomes corresponding to event A and B}}{\text{number of outcomes of B}}.$$

Thus,

$$\Pr(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(B)}{n(S)}} = \frac{\Pr(A \cap B)}{\Pr(B)}$$

provided that $\Pr(B) > 0$.

Example 2.4
Let A denote the event "an immigrant is male" and let B denote the event "an immigrant is Brazilian". In a group of 100 immigrants, suppose 60 are Brazilians, and suppose that 10 of the Brazilians are males. Find the probability that if I pick a Brazilian immigrant, it will be a male, that is, find $\Pr(A|B)$.

Solution.
Since 10 out of 100 in the group are both Brazilians and male, $\Pr(A \cap B) = \frac{10}{100} = 0.1$. Also, 60 out of the 100 are Brazilians, so $\Pr(B) = \frac{60}{100} = 0.6$. Hence, $\Pr(A|B) = \frac{0.1}{0.6} = \frac{1}{6}$ ■

It is often the case that we know the probabilities of certain events conditional on other events, but what we would like to know is the "reverse". That is, given $\Pr(A|B)$ we would like to find $\Pr(B|A)$.

Bayes' formula is a simple mathematical formula used for calculating $\Pr(B|A)$ given $\Pr(A|B)$. We derive this formula as follows. Let A and B be two events. Then

$$A = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

Since the events $A \cap B$ and $A \cap B^c$ are mutually exclusive, we can write

$$\begin{aligned} \Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c) \qquad (2.1) \end{aligned}$$

Example 2.5
A soccer match may be delayed because of bad weather. The probabilities are 0.60 that there will be bad weather, 0.85 that the game will take place if there is no bad weather, and 0.35 that the game will be played if there is bad weather. What is the probability that the match will occur?

Solution.
Let A be the event that the game will be played and B is the event that there will be a bad weather. We are given $\Pr(B) = 0.60,\ \Pr(A|B^c) = 0.85,$ and $\Pr(A|B) = 0.35$. From Equation (2.1) we find

$$\Pr(A) = \Pr(B)\Pr(A|B) + \Pr(B^c)\Pr(A|B^c) = (0.60)(0.35) + (0.4)(0.85) = 0.55 \blacksquare$$

From Equation (2.1) we can get Bayes' formula:

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{\Pr(A|B)\Pr(B)}{\Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c)}. \tag{2.2}$$

Example 2.6
A factory uses two machines A and B for making socks. Machine A produces 10% of the total production of socks while machine B produces the remaining 90%. Now, 1% of all the socks produced by A are defective while 5% of all the socks produced by B are defective. Find the probability that a sock taken at random from a day's production was made by the machine A, given that it is defective?

Solution.
We are given $\Pr(A) = 0.1,\ \Pr(B) = 0.9,\ \Pr(D|A) = 0.01,$ and $\Pr(D|B) = 0.05$. We want to find $\Pr(A|D)$. Using Bayes' formula we find

$$\begin{aligned}\Pr(A|D) =& \frac{\Pr(A \cap D)}{\Pr(D)} = \frac{\Pr(D|A)\Pr(A)}{\Pr(D|A)\Pr(A) + \Pr(D|B)\Pr(B)} \\ =& \frac{(0.01)(0.1)}{(0.01)(0.1) + (0.05)(0.9)} \approx 0.0217 \blacksquare\end{aligned}$$

Formula 2.2 is a special case of the more general result:

Theorem 2.1 *(Bayes' Theorem)*
Suppose that the sample space Ω is the union of mutually exclusive events $H_1, H_2, \cdots, H_n$ with $P(H_i) > 0$ for each i. Then for any event A and $1 \leq i \leq n$ we have

$$\Pr(H_i|A) = \frac{\Pr(A|H_i)\Pr(H_i)}{\Pr(A)}$$

where

$$\Pr(A) = \Pr(H_1)\Pr(A|H_1) + \Pr(H_2)\Pr(A|H_2) + \cdots + \Pr(H_n)\Pr(A|H_n).$$

Example 2.7
A survey is taken in Oklahoma, Kansas, and Arkansas. In Oklahoma, 50% of surveyed support raising tax, in Kansas, 60% support a tax increase, and in Arkansas only 35% favor the increase. Of the total population of the three states, 40% live in Oklahoma, 25% live in Kansas, and 35% live in Arkansas. Given that a surveyed person is in favor of raising taxes, what is the probability that he/she lives in Kansas?

Solution.
Let L_I denote the event that a surveyed person lives in state I, where I = OK, KS, AR. Let S denote the event that a surveyed person favors tax increase. We want to find $\Pr(L_{KS}|S)$. By Bayes' formula we have

$$\begin{aligned}\Pr(L_{KS}|S) =& \frac{\Pr(S|L_{KS})\Pr(L_{KS})}{\Pr(S|L_{OK})\Pr(L_{OK}) + \Pr(S|L_{KS})\Pr(L_{KS}) + \Pr(S|L_{AR})\Pr(L_{AR})}\\ =& \frac{(0.6)(0.25)}{(0.5)(0.4) + (0.6)(0.25) + (0.35)(0.35)} \approx 0.3175 \blacksquare\end{aligned}$$

Practice Problems

Problem 2.1
Consider the sample space of rolling a die. Let A be the event of rolling an even number, B the event of rolling an odd number, and C the event of rolling a 2.

Find
(a) A^c, B^c and C^c.
(b) $A \cup B, A \cup C$, and $B \cup C$.
(c) $A \cap B, A \cap C$, and $B \cap C$.
(d) Which events are mutually exclusive?

Problem 2.2
If, for a given experiment, $O_1, O_2, O_3, \cdots$ is an infinite sequence of outcomes, verify that

$$\Pr(O_i) = \left(\frac{1}{2}\right)^i, \ i = 1, 2, 3, \cdots$$

is a probability measure.

Problem 2.3 ‡
An insurer offers a health plan to the employees of a large company. As part of this plan, the individual employees may choose exactly two of the supplementary coverages A, B, and C, or they may choose no supplementary coverage. The proportions of the company's employees that choose coverages A, B, and C are $\frac{1}{4}, \frac{1}{3}$, and , $\frac{5}{12}$ respectively.

Determine the probability that a randomly chosen employee will choose no supplementary coverage.

Problem 2.4
A toll has two crossing lanes. Let A be the event that the first lane is busy, and let B be the event the second lane is busy. Assume that $\Pr(A) = 0.2,\ \Pr(B) = 0.3$ and $\Pr(A \cap B) = 0.06$.

Find the probability that both lanes are not busy.
Hint: Recall the identity

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

Problem 2.5
If a person visits a car service center, suppose that the probability that he will have his oil changed is 0.44, the probability that he will have a tire replacement is 0.24, the probability that he will have airfilter replacement is 0.21, the probability that he will have oil changed and a tire replaced is 0.08, the probability that he will have oil changed and air filter changed is 0.11, the probability that he will have a tire and air filter replaced is 0.07, and the probability that he will have oil changed, a tire replacement, and an air filter changed is 0.03.

What is the probability that at least one of these things done to the car? Recall that

$$\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) + \Pr(A \cap B \cap C)$$

Problem 2.6 ‡
A survey of a group's viewing habits over the last year revealed the following information

(i)	28% watched gymnastics
(ii)	29% watched baseball
(iii)	19% watched soccer
(iv)	14% watched gymnastics and baseball
(v)	12% watched baseball and soccer
(vi)	10% watched gymnastics and soccer
(vii)	8% watched all three sports.

Find the probability of a viewer that watched none of the three sports during the last year.

Problem 2.7 ‡
The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 35% . Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work.

Determine the probability that a visit to a PCP's office results in both lab work and referral to a specialist.

Problem 2.8 ‡
You are given $\Pr(A \cup B) = 0.7$ and $\Pr(A \cup B^c) = 0.9$.

Determine $\Pr(A)$.

Problem 2.9 ‡
Among a large group of patients recovering from shoulder injuries, it is found that 22% visit both a physical therapist and a chiropractor, whereas 12% visit neither of these. The probability that a patient visits a chiropractor exceeds by 14% the probability that a patient visits a physical therapist.

Determine the probability that a randomly chosen member of this group visits a physical therapist.

Problem 2.10 ‡
In modeling the number of claims filed by an individual under an automobile policy during a three-year period, an actuary makes the simplifying assumption that for all integers $n \geq 0, p_{n+1} = \frac{1}{5}p_n$, where p_n represents the probability that the policyholder files n claims during the period.

Under this assumption, what is the probability that a policyholder files more than one claim during the period?

Problem 2.11
An urn contains three red balls and two blue balls. You draw two balls without replacement. Construct a probability tree diagram that represents the various outcomes that can occur.

What is the probability that the first ball is red and the second ball is blue?

Problem 2.12
Repeat the previous exercise but this time replace the first ball before drawing the second.

Problem 2.13 ‡
A public health researcher examines the medical records of a group of 937 men who died in 1999 and discovers that 210 of the men died from causes related to heart disease. Moreover, 312 of the 937 men had at least one parent who suffered from heart disease, and, of these 312 men, 102 died from causes related to heart disease.

Determine the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.

Problem 2.14 ‡
An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability that a woman has all three risk factors, given that she has A and B, is $\frac{1}{3}$.

What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?

Problem 2.15 ‡
An auto insurance company insures drivers of all ages. An actuary compiled the following statistics on the company's insured drivers:

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16 - 20	0.06	0.08
21 - 30	0.03	0.15
31 - 65	0.02	0.49
66 - 99	0.04	0.28

A randomly selected driver that the company insures has an accident.

Calculate the probability that the driver was age 16-20.

Problem 2.16 ‡
An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. Each standard policyholder has probability 0.010 of dying in the next year, each preferred policyholder has probability 0.005 of dying in the next year, and each ultra-preferred policyholder has probability 0.001 of dying in the next year.
A policyholder dies in the next year.

What is the probability that the deceased policyholder was ultra-preferred?

Problem 2.17 ‡
Upon arrival at a hospital's emergency room, patients are categorized according to their condition as critical, serious, or stable. In the past year:

(i) 10% of the emergency room patients were critical;
(ii) 30% of the emergency room patients were serious;
(iii) the rest of the emergency room patients were stable;
(iv) 40% of the critical patients died;
(vi) 10% of the serious patients died; and
(vii) 1% of the stable patients died.

Given that a patient survived, what is the probability that the patient was categorized as serious upon arrival?

Problem 2.18 ‡
A health study tracked a group of persons for five years. At the beginning of the study, 20% were classified as heavy smokers, 30% as light smokers, and 50% as nonsmokers.
Results of the study showed that light smokers were twice as likely as nonsmokers to die during the five-year study, but only half as likely as heavy smokers.
A randomly selected participant from the study died over the five-year period.

Calculate the probability that the participant was a heavy smoker.

Problem 2.19 ‡
An actuary studied the likelihood that different types of drivers would be involved in at least one collision during any one-year period. The results of the study are presented below.

Type of driver	Percentage of all drivers	Probability of at least one collision
Teen	8%	0.15
Young adult	16%	0.08
Midlife	45%	0.04
Senior	31%	0.05
Total	100%	

Given that a driver has been involved in at least one collision in the past year, what is the probability that the driver is a young adult driver?

Problem 2.20 ‡
A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence

of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease.

Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

3 A Review of Random Variables

Let Ω be the sample space of an experiment. Any function $X : \Omega \longrightarrow \mathbb{R}$ is called a **random variable** with **support** the range of X. The function notation $X(s) = x$ means that the random variable X assigns the value x to the outcome s.

We consider three types of random variables: Discrete, continuous, and mixed random variables.

A random variable is called **discrete** if either its support is finite or a countably infinite. For example, in the experiment of rolling two dice, let X be the random variable that adds the two faces. The support of X is the finite set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. An example of an infinite discrete random variable is the random variable that counts the number of times you play a lottery until you win. For such a random variable, the support is the set $\mathbb{N}$.

A random variable is called **continuous** if its support is uncountable. An example of a continuous random variable is the random variable X that gives a randomnly chosen number between 2 and 3 inclusively. For such a random variable the support is the interval $[2, 3]$.

A **mixed** random variable is partly discrete and partly continuous. An example of a mixed random variable is the random variable $X : (0, 1) \longrightarrow \mathbb{R}$ defined by

$$X(s) = \begin{cases} 1 - s, & 0 < s < \frac{1}{2} \\ \frac{1}{2}, & \frac{1}{2} \leq s < 1. \end{cases}$$

We use upper-case letters X, Y, Z, etc. to represent random variables. We use small letters x, y, z, etc to represent possible values that the corresponding random variables X, Y, Z, etc. can take. The statement $X = x$ defines an event consisting of all outcomes with $X-$measurement equal to x which is the set $\{s \in \Omega : X(s) = x\}$.

Example 3.1
State whether the random variables are discrete, continuous, or mixed.
(a) A coin is tossed ten times. The random variable X is the number of heads that are noted.
(b) A coin is tossed repeatedly. The random variable X is the number of times needed to get the first head.
(c) $X : (0, 1) \longrightarrow \mathbb{R}$ defined by $X(s) = 2s - 1$.

(d) $X : (0,1) \longrightarrow \mathbb{R}$ defined by $X(s) = 2s - 1$ for $0 < s < \frac{1}{2}$ and $X(s) = 1$ for $\frac{1}{2} \le s < 1$.

Solution.
(a) The support of X is $\{1, 2, 3, \cdots, 10\}$. X is an example of a finite discrete random variable.
(b) The support of X is $\mathbb{N}$. X is an example of a countably infinite discrete random variable.
(c) The support of X is the open interval $(-1, 1)$. X is an example of a continuous random variable.
(d) X is continuous on $(0, \frac{1}{2})$ and discrete on $[\frac{1}{2}, 1)$ $\blacksquare$

Because the value of a random variable is determined by the outcome of the experiment, we can find the probability that the random variable takes on each value. The notation $\Pr(X = x)$ stands for the probability of the event $\{s \in \Omega : X(s) = x\}$.

There are five key functions used in describing a random variable:

Probability Mass Function (PMF)
For a discrete random variable X, the distribution of X is described by the **probability function**(pf) or the **probability mass function**(pmf) given by the equation

$$p(x) = \Pr(X = x).$$

That is, a probability mass function (pmf) gives the probability that a discrete random variable is exactly equal to some value. Note that the domain of the pf is the support of the corresponding random variable. The pmf can be an equation, a table, or a graph that shows how probability is assigned to possible values of the random variable.

Example 3.2
Suppose a variable X can take the values 1, 2, 3, or 4. The probabilities associated with each outcome are described by the following table:

x	1	2	3	4
$p(x)$	0.1	0.3	0.4	0.2

Draw the probability histogram.

Solution.
The probability histogram is shown in Figure 3.1 ■

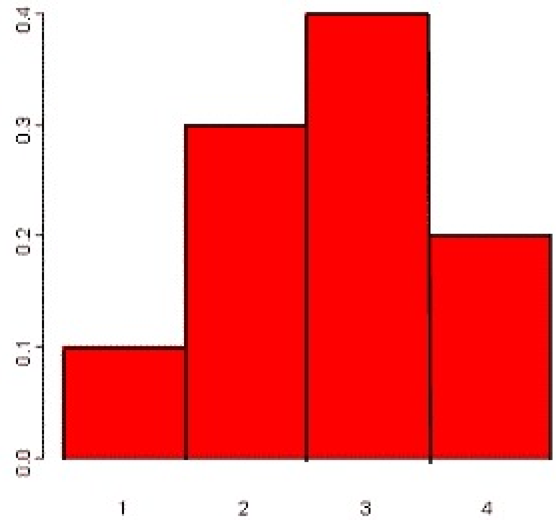


Figure 3.1

Example 3.3
A committee of m is to be selected from a group consisting of x men and y women. Let X be the random variable that represents the number of men in the committee. Find $p(n)$ for $0 \leq n \leq m$.

Solution.
For $0 \leq n \leq m$, we have

$$p(n) = \frac{\binom{x}{n}\binom{y}{m-n}}{\binom{x+y}{m}} \quad \blacksquare$$

Note that if the support of a discrete random variable is Support $= \{x_1, x_2, \cdots\}$ then

$$\begin{array}{rcll} p(x) & \geq & 0, & x \in \text{Support} \\ p(x) & = & 0, & x \notin \text{Support} \end{array}$$

Moreover,

$$\sum_{x \in \text{Support}} p(x) = 1.$$

Probability Density Function

Associated with a continuous random variable X is a nonnegative function f (not necessarily continuous) defined for all real numbers and having the property that for any set B of real numbers we have

$$\Pr(X \in B) = \int_B f(x)dx.$$

We call the function f the **probability density function** (abbreviated pdf) of the random variable X.

If we let $B = (-\infty, \infty) = \mathbb{R}$ then

$$\int_{-\infty}^{\infty} f(x)dx = \Pr[X \in (-\infty, \infty)] = 1.$$

Now, if we let $B = [a, b]$ then

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx.$$

That is, areas under the probability density function represent probabilities as illustrated in Figure 3.2.

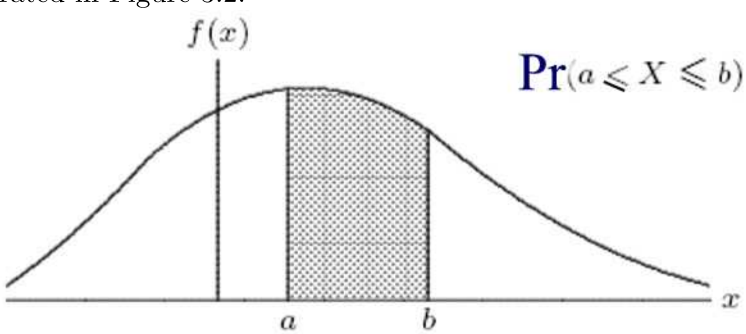


Figure 3.2

Now, if we let $a = b$ in the previous formula we find

$$\Pr(X = a) = \int_a^a f(x)dx = 0.$$

It follows from this result that

$$\Pr(a \leq X < b) = \Pr(a < X \leq b) = \Pr(a < X < b) = \Pr(a \leq X \leq b)$$

and

$$\Pr(X \leq a) = \Pr(X < a) \text{ and } \Pr(X \geq a) = \Pr(X > a).$$

Example 3.4
Suppose that the function $f(t)$ defined below is the density function of some random variable X.

$$f(t) = \begin{cases} e^{-t} & t \geq 0, \\ 0 & t < 0. \end{cases}$$

Compute $P(-10 \leq X \leq 10)$.

Solution.

$$\begin{aligned} P(-10 \leq X \leq 10) &= \int_{-10}^{10} f(t)dt = \int_{-10}^{0} f(t)dt + \int_{0}^{10} f(t)dt \\ &= \int_{0}^{10} e^{-t}dt = -e^{-t}\Big|_0^{10} = 1 - e^{-10} \quad \blacksquare \end{aligned}$$

<u>**Cumulative Distribution Function**</u>
The **cumulative distribution function** (abbreviated cdf) $F(t)$ of a random variable X is defined as follows

$$F(t) = \Pr(X \leq t)$$

i.e., $F(t)$ is equal to the probability that the variable X assumes values, which are less than or equal to t.

For a discrete random variable, the cumulative distribution function is found by summing up the probabilities. That is,

$$F(t) = \Pr(X \leq t) = \sum_{x \leq t} p(x).$$

Example 3.5
Given the following pmf

$$p(x) = \begin{cases} 1, & \text{if } x = a \\ 0, & \text{otherwise.} \end{cases}$$

Find a formula for $F(x)$ and sketch its graph.

Solution.
A formula for $F(x)$ is given by

$$F(x) = \begin{cases} 0, & \text{if } x < a \\ 1, & \text{otherwise} \end{cases}$$

Its graph is given in Figure 3.3 ■

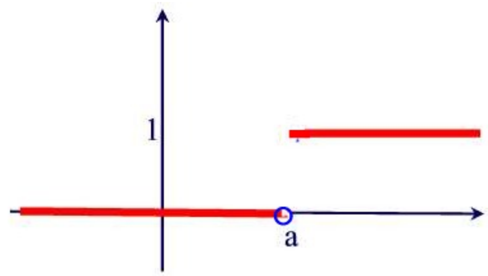


Figure 3.3

For discrete random variables the cumulative distribution function will always be a step function with jumps at each value of x that has probability greater than 0. Note that the value of $F(x)$ is assigned to the top of the jump.

For a continuous random variable, the cumulative distribution function is given by

$$F(t) = \int_{-\infty}^{t} f(y)dy.$$

Geometrically, $F(t)$ is the area under the graph of f to the left of t.

Example 3.6
Find the distribution functions corresponding to the following density functions:

$$(a) f(x) = \frac{1}{\pi(1+x^2)}, \qquad -\infty < x < \infty.$$

$$(b) f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad -\infty < x < \infty.$$

Solution.
(a)

$$F(x) = \int_{-\infty}^{x} \frac{1}{\pi(1+y^2)} dy = \left[\frac{1}{\pi}\arctan y\right]_{-\infty}^{x}$$
$$= \frac{1}{\pi}\arctan x - \frac{1}{\pi}\cdot\frac{-\pi}{2} = \frac{1}{\pi}\arctan x + \frac{1}{2}.$$

(b)

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} \frac{e^{-y}}{(1+e^{-y})^2} dy \\
&= \left[\frac{1}{1+e^{-y}}\right]_{-\infty}^{x} = \frac{1}{1+e^{-x}} \blacksquare
\end{aligned}
$$

Next, we list the properties of the cumulative distribution function $F(x)$ for any random variable X.

Theorem 3.1
The cumulative distribution function of a random variable X satisfies the following properties:
(a) $0 \le F(x) \le 1$.
(b) $F(x)$ is a non-decreasing function, i.e. if $a < b$ then $F(a) \le F(b)$.
(c) $F(x) \to 0$ as $x \to -\infty$ and $F(x) \to 1$ as $x \to \infty$.
(d) F is right-continuous.

In addition to the above properties, a continuous random variable satisfies these properties:
(e) $F'(x) = f(x)$.
(f) $F(x)$ is continuous.

For a discrete random variable with support $\{x_1, x_2, \cdots,\}$ we have

$$
p(x_i) = F(x_i) - F(x_{i-1}). \tag{3.1}
$$

Example 3.7
If the distribution function of X is given by

$$
F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{16} & 0 \le x < 1 \\ \frac{5}{16} & 1 \le x < 2 \\ \frac{11}{16} & 2 \le x < 3 \\ \frac{15}{16} & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}
$$

find the pmf of X.

Solution.
Using 3.1, we get $p(0) = \frac{1}{16}, p(1) = \frac{1}{4}, p(2) = \frac{3}{8}, p(3) = \frac{1}{4}$, and $p(4) = \frac{1}{16}$ and

0 otherwise ■

The Survival Distribution Function of X

The **survival function** (abbreviated SDF), also known as a **reliability function** is a property of any random variable that maps a set of events, usually associated with mortality or failure of some system, onto time. It captures the probability that the system will survive beyond a specified time. Thus, we define the survival distribution function by

$$S(x) = \Pr(X > x) = 1 - F(x).$$

It follows from Theorem 3.1, that any random variable satisfies the properties: $S(-\infty) = 1,\ \ S(\infty) = 0$, $S(x)$ is right-continuous, and that $S(x)$ is nonincreasing.

For a discrete random variable, the survival function is given by

$$S(x) = \sum_{t>x} p(t)$$

and for a continuous random variable, we have

$$S(x) = \int_x^\infty f(t)dt.$$

Note that $S'(t) = -f(t)$.

Remark 3.1

For a discrete random variable, the survival function need not be left-continuous, that is, it is possible for its graph to jump down. When it jumps, the value is assigned to the bottom of the jump.

Example 3.8 ‡

For watches produced by a certain manufacturer:

(i) Lifetimes follow a single-parameter Pareto distribution with α¿ 1 and $\theta = 4$.

(ii) The expected lifetime of a watch is 8 years.

Calculate the probability that the lifetime of a watch is at least 6 years.

Solution.

From Table C, we have

$$E(X) = \frac{\alpha\theta}{\alpha - 1} = \frac{4\alpha}{\alpha - 1} = 8 \Longrightarrow \alpha = 2.$$

Also, from Table C, we have

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha} \Longrightarrow F(6) = 0.555.$$

Hence, $S(6) = 1 - F(6) = 1 - 0.555 = 0.444$ ■

The Hazard Rate Function

The **hazard rate function**, also known as the **force of mortality** or the **failue rate function**, is defined to be the ratio of the density and the survival functions:

$$h(x) = h_X(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)}.$$

Example 3.9

Show that

$$h(x) = -\frac{S'(x)}{S(x)} = -\frac{d}{dx}[\ln S(x)]. \tag{3.2}$$

Solution.

The equation follows from $f(x) = -S'(x)$ and $\frac{d}{dx}[\ln S(x)] = \frac{S'(x)}{S(x)}$ ■

Example 3.10

Find the hazard rate function of a random variable with pdf given by $f(x) = e^{-ax}, a > 0$.

Solution.

We have

$$h(x) = \frac{f(x)}{S(x)} = \frac{ae^{-ax}}{e^{-ax}} = a \ ■$$

Example 3.11

Let X be a random variable with support $[0, \infty)$. Show that

$$S(x) = e^{-\Lambda(x)}$$

where

$$\Lambda(x) = \int_0^x h(s)ds.$$

$\Lambda(x)$ is called the **cumulative hazard function**

Solution.
Integrating equation (3.2) from 0 to x, we have

$$\int_0^x h(s)ds = -\int_0^x \frac{d}{ds}[\ln S(s)]ds = \ln S(0) - \ln S(x) = \ln 1 - \ln S(x) = -\ln S(x).$$

Now the result follows upon exponentiation ■

Some Additional Results
For a given event A, we define the indicator of A by

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

Let X and Y be two random variables. It is proven in advanced probability theory that the conditional expectation $E(X|Y)$ is given by the formula

$$E(X|Y) = \frac{E(XI_Y)}{\Pr(Y)}. \tag{3.3}$$

Example 3.12
Let X and Y be two random variables. Find a formula of $E[(X-d)^k|X > d]$ in the
(a) discrete case
(b) continuous case.

Solution.
(a) We have

$$E[(X-d)^k|X > d] = \frac{\sum_{x_j > d}(x-d)^k p(x_j)}{\Pr(X > d)}.$$

(b) We have

$$E[(X-d)^k|X > d] = \frac{1}{\Pr(X > d)}\int_d^\infty (x-d)^k f_X(x)dx \ ■$$

If $\Omega = A \cup B$ then for any random variable X on Ω, we have by the double expectation property[1]

$$E(X) = E(X|A)\Pr(A) + E(X|B)\Pr(B).$$

[1] See Section 37 of [2]

Probabilities can be found by conditioning:

$$\Pr(A) = \sum_y \Pr(A|Y=y)\Pr(Y=y)$$

in the discrete case and

$$\Pr(A) = \int_{-\infty}^{\infty} \Pr(A|Y=y) f_Y(y)dy$$

in the continuous case.

Weighted mean and variance
Given a set of data $x_1, x_2, \cdots, x_n$ with weights $w_1, w_2, \cdots, w_n$. The weighted mean is given by

$$\overline{X} = \frac{w_1x_1 + \cdots + x_nw_n}{w_1 + \cdots + w_n}$$

and the weighted variance is

$$\text{Var}(X) = \frac{1}{\sum_{i=1}^n w_i - 1} \sum_{i=1}^n w_i(x_i - \overline{X})^2.$$

Example 3.13
You are given the following information

x_i	0	1	2	3	4	5
w_i	512	307	123	41	11	6

Determine the weighted variance.

Solution.
The weighted mean is

$$\overline{X} = \frac{0(512) + 1(307) + 2(123) + 3(41) + 4(11) + 5(6)}{512 + 307 + 123 + 41 + 11 + 6} = 0.75.$$

The weighted variance is

$$\begin{aligned}\text{Var}(X) =& \frac{1}{1000-1}[512(0-0.75)^2 + 307(1-0.75)^2 + 123(2-0.75)^2 + 41(3-0.75)^2 \\ &+11(4-0.75)^2 + 6(5-0.75)^2] = 0.93243 \blacksquare\end{aligned}$$

Practice Problems

Problem 3.1
State whether the random variables are discrete, continuous, or mixed.

(a) In two tossing of a coin, let X be the number of heads in the two tosses.
(b) An urn contains one red ball and one green ball. Let X be the number of picks necessary in getting the first red ball.
(c) X is a random number in the interval $[4, 7]$.
(d) $X : \mathbb{R} \longrightarrow \mathbb{R}$ such that $X(s) = s$ if s is irrational and $X(s) = 1$ if s is rational.

Problem 3.2
Toss a pair of fair dice. Let X denote the sum of the dots on the two faces.

Find the probability mass function.

Problem 3.3
Consider the random variable $X : \{S, F\} \longrightarrow \mathbb{R}$ defined by $X(S) = 1$ and $X(F) = 0$. Suppose that $p = \Pr(X = 1)$.

Find the probability mass function of X.

Problem 3.4 ‡
The loss due to a fire in a commercial building is modeled by a random variable X with density function

$$f(x) = \begin{cases} 0.005(20 - x) & 0 < x < 20 \\ 0 & \text{otherwise.} \end{cases}$$

Given that a fire loss exceeds 8, what is the probability that it exceeds 16 ?

Problem 3.5 ‡
The lifetime of a machine part has a continuous distribution on the interval $(0, 40)$ with probability density function f, where $f(x)$ is proportional to $(10 + x)^{-2}$.

Calculate the probability that the lifetime of the machine part is less than 6.

Problem 3.6 ‡
A group insurance policy covers the medical claims of the employees of a small company. The value, V, of the claims made in one year is described by

$$V = 100000Y$$

where Y is a random variable with density function

$$f(x) = \begin{cases} k(1-y)^4 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

What is the conditional probability that V exceeds 40,000, given that V exceeds 10,000?

Problem 3.7 ‡
An insurance policy pays for a random loss X subject to a deductible of C, where $0 < C < 1$. The loss amount is modeled as a continuous random variable with density function

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Given a random loss X, the probability that the insurance payment is less than 0.5 is equal to 0.64 .

Calculate C.

Problem 3.8
Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} \alpha x e^{-x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Determine the value of α.

Problem 3.9
Consider the following probability distribution

x	1	2	3	4
$p(x)$	0.25	0.5	0.125	0.125

Find a formula for $F(x)$ and sketch its graph.

Problem 3.10
Find the distribution functions corresponding to the following density functions:

$$\begin{aligned}(a)f(x) =&\frac{a-1}{(1+x)^a}, && 0 < x < \infty, 0 \text{ otherwise.}\\ (b)f(x) =&k\alpha x^{\alpha-1}e^{-kx^\alpha}, && k > 0, \alpha, 0 < x < \infty, 0 \text{ otherwise.}\end{aligned}$$

Problem 3.11
Let X be a random variable with pmf

$$p(n) = \frac{1}{3}\left(\frac{2}{3}\right)^n, \quad n = 0, 1, 2, \cdots .$$

Find a formula for $F(n)$.

Problem 3.12
Given the pdf of a continuous random variable X.

$$f(x) = \begin{cases} \frac{1}{5}e^{-\frac{x}{5}} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $\Pr(X > 10)$.
(b) Find $\Pr(5 < X < 10)$.
(c) Find $F(x)$.

Problem 3.13
A random variable X has the cumulative distribution function

$$F(x) = \frac{e^x}{e^x + 1}.$$

Find the probability density function.

Problem 3.14
Consider an age-at-death random variable X with survival distribution defined by

$$S(x) = \frac{1}{10}(100 - x)^{\frac{1}{2}}, \; 0 \leq x \leq 100.$$

(a) Explain why this is a suitable survival function.
(b) Find the corresponding expression for the cumulative probability function.
(c) Compute the probability that a newborn with survival function defined above will die between the ages of 65 and 75.

Problem 3.15
Consider an age-at-death random variable X with survival distribution defined by

$$S(x) = e^{-0.34x}, \ x \geq 0.$$

Compute $\Pr(5 < X < 10)$.

Problem 3.16
Consider an age-at-death random variable X with survival distribution $S(x) = 1 - \frac{x^2}{100}$ for $x \geq 0$.

Find $F(x)$.

Problem 3.17
Consider an age-at-death random variable X. The survival distribution is given by $S(x) = 1 - \frac{x}{100}$ for $0 \leq x \leq 100$ and 0 for $x > 100$.

(a) Find the probability that a person dies before reaching the age of 30.
(b) Find the probability that a person lives more than 70 years.

Problem 3.18
An age-at-death random variable has a survival function

$$S(x) = \frac{1}{10}(100 - x)^{\frac{1}{2}}, \ 0 \leq x \leq 100$$

and 0 otherwise.

Find the hazard rate function of this random variable.

Problem 3.19
Consider an age-at-death random variable X with force of mortality $h(x) = \mu > 0$.

Find $S(x), f(x)$, and $F(x)$.

Problem 3.20
Let

$$F(x) = 1 - \left(1 - \frac{x}{120}\right)^{\frac{1}{6}},\ 0 \le x \le 120.$$

Find $h(40)$.

4 Raw and Central Moments

Several quantities can be computed from the pdf that describe simple characteristics of the distribution. These are called **moments.** The most common is the mean, the first moment about the origin, and the variance, the second moment about the mean. The mean is a measure of the centrality of the distribution and the variance is a measure of the spread of the distribution about the mean.

The n^{th} moment $= \mu'_n = E(X^n)$ of a random variable X is also known as the n^{th} **moment about the origin** or the n^{th} **raw moment.** For a continuous random variable X we have

$$\mu'_n = \int_{-\infty}^{\infty} x^n f(x)dx$$

and for a discrete random variable we have

$$\mu'_n = \sum_x x^n p(x).$$

By contrast, the quantity $\mu_n = E[(X - E(X))^n]$ is called the n^{th} **central moment** of X or the n^{th} **moment about the mean.** For a continuous random variable X we have

$$\mu_n = \int_{-\infty}^{\infty} (x - E(X))^n f(x)dx$$

and for a discrete random variable we have

$$\mu_n = \sum_x (x - E(X))^n p(x).$$

Note that $\text{Var}(X)$ is the second central moment of X.

Example 4.1
Let X be a continuous random variable with pdf given by $f(x) = \frac{3}{8}x^2$ for $0 < x < 2$ and 0 otherwise. Find the second central moment of X.

Solution.
We first find the mean of X. We have

$$E(X) = \int_0^2 xf(x)dx = \int_0^2 \frac{3}{8}x^3 dx = \frac{3}{32}x^4\Big|_0^2 = 1.5.$$

The second central moment is

$$\begin{aligned}\mu_2 =& \int_0^2 (x-1.5)^2 f(x)dx \\ =& \int_0^2 \frac{3}{8}x^2(x-1.5)^2 dx \\ =& \frac{3}{8}\left[\frac{x^5}{5} - 0.75x^4 + 0.75x^3\right]_0^2 = 0.15 \blacksquare\end{aligned}$$

The importance of moments is that they are used to define quantities that characterize the shape of a distribution. These quantities which will be discussed below are: skewness, kurtosis and coefficient of variation.

Departure from Normality: Coefficient of Skewness

The third central moment, μ_3, is called the **skewness** and is a measure of the symmetry of the pdf. A distribution, or data set, is symmetric if it looks the same to the left and right of the mean.

A measure of skewness is given by the **coefficient of skewness** γ_1 :

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{E(X^3) - 3E(X)E(X^2) + 2[E(X)]^3}{[E(X^2) - E(X)^2]^{\frac{3}{2}}}.$$

That is, γ_1 is the ratio of the third central moment to the cube of the standard deviation. Equivalently, γ_1 is the third central moment of the standardized variable

$$X^* = \frac{X - \mu}{\sigma}.$$

If γ_1 is close to zero then the distribution is symmetric about its mean such as the normal distribution. A positively skewed distribution has a "tail" which is pulled in the positive direction. A negatively skewed distribution has a "tail" which is pulled in the negative direction (see Figure 4.1).

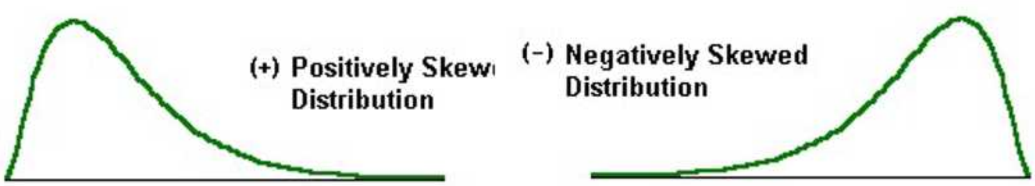


Figure 4.1

Example 4.2
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the coefficient of skewness of X.

Solution.
We first find the mean of X :

$$\mu = E(X) = 120\times\frac{1}{4}+122\times\frac{1}{12}+124\times\frac{1}{6}+150\times\frac{1}{12}+167\times\frac{1}{12}+245\times\frac{1}{3} = \frac{2027}{12}.$$

The second raw moment is

$$E(X^2) = 120^2\times\frac{1}{4}+122^2\times\frac{1}{12}+124^2\times\frac{1}{6}+150^2\times\frac{1}{12}+167^2\times\frac{1}{12}+245^2\times\frac{1}{3} = \frac{379325}{12}.$$

Thus, the variance of X is

$$\text{Var}(X) = \frac{379325}{12} - \frac{4108729}{144} = \frac{443171}{144}$$

and the standard deviation is

$$\sigma = \sqrt{\frac{443171}{144}} = 55.475908183.$$

The third central moment is

$$\begin{aligned}\mu_3 =& \left(120 - \frac{2027}{12}\right)^3 \times \frac{1}{4} + \left(122 - \frac{2027}{12}\right)^3 \times \frac{1}{12} + \left(124 - \frac{2027}{12}\right)^3 \times \frac{1}{6} \\ &+ \left(150 - \frac{2027}{12}\right)^3 \times \frac{1}{12} + \left(167 - \frac{2027}{12}\right)^3 \times \frac{1}{12} + \left(245 - \frac{2027}{12}\right)^3 \times \frac{1}{3} \\ &=93270.81134.\end{aligned}$$

Thus,

$$\gamma_1 = \frac{93270.81134}{55.475908183^3} = 0.5463016252 \blacksquare$$

Example 4.3
Let X be a random variable with density $f(x) = e^{-x}$ on $(0, \infty)$ and 0 otherwise. Find the coefficient of skewness of X.

Solution.
Since

$$E(X) = \int_0^\infty xe^{-x}dx = -e^{-x}(1+x)\Big|_0^\infty = 1$$
$$E(X^2) = \int_0^\infty x^2e^{-x}dx = -e^{-x}(x^2+2x+2)\Big|_0^\infty = 2$$
$$E(X^3) = \int_0^\infty x^3e^{-x}dx = -e^{-x}(x^3+3x^2+6x+6)\Big|_0^\infty = 6$$

we find

$$\gamma_1 = \frac{6-3(1)(2)+2(1)^3}{(2-1^2)^{\frac{3}{2}}} = 2 \blacksquare$$

<u>**Coefficient of Kurtosis**</u>
The fourth central moment, μ_4, is called the **kurtosis** and is a measure of peakedness/flatness of a distribution with respect to the normal distribution.

A measure of kurtosis is given by the **coefficient of kurtosis:**

$$\gamma_2 = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{E(X^4) - 4E(X^3)E(X) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[E(X^2) - (E(X))^2]^2}.$$

The coefficient of kurtosis of the normal distribution is 3. The condition $\gamma_2 < 3$ indicates that the distribution is flatter compared to the normal distribution, and the condition $\gamma_2 > 3$ indicates a higher peak (relative to the normal distribution) around the mean value.(See Figure 4.2)

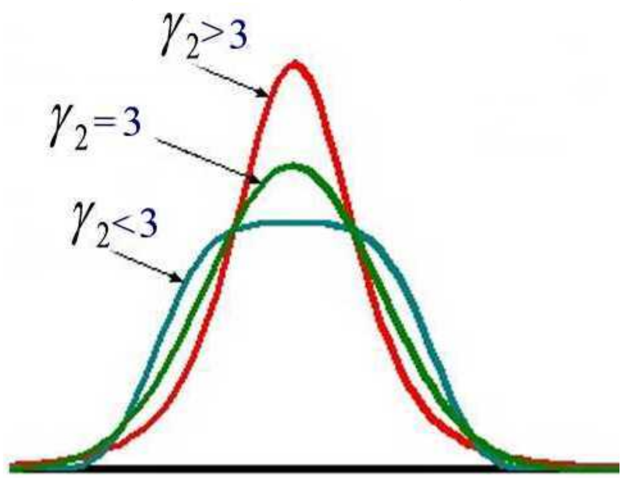


Figure 4.2

Example 4.4
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the coefficient of kurtosis of X.

Solution.
We first find the fourth central moment.

$$\begin{aligned}\mu_4=&\left(120-\frac{2027}{12}\right)^4\times\frac{1}{4}+\left(122-\frac{2027}{12}\right)^4\times\frac{1}{12}+\left(124-\frac{2027}{12}\right)^4\times\frac{1}{6}\\&+\left(150-\frac{2027}{12}\right)^4\times\frac{1}{12}+\left(167-\frac{2027}{12}\right)^4\times\frac{1}{12}+\left(245-\frac{2027}{12}\right)^4\times\frac{1}{3}\\=&13693826.62.\end{aligned}$$

Thus,

$$\gamma_2=\frac{13693826.62}{55.475908183^4}=1.44579641 \blacksquare$$

Example 4.5
Find the coefficient of kurtosis of the random variable X with density function $f(x)=1$ on $(0,1)$ and 0 elsewhere.

Solution.
Since

$$E(X^k)=\int_0^1 x^k dx=\frac{1}{k+1}.$$

we obtain,

$$\gamma_2=\frac{\frac{1}{5}-4\left(\frac{1}{4}\right)\left(\frac{1}{2}\right)+6\left(\frac{1}{3}\right)\left(\frac{1}{2}\right)^2-3\left(\frac{1}{2}\right)^4}{\left(\frac{1}{3}-\frac{1}{4}\right)^2}=\frac{9}{5} \blacksquare$$

Coefficient of Variation
Some combinations of the raw moments and central moments that are also commonly used. One such combination is the **coefficient of variation**, denoted by $CV(X)$, of a random variable X which is defined as the ratio of the standard deviation to the mean:

$$CV(X)=\frac{\sigma}{\mu},\quad \mu=\mu_1'=E(X).$$

It is an indication of the size of the standard deviation relative to the mean, for the given random variable.

Often the coefficient of variation is expressed as a percentage. Thus, it expresses the standard deviation as a percentage of the sample mean and it is unitless. Statistically, the coefficient of variation is very useful when comparing two or more sets of data that are measured in different units of measurement.

Example 4.6
Let X be a random variable with mean of 4 meters and standard deviation of 0.7 millimeters. Find the coefficient of variation of X.

Solution.
The coefficient of variation is

$$CV(X) = \frac{0.7}{4000} = 0.0175\% \blacksquare$$

Example 4.7
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the coefficient of variation of X.

Solution.
We know that $\mu = \frac{2027}{12} = 168.9166667$ and $\sigma = 55.47590818$. Thus, the coefficient of variation of X is

$$CV(X) = \frac{55.47590818}{168.9166667} = 0.3284217754 \blacksquare$$

Example 4.8
Find the coefficient of variation of the random variable X with density function $f(x) = e^{-x}$ on $(0, \infty)$ and 0 otherwise.

Solution.
We have

$$\mu = E(X) = \int_0^\infty xe^{-x}dx = -e^{-x}(1+x)\Big|_0^\infty = 1$$

and

$$\sigma = (E(X^2) - (E(X))^2)^{\frac{1}{2}} = (2-1)^{\frac{1}{2}} = 1.$$

Hence,

$$CV(X) = 1 \blacksquare$$

Practice Problems

Problem 4.1
Consider n independent trials. Let X denote the number of successes in n trials. We call X a **binomial random variable.** Its pmf is given by

$$p(r) = C(n,r)p^r(1-p)^{n-r}$$

where p is the probability of a success.

(a) Show that $E(X) = np$ and $E[X(X-1)] = n(n-1)p^2$. Hint: $(a+b)^n = \sum_{k=0}^{n} C(n,k)a^k b^{n-k}$.
(b) Find the variance of X.

Problem 4.2
A random variable X is said to be a **Poisson** random variable with parameter $\lambda > 0$ if its probability mass function has the form

$$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots$$

where λ indicates the average number of successes per unit time or space.

(a) Show that $E(X) = \lambda$ and $E[X(X-1)] = \lambda^2$.
(b) Find the variance of X.

Problem 4.3
A **geometric** random variable with parameter $p, 0 < p < 1$ has a probability mass function

$$p(n) = p(1-p)^{n-1}, \quad n = 1, 2, \cdots .$$

(a) By differentiating the geometric series $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ twice and using $x = 1-p$ is each equation, show that

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} = p^{-2} \text{ and } \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-2} = 2p^{-3}.$$

(b) Show that $E(X) = \frac{1}{p}$ and $E[X(X-1)] = 2p^{-2}(1-p)$.
(c) Find the variance of X.

Problem 4.4
A **normal** random variable with parameters μ and σ^2 has a pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Show that $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$. Hint: $E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = 0$ where Z is the standard normal distribution with parameters (0,1).

Problem 4.5
An **exponential** random variable with parameter $\lambda > 0$ is a random variable with pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

(a) Show that $E(X) = \frac{1}{\lambda}$ and $E(X^2) = \frac{2}{\lambda^2}$.
(b) Find $\text{Var}(X)$.

Problem 4.6
A **Gamma** random variable with parameters $\alpha > 0$ and $\theta > 0$ has a pdf

$$f(x) = \begin{cases} \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\theta}} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

where

$$\int_0^\infty e^{-y} y^{\alpha-1} dy = \Gamma(\alpha) = \alpha\Gamma(\alpha - 1).$$

Show:
(a) $E(X) = \alpha\theta$
(b) $Var(X) = \alpha\theta^2$.

Problem 4.7
Let X be a continuous random variable with pdf given by $f(x) = \frac{3}{8}x^2$ for $0 \leq x \leq 2$ and 0 otherwise.

Find the third raw moment of X.

Problem 4.8
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the fourth raw moment.

Problem 4.9
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the fifth central moment of X.

Problem 4.10
Compute the coefficient of skewness of a uniform random variable, X, on $[0,1]$.

Problem 4.11
Let X be a random variable with density $f(x) = e^{-x}$ and 0 otherwise.

Find the coefficient of kurtosis.

Problem 4.12
A random variable X has the following pmf:

x	120	122	124	150	167	245
$p(x)$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$

Find the coefficient of variation.

Problem 4.13
Let X be a continuous random variable with density function $f(x) = Ax^b e^{-Cx}$ for $x \geq 0$ and 0 otherwise. The parameters $A, B,$ and C satisfy

$$A = \frac{1}{\int_0^\infty x^B e^{-Cx} dx}, \; B \geq -\frac{1}{2}, \; C > 0.$$

Show that

$$E(X^n) = \frac{B+r}{C} E(X^{n-1}).$$

Problem 4.14
Let X be a continuous random variable with density function $f(x) = Ax^b e^{-Cx}$ for $x \geq 0$ and 0 otherwise. The parameters $A, B,$ and C satisfy

$$A = \frac{1}{\int_0^\infty x^B e^{-Cx} dx}, \; B \geq -\frac{1}{2}, \; C > 0.$$

Find the first and second raw moments.

Problem 4.15
Let X be a continuous random variable with density function $f(x) = Ax^b e^{-Cx}$ for $x \geq 0$ and 0 otherwise. The parameters $A, B,$ and C satisfy

$$A = \frac{1}{\int_0^\infty x^B e^{-Cx} dx}, \; B \geq -\frac{1}{2}, \; C > 0.$$

Find the coefficient of skewness.

Problem 4.16
Let X be a continuous random variable with density function $f(x) = Ax^be^{-Cx}$ for $x \geq 0$ and 0 otherwise. The parameters A, B, and C satisfy

$$A = \frac{1}{\int_0^\infty x^B e^{-Cx} dx}, \; B \geq -\frac{1}{2}, \; C > 0.$$

Find the coefficient of kurtosis.

Problem 4.17
You are given: $E(X) = 2, \; CV(X) = 2$, and $\mu_3' = 136$. Calculate γ_1.

Problem 4.18
Let X be a random variable with pdf $f(x) = 0.005x$ for $0 \leq x \leq 20$ and 0 otherwise.

(a) Find the cdf of X.
(b) Find the mean and the variance of X.
(c) Find the coefficient of variation.

Problem 4.19
Let X be the Gamma random variable with pdf $f(x) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\theta}}$ for $x > 0$ and 0 otherwise. Suppose $E(X) = 8$ and $\gamma_1 = 1$.

Find the variance of X.

Problem 4.20
Let X be a Pareto random variable in one parameter and with a pdf $f(x) = \frac{a}{x^{a+1}}, x \geq 1$ and 0 otherwise.

(a) Show that $E(X^k) = \frac{a}{a-k}$ for $0 < k < a$.
(b) Find the coefficient of variation of X.

Problem 4.21
For the random variable X you are given:
(i) $E(X) = 4$
(ii) $\text{Var}(X) = 64$
(iii) $E(X^3) = 15$.

Calculate the skewness of X.

Problem 4.22
Let X be a Pareto random variable with two parameters α and θ, i.e., X has the pdf

$$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \ \alpha > 1, \theta > 0, x > 0$$

and 0 otherwise.

Calculate the mean and the variance of X.

Problem 4.23
Let X be a Pareto random variable with two parameters α and θ, i.e., X has the pdf

$$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \ \alpha > 1, \theta > 0, x > 0$$

and 0 otherwise.

Calculate the coefficient of variation.

Problem 4.24
Let X be the Gamma random variable with pdf $f(x) = \frac{1}{\theta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\theta}}$ for $x > 0$ and 0 otherwise. Suppose $CV(X) = 1$.

Determine γ_1.

Problem 4.25
You are given the following times of first claim for five randomly selected auto insurance policies observed from time $t = 0$:

$$1 \quad 2 \quad 3 \quad 4 \quad 5$$

Calculate the kurtosis of this sample.

5 Empirical Models, Excess and Limited Loss variables

Empirical models are those that are based entirely on data. Consider a statistical model that results in a sample Ω of size n. Data points in the sample are assigned equal probability of $\frac{1}{n}$. Let X be a random variable defined on Ω. We refer to this model as an **empirical model.**

Example 5.1 ‡
You are given the following for a sample of five observations from a bivariate distribution:
(i)

x	y
1	4
2	2
4	3
5	6
6	4

(ii) $\overline{x} = 3.6$ and $\overline{y} = 3.8$.
A is the covariance of the empirical distribution F_e as defined by these five observations. B is the maximum possible covariance of an empirical distribution with identical marginal distributions to F_e.
Determine $B - A$.

Solution.
We have

$$\begin{aligned}A =&\text{Cov}(X,Y) = E(XY) - E(X)E(Y)\\ =&\frac{4+4+12+30+24}{5} - 3.6(3.8) = 1.12.\end{aligned}$$

Now, since $E(X)$ and $E(Y)$ are fixed, we want to create a new bivariate distribution from the given one with maximum $E(XY)$. Clearly, this occurs if largest values of X are paired with largest values of Y. Hence, the following bivariate distribution has the same marginal distributions as the original bivariate distribution:

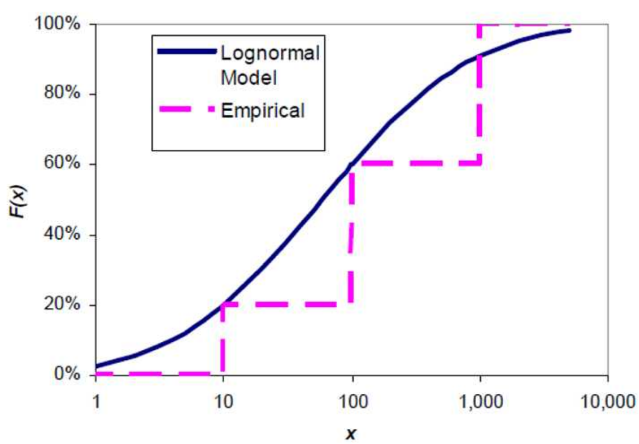
x	y
6	6
5	4
4	4
2	3
1	2

The covariance of this distribution is

$$B = \frac{36+20+16+6+2}{5} - 3.6(3.8) = 2.32.$$

The final answer is $B - A = 2.32 - 1.12 = 1.2$ ∎

Example 5.2 ‡
You are given the following graph of cumulative distribution functions:



Determine the difference between the mean of the lognormal model and the mean of the data.

Solution.
The empirical distribution is given by

$$\begin{aligned}
p(10) =& F(10) - F(0) = 0.20 - 0 = 0.20\\
p(100) =& F(100) - F(10) = 0.60 - 0.2 = 0.4\\
p(1000) =& F(1000) - F(100) = 1 - 0.6 = 0.4.
\end{aligned}$$

The mean of the data is

$$0.2(10) + 0.4(100) + 0.4(1000) = 442.$$

Now, from the graph we see that the 20^{th} and 60^{th} percentiles of the lognormal distribution are 10 and 100 respectively. That is,

$$0.2 = \Phi\left(\frac{\ln 10 - \mu}{\sigma}\right) \text{ and } 0.6 = \Phi\left(\frac{\ln 100 - \mu}{\sigma}\right)$$

Using the table of standard normal distribution, we find

$$-0.84 = \frac{\ln 10 - \mu}{\sigma} \text{ and } 0.25 = \frac{\ln 100 - \mu}{\sigma}$$

Solving this system, we find $\mu = 4.0771$ and $\sigma = 2.1125$. Thus, the mean of the lognormal distribution is (Table C)

$$e^{\mu + 0.5\sigma^2} = e^{4.0771 + 0.5(2.1125^2)} = 549.18.$$

The final answer is $549.18 - 442 = 107.18$ ■

Example 5.3
In a fitness club monthly new memberships are recorded in the table below.

January	February	March	April	May	June
100	102	84	84	100	100

July	August	September	October	November	December
67	45	45	45	45	93

Use an empirical model to construct a discrete probability mass function for X, the number of new memberships per month.

Solution.
The sample under consideration has 12 data points which are the months of the year. For our empirical model, each data point is assigned a probability of $\frac{1}{12}$. The pmf of the random variable X is given by:

x	45	67	84	93	100	102
$p(x)$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{12}$

■

Excess Loss Random Variable

Consider an insurance policy with a deductible. The insurer's interest are usually in the losses that resulted in a payment, and the actual amount paid by the insurance. The insurer pays the insuree the amount of the loss that was in excess of the deductible[2]. Any amount of losses that are below the deductible are ignored by the insurance since they do not result in an insurance payment being made. Hence, the insurer would be considering the conditional distribution of amount paid, given that a payment was actually made. This is what is referred to as the excess loss random variable.

Let X be the random variable representing the amount of a single loss. In insurance terms, X is known as the **loss random variable** or the **severity random variable**. For a given threshold d such that $\Pr(X > d) > 0$, the random variable

$$Y^P = (X-d|X>d) = \begin{cases} \text{undefined}, & X \le d \\ X-d, & X > d \end{cases}$$

is called the **excess loss variable**, the **cost per payment**, or the **left truncated and shifted variable**. It stands for the amount paid by the insurance which is also known as **claim amount**.

We can find the k^{th} moment of the excess loss variable as follows. For a continuous distribution with probability density function $f(x)$ and cumulative distribution function $F(x)$, we have[3]

$$\begin{aligned} e_X^k(d) =& E[(X-d)^k|X>d] = \frac{1}{\Pr(X>d)}\int_d^\infty (x-d)^k f(x)dx \\ =& \frac{1}{1-F(d)}\int_d^\infty (x-d)^k f(x)dx \end{aligned}$$

provided that the improper integral is convergent.

For a discrete distribution with probability density function $p(x)$ and a cumulative distribution function $F(x)$, we have

$$e_X^k(d) = \frac{1}{1-F(d)}\sum_{x_j>d}(x_j-d)^k p(x_j)$$

provided that the sum is convergent.

[2] The deductible is referred to as **ordinary deductible**. Another type of deductible is called **franchise deductible** and will be discussed in Section 32.

[3] See (3.3)

When $k = 1$, the expected value

$$e_X(d) = E(Y^P) = E(X - d|X > d)$$

is called the **mean excess loss function**. Other names used have been **mean residual life function** and **complete expectation of life**.

If X denotes payment, then $e_X(d)$ stands for the expected amount paid given that there has been a payment in excess of the deductible d. If X denotes age at death, then $e_X(d)$ stands for the expected future lifetime given that the person is alive at age d.

Example 5.4
Show that for a continuous random variable X, we have

$$e_X(d) = \frac{1}{1 - F(d)} \int_d^\infty (1 - F(x))dx = \frac{1}{S(d)} \int_d^\infty S(x)dx.$$

Solution.
Using integration by parts with $u = x - d$ and $v' = f(x)$, we have

$$\begin{aligned}
e_X(d) =& \frac{1}{1 - F(d)} \int_d^\infty (x - d)f(x)dx \\
=& -\frac{(x - d)(1 - F(x))}{1 - F(d)}\bigg|_d^\infty + \frac{1}{1 - F(d)} \int_d^\infty (1 - F(x))dx \\
=& \frac{1}{1 - F(d)} \int_d^\infty (1 - F(x))dx = \frac{1}{S(d)} \int_d^\infty S(x)dx.
\end{aligned}$$

Note that

$$0 \le xS(x) = x \int_x^\infty f(t)dt \le \int_x^\infty tf(t)dt \Longrightarrow \lim_{x\to\infty} xS(x) = 0 \blacksquare$$

Example 5.5
Let X be an excess loss random variable with pdf given by $f(x) = \frac{1}{3}(1 + 2x)e^{-x}$ for $x > 0$ and 0 otherwise. Calculate the mean excess loss function with deductible amount x.

Solution.
The cdf of X is given by

$$F(x) = \int_0^x \frac{1}{3}(1 + 2t)e^{-t}dt = -\frac{1}{3}e^{-t}(2t + 3)\bigg|_0^x = 1 - \frac{e^{-x}(2x + 3)}{3}$$

where we used integration by parts. The mean excess loss function is

$$e_X(x) = \frac{1}{\frac{e^{-x}(2x+3)}{3}} \int_x^\infty \left[1 - \frac{e^{-t}(2t+3)}{3}\right] dt = \frac{\frac{e^{-x}(2x+5)}{3}}{\frac{e^{-x}(2x+3)}{3}} = \frac{2x+5}{2x+3} \blacksquare$$

Example 5.6 ‡
For an industry-wide study of patients admitted to hospitals for treatment of cardiovascular illness in 1998, you are given:
(i)

Duration In Days	Number of Patients Remaining Hospitalized
0	4,386,000
5	1,461,554
10	486,739
15	161,801
20	53,488
25	17,384
30	5,349
35	1,337
40	0

(ii) Discharges from the hospital are uniformly distributed between the durations shown in the table.
Calculate the mean residual time remaining hospitalized, in days, for a patient who has been hospitalized for 21 days.

Solution.
Let X denote the number of days at the hospital measured from time 0. We are asked to find $E(X-21|X>21)$ which by Example 5.4 can be expressed as

$$E(X-21|X>21) = \int_{21}^\infty \frac{S_X(x)}{S_X(21)} dx.$$

In life contingency theory, the assumption that discharges are uniform on a given interval means that the graph of survival function is a linear function on that interval. Hence, the graph of $\frac{S_X(x)}{S_X(21)}$ consists of line segments on the intervals $[0,5], [5,10], etc.$ Thus, $E(X-21|X>21)$ is just the area under the graph from 21 to 40. The area under the graph is just the sum of areas of the trapezoids with bases $[21, 25], [25, 30], [30, 35]$, and $[35, 40]$. For instance, the area under the first trapezoid is

$$\frac{1}{2}(25-21)\left(\frac{S_X(21)}{S_X(21)}+\frac{S_X(25)}{S_X(21)}\right).$$

Again, from the theory of life contingencies (see [3]), we have

$$S_X(x)=\frac{\ell_x}{\ell_0}$$

where ℓ_0 is the number of patients at the hospital at time 0 and ℓ_x is the expected number of patients in the hospital at time x. Thus,

$$\frac{S_X(x)}{S_X(21)}=\frac{\ell_x}{\ell_{21}}.$$

By linear interpolation, we have

$$\ell_{21}=53,488-\frac{53488-17384}{25-20}(21-20)=46,267.2.$$

Now, we go back to finding the area of the first trapezoid mentioned above, we find

$$\frac{1}{2}(25-21)\left(\frac{S_X(21)}{S_X(21)}+\frac{S_X(25)}{S_X(21)}\right)=\frac{1}{2}(4)\left(1+\frac{17384}{46,267.20}\right)=2.751.$$

We repeat the same calculation with the remaining three trapezoids, we find

$$E(X-21|X>21)=2.751+1.228+0.361+0.072=4.412 \blacksquare$$

Example 5.7
Show that

$$F_{Y^P}(y)=\frac{F_X(y+d)-F_X(d)}{1-F_X(d)}.$$

Solution.
We have

$$\begin{aligned}F_{Y^P}(y)=&\Pr(Y^P\le y)=\Pr(X-d\le y|X>d)=\frac{\Pr(d<X\le y+d)}{\Pr(X>d)}\\=&\frac{F_X(y+d)-F_X(d)}{1-F_X(d)} \blacksquare\end{aligned}$$

Left-Censored and Shifted Random Variable

Note that in the excess loss situation, losses below or at the value d are not recorded in any way, that is, the excess loss random variable is left-truncated and it is shifted because of a number subtracted from X. However, when small losses at or below d are recorded as 0 then we are led to a new random variable which we call a **left-censored and shifted** random variable or the **cost per loss** random variable. It is defined as

$$Y^L = (X-d)_+ = \begin{cases} 0, & X \le d \\ X-d, & X > d. \end{cases}$$

The k^{th}-moment of this random variable is given by

$$E[(X-d)_+^k] = \int_d^\infty (x-d)^k f(x)dx$$

in the continuous case and

$$E[(X-d)_+^k] = \sum_{x_j > d} (x_j - d)^k p(x_j)$$

in the discrete case. Note the relationship between the moments of Y^P and Y^L given by

$$E[(X-d)_+^k] = e_X^k(d)[1-F(d)] = e_X^k(d)S(d).$$

Setting $k=1$ and using the formula for $e_X(d)$ we see that

$$E(Y^L) = \int_d^\infty S(x)dx.$$

This expected value is sometimes called the **stop loss premium.**[4]

We can think of the excess loss random variable as of a random variable that exists only if a payment has been made. Alternatively, the left censored and shifted random variable is equal to zero every time a loss does not produce payment.

Example 5.8

For a house insurance policy, the loss amount (expressed in thousands), in the event of a fire, is being modeled by a distribution with density

$$f(x) = \frac{3}{56}x(5-x), \; 0 < x < 4.$$

[4] See Section 39.

For a policy with a deductible amount of \$1,000, calculate the expected amount per loss.

Solution.
We first calculate the survival function:

$$S(x) = 1 - F(x) = 1 - \int_0^x \frac{3}{56}t(5-t)dt = 1 - \frac{3}{56}\left(\frac{5}{2}x^2 - \frac{1}{3}x^3\right).$$

Thus,

$$E(Y^L) = \int_1^4 \left[1 - \frac{3}{56}\left(\frac{5}{2}x^2 - \frac{1}{3}x^3\right)\right] dx = 1325.893 \blacksquare$$

Note that Y^L is a mixed random variable. The discrete part is represented by $Y^L = 0$ for $X \le d$ where

$$p_{Y^L}(0) = \Pr(Y^L = 0) = \Pr(X \le d) = F_X(d) = F_X(d) - F_X(d^-).$$

The continuous part of the distribution of Y^L is given by

$$f_{Y^L}(y) = f_X(y + d)$$

for $y > 0$ since $Y^L = y$ for $X = y + d$.

Limited Loss Variable or Policy Limits

Many insurance policies are covered up to a certain limit which we refer to as **policy limit**. Let's say the limit is u. That is, the insurer covers all losses up to u fully but pays u for losses greater than u. Thus, if X is a loss random variable then the amount paid by the insurer is $X \wedge u$. We call $X \wedge u$ the **limited loss variable** and is defined by

$$X \wedge u = \min(X, u) = \begin{cases} X, & X \le u \\ u, & X > u. \end{cases}$$

Notice that the distribution of X is censored on the right and that is why the limit loss variable is also known as the **right-censored random variable.**

The expected value of the limited loss value is $E(X \wedge u)$ and is called the **limited expected value**.

For a discrete distribution with probability density function $p(x_j)$ and a cumulative distribution function $F(x_j)$ for all relevant index values j, the k^{th}moment of the limited loss random variable is given by

$$E[(X \wedge u)^k] = \sum_{x_j \le u} x_j^k p(x_j) + u^k[1 - F(u)].$$

For a continuous distribution with probability density function $f(x)$ and cumulative distribution function $F(x)$, the k^{th} moment is given by

$$E[(X \wedge u)^k] = \int_{-\infty}^{u} x^k f(x)dx + \int_{u}^{\infty} u^k f(x)dx = \int_{-\infty}^{u} x^k f(x)dx + u^k[1 - F(u)].$$

Using integration by parts, we can derive an alternative formula for the k^{th}moment:

$$\begin{aligned}
E[(X \wedge u)^k] &= \int_{-\infty}^{u} x^k f(x)dx + u^k[1 - F(u)] \\
&= \int_{-\infty}^{0} x^k f(x)dx + \int_{0}^{u} x^k f(x)dx + u^k[1 - F(u)] \\
&= x^k F(x)\Big|_{-\infty}^{0} - \int_{-\infty}^{0} kx^{k-1} F(x)dx \\
&\quad - x^k S(x)\Big|_{0}^{u} + \int_{0}^{u} kx^{k-1} S(x)dx + u^k S(u) \\
&= -\int_{-\infty}^{0} kx^{k-1} F(x)dx + \int_{0}^{u} kx^{k-1} S(x)dx.
\end{aligned}$$

Note that for $x < 0$ and k odd we have

$$\int_{-\infty}^{x} t^k f(t)dt \le x^k \int_{-\infty}^{x} f(t)dt = x^k F(x) \le 0$$

so that

$$\lim_{x \to -\infty} x^k F(x) = 0.$$

A similar argument for $x < 0$ and k even.
In particular, for $k = 1$, we obtain

$$E(X \wedge u) = -\int_{-\infty}^{0} F(x)dx + \int_{0}^{u} S(x)dx.$$

One can easily see that

$$(X - d)_+ + X \wedge u = \begin{cases} 0, & X \le d \\ X - d, & X > d. \end{cases} + \begin{cases} X, & X \le u \\ u, & X > u. \end{cases} = X.$$

That is, buying one policy with a deductible d and another one with a limit d is equivalent to purchasing full cover.

Remark 5.1
Usually the random variable X is non-negative (for instance, when X represents a loss or a time until death), and the lower integration limit $-\infty$ is replaced by 0. Thus, for $X \geq 0$, we whave

$$E(X \wedge u) = \int_0^u S(x)dx.$$

Example 5.9
A continuous random variable X has a pdf $f(x) = 0.005x$ for $0 \leq x \leq 20$ and 0 otherwise. Find the mean and the variance of $X \wedge 10$.

Solution.
The cdf of X is

$$F(x) = \int_{-\infty}^x 0.005tdt = \int_0^x 0.005tdt = \begin{cases} 0, & x < 0 \\ 0.0025x^2, & 0 \leq x \leq 20 \\ 1, & x > 20. \end{cases}$$

Thus,

$$E(X \wedge 10) = \int_0^{10} [1 - F(x)]dx = \int_0^{10} [1 - 0.0025x^2]dx = \frac{55}{6}.$$

Now,

$$E[(X \wedge 10)^2] = \int_0^{10} x^2(0.005x)dx + 10^2[1 - F(10)] = 87.5.$$

Finally,

$$\mathrm{Var}(X \wedge 10) = 87.5 - \left(\frac{55}{6}\right)^2 = \frac{125}{36} \blacksquare$$

Example 5.10 ‡
The unlimited severity distribution for claim amounts under an auto liability insurance policy is given by the cumulative distribution:

$$F(x) = 1 - 0.8e^{-0.02x} - 0.2e^{-0.001x}, \; x \geq 0.$$

The insurance policy pays amounts up to a limit of 1000 per claim. Calculate the expected payment under this policy for one claim.

Solution.
We are asked to find the limited expected value $E(X \wedge 1000)$. We have

$$E(X \wedge 1000) = \int_0^{1000} S(x)dx = \int_0^{1000} [0.8e^{-0.02x} + 0.2e^{-0.001x}]dx = 166.4 \blacksquare$$

Example 5.11 ‡
A health plan implements an incentive to physicians to control hospitalization under which the physicians will be paid a bonus B equal to c times the amount by which total hospital claims are under 400 $(0 \le c \le 1)$.
The effect the incentive plan will have on underlying hospital claims is modeled by assuming that the new total hospital claims will follow a two-parameter Pareto distribution with $\alpha = 2$ and $\theta = 300$.
Suppose that $E(B) = 100$. Calculate the value of c.

Solution.
Let X denote the number of hospital claims. We are told that

$$B = \begin{cases} c(400 - x), & x < 400 \\ 0, & x \ge 400 \end{cases} = \begin{cases} 400c - cx, & x < 400 \\ 400c - 400c, & x \ge 400 \end{cases} = 400c - X \wedge 400.$$

Thus,

$$100 = E(B) = E[400c - X \wedge 400] = 400c - cE(X \wedge 400).$$

We are told that X has a Pareto distribution with parameters $\alpha = 2$ and $\theta = 300$. Using Table C, we have

$$E(X \wedge u) = \frac{\theta}{\alpha - 1}\left[1 - \left(\frac{\theta}{u + \theta}\right)^{\alpha - 1}\right].$$

With $u = 400$, we find $E(X \wedge 400) = \frac{1200}{7}$. Finally,

$$100 = 400c - c\frac{1200}{7} \Longrightarrow c \approx 0.44 \blacksquare$$

Example 5.12
Show that, for $X \ge 0$, we have

$$e_X(d) = \frac{E(X) - E(X \wedge d)}{S(d)}.$$

Solution.
We have

$$\begin{aligned}
E(X) &= \int_0^\infty xf(x)dx = -xS(x)|_0^\infty + \int_0^\infty S(x)dx \\
&= \int_0^\infty S(x)dx \\
e_X(d) &= \frac{\int_d^\infty S(x)dx}{S(x)} = \frac{\int_0^\infty S(x)dx - \int_0^d S(x)dx}{S(d)} \\
&= \frac{E(X) - E(X \wedge d)}{S(d)} \blacksquare
\end{aligned}$$

Example 5.13 ‡
The random variable for a loss, X, has the following characteristics:

x	$F(x)$	$E(X \wedge x)$
0	0.0	0
100	0.2	91
200	0.6	153
1000	1.0	331

Calculate the mean excess loss for a deductible of 100.

Solution.
We are asked to find

$$e_X(100) = \frac{E(X) - E(X \wedge 100)}{1 - F_X(100)}.$$

The only term unknown in this formula is $E(X)$. Now, $\Pr(X > 1000) = 1 - F(1000) = 0$. This shows that $X \leq 1000$ so that $X \wedge 1000 = X$. It follows that $E(X) = E(X \wedge 1000) = 331$.
The mean excess loss is

$$e_X(100) = \frac{E(X) - E(X \wedge 100)}{1 - F_X(100)} = \frac{331 - 91}{1 - 0.2} = 300 \blacksquare$$

Remark 5.2
Just as in the case of a deductible, the random variable $Y = X \wedge u$ has a mixed distribution with continuous part $f_Y(y) = f_X(y)$ for $y < u$ and a discrete part $p_Y(u) = 1 - F_X(u)$.

Practice Problems

Problem 5.1
Suppose that a policy has a deductible of \$500. Complete the following table.

Amount of loss	750	500	1200
Insurance payment			

Problem 5.2
Referring to Example 5.3, find the cumulative distribution function of X.

Problem 5.3
Referring to Example 5.3, find the first and second raw moments of X.

Problem 5.4
Suppose you observe 8 claims with amounts

$$5 \quad 10 \quad 15 \quad 20 \quad 25 \quad 30 \quad 35 \quad 40$$

Calculate the empirical coefficient of variation.

Problem 5.5
Let X be uniform on the interval $[0, 100]$. Find $e_X(d)$ for $d > 0$.

Problem 5.6
Let X be uniform on $[0, 100]$ and Y be uniform on $[0, \alpha]$. Suppose that $e_Y(30) = e_X(30) + 4$.

Calculate the value of α.

Problem 5.7
Let X be the exponential random variable with mean λ. Its pdf is $f(x) = \lambda e^{-\lambda x}$ for $x > 0$ and 0 otherwise.

Find the expected cost per payment (i.e., mean excess loss function).

Problem 5.8
For an automobile insurance policy, the loss amount (expressed in thousands), in the event of an accident, is being modeled by a distribution with density

$$f(x) = \frac{3}{56}x(5-x),\ 0 < x < 4$$

and 0 otherwise.

For a policy with a deductible amount of \$2,500, calculate the expected amount per loss.

Problem 5.9
The loss random variable X has an exponential distribution with mean $\frac{1}{\lambda}$ and an ordinary deductible is applied to all losses.

Find the expected cost per loss.

Problem 5.10
The loss random variable X has an exponential distribution with mean $\frac{1}{\lambda}$ and an ordinary deductible is applied to all losses.

Find the variance of the cost per loss random variable.

Problem 5.11
The loss random variable X has an exponential distribution with mean $\frac{1}{\lambda}$ and an ordinary deductible is applied to all losses. The variance of the cost per payment random variable (excess loss random variable) is 25,600.

Find λ.

Problem 5.12
The loss random variable X has an exponential distribution with mean $\frac{1}{\lambda}$ and an ordinary deductible is applied to all losses. The variance of the cost per payment random variable (excess loss random variable) is 25,600. The variance of the cost per loss random variable is 20,480.

Find the amount of the deductible d.

Problem 5.13
The loss random variable X has an exponential distribution with mean $\frac{1}{\lambda}$ and an ordinary deductible is applied to all losses. The variance of the cost per payment random variable (excess loss random variable) is 25,600. The variance of the cost per loss random variable is 20,480.

Find expected cost of loss.

Problem 5.14
For the loss random variable with cdf $F(x) = \left(\frac{x}{\theta}\right)^{\phi}$, $0 < x < \theta$, and 0 otherwise, determine the mean residual lifetime $e_X(x)$.

Problem 5.15
Let X be a loss random variable with pdf $f(x) = (1 + 2x^2)e^{-2x}$ for $x > 0$ and 0 otherwise.

(a) Find the survival function $S(x)$.
(b) Determine $e_X(x)$.

Problem 5.16
Show that

$$S_{Y^P}(y) = \frac{S_X(y+d)}{S_X(d)}.$$

Problem 5.17
Let X be a loss random variable with cdf $F(x) = 1 - e^{-0.005x} - 0.004e^{-0.005x}$ for $x \geq 0$ and 0 otherwise.

(a) If an ordinary deductible of 100 is applied to each loss, find the pdf of the per payment random variable Y^P.
(b) Calculate the mean and the variance of the per payment random variable.

Problem 5.18
A continuous random variable X has a pdf $f(x) = 0.005x$ for $0 \leq x \leq 20$ and 0 otherwise.

Find the mean and the variance of $(X - 10)_+$.

Problem 5.19 ‡
For a random loss X, you are given: $\Pr(X = 3) = \Pr(X = 12) = 0.5$ and $E[(X - d)_+] = 3$.

Calculate the value of d.

Problem 5.20 ‡
A loss, X, follows a 2-parameter Pareto distribution with $\alpha = 2$ and unspecified parameter θ. You are given:

$$E(X - 100|X > 100) = \frac{5}{3}E(X - 50|X > 50).$$

Calculate $E(X - 150|X > 150)$.

Problem 5.21 ‡
For an insurance:
(i) Losses can be 100, 200 or 300 with respective probabilities 0.2, 0.2, and 0.6.
(ii) The insurance has an ordinary deductible of 150 per loss.
(iii) Y^P is the claim payment per payment random variable.

Calculate $\mathrm{Var}(Y^P)$.

Problem 5.22 ‡
For an insurance:
(i) Losses have density function

$$f(x) = \begin{cases} 0.02x, & 0 << 10 \\ 0, & \text{otherwise.} \end{cases}$$

(ii) The insurance has an ordinary deductible of 4 per loss.
(iii) Y^P is the claim payment per payment random variable.

Calculate $E[Y^P]$.

Problem 5.23 ‡
The loss severity random variable X follows the exponential distribution with mean 10,000.

Determine the coefficient of variation of the excess loss variable $Y = \max\{(X - 30000, 0)\}$.

6 Median, Mode, Percentiles, and Quantiles

In addition to the information provided by the moments of a distribution, some other metrics such as the median, the mode, the percentile, and the quantile provide useful information.

Median of a Random Variable

In probability theory, **median** is described as the numerical value separating the higher half of a probability distribution, from the lower half. Thus, the median of a discrete random variable X is the number M such that $\Pr(X \leq M) \geq 0.50$ and $\Pr(X \geq M) \geq 0.50$.

Example 6.1

Given the pmf of a discrete random variable X.

x	0	1	2	3	4	5
$p(x)$	0.35	0.20	0.15	0.15	0.10	0.05

Find the median of X.

Solution.

Since $\Pr(X \leq 1) = 0.55$ and $\Pr(X \geq 1) = 0.65$, 1 is the median of X ■

In the case of a continuous random variable X, the median is the number M such that $\Pr(X \leq M) = \Pr(X \geq M) = 0.5$. Generally, M is found by solving the equation $F(M) = 0.5$ where F is the cdf of X.

Example 6.2

Let X be a continuous random variable with pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$ and 0 otherwise. Find the median of X.

Solution.

We must find a number M such that $\int_a^M \frac{dx}{b-a} = 0.5$. This leads to the equation $\frac{M-a}{b-a} = 0.5$. Solving this equation we find $M = \frac{a+b}{2}$ ■

Remark 6.1

A discrete random variable might have many medians. For example, let X be the discrete random variable with pmf given by $p(x) = \left(\frac{1}{2}\right)^x, \ x = 1, 2, \cdots$ and 0 otherwise. Then any number $1 < M < 2$ satisfies $\Pr(X \leq M) = \Pr(X \geq M) = 0.5$.

Mode of a Random Variable

The **mode** is defined as the value that maximizes the probability mass function $p(x)$ (discrete case) or the probability density function $f(x)$ (continuous case.) In the discrete case, the mode is the value that is most likely to be sampled. In the continuous case, the mode is where $f(x)$ is at its peak.

Example 6.3
Let X be the discrete random variable with pmf given by $p(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, \cdots$ and 0 otherwise. Find the mode of X.

Solution.
The value of x that maximizes $p(x)$ is $x = 1$. Thus, the mode of X is 1 ■

Example 6.4
Let X be the continuous random variable with pdf given by $f(x) = 0.75(1 - x^2)$ for $-1 \leq x \leq 1$ and 0 otherwise. Find the mode of X.

Solution.
The pdf is maximum for $x = 0$. Thus, the mode of X is 0 ■

Percentiles and Quantiles

In statistics, a percentile is the value of a variable below which a certain percent of observations[5] fall. For example, if a score is in the 85$^{\text{th}}$ percentile, it is higher than 85% of the other scores. For a random variable X and $0 < p < 1$, the 100p$^{\text{th}}$ **percentile** (or the p^{th} **quantile**) is the number x such

$$\Pr(X < x) \leq p \leq \Pr(X \leq x).$$

For a continuous random variable, this is the solution to the equation $F(x) = p$. The 25$^{\text{th}}$ percentile is also known as the **first quartile**, the 50$^{\text{th}}$ percentile as the median or second quartile, and the 75$^{\text{th}}$ percentile as the third quartile.

Example 6.5
A loss random variable X has the density function

$$f(x) = \begin{cases} \frac{2.5(200)^{2.5}}{x^{3.5}} & x > 200 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the difference between the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of X.

[5] Another term for an "observation" in this text is "exposure".

Solution.
First, the cdf is given by

$$F(x) = \int_{200}^{x} \frac{2.5(200)^{2.5}}{t^{3.5}} dt.$$

If Q_1 is the 25^{th} percentile then it satisfies the equation

$$F(Q_1) = \frac{1}{4}$$

or equivalently

$$1 - F(Q_1) = \frac{3}{4}.$$

This leads to

$$\frac{3}{4} = \int_{Q_1}^{\infty} \frac{2.5(200)^{2.5}}{t^{3.5}} dt = -\left(\frac{200}{t}\right)^{2.5}\Bigg|_{Q_1}^{\infty} = \left(\frac{200}{Q_1}\right)^{2.5}.$$

Solving for Q_1 we find $Q_1 = 200(4/3)^{0.4} \approx 224.4$. Similarly, the third quartile (i.e. 75^{th} percentile) is given by $Q_3 = 348.2$, The **interquartile range** (i.e., the difference between the 25^{th} and 75^{th} percentiles) is $Q_3 - Q_1 = 348.2 - 224.4 = 123.8$ ■

Example 6.6
Let X be the random variable with pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$ and 0 otherwise. Find the p^{th} quantile of X.

Solution.
We have

$$p = \Pr(X \le x) = \int_a^x \frac{dt}{b-a} = \frac{x-a}{b-a}.$$

Solving this equation for x, we find $x = a + (b-a)p$ ■

Example 6.7
What percentile is 0.63 quantile?

Solution.
0.63 quantile is 63^{rd} percentile ■

Practice Problems

Problem 6.1
Using words, explain the meaning of $F(1120) = 0.2$ in terms of percentiles and quantiles.

Problem 6.2
Let X be a discrete random variable with pmf $p(n) = (n-1)(0.4)^2(0.6)^{n-2}$, $n \geq 2$ and 0 otherwise.

Find the mode of X.

Problem 6.3
Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \lambda \frac{1}{9} x(4-x), & 0 < x < 3 \\ 0, & \text{otherwise.} \end{cases}$$

Find the mode of X.

Problem 6.4
Suppose the random variable X has pmf

$$p(n) = \frac{1}{3}\left(\frac{2}{3}\right)^n, \quad n = 0, 1, 2, \cdots$$

and 0 otherwise.

Find the median and the 70^{th} percentile.

Problem 6.5
The time in minutes between individuals joining the line at an Ottawa Post Office is a random variable with the density function

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Find the median time between individuals joining the line and interpret your answer.

Problem 6.6
Suppose the random variable X has pdf

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Find the 100p^{th} percentile.

Problem 6.7 ‡
An insurance company sells an auto insurance policy that covers losses incurred by a policyholder, subject to a deductible of 100 . Losses incurred follow an exponential distribution with mean 300.

What is the 95^{th} percentile of actual losses that exceed the deductible?

Problem 6.8
Let X be a randon variable with density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Find λ if the median of X is $\frac{1}{3}$.

Problem 6.9
People are dispersed on a linear beach with a density function $f(y) = 4y^3,\ 0 < y < 1$, and 0 elsewhere. An ice cream vendor wishes to locate her cart at the median of the locations (where half of the people will be on each side of her).

Where will she locate her cart?

Problem 6.10 ‡
An automobile insurance company issues a one-year policy with a deductible of 500. The probability is 0.8 that the insured automobile has no accident and 0.0 that the automobile has more than one accident. If there is an accident, the loss before application of the deductible is exponentially distributed with mean 3000.

Calculate the 95^{th} percentile of the insurance company payout on this policy.

Problem 6.11
Let Y be a continuous random variable with cumulative distribution function

$$F(y) = \begin{cases} 0, & y \le a \\ 1 - e^{-\frac{1}{2}(y-a)^2}, & \text{otherwise} \end{cases}$$

where a is a constant.

Find the 75^{th} percentile of Y.

Problem 6.12
Find the p^{th} quantile of the exponential distribution defined by the distribution function $F(x) = 1 - e^{-x}$ for $x \geq 0$ and 0 otherwise.

Problem 6.13
A continuous random variable has the pdf $f(x) = e^{-|x|}$ for $x \in \mathbb{R}$.

Find the p^{th} quantile of X.

Problem 6.14
Let X be a loss random variable with cdf

$$F(x) = \begin{cases} 1 - \left(\frac{\theta}{\theta+x}\right)^{\alpha}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

The 10^{th} percentile is $\theta - k$. The 90^{th} percentile is $3\theta - 3k$.

Determine the value of α.

Problem 6.15
A random variable X follows a normal distribution with $\mu = 1$ and $\sigma^2 = 4$. Define a random variable $Y = e^X$, then Y follows a lognormal distribution. It is known that the 95^{th} percentile of a standard normal distribution is 1.645.

Calculate the 95^{th} percentile of Y.

Problem 6.16
Let X be a random variable with density function $f(x) = \frac{4x}{(1+x^2)^3}$ for $x > 0$ and 0 otherwise.

Calculate the mode of X.

Problem 6.17
Let X be a random variable with pdf $f(x) = \left(\frac{3}{5000}\right)\left(\frac{5000}{x}\right)^4$ for $x > 0$ and 0 otherwise.

Determine the median of X.

Problem 6.18
Let X be a random variable with cdf

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{27}, & 0 \leq x \leq 3 \\ 1, & x > 3. \end{cases}$$

Find the median of X.

Problem 6.19
Consider a sample of size 9 and observed data

$$45, 50, 50, 50, 60, 75, 80.120, 230.$$

Using this data as an empirical distribution, calculate the empirical mode.

Problem 6.20
A distribution has a pdf $f(x) = \frac{3}{x^4}$ for $x > 1$ and 0 otherwise.

Calculate the 0.95^{th} quantile of this distribution.

7 Sum of Random Variables and the Central Limit Theorem

Random variables of the form

$$S_n = X_1 + X_2 + \cdots + X_n$$

appear repeatedly in probability theory and applications. For example, in the insurance context, S_n can represent the total claims paid on all policies where X_i is the i^{th} claim. Thus, it is useful to be able to determine properties of S_n.

For the expected value of S_n, we have

$$E(S_n) = E(X_1) + E(X_2) + \cdots + E(X_{n-1}) + E(X_n).$$

A similar formula holds for the variance provided that the $X_i's$ are independent[6] random variables. In this case,

$$\text{Var}(S_n) = \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n).$$

Example 7.1 ‡
The random variables $X_1, X_2, \cdots, X_n$ are independent and identically distributed with probability density function

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}.$$

Determine $E[\overline{X}^2]$.

Solution.
The random variable X_i has an exponential distribution with mean θ. Thus,

[6] We say that X and Y are independent random variables if and only if for any two sets of real numbers A and B we have

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

$E(X_i) = \theta$ and $\text{Var}(X_i) = \theta^2$. Thus,

$$\begin{aligned}
E[\overline{X}] =& \frac{E(X_1) + \cdots + E(X_n)}{n} = \theta \\
\text{Var}[\overline{X}] =& \frac{\text{Var}(X_1) + \cdots + \text{Var}(X_n)}{n^2} \\
=& \frac{\theta^2}{n} \\
E[\overline{X}^2] =& \text{Var}[\overline{X}] + E[\overline{X}]^2 \\
=& \frac{\theta^2}{n} + \theta^2 = \left(\frac{n+1}{n}\right)\theta^2 \blacksquare
\end{aligned}$$

The central limit theorem reveals a fascinating property of the sum of independent random variables. It states that the CDF of the sum converges to the standard normal CDF as the number of terms grows without limit. This theorem allows us to use the properties of the standard normal distribution to obtain accurate estimates of probabilities associated with sums of random variables.

Theorem 7.1
Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables, each with mean μ and variance σ^2. Then,

$$P\left(\frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right) \le a\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

as $n \to \infty$.

The Central Limit Theorem says that regardless of the underlying distribution of the variables X_i, so long as they are independent, the distribution of $\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right)$ converges to the same, normal, distribution.

Example 7.2
The weight of a typewriter has a mean of 20 pounds and a variance of 9 pounds. Consider a train that carries 200 of these typewriters. Estimate the probability that the total weight of typewriters carried in the train exceeds 4050 pounds.

Solution.
Label the typewriters as Typewriter 1, Typewriter 2, etc. Let X_i be the

weight of Typewriter i. Thus,

$$\begin{aligned}P\left(\sum_{i=1}^{200} X_i > 4050\right) =& P\left(\frac{\sum_{i=1}^{200} X_i - 200(20)}{3\sqrt{200}} > \frac{4050 - 20(200)}{3\sqrt{200}}\right)\\ \approx& P(Z > 1.179) = 1 - P(Z \le 1.179)\\ =& 1 - \Phi(1.179) = 1 - 0.8810 = 0.119\end{aligned}$$

where Φ is the CDF of the standard normal distribution ■

Example 7.3 ‡
In an analysis of healthcare data, ages have been rounded to the nearest multiple of 5 years. The difference between the true age and the rounded age is assumed to be uniformly distributed on the interval from -2.5 years to 2.5 years. The healthcare data are based on a random sample of 48 people. What is the approximate probability that the mean of the rounded ages is within 0.25 years of the mean of the true ages?

Solution.
Let X denote the difference between true and reported age. We are given X is uniformly distributed on $(-2.5, 2.5)$. That is, X has pdf $f(x) = 1/5, -2.5 < x < 2.5$. It follows that $E(X) = 0$ and

$$\sigma_X^2 = E(X^2) = \int_{-2.5}^{2.5} \frac{x^2}{5} dx \approx 2.083$$

so that $SD(X) = \sqrt{2.083} \approx 1.443$.
Now $\overline{X}_{48}$ the difference between the means of the true and rounded ages, has a distribution that is approximately normal with mean 0 and standard deviation $\frac{1.443}{\sqrt{48}} \approx 0.2083$. Therefore,

$$\begin{aligned}P\left(-\frac{1}{4} \le \overline{X}_{48} \le \frac{1}{4}\right) =& P\left(\frac{-0.25}{0.2083} \le \frac{X_{48}}{0.2083} \le \frac{0.25}{0.2083}\right)\\ =& P(-1.2 \le Z \le 1.2) = 2\Phi(1.2) - 1\\ \approx& 2(0.8849) - 1 = 0.77\end{aligned}$$ ■

Example 7.4
Let X_1, X_2, X_3, X_4 be a random sample of size 4 from a normal distribution with mean 2 and variance 10, and let $\overline{X}$ be the sample mean. Determine a such that $P(\overline{X} \le a) = 0.90$.

Solution.
The sample mean $\overline{X}$ is normal with mean $\mu = 2$ and variance $\frac{\sigma^2}{n} = \frac{10}{4} = 2.5$, and standard deviation $\sqrt{2.5} \approx 1.58$, so

$$0.90 = P(\overline{X} \leq a) = P\left(\frac{\overline{X} - 2}{1.58} < \frac{a-2}{1.58}\right) = \Phi\left(\frac{a-2}{1.58}\right).$$

Using Excel, we get $\frac{a-2}{1.58} = 1.28$, so $a = 4.02$ ■

Practice Problems

Problem 7.1
A shipping agency ships boxes of booklets with each box containing 100 booklets. Suppose that the average weight of a booklet is 1 ounce and the standard deviation is 0.05 ounces.

What is the probability that 1 box of booklets weighs more than 100.4 ounces?

Problem 7.2
In the National Hockey League, the standard deviation in the distribution of players' height is 2 inches. The heights of 25 players selected at random were measured.

Estimate the probability that the average height of the players in this sample is within 1 inch of the league average height.

Problem 7.3
A battery manufacturer claims that the lifespan of its batteries has a mean of 54 hours and a standard deviation of 6 hours. A sample of 60 batteries were tested.

What is the probability that the mean lifetime is less than 52 hours?

Problem 7.4
Roll a dice 10 times. Estimate the probability that the sum obtained is between 30 and 40, inclusive.

Problem 7.5
Consider 10 independently random variables each uniformly distributed over (0,1).

Estimate the probability that the sum of the variables exceeds 6.

Problem 7.6
The Chicago Cubs play 100 independent baseball games in a given season. Suppose that the probability of winning a game in 0.8.

What's the probability that they win at least 90?

Problem 7.7
An insurance company has 10,000 home policyholders. The average annual claim per policyholder is found to be \$240 with a standard deviation of \$800.

Estimate the probability that the total annual claim is at least \$2.7 million.

Problem 7.8
A certain component is critical to the operation of a laptop and must be replaced immediately upon failure. It is known that the average life of this type of component is 100 hours and its standard deviation is 30 hours.

Estimate the number of such components that must be available in stock so that the system remains in continual operation for the next 2000 hours with probability of at least 0.95?

Problem 7.9
An instructor found that the average student score on class exams is 74 and the standard deviation is 14. This instructor gives two exams: One to a class of size 25 and the other to a class of 64.

Using the Central Limit Theorem, estimate the probability that the average test score in the class of size 25 is at least 80.

Problem 7.10
The Salvation Army received 2025 in contributions. Assuming the contributions to be independent and identically distributed with mean 3125 and standard deviation 250.

Estimate the 90^{th} percentile for the distribution of the total contributions received.

Problem 7.11 ‡
An insurance company issues 1250 vision care insurance policies. The number of claims filed by a policyholder under a vision care insurance policy during one year is a Poisson random variable with mean 2. Assume the numbers of claims filed by distinct policyholders are independent of one another.

What is the approximate probability that there is a total of between 2450 and 2600 claims during a one-year period?

Problem 7.12
A battery manufacturer finds that the lifetime of a battery, expressed in months, follows a normal distribution with mean 3 and standard deviation 1 . Suppose that you want to buy a number of these batteries with the intention of replacing them successively into your radio as they burn out.

Assuming that the batteries' lifetimes are independent, what is the smallest number of batteries to be purchased so that the succession of batteries keeps your radio working for at least 40 months with probability exceeding 0.9772?

Problem 7.13
The total claim amount for a home insurance policy has a pdf

$$f(x) = \begin{cases} \frac{1}{1000}e^{-\frac{x}{1000}} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

An actuary sets the premium for the policy at 100 over the expected total claim amount.

If 100 policies are sold, estimate the probability that the insurance company will have claims exceeding the premiums collected.

Problem 7.14 ‡
A city has just added 100 new female recruits to its police force. The city will provide a pension to each new hire who remains with the force until retirement. In addition, if the new hire is married at the time of her retirement, a second pension will be provided for her husband. A consulting actuary makes the following assumptions:

(i) Each new recruit has a 0.4 probability of remaining with the police force until retirement.
(ii) Given that a new recruit reaches retirement with the police force, the probability that she is not married at the time of retirement is 0.25.
(iii) The number of pensions that the city will provide on behalf of each new hire is independent of the number of pensions it will provide on behalf of any other new hire.

Determine the probability that the city will provide at most 90 pensions to the 100 new hires and their husbands.

Problem 7.15
The amount of an individual claim has a two-parameter Pareto distribution with $\theta = 8000$ and $\alpha = 9$. Consider a sample of 500 claims.

Estimate the probability that the total sum of the claims is at least 550,000.

Problem 7.16
Suppose that the current profit from selling a share of a stock is found to follow a uniform distribution on $[-45, 72]$.

Using the central limit theorem, approximate the probability of making a profit from the sale of 55 stocks.

Problem 7.17
The severities of individual claims have the Pareto distribution with parameters $\alpha = \frac{8}{3}$ and $\theta = 8000$.

Use the central limit theorem to approximate the probability that the sum of 100 independent claims will exceed 600,000.

Problem 7.18 ‡
Let X and Y be the number of hours that a randomly selected person watches movies and sporting events, respectively, during a three-month period. The following information is known about X and Y :

E(X) =	50
E(Y) =	20
Var(X) =	50
Var(Y) =	30
Cov (X,Y) =	10

One hundred people are randomly selected and observed for these three months. Let T be the total number of hours that these one hundred people watch movies or sporting events during this three-month period.

Approximate the value of $P(T < 7100)$.

Problem 7.19 ‡
Automobile losses reported to an insurance company are independent and uniformly distributed between 0 and 20,000. The company covers each such loss subject to a deductible of 5,000.

Calculate the probability that the total payout on 200 reported losses is between 1,000,000 and 1,200,000.

Problem 7.20 ‡
For Company A there is a 60% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 10,000 and standard deviation 2, 000.
For Company B there is a 70% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 9,000 and standard deviation 2,000.
Assume that the total claim amounts of the two companies are independent.

What is the probability that, in the coming year, Company B's total claim amount will exceed Company A's total claim amount?

8 Moment Generating Functions and Probability Generating Functions

A useful way to analyze the sum of independent random variables is to transform the PDF or PMF of each random variable to a **moment generating function**, abbreviated mgf.

The moment generating function of a continuous random variable X with a density function $f(x)$ is denoted by $M_X(t)$ and is given by

$$M_X(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x)dx.$$

The moment generating function of a discrete random variable X with a probability mass function $p(x)$ is denoted by $M_X(t)$ and is given by

$$M_X(t) = E[e^{tx}] = \sum_{x \in \text{support}(X)} e^{tx} p(x).$$

Example 8.1
Calculate the moment generating function for the exponential distribution with parameter λ, i.e. $f(x) = \lambda e^{-\lambda x}$ for $x > 0$ and 0 otherwise.

Solution.
We have

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{\infty} \lambda e^{-x(\lambda - t)} dx = -\frac{\lambda}{\lambda - t} e^{-x(\lambda - t)} \Big|_0^{\infty} = \frac{\lambda}{\lambda - t}, \; t < \lambda \blacksquare$$

Example 8.2
Let X be a discrete random variable with pmf given by the following table

x	1	2	3	4	5
$p(x)$	0.15	0.20	0.40	0.15	0.10

and 0 otherwise. Calculate $M_X(t)$.

Solution.
We have

$$M_X(t) = 0.15e^t + 0.20e^{2t} + 0.40e^{3t} + 0.15e^{4t} + 0.10e^{5t} \blacksquare$$

As the name suggests, the moment generating function can be used to generate moments $E(X^n)$ for $n = 1, 2, \cdots$. The next result shows how to use the moment generating function to calculate moments.

Theorem 8.1
For any random variable X, we have

$$E(X^n) = M_X^n(0) \quad \text{where} \quad M_X^n(0) = \tfrac{d^n}{dt^n} M_X(t)\big|_{t=0}.$$

Example 8.3
Let X be a binomial random variable with parameters n and p. Find the expected value and the variance of X using moment generating functions.

Solution.
We can write

$$M_X(t) = \sum_{k=0}^{n} e^{tk} C(n,k) p^k (1-p)^{n-k} = \sum_{k=0}^{n} C(n,k)(pe^t)^k (1-p)^{n-k} = (pe^t+1-p)^n.$$

Differentiating yields

$$\frac{d}{dt} M_X(t) = npe^t(pe^t+1-p)^{n-1} \Longrightarrow E(X) = \frac{d}{dt} M_X(t) \mid_{t=0} = np.$$

To find $E(X^2)$, we differentiate a second time to obtain

$$\frac{d^2}{dt^2} M_X(t) = n(n-1)p^2 e^{2t}(pe^t+1-p)^{n-2} + npe^t(pe^t+1-p)^{n-1}.$$

Evaluating at $t=0$ we find

$$E(X^2) = M_X''(0) = n(n-1)p^2 + np.$$

Observe that this implies the variance of X is

$$Var(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p) \blacksquare$$

Example 8.4
Let X be a Poisson random variable with parameter λ. Find the expected value and the variance of X using moment generating functions.

Solution.
We can write

$$M_X(t) = \sum_{n=0}^{\infty} \frac{e^{tn} e^{-\lambda} \lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{e^{tn}\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}.$$

Differentiating for the first time we find

$$M_X'(t) = \lambda e^t e^{\lambda(e^t-1)} \Longrightarrow E(X) = M_X'(0) = \lambda.$$

Differentiating a second time we find

$$M''_X(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)} \Longrightarrow E(X^2) = M''_X(0) = \lambda^2 + \lambda.$$

The variance is then

$$Var(X) = E(X^2) - (E(X))^2 = \lambda \blacksquare$$

Example 8.5
Let X be a normal random variable with parameters μ and σ^2. Find the expected value and the variance of X using moment generating functions.

Solution.
First we find the moment of a standard normal random variable with parameters 0 and 1. We can write

$$\begin{aligned} M_Z(t) =& E(e^{tZ}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tz}e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left\{-\frac{(z^2-2tz)}{2}\right\}dz \\ =& \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left\{-\frac{(z-t)^2}{2}+\frac{t^2}{2}\right\}dz = e^{\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2}}dz = e^{\frac{t^2}{2}} \end{aligned}$$

Now, since $X = \mu + \sigma Z$ we have

$$\begin{aligned} M_X(t) =& E(e^{tX}) = E(e^{t\mu + t\sigma Z}) = E(e^{t\mu}e^{t\sigma Z}) = e^{t\mu}E(e^{t\sigma Z}) \\ =& e^{t\mu}M_Z(t\sigma) = e^{t\mu}e^{\frac{\sigma^2 t^2}{2}} = exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} \end{aligned}$$

By differentiation we obtain

$$M'_X(t) = (\mu + t\sigma^2)exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

and

$$M''_X(t) = (\mu + t\sigma^2)^2 exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} + \sigma^2 exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

and thus

$$E(X) = M'_X(0) = \mu \quad \text{and} \quad E(X^2) = M''_X(0) = \mu^2 + \sigma^2$$

The variance of X is

$$Var(X) = E(X^2) - (E(X))^2 = \sigma^2 \blacksquare$$

Moment generating functions are also useful in establishing the distribution of sums of independent random variables. Suppose $X_1, X_2, \cdots, X_N$ are independent random variables. Then, the moment generating function of $Y = X_1 + \cdots + X_N$ is

$$M_Y(t) = E(e^{t(X_1+X_2+\cdots+X_N)}) = E(e^{X_1 t}\cdots e^{X_N t}) = \prod_{k=1}^{N} E(e^{X_k t}) = \prod_{k=1}^{N} M_{X_k}(t).$$

Another important property is that the moment generating function uniquely determines the distribution. That is, if random variables X and Y both have moment generating functions $M_X(t)$ and $M_Y(t)$ that exist in some neighborhood of zero and if $M_X(t) = M_Y(t)$ for all t in this neighborhood, then X and Y have the same distributions.

Example 8.6
If X and Y are independent binomial random variables with parameters (n, p) and (m, p), respectively, what is the pmf of $X + Y$?

Solution.
We have

$$M_{X+Y}(t) = M_X(t)M_Y(t) = (pe^t + 1 - p)^n (pe^t + 1 - p)^m = (pe^t + 1 - p)^{n+m}.$$

Since $(pe^t + 1 - p)^{n+m}$ is the moment generating function of a binomial random variable having parameters $m + n$ and p, $X + Y$ is a binomial random variable with this same pmf ■

Example 8.7
If X and Y are independent Poisson random variables with parameters λ_1 and λ_2, respectively, what is the pmf of $X + Y$?

Solution.
We have

$$M_{X+Y}(t) = M_X(t)M_Y(t) = e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)}.$$

Since $e^{(\lambda_1+\lambda_2)(e^t-1)}$ is the moment generating function of a Poisson random variable having parameter $\lambda_1 + \lambda_2$, $X + Y$ is a Poisson random variable with this same pmf ■

Probability Generating Function

Another useful tool for dealing with the distribution of a sum of discrete random variables is the **probability generating function** or the $z-$transform of the probability mass function, abbreviated pgf. For a discrete random variable X, we define the probability generating function by

$$P_X(t) = E(t^x) = \sum_{x \in \text{Support(X)}} t^x p(x), \ \forall t \in \mathbb{R} \text{ for which the sum converges.}$$

Note that $P_X(t) = M_X[e^{x \ln t}]$ and $M_X(t) = P_X(e^t)$. The pgf transforms a sum into a product and enables it to be handled much more easily: Let $X_1, X_2, \cdots, X_n$ be independent random variables and $S_n = X_1 + X_2 + \cdots + X_n$. It can be shown that

$$P_{S_n}(t) = P_{X_1}(t)P_{X_2}(t) \cdots P_{X_n}(t).$$

Example 8.8

Find the pgf of the Poisson distribution of parameter λ.

Solution.

Recall that the Poisson random variable has a pmf $p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$. Hence,

$$P_X(t) = \sum_{x=0}^{\infty} t^x \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)} \blacksquare$$

The probability generating function gets its name because the power series can be expanded and differentiated to reveal the individual probabilities. Thus, given only the pgf $P_X(t) = E(t^x)$, we can recover all probabilities $\Pr(X = x)$.

It can be shown that

$$p(n) = \frac{1}{n!} \frac{d^n}{dt^n} P_X(t) \Big|_{t=0}.$$

Example 8.9

Let X be a discrete random variable with pgf $P_X(t) = \frac{t}{5}(2 + 3t^2)$. Find the distribution of X.

Solution.

We have, $p(0) = P_X(0) = 0; p(1) = P'_X(0) = \frac{2}{5}; p(2) = \frac{P''_X(0)}{2!} = 0, p(3) = \frac{PX'''(0)}{3!} = \frac{3}{5}$; and $p(n) = 0, \ \forall n \geq 4 \blacksquare$

Theorem 8.2
For any discrete random variable X, we have

$$E[X(X-1)(X-2)\cdots(X-k+1)] = \left.\frac{d^k}{dt^k}P_X(t)\right|_{t=1}.$$

Example 8.10
Let X be a Poisson random variable with parameter λ. Find the mean and the variance using probability generating functions.

Solution.
We know that the pgf of X is $P_X(t) = e^{\lambda(t-1)}$. We have

$$\begin{aligned} E(X) =& P'_X(1) = \lambda \\ E[X(X-1)] =& \lambda^2 \\ E(X^2) =& \lambda^2 + \lambda \\ \text{Var}(X) =& \lambda \ \blacksquare \end{aligned}$$

Like moment generating functions, the probability generating function uniquely determines the distribution. That is, if we can show that two random variables have the same pgf in some interval containing 0, then we have shown that the two random variables have the same distribution.

Example 8.11
Let X be a Poisson random variable with parameter λ and Y is Poisson with parameter μ. Find the distribution of $X+Y$, assuming X and Y are independent.

Solution.
We have

$$P_{X+Y}(t) = P_X(t)P_Y(t) = e^{\lambda(t-1)}e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)}.$$

This is the pgf of a Poisson random variable with parameter $\lambda+\mu$. So, by the uniqueness of pgfs, $X+Y$ is a Poisson random variable with parameter $\lambda+\mu$ $\blacksquare$

Practice Problems

Problem 8.1
Let X be an exponential random variable with parameter λ.

Find the expected value and the variance of X using moment generating functions.

Problem 8.2
Let X and Y be independent normal random variables with parameters (μ_1, σ_1^2) and (μ_2, σ_2^2), respectively.

Find the distribution of $X + Y$.

Problem 8.3 ‡
Let X and Y be identically distributed independent random variables such that the moment generating function of $X + Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t}, \quad -\infty < t < \infty.$$

Calculate $\Pr(X \leq 0)$.

Problem 8.4 ‡
The value of a piece of factory equipment after three years of use is $100(0.5)^X$ where X is a random variable having moment generating function

$$M_X(t) = \frac{1}{1-2t} \text{ for } t < \frac{1}{2}.$$

Calculate the expected value of this piece of equipment after three years of use.

Problem 8.5
Let X and Y be two independent random variables with moment generating functions

$$M_X(t) = e^{t^2+2t} \text{ and } M_Y(t) = e^{3t^2+t}.$$

Determine the moment generating function of $X + 2Y$.

Problem 8.6
The random variable X has an exponential distribution with parameter b. It is found that $M_X(-b^2) = 0.2$.

Find b.

Problem 8.7
If the moment generating function for the random variable X is $M_X(t) = \frac{1}{t+1}$, find $E[(X-2)^3]$.

Problem 8.8
Suppose a random variable X has moment generating function

$$M_X(t) = \left(\frac{2+e^t}{3}\right)^9.$$

Find the variance of X.

Problem 8.9
A random variable X has the moment generating function

$$M_X(t) = \frac{1}{(1-2500t)^4}.$$

Determine the standard deviation of X.

Problem 8.10 ‡
A company insures homes in three cities, J, K, and L . Since sufficient distance separates the cities, it is reasonable to assume that the losses occurring in these cities are independent.
The moment generating functions for the loss distributions of the cities are:

$$\begin{aligned}
M_J(t) =&(1-2t)^{-3}\\
M_K(t) =&(1-2t)^{-2.5}\\
M_J(t) =&(1-2t)^{-4.5}
\end{aligned}$$

Let X represent the combined losses from the three cities.

Calculate $E(X^3)$.

Problem 8.11
Let X be a binomial random variable with pmf $p(k) = C(n,k)p^k(1-p)^{n-k}$.

Find the pgf of X.

Problem 8.12
Let X be a geometric random variable with pmf $p(n) = p(1-p)^{n-1},\ n = 1, 2, \cdots$, where $0 < p < 1$.

Find the pgf of X.

Problem 8.13
Let X be a random variable with pgf $P_X(t) = e^{\lambda(t-1)}$. True or false: X is a Poisson random variable with parameter λ.

Problem 8.14
Let X be a random variable and $Y = a + bX$. Express $P_Y(t)$ in terms of $P_X(t)$.

Problem 8.15
Let X have the distribution of a geometric random variable with parameter p. That is, $p(x) = p(1-p)^{x-1},\ x = 1, 2, 3, \cdots$.

Find the mean and the variance of X using probability generating functions.

Problem 8.16
You are given a sample of size 4 with observed data

$$2 \quad 2 \quad 3 \quad 5 \quad 8$$

Using empirical distribution framework, calculate the probability generating function.

Problem 8.17
Let X be a discrete random variable with the pmf given below

x	-2	3	π	$\frac{7}{2}$
$p(x)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{3}{8}$

and 0 otherwise. Find the probability generating function $P_X(t)$.

Problem 8.18
Suppose $p(n) = \frac{1}{2^{n-2}},\ n = 3, 4, 5, \cdots$ and 0 otherwise. Find the probability generating function $P_X(t)$.

Problem 8.19
Suppose $P_X(t) = \frac{1}{t}\left(\frac{1}{3}t + \frac{2}{3}\right)^4$. Find the pmf of X.

Problem 8.20
Let X be a random variable with probability generating function $P_X(t) = \frac{t(1+t)}{2(3-2t)}$.

Using $P_X(t)$, find $E(X)$ and $\mathrm{Var}(X)$.

Tail Weight of a Distribution

The (**right-**)**tail** of a distribution is the portion of the distribution corresponding to large values of the random variable. Alternatively, we can define the tail of a random variable X as the interval (x, ∞) with probability

$$\Pr(X > x) = S_X(x) = 1 - F_X(x) = \int_x^\infty f_X(x)dx$$

where $S_X(x)$ is the survival function of X.

A distribution is said to be a **heavy-tailed distribution** if it significantly puts more probability on larger values of the random variable. We also say that the distribution has a larger tail weight. In contrast, a distribution that puts less and less probability for larger values of the random variable is said to be **light-tailed distribution.** According to [1], there are four ways to look for indication that a distribution is heavy-tailed. The purpose of this chapter is to discuss these various ways.

9 Tail Weight Measures: Moments and the Speed of Decay of $S(x)$

There are four ways of measuring heavy-tailed distributions as suggested by [1]:
- Existence of non-central moments.
- The speed for which the survival function decays to zero.
- The hazard rate function.
- The mean excess loss function.

In this section we cover the first two and the last two will be covered in the next section.

The Existence of Moments
A distribution $f_X(x)$ is said to be **light-tailed** if $E(x^k) < \infty$ for all $k > 0$. The distribution $f_X(x)$ is said to be **heavy-tailed** if either $E(x^k)$ does not exist for all $k > 0$ or the moments exist only up to a certain value of a positive integer k.

Example 9.1
Show that the exponential distribution with parameter $\lambda > 0$ is light-tailed according to the above definition. Refer to Table C.

Solution.
Using Table C, for all positive integers k, we have

$$E(X^k) = \frac{\Gamma(k+1)}{\lambda^k}.$$

Hence, the exponential distribution is light-tailed ■

Example 9.2
Show that the Pareto distribution with parameters α and θ is heavy-tailed. Refer to Table C.

Solution.
Using Table C, we have

$$E(X^k) = \frac{\theta^k \Gamma(k+1)\Gamma(\alpha - k)}{\Gamma(\alpha)}$$

provided that $-1 < k < \alpha$. Since the moments are not finite for all positive k, the Pareto distribution is heavy-tailed ■

Example 9.3
Let X be a continuous random variable with pdf $f_X(x)$ defined for $x > 0$ and 0 otherwise. Suppose that there is a constant $M > 0$ such that $f_X(x) = \frac{C}{x^n}$ for all $x \geq M$ and 0 otherwise, where $n > 1$ and $C = \frac{n-1}{M^{1-n}}$. Show that X has a heavy-tailed distribution.

Solution.
We have

$$\begin{aligned}E(X^k) =& \int_0^M x^k f_X(x)dx + C\int_M^\infty x^{k-n}dx\\ =& \int_0^M x^k f_X(x)dx + C\left.\frac{x^{k-n+1}}{k-n+1}\right|_M^\infty\\ =&\infty\end{aligned}$$

for all $k > n-1$ ■

<u>**Classification Based on the Speed of Decay of the Survival Function**</u>
The survival function $S_X(x) = P(X > x)$ captures the probability of the tail of a distribution. Recall that $S_X(x)$ decreases to 0 as $x \to \infty$. The question is how fast the function decays to zero. If the survival function of a distribution decays slowly to zero (equivalently the cdf goes slowly to one), it is another indication that the distribution is heavy-tailed. In contrast, when the survival function decays to 0 very rapidly then this is indication that the distribution is light-tailed.

Next, we consider comparing the tail weights of two distributions with the same mean. This is done by comparing the survival functions of the two distributions. Algebraically, we compute the ratio of the tail probabilities or the survival functions which we will refer to as the **relative tail weight**:

$$\lim_{x\to\infty}\frac{S_X(x)}{S_Y(x)} = \lim_{x\to\infty}\frac{-S'_X(x)}{-S'_Y(x)} = \lim_{x\to\infty}\frac{f_X(x)}{f_Y(x)} \geq 0.$$

Note that in the middle limit we used L'Hôpital's rule since $\lim_{x\to\infty} S_X(x) = \lim_{x\to\infty} S_Y(x) = 0$.

Now, if the above limit is 0, then this happens only when the numerator is 0 and the denominator is positive. In this case, we say that the distribution of X has **lighter tail** than Y. If the limit is finite positive number then we say that the distributions have similar or proportional tails. If the limit diverges to infinity, then more probabilities on large values of X are assigned to the numerator, In this case, we say that the distribution X has **heavier tail** than the distribution Y.

Example 9.4
Compare the tail weight of the inverse Pareto distribution with pdf $f_X(x) = \frac{\tau\theta_1 x^{\tau-1}}{(x+\theta_1)^{\tau+1}}$ with the inverse Gamma distribution with pdf $f_Y(x) = \frac{\theta_2^\alpha e^{-\frac{\theta_2}{x}}}{x^{\alpha+1}\Gamma(\alpha)}$ where $\theta_1, \theta_2, \tau > 0$ and $\alpha > 1$.

Solution.
We have

$$\begin{aligned}\lim_{x\to\infty}\frac{f_X(x)}{f_Y(x)} =& \lim_{x\to\infty}\frac{\tau\theta_1 x^{\tau-1}}{(x+\theta_1)^{\tau+1}}\cdot\frac{x^{\alpha+1}\Gamma(\alpha)}{\theta_2^\alpha e^{-\frac{\theta_2}{x}}}\\ =& \lim_{x\to\infty}\frac{\tau\theta_1\Gamma(\alpha)}{\theta_2^\alpha}e^{\frac{\theta_2}{x}}\left(\frac{x}{x+\theta_1}\right)^{\tau+1}x^{\alpha-1}\\ =&\frac{\tau\theta_1\Gamma(\alpha)}{\theta_2^\alpha}\cdot e^0\cdot\infty=\infty.\end{aligned}$$

Thus, X has a heavier tail than Y ∎

Example 9.5
Let$X$ be the exponential distribution with survival function $S_X = e^{-x}$ for $x \geq 0$ and 0 otherwise, and Y be the distribution with survival function $S_Y(x) = \frac{1}{x}$ for $x \geq 1$ and 0 otherwise. Compare the tail weight of these distributions.

Solution.
We have

$$\lim_{x\to\infty}\frac{S_X(x)}{S_Y(x)} = \lim_{x\to\infty} xe^{-x} = 0.$$

Hence, X has a lighter tail than Y ∎

Practice Problems

Problem 9.1
Let X be a random variable with pdf $f_X(x) = Cx^n e^{-bx}$ for $x > 0$ and 0 otherwise, where $b, n > 0$ and $C = \left[\int_0^\infty x^n e^{-bx} dx\right]^{-1}$.

Show that X has a light tail distribution.

Problem 9.2
Suppose X has a heavy-tailed distribution. Let $t > 0$. Show that $\int_N^\infty e^{tx} f_X(x)dx = \infty$ for some $N = N(t) > 0$.

Problem 9.3
Suppose X has a heavy-tailed distribution. Show that $M_X(t) = \infty$ for all $t > 0$.

Problem 9.4
Determine whether the Γ distribution with parameters $\alpha > 0$ and $\theta > 0$ is light-tailed or heavy-tailed. Refer to Table C.

Problem 9.5
Let X be the inverse Weibull random variables with parameters θ and τ.

Determine whether the distribution is light-tailed or heavy-tailed. Refer to Table C.

Problem 9.6
Compare the tail weight of the Pareto distribution with pdf $f_X(x) = \frac{\alpha\theta_1^\alpha}{(x+\theta_1)^{\alpha+1}}$ with the Gamma distribution with pdf $f_Y(x) = \frac{x^{\tau-1}e^{-\frac{x}{\theta_2}}}{\theta_2^\tau\Gamma(\tau)}$ where $\theta_1, \theta_2, \tau > 0$ and $\alpha > 1$.

Problem 9.7
Compare the tail weight of the Weibull distribution with pdf $f_X(x) = \frac{\tau}{x}\left(\frac{x}{\theta}\right)^\tau e^{-\left(\frac{x}{\theta}\right)^\tau}$ and the inverse Weibull distribution with pdf $f_Y(x) = \frac{\tau}{x}\left(\frac{\theta}{x}\right)^\tau e^{-\left(\frac{\theta}{x}\right)^\tau}$ where $\tau, \theta > 0$.

Problem 9.8
Let X be a random variable with pdf $f_X(x) = \frac{2}{\pi(1+x^2)}$ for $x > 0$ and 0 otherwise. Let Y be a random variable with pdf $f_Y(x) = \frac{\alpha}{(1+x)^{\alpha+1}}$ for $x > 0$ and 0 otherwise, where $\alpha > 1$.

Compare the tail weight of these distributions.

Problem 9.9
Let X be a random variable with pdf $f_X(x) = \frac{2}{\pi(1+x^2)}$ for $x > 0$ and 0 otherwise. Let Y be a random variable with pdf $f_Y(x) = \frac{1}{(1+x)^2}$ for $x > 0$ and 0 otherwise.

Compare the tail weight of these distributions.

Problem 9.10
Let X be a random variable with pdf $f_X(x) = \frac{2}{\pi(1+x^2)}$ for $x > 0$ and 0 otherwise. Let Y be a random variable with pdf $f_Y(x) = \frac{\alpha}{(1+x)^{\alpha+1}}$ for $x > 0$ and 0 otherwise, where $0 < \alpha < 1$.

Compare the tail weight of these distributions.

Problem 9.11
The distribution of X has the survival function

$$S_X(x) = 1 - \frac{\theta x^\gamma}{1 + \theta x^\gamma}, \ \theta, \gamma > 0.$$

and 0 otherwise. The distribution of Y has pdf

$$S_Y(x) = \frac{x^{\gamma-1} e^{-\frac{x}{\theta}}}{\theta^\gamma \Gamma(\gamma)}.$$

and 0 otherwise.

Compare the tail behavior of these distributions.

Problem 9.12
Using the criterion of existence of moments, complete the following. Refer to Table C.

Distribution	Heavy-Tail	Light-Tail
Weibull		
Inverse Pareto		
Normal		
Loglogistic		

Problem 9.13
Using the criterion of existence of moments, complete the following. Refer to Table C.

Distribution	Heavy-Tail	Light-Tail
Paralogistic		
Lognormal		
Inverse Gamma		
Inverse Gaussian		

Problem 9.14
Using the criterion of existence of moments, complete the following. Refer to Table C.

Distribution	Heavy-Tail	Light-Tail
Inverse Paralogistic		
Inverse Exponential		

Problem 9.15
Show that the Loglogistic distribution has a heavier tail than the Gamma distribution.

Problem 9.16
Show that the Paraloglogistic distribution has a heavier tail than the Lognormal distribution.

Problem 9.17
Show that the inverse exponential distribution has a heavier tail than the exponential distribution.

Problem 9.18
Let X and Y have similar (proportional) right tails and $\lim_{x\to\infty}\frac{S_X(x)}{S_Y(x)}=c.$ Which of the following is a possible value of c?

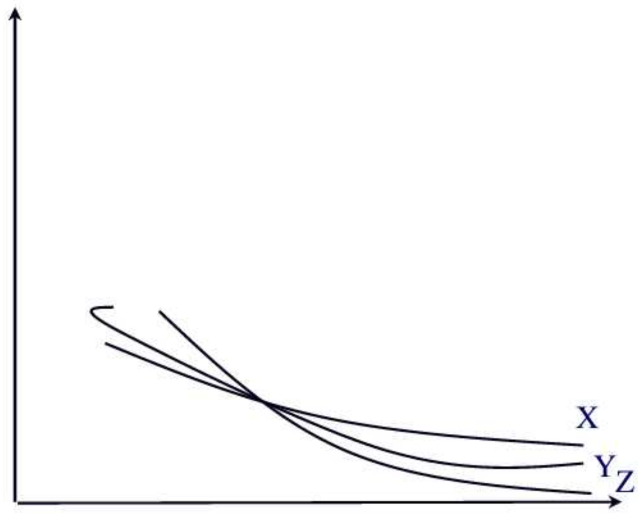
(i) $c=\infty$ (ii) $c=0$ (c) $c>0.$

Problem 9.19
Let X be a Pareto distribution with parameters $\alpha=4$ and $\theta=340$. Let Y be a Pareto distribution with parameters $\alpha=6$ and $\theta=340$.

Which of these has a heavier right tail relative to the other?

Problem 9.20
You are given the right-tails of the survival functions of three distributions

X, Y and Z. Order these distributions according to tail weight.



10 Tail Weight Measures: Hazard Rate Function and Mean Excess Loss Function

In this section we classify the tail weight of a distribution based on the hazard rate function and the mean excess loss function.

Classification Based on the Hazard Rate Function

Another way to classify the tail weight of a distribution is by using the hazard rate function:

$$h(x) = \frac{f(x)}{S(x)} = \frac{F'(x)}{1-F(x)} = -\frac{d}{dx}[\ln S(x)] = -\frac{S'(x)}{S(x)}.$$

By the existence of moments, the Pareto distribution is considered heavy-tailed. Its hazard rate function is

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha}{x+\theta}.$$

Note that $h'(x) = -\frac{\alpha}{(x+\theta)^2} < 0$ so that $h(x)$ is nonincreasing. Thus, it makes sense to say that a distribution is considered to be heavy-tailed if the hazard rate function is nonincreasing. Likewise, the random variable X with pdf $f(x) = xe^{-x}$ for $x > 0$ and 0 otherwise has a light-tailed distribution according to the existence of moments (See Problem 9.1). Its hazrad function is $h(x) = \frac{x}{x+1}$ which is a nondecreasing function. Hence, a nondecreasing hazard rate function is an indication of a light-tailed distribution.

Example 10.1

Let X be a random variable with survival function $f(x) = \frac{1}{x^2}$ if $x \geq 1$ and 0 otherwise. Based on the hazard rate function of the distribution, decide whether the distribution is heavy-tailed or light-tailed.

Solution.

The hazard rate function is

$$h(x) = -\frac{S'(x)}{S(x)} = -\frac{-\frac{2}{x^3}}{\frac{1}{x^2}} = \frac{2}{x}.$$

Hence, for $x \geq 1$,

$$h'(x) = -\frac{2}{x^2} < 0$$

which shows that $h(x)$ is nonincreasing. We conclude that the distribution of X is heavy-tailed ■

Remark 10.1
Under this definition, a constant hazard function can be called both nonincreasing and nondecreasing. We will refer to distributions with constant hazard function as **medium-tailed distribution**. Thus, the exponential random variable which was classified as light-tailed in Example 9.1, will be referred to as a medium-tailed distribution.

The next result provides a criterion for testing tail weight based on the probability density function.

Theorem 10.1
If for a fixed $y \geq 0$, the function $\frac{f(x+y)}{f(x)}$ is nonincreasing (resp. nondecreasing) in x then the hazard rate function is nondecreasing (resp. nonincreasing).

Proof.
We have

$$[h(x)]^{-1} = \frac{\int_x^\infty f(t)dt}{f(x)} = \frac{\int_0^\infty f(x+y)dy}{f(x)} = \int_0^\infty \left[\frac{f(x+y)}{f(x)}\right] dy.$$

Thus, if $\frac{f(x+y)}{f(x)}$ is nondecreasing in x for a fixed y, then $h(x)$ is a nonincreasing in x. Likewise, if $\frac{f(x+y)}{f(x)}$ is nonincreasing in x for a fixed y, then $h(x)$ is a nondecreasing in x ■

Example 10.2
Using the above theorem, show that the Gamma distribution with parameters $\theta > 0$ and $0 < \alpha < 1$ is heavy-tailed.

Solution.
We have

$$\frac{f(x+y)}{f(x)} = \left(1 + \frac{y}{x}\right)^{\alpha-1} e^{-\frac{y}{\theta}}$$

and

$$\frac{d}{dx}\left[\frac{f(x+y)}{f(x)}\right] = \frac{y(1-\alpha)}{x^2}\left(1 + \frac{y}{x}\right)^{\alpha-2} e^{-\frac{y}{\theta}} > 0$$

for $0 < \alpha < 1$. Thus, the hazard rate function is nonincreasing and the distribution is heavy-tailed ■

Next, the hazard rate function can be used to compare the tail weight of two distributions. For example, if X and Y are two distributions with increasing (resp. decreasing) hazard rate functions, the distribution of X has a lighter (resp. heavier) tail than the distribution of Y if $h_X(x)$ is increasing (resp. decreasing) at a faster rate than $h_Y(x)$ for a large value of the argument.

Example 10.3
Let X be the Pareto distribution with $\alpha = 2$ and $\theta = 150$ and Y be the Pareto distribution with $\alpha = 3$ and $\theta = 150$. Compare the tail weight of these distributions using
(a) the relative tail weight measure;
(b) the hazard rate measure.
Compare your results in (a) and (b).

Solution.
(a) Note that both distributions are heavy-tailed using the hazard rate analysis. However, $h'_Y(x) = -\frac{2}{(x+150)^2} < h'_X(x) = -\frac{1}{(x+150)^2}$ so that $h_Y(x)$ decreases at a faster rate than $h_X(x)$. Thus, X has a lighter tail than X.
(b) Using the relative tail weight, we find

$$\lim_{x\to\infty} \frac{f_X(x)}{f_Y(x)} = \lim_{x\to\infty} \frac{2(150)^2}{(x+150)^2} \cdot \frac{(x+150)^4}{3(150)^4} = \infty.$$

Hence, X has a heavier tail than Y which is different from the result in (a)! ■

Remark 10.2
Note that the Gamma distribution is light-tailed for all $\alpha > 0$ and $\theta > 0$ by the existence of moments analysis. However, the Gamma distribution is heavy-tailed for $0 < \alpha < 1$ by the hazard rate analysis. Thus, the concept of light/heavy right tailed is somewhat vague in this case.

<u>Classification Based on the Mean Excess Loss Function</u>
A fourth measure of tail weight is the mean excess loss function as introduced in Section 5. For a loss random variable X, the expected amount by which loss exceeds x, given that it does exceed x is

$$e(x) = e_X(x) = E[X - x|X > x] = \frac{E(X) - E(X \wedge x)}{1 - F(x)}.$$

In the context of life contingency models (See [3]), if X is the random variable representing the age at death and if $T(x)$ is the continuous random variable representing time until death of someone now alive at age x then $e(x)$ is denoted by $\mathring{e}_(x) = E[T(x)] = E[X - x|X > x]$. In words, for a newborm alive at age x, $\mathring{e}_(x)$ is the average additional number of years until death from age x, given that an individual has survived to age x. We call $\mathring{e}_(x)$ the **complete expectation of life** or the **residual mean lifetime**.

Viewed as a function of x, an increasing mean excess loss function is an indication of a heavy-tailed distribution. On the other hand, a decreasing mean excess loss function indicates a light-tailed distribution.

Next, we establish a relationship between $e(x)$ and the hazard rate function. We have

$$\begin{aligned} e(x) =&\frac{E(X) - E(X \wedge x)}{1 - F(x)} = \frac{\int_0^\infty S_X(y)dy - \int_0^x S_X(y)dy}{S_X(x)} \\ =&\frac{\int_x^\infty S_X(y)dy}{S_X(x)} = \int_0^\infty \left(\frac{S_X(x+y)}{S_X(x)}\right) dy. \end{aligned}$$

But one of the characteristics of the hazard rate function is that it can generate the survival function:

$$S_X(x) = e^{-\int_0^x h(t)dt}.$$

Hence, we can write

$$e(x) = \int_0^\infty \frac{e^{-\int_0^{x+y} h(u)du}}{e^{-\int_0^x h(u)du}} dy = \int_0^\infty e^{-\int_x^{x+y} h(u)du} dy, \ y > 0.$$

From the above discussion, we see that for a fixed $y > 0$, if $\frac{S_X(x+y)}{S_X(x)}$ is an increasing function of x (and therefore $e(x)$ is increasing) then the hazard rate function is decreasing and consequently the distribution is heavy-tailed. Likewise, if the $\frac{S_X(x+y)}{S_X(x)}$ is a decreasing function of x (and therefore $e(x)$ is decreasing) then the hazard rate function is increasing and consequently the distribution is light-tailed.

Example 10.4
Let X be a random variable with pdf $f(x) = 2xe^{-x^2}$ for $x > 0$ and 0 otherwise. Show that the distribution is light-tailed by showing $\frac{S_X(x+y)}{S_X(x)}$ is a decreasing function of x.

Solution.
We have $S_X(x) = \int_x^\infty 2te^{-t^2}dt = e^{-x^2}$. Thus, for a fixed $y > 0$, we have

$$\frac{S_X(x+y)}{S_X(x)} = e^{-2xy} - y^2$$

whose derivative with respect to x is

$$\frac{d}{dx}\left[\frac{S_X(x+y)}{S_X(x)}\right] = -2ye^{-2xy-y^2} < 0.$$

That is, $\frac{S_X(x+y)}{S_X(x)}$ is a decreasing function of x ■

Practice Problems

Problem 10.1
Show that the Gamma distribution with parameters $\theta > 0$ and $\alpha > 1$ is light-tailed by showing that $\frac{f(x+y)}{f(x)}$ is nonincreasing.

Problem 10.2
Show that the Gamma distribution with parameters $\theta > 0$ and $\alpha = 1$ is medium-tailed.

Problem 10.3
Let X be the Weibull distribution with probability density function $f(x) = \frac{\tau x^{\tau-1} e^{-\left(\frac{x}{\theta}\right)^\tau}}{\theta^\tau}$.

Using hazard rate analysis, show that the distribution is heavy-tailed for $0 < \tau < 1$ and light-tailed for $\tau > 1$.

Problem 10.4
Let X be a random variable with pdf $f(x) = 2xe^{-x^2}$ for $x > 0$ and 0 otherwise.

Determine the tail weight of this distributions using Theorem 10.1.

Problem 10.5
Using Theorem 10.1, show that the Pareto distribution is heavy-tailed.

Problem 10.6
Show that the hazard rate function of the Gamma distribution approaches $\frac{1}{\theta}$ as $x \to \infty$.

Problem 10.7
Show that $\lim_{x\to\infty} e(x) = \lim_{x\to\infty} \frac{1}{h(x)}$.

Problem 10.8
Find $\lim_{x\to\infty} e(x)$ where X is the Gamma distribution.

Problem 10.9
Let X be the Gamma distribution with $0 < \alpha < 1$ and $\theta > 0$. Show that $e(x)$ increases from $\alpha\theta$ to θ.

Problem 10.10
Let X be the Gamma distribution with $\alpha > 1$ and $\theta > 0$. Show that $e(x)$ decreases from $\alpha\theta$ to θ.

Problem 10.11
Find $\lim_{x\to\infty} e(x)$ where X is the Pareto distribution with parameters α and θ and conclude that the distribution is heavy-tailed.

Problem 10.12
Let X be a random variable with pdf $f(x) = \frac{1}{(1+x)^2}$ for $x > 0$ and 0 otherwise.

Find an expression of $\lim_{x\to\infty} e(x)$.

Problem 10.13
Let X be a random variable with mean excess loss function $e(x) = x + 1$.

(a) Find $S(x), f(x)$ and $h(x)$.
(b) Determine the tail behavior of X by using the moment criterion for tail weight.

Problem 10.14
Let X be a random variable with mean excess loss function $e(x) = x + 1$. Determine the tail behavior of X by using the hazard rate analysis.

Problem 10.15
Let X be a random variable with mean excess loss function $e(x) = x + 1$. Determine the tail behavior of X by using the mean excess loss function analysis.

Problem 10.16
Let X be a random variable with cdf $S(x) = e^{-\left(\frac{x}{\theta}\right)^2}$. Determine the tail behavior of X by using the mean excess loss function analysis.

Problem 10.17
Let X be the single-Pareto distribution with pdf

$$f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}.$$

Use Theorem 10.1, to show that X is heavy-tailed.

11 Equilibrium Distributions and Tail Weight

In this section, we shed more insight into the mean residual lifetime. Let X be a random variable such that $S(0) = 1$. Using Example 5.4 with $d = 0$, we can write $e(0) = E(X) = \int_0^\infty S(x)dx$. We define the random variable X_e with probability density function

$$f_e(x) = \frac{S(x)}{E(X)},\ x \geq 0$$

and 0 otherwise. We call the distribution of X_e, the **equilibrium distribution** or **integrated tail distribution.**

The corresponding survival function is

$$S_e(x) = \int_x^\infty f_e(t)dt = \int_x^\infty \left(\frac{S(t)}{E(X)}\right) dt,\ x \geq 0.$$

The corresponding hazard rate function is

$$h_e(x) = \frac{f_e(x)}{S_e(x)} = \frac{S(x)}{\int_x^\infty S(t)dt} = \frac{1}{e(x)}.$$

Thus, if $h_e(x)$ is increasing then the distribution X_e (and thus X) is ligh-tailed. If $h_e(x)$ is decreasing then the distribution X_e (or X) is heavy-tailed.

Example 11.1
Show that the equilibrium mean is given by

$$E(X_e) = \frac{E(X^2)}{2E(X)}.$$

Solution.
Using integration by parts , we find

$$\begin{aligned}
E(X^2) =& \int_0^\infty x^2 f(x)dx \\
=& -x^2 S(x)\Big|_0^\infty + 2\int_0^\infty xS(x)dx \\
=&2\int_0^\infty xS(x)dx
\end{aligned}$$

since

$$0 \leq x^2 S(x) = x^2 \int_x^\infty f(t)dt \leq \int_x^\infty t^2 f(t)dt$$

which implies

$$\lim_{x\to\infty} x^2S(x) = 0.$$

Now,

$$E(X_e) = \int_0^\infty xf_e(x)dx = \frac{1}{E(X)}\int_0^\infty xS(x)dx = \frac{E(X^2)}{2E(X)} \blacksquare$$

Example 11.2
Show that

$$S(x) = \frac{e(0)}{e(x)}e^{-\int_0^x\left[\frac{1}{e(t)}\right]dt}.$$

Solution.
Using $e(0) = E(X)$, We have

$$\begin{aligned}S(x) =&E(X)f_e(x) = e(0)f_e(x) = e(0)S_e(x)h_e(x)\\ =&e(0)h_e(x)e^{-\int_0^x\left[\frac{1}{e(t)}\right]dt}\\ =&\frac{e(0)}{e(x)}e^{-\int_0^x\left[\frac{1}{e(t)}\right]dt} \blacksquare\end{aligned}$$

Example 11.3
Show that

$$\frac{e(x)}{e(0)} = \frac{S_e(x)}{S(x)}.$$

Solution.
Since $S_e(x) = \frac{1}{E(X)}\int_x^\infty S(t)dt$, we have $\int_x^\infty S(t)dt = e(0)S_e(x)$. Since $\frac{S(x)}{\int_x^\infty S(t)dt} = \frac{1}{e(x)}$, we obtain $\int_x^\infty S(t)dt = e(x)S(x)$. Thus, $e(x)S(x) = e(0)S_e(x)$ or equivalently

$$\frac{e(x)}{e(0)} = \frac{S_e(x)}{S(x)} \blacksquare$$

If the mean residual life function is increasing (implied if the hazard rate function of X is decreasing by Section 10) then $e(x) \geq e(0)$ and

$$S_e(x) \geq S(x).$$

Integrating both sides of this inequality, we find

$$\int_0^\infty S_e(x)dx \geq \int_0^\infty S(x)dx$$

which implies

$$\frac{E(X^2)}{2E(X)} \geq E(X)$$

and this can be rewritten as

$$E(X^2) - [E(X)]^2 \geq [E(X)]^2$$

which gives

$$\mathrm{Var}(X) \geq [E(X)]^2.$$

Also,

$$[CV(x)]^2 = \frac{\mathrm{Var}(X)}{[E(X)]^2} \geq 1.$$

Example 11.4
Let X be the random variable with pdf $f(x) = \frac{2}{(1+x)^3}$ for $x \geq 0$ and 0 otherwise.
(a) Determine the survival function $S(x)$.
(b) Determine the hazard rate function $h(x)$.
(c) Determine $E(X)$.
(d) Determine the pdf of the equilibrium distribution.
(e) Determine the survival function $S_e(x)$ of the equilibrium distribution.
(f) Determine the hazard function of the equilibrium distribution.
(g) Determine the mean residual lifetime of X.

Solution.
(a) The survival function is

$$S(x) = \int_x^\infty \frac{2dt}{(1+t)^3} = -\frac{1}{(1+t)^2}\bigg|_x^\infty = \frac{1}{(1+x)^2}.$$

(b) The hazard rate function is

$$h(x) = \frac{f(x)}{S(x)} = \frac{2}{1+x}.$$

(c) We have

$$E(X) = \int_0^\infty xf(x)dx = \int_0^\infty \frac{2x}{(1+x)^3}dx = 1.$$

(d) We have

$$f_e(x) = \frac{S(x)}{E(X)} = \frac{1}{(1+x)^2}$$

for $x > 0$ and 0 otherwise.
(e) We have

$$S_e(x) = \int_x^\infty \frac{dt}{(1+t)^2} = -\frac{1}{1+t}\bigg|_x^\infty = \frac{1}{1+x}.$$

(f) We have

$$h_e(x) = \frac{f_e(x)}{S_e(x)} = \frac{1}{1+x}.$$

(g) We have

$$e(x) = \frac{1}{h_e(x)} = x+1, \ x \geq 0 \blacksquare$$

Practice Problems

Problem 11.1
Let X be the random variable with pdf $f(x) = 2xe^{-x^2}$ for $x > 0$ and 0 otherwise. Recall $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.
(a) Determine the survival function $S(x)$.
(b) Determine the equilibrium distribution $f_e(x)$.

Problem 11.2
Let X be a random variable with pdf $f(x) = \frac{1}{3}(1+2x)e^{-x}$ for $x > 0$ and 0 otherwise. Determine the hazard rate function of the equilibrium distribution. Hint: Example 5.3.

Problem 11.3
Let X be a random variable with mean excees loss function

$$e(x) = \frac{1}{1+x}, \ x > 0.$$

Determine the survival funtion of the distribution X.

Problem 11.4
Let X be a random variable with mean excees loss function

$$e(x) = \frac{1}{1+x}, \ x > 0.$$

Determine the survival function of the equilibrium distribution.

Problem 11.5
Let X be a random variable with mean excees loss function

$$e(x) = \frac{1}{1+x}, \ x > 0.$$

Determine the mean of the equilibrium distribution.

Problem 11.6
Let X be a random variable with pdf $f(x) = \frac{3}{8}x^2$ for $0 < x < 2$ and 0 otherwise.
(a) Find $E(X)$ and $E(X^2)$.
(b) Find the equilibrium mean.

Problem 11.7
Let X be a loss random variable with mean excess loss function

$$e(x) = 10 + 9x, \ \ x > 0.$$

Determine the survival function $S(x)$.

Problem 11.8
A random variable X has an exponential distribution with parameter λ. Calculate the equilibrium mean.

Risk Measures

Most insurance models are stochastic or probabilistic models, i.e. involve future uncertainty. As such, insurance companies are exposed to risks. Actuaries and risk managers job is to try to find the degree at which the insurance companies are subject to a particular aspects of risk. In this chapter, we provide a definition of risk measure and discuss a couple of ways of measuring risk.

12 Coherent Risk Measurement

In financial terms, a **risk measure** is the necessary capital to be put on reserve to support future risks associated, say to a portfolio. **Risk management** is about understanding and managing the potential losses in the total portfolio. One of the key tasks of risk management is to quantify the risk of the uncertainty in the future value of a portfolio. This quantification is usually achieved by modeling the uncertain payoff as a random variable, to which then a certain functional is applied. Such functionals are usually called **risk measures.** This functional gives a single value that is intended to provide a magnitude of the level of risk exposure.

Mathematically, we assign a random variable, defined on an appropriate probability space, to each portfolio loss over some fixed time interval. Let $\mathcal{L}$ be the collection of all such random variables. We will assume that $\mathcal{L}$ is a **convex cone** so that if L_1 and L_2 are members of $\mathcal{L}$ then $L_1 + L_2$ and cL_1 belong to $\mathcal{L}$ as well, where c is a constant. We define the **coherent risk measure** to be the functional $\rho : \mathcal{L} \longrightarrow [0, \infty)$ that satisfies the following properties:

(P1) *Subadditivity:* For any $L_1, L_2 \in \mathcal{L}$, we have $\rho(L_1+L_2) \le \rho(L_1)+\rho(L_2)$. This property says that the risk of two positions cannot get any worse than adding the two risks separately. This property reflects the idea that pooling risks helps to diversify a portfolio.

(P2) *Monotonicity:* If $L_1 \le L_2$ then $\rho(L_1) \le \rho(L_2)$.
From an economic viewpoint, this property is obvious−positions leading always to higher losses require more risk capital.

(P3) *Positive homogeneity:* $\rho(\alpha L) = \alpha\rho(L)$, $\alpha > 0$.
This property reflects the fact that there are no diversification benefits when we hold multiples of the same portfolio, L.

(P4) *Translation invariance:* For any real number α, $\rho(L + \alpha) = \rho(L) + \alpha$. That property states that by adding or subtracting a deterministic quantity α to a position leading to the loss L we alter our capital requirements by exactly that amount.

Remark 12.1
If L is a loss random variable then $\rho(L)$ may be interpreted as the riskiness

of a portfolio or the amount of capital that should be added to a portfolio with a loss given by L, so that the portfolio can then be deemed acceptable from a risk point of view. Coherent risk measures are important when the risk comes from separate risk-taking departments within the same company.

Example 12.1
Show that the expectation function $E(\cdot)$ is a coherent risk measure on $\mathcal{L}$.

Solution.
The expectation function $E(\cdot)$ satisfies the following properties:

(P1) $E(L_1 + L_2) = E(L_1) + E(L_2)$.
(P2) If $L_1 \leq L_2$ then $E(L_1) \leq E(L_2)$.
(P3) $E(\alpha L) = \alpha E(L),\ \alpha > 0$.
(P4) $E(L + \alpha) = E(L) + \alpha$ ■

Example 12.2
Show that the variance is not a cohorent risk measure.

Solution.
Since $\text{Var}(L + a) = \text{Var}(L) \neq \text{Var}(L) + a$, the variance of a distribution is not a cohorent risk measure ■

Example 12.3
Show that $\rho(L) = E(L) + \beta\text{Var}(L)$, where $\beta > 0$, satisfies the property of translation invariant but not positive homogeneity. We refer to this risk measure as the **variance premium principle**

Solution.
We have

$$\begin{aligned}
\rho(L + \alpha) =& E(L + \alpha) + \beta\text{Var}(L + \alpha)\\
=& E(L) + \alpha + \beta\text{Var}(L)\\
=& \rho(L) + \alpha\\
\rho(\alpha L) =& E(\alpha L) + \beta\text{Var}(\alpha L)\\
=& \alpha E(L) + \alpha^2\beta\text{Var}(L)\\
\neq& \alpha\rho(L)
\end{aligned}$$

where $\alpha > 0$ ■

Example 12.4
Show that $\rho(L) = \frac{1}{\alpha}\ln[E(e^{\alpha L})]$, where $\alpha, t > 0$, satisfies the properties of translation invariant and monotonicity. We refer to this risk measure as the **exponential premium principle**

Solution.
We have

$$\begin{aligned}\rho(L+\beta) =& \frac{1}{\alpha}\ln[E(e^{\alpha(L+\beta)})]\\ =& \frac{1}{\alpha}\ln[E(e^{\alpha L}e^{\alpha\beta})]\\ =& \frac{1}{\alpha}\ln[e^{\alpha\beta}E(e^{\alpha L})]\\ =& \frac{1}{\alpha}[\alpha\beta + \ln[E(e^{\alpha L})]\\ =& \rho(L) + \beta.\end{aligned}$$

Next, suppose that $L_1 \le L_2$. We have

$$\begin{aligned}e^{\alpha L_1} \le& e^{\alpha L_2}\\ E(e^{\alpha L_1}) \le& E(e^{\alpha L_2})\\ \ln[E(e^{\alpha L_1})] \le& \ln[E(e^{\alpha L_2})]\\ \rho(L_1) \le& \rho(L_2) \blacksquare\end{aligned}$$

Practice Problems

Problem 12.1
Show that $\rho(0) = 0$ and interpret this result.

Problem 12.2
Show that $\rho(\alpha L + \beta) = \alpha\rho(L) + \beta$, where $\alpha > 0$ and $\beta \in \mathbb{R}$.

Problem 12.3
Show that $\rho(L) = (1 + \alpha)E(L)$ is a coherent risk measure, where $\alpha \geq 0$. This risk measure is known as the **expected value premium principle**.

Problem 12.4
Which of the following is an implication of the subadditivity requirement for a coherent risk measure?

(a) If the subadditivity requirement is met, then a merger of positions creates extra risk.
(b) If the subadditivity requirement is met, then a merger of positions does not create extra risk.
(c) If the subadditivity requirement is met, then a merger of positions does not affect risk.

Problem 12.5
Which of the following is an implication of the monotonicity requirement for a coherent risk measure?

(a) Increasing the value of a portfolio increases risk.
(b) Increasing the value of a portfolio reduces risk.
(c) Increasing the value of a portfolio does not affect risk.

Problem 12.6
Which of the following is an implication of the positive homogeneity requirement for a coherent risk measure? More than one answer may be correct.

(a) If one assumes twice the amount of risk formerly assumed, one will need twice the capital.
(b) As the size of a position doubles, the risk stays unchanged.
(c) The risk of the position increases in a linear way with the size of the position.

Problem 12.7
Which of the following is an implication of the translation invariant requirement for a coherent risk measure? More than one answer may be correct.

(a) Adding a fixed amount to the initial financial position should increase the risk by the same amount.
(b) Subtracting a fixed amount to a portfolio decreases the required risk capital by the same amount.
(c) Getting additional capital, if it is from a risk-free source, cannot fundamentally alter the riskiness of a position.

Problem 12.8
Show that $\rho(L) = E(L) + \alpha E[L - E(L)]$ satisfies (P1), (P3), and (P4).

Problem 12.9
Find the numerical value of $\rho(L - \rho(L))$.

Problem 12.10
Show that $\rho(L) = E(L) + \sqrt{\text{Var}(L)}$ satisfies the properties of translation invariant and positive homogeneity. We refer to this risk measure as the **standard deviation principle**.

13 Value-at-Risk

A standard risk measure used to evaluate exposure to risk is the value-at-risk, abbreviated VaR. In general terms, the **value-at-risk** measures the potential loss of value of an asset or a portfolio over a defined time with a high level of certainty. For example, if the VaR is \$1 million at one-month, 99% confidence level, then there is 1% chance that under normal market movements the monthly loss will exceed \$1 million. Bankers use VaR to capture the potenetial losses in their traded portfolios from adverse market movements over a period of time; then they use it to compare with their available capital and cash reserves to ensure that the losses can be covered withoud putting the firm at risk.

Mathematically, the value-at-risk is found as follows: Let L be a loss random variable. The Value-at-risk at the $100p\%$ level, denoted by $\text{VaR}_p(L)$ or π_p, is the $100p$th percentile or the p quantile of the distribution L. That is, π_p is the solution of $F_L(\pi_p) = p$ or equivalently, $S_L(\pi_p) = 1 - p$.

Example 13.1
Let L be an exponential loss random variable with mean $\lambda > 0$. Find π_p.

Solution.
The pdf of L is $f(x) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}$ for $x > 0$ and 0 otherwise. Thus, $F(x) = 1 - e^{-\frac{x}{\lambda}}$. Now, solving the equation $F(\pi_p) = p$, that is, $1 - e^{-\frac{\pi_p}{\lambda}} = p$, we obtain $\pi_p = -\lambda \ln(1 - p)$ ■

Example 13.2
The loss random variable L has a Pareto distribution with parameters α and θ. Find π_p.

Solution.
The pdf of L is $f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}$ for $x > 0$ and 0 otherwise. The cdf is $F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha$. Solving the equation $F(\pi_p) = p$, we find

$$\pi_p = \theta[(1 - p)^{-\frac{1}{\alpha}} - 1] \blacksquare$$

Example 13.3
The loss random variable L follows a normal distribution with mean μ and standard deviation σ. Find π_p.

Solution.
Let $Z = \frac{L-\mu}{\sigma}$. Then Z is the standard normal distribution. The $p-$ quantile of Z satisfies the equation $\Phi(z) = p$. Thus, $z = \Phi^{-1}(p)$. Hence,

$$\pi_p = \mu + \sigma z = \mu + \sigma\Phi^{-1}(p) \blacksquare$$

Example 13.4
Consider a sample of size 8 in which the observed data points were 3,5,6,6,6,7,7, and 10. Find $\mathrm{VaR}_{0.90}(L)$ for this empirical distribution.

Solution.
The pmf of L is given below.

x	3	5	6	7	10
$p(x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

We want to find $\pi_{0.90}$ such that

$$\Pr(L < \pi_{0.90}) < 0.90 \le \Pr(L \le 0.90).$$

Thus, $\pi_{0.90} = 10$ $\blacksquare$

Remark 13.1
According to [1], $\mathrm{VaR}_p(L)$ is monotone, positive homogeneous, and translation invariant but not subadditive. Thus, $\mathrm{VaR}_p(L)$ is not a coherent risk measure.

Practice Problems

Problem 13.1
The loss random variable L has a uniform distribution in $[a, b]$. Find $\text{VaR}_p(L)$.

Problem 13.2
The cdf of a loss random variable L is given by

$$F_L(x) = \begin{cases} \frac{x^2}{4}, & 0 < x \leq 2 \\ 1, & x > 2. \end{cases}$$

Find $\pi_{0.90}$.

Problem 13.3
You are given the following empirical distribution

$$3, 5, 6, 6, 6, 7, 7, 10.$$

The risk measure under the standard deviation principle is $\rho(L) = E(L) + \alpha\sigma(L)$. Determine the value of α so that $\rho(L) = \pi_{0.90}$.

Problem 13.4
Losses represented by L are distributed as a Pareto distribution with parameters $\alpha = 2$ and $\theta = 60$. Find $\text{VaR}_{0.75}(L)$.

Problem 13.5
Losses represented by L are distributed as a single Pareto distribution with a pdf $f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, x > \theta$ and 0 otherwise. Find π_p.

Problem 13.6
A loss random variable X has a survival function

$$S(x) = \left(\frac{\theta}{x + \theta}\right)^2, \ x > 0.$$

Find θ given that $\pi_{0.75} = 40$.

Problem 13.7
Let L be a random variable with discrete loss distribution given by

x	0	100	1000	10000	100000
$p(x)$	0.65	0.20	0.07	0.05	0.03

Calculate the Value-at-Risk of L at the 90% level.

Problem 13.8
A loss random variable L has a two-parameter Pareto distribution satisfying:

$$\mathrm{VaR}_{0.90}(L) = 216.23 \text{ and } \mathrm{VaR}_{0.99}(L) = 900.$$

Calculate $\mathrm{VaR}_{0.95}(L)$.

Problem 13.9
Let L be a loss random variable with probability generating function

$$P_L(x) = 0.4x^2 + 0.2x^3 + 0.2x^5 + 0.2x^8.$$

Determine $\mathrm{VaR}_{0.80}(L)$.

Problem 13.10
A loss random variable L has a survival function

$$S(x) = \left(\frac{100}{x+100}\right)^2, \ x > 0.$$

Calculate $\mathrm{VaR}_{0.96}$ and interpret this result.

14 Tail-Value-at-Risk

The quantile risk meaure discussed in the previous section provides us only with the probability that a loss random variable L will exceed the $\mathrm{VaR}_p(L)$ for a certain confidence level. It does not provide any information about how large the losses are beyond a particular percentile. The **Tail-Value-at-Risk** (TVaR) measure does consider losses above a percentile. Other names used for TVaR are **Tail Conditional Expectation** and **Expected Shortfall**.

The Tail-Value-at-Risk of a random variable L at the 100p% security level is defined as

$$TVaR_p(L) = E[L|L > VaR_p(L)] = E[L|L > F_L^{-1}(p)]$$

where $F_L(x)$ is the distribution of L. This is the expected value of the loss, conditional on the loss exceeding π_p. Note that TVaR is also the expected cost per payment with a franchise deductible[7] of π_p.

We can write TVaR as

$$TVaR_p(L) = E[L|L > VaR_p(L)] = \frac{\int_{\pi_p}^{\infty} xf(x)dx}{1 - F(\pi_p)} = \frac{\int_{\pi_p}^{\infty} xf(x)dx}{1-p}.$$

Now, using the substitution $u = F(x)$, we can write

$$\mathrm{TVaR}_p(L) = \frac{\int_p^1 xf(x)dx}{1-p} = \frac{\int_p^1 \mathrm{VaR}_u(L)du}{1-p}.$$

TVaR can also be written as

$$\mathrm{TVaR}_p(L) = \frac{\pi_p}{1-p}\int_{\pi_p}^{\infty} f(x)dx + \frac{1}{1-p}\int_{\pi_p}^{\infty}(x-\pi_p)f(x)dx = \pi_p + E[X-\pi_p|X > \pi_p].$$

Since

$$E[X - \pi_p|X > \pi_p] = \frac{E(L) - E(L \wedge \pi_p)}{1-p},$$

we can write

$$\mathrm{TVaR}_p(L) = \mathrm{VaR}_p(L) + \frac{E(L) - E(L \wedge \pi_p)}{1-p}.$$

[7] See Section 32.

Remark 14.1
Unlike the VaR risk measure, TVaR risk measure is shown to be coherent.

Example 14.1
Find the Tail-Value-at-Risk of an exponential distribution with mean $\lambda > 0$.

Solution.
From Problem 5.7, we have $e(\pi_p) = \lambda$. This and Example 13.1 give

$$TVaR_p(L) = \lambda - \lambda \ln(1-p) \blacksquare$$

Example 14.2
Find the Tail-Value-at-Risk of a Pareto distribution with parameters $\alpha > 1$ and $\theta > 0$.

Solution.
The survival function of the Pareto distribution is

$$S(x) = \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \ x > 0.$$

Thus,

$$\begin{aligned}
e(\pi_p) =& \frac{1}{S(\pi_p)} \int_{\pi_p}^{\infty} S(x)dx \\
=& (\pi_p + \theta)^{\alpha} \int_{\pi_p}^{\infty} (x+\theta)^{-\alpha} dx \\
=& \frac{(\pi_p+\theta)^{\alpha}}{\alpha - 1} \left. (x+\theta)^{1-\alpha} \right|_{\pi_p}^{\infty} \\
=& \frac{\pi_p + \theta}{\alpha - 1}.
\end{aligned}$$

On the other hand, using Example 13.2, we have

$$\pi_p = \theta[(1-p)^{-\frac{1}{\alpha}} - 1].$$

Hence,

$$\text{TVaR}_p(L) = \frac{\pi_p + \theta}{\alpha - 1} + \theta[(1-p)^{-\frac{1}{\alpha}} - 1] \blacksquare$$

Example 14.3
Let Z be the standard normal distribution with pdf $f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Find $\text{TVaR}_p(Z)$.

Solution.
Notice first that $f_Z(x)$ satisfies the differential equation $xf_X(x) = -f'_Z(x)$. Using the Fundamental Theorem of Calculus, we can write

$$\begin{aligned}
\mathrm{TVaR}_p(Z) =& \frac{1}{1-p}\int_{\Phi^{-1}(p)}^{\infty} xf_Z(x)dx \\
=& -\frac{1}{1-p}\int_{\Phi^{-1}(p)}^{\infty} f'_Z(x)dx \\
=& -\frac{1}{1-p} f_Z(x)\Big|_{\Phi^{-1}(p)}^{\infty} \\
=& \frac{1}{1-p} f_Z[\Phi^{-1}(p)] \blacksquare
\end{aligned}$$

Example 14.4
Let L be a loss random variable having a normal distribution with mean μ and standard deviation σ. Find $\mathrm{TVaR}_p(L)$.

Solution.
Since $\mathrm{TVaR}_p(L)$ is a coherent risk measure, it is positive homogeneous and translation invariant. Thus, we have

$$\mathrm{TVaR}_p(L) = \mathrm{TVaR}_p(\mu + \sigma Z) = \mu + \sigma \mathrm{TVaR}_p(Z) = \mu + \frac{\sigma}{1-p} f_Z[\Phi^{-1}(p)] \blacksquare$$

Practice Problems

Problem 14.1
Let L be a loss random variable with uniform distribution in (a, b).

(a) Find the mean residual life $e(x)$.
(b) Find $\mathrm{TVaR}_p(L)$.

Problem 14.2
The cdf of a loss random variable L is given by

$$F_L(x) = \begin{cases} \frac{x^2}{4}, & 0 < x \leq 2 \\ 1, & x > 2. \end{cases}$$

(a) Find $\pi_{0.90}$ and $e(\pi_{0.90})$.
(b) Find $\mathrm{TVaR}_{0.90}(L)$.

Problem 14.3
You are given the following empirical distribution

$$3, 5, 6, 6, 6, 7, 7, 10.$$

Find $\pi_{0.85}$ and $\mathrm{TVaR}_{0.85}(L)$.

Problem 14.4
Losses are distributed as Pareto distributions with mean of 200 and variance of 60000.

(a) Find the values of the parameters α and β.
(b) Find $e(100)$.
(c) Find $\pi_{0.95}$.
(d) Find $\mathrm{TRaV}_{0.95}(L)$.

Problem 14.5
Losses represented by the random variable L are uniformly distributed from 0 to the maximum loss. You are given that $\mathrm{Var}(L) = 62,208$.

Find $\mathrm{TVaR}_{0.75}(L)$.

Problem 14.6
Losses represented by the random variable L are uniformly distributed in $(0, 864)$. Determine β so that the standard deviation principle is equal to $\mathrm{TVaR}_{0.75}(L)$.

Problem 14.7
Diabetes claims follow an exponential distribution with parameter $\lambda = 2$. Find $\text{TVaR}_{0.90}(L)$.

Problem 14.8
You are given the following empirical distribution

$$3, 5, 6, 6, 6, 7, 7, 10.$$

Let β_1 be the value of β in the standard deviation principle such that $\mu + \beta_1\sigma = \text{VaR}_{0.85}$. Let β_2 be the value of β in the standard deviation principle such that $\mu + \beta_2\sigma = \text{TVaR}_{0.85}(L)$.

Calculate $\beta_2 - \beta_1$.

Problem 14.9
Let L_1 be a Pareto random variable with parameters $\alpha = 2$ and $\theta = 100$. Let L_2 be a random variable with uniform distribution on $(0, 864)$. Find p such that

$$\frac{\text{TVaR}_{0.99}(L_1)}{\text{VaR}_{0.99}(L_1)} = \frac{\text{TVaR}_p(L_2)}{\text{VaR}_p(L_2)}.$$

Problem 14.10
Let L be a random variable with discrete loss distribution given by

x	0	100	1000	10000	100000
$p(x)$	0.65	0.20	0.07	0.05	0.03

Calculate the Tail-Value-at-Risk of L and the 90% level.

Problem 14.11
Find $\text{TVaR}_{0.95}(L)$ when L has a normal distribution with mean of 100 and standard deviation of 10.

Characteristics of Actuarial Models

In the previous chapter, a characteristic of an actuarial model is the tail weight of the corresponding distribution. In this chapter we look at other factors that characterize a model from another model. One such a factor is the number of parameters needed in the determination of the model's distribution. More parameters involved in a model means that more information is required and in this case the model is categorized as a complex model. We start this chapter by discussing first simple models and then move toward more complex models.

15 Parametric and Scale Distributions

These are considered the simplest families of actuarial models. We will consider a model that requires less parameters than another model as less complex.

A **parametric distribution** is one that is completely determined by a set of quantities called **parameters**. Examples of commonly used parametric distributions are listed below.

Name	PDF	Parameters
Exponential	$f(x) = \theta e^{-\theta x}$	$\theta > 0$
Pareto	$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}$	$\alpha > 0, \theta > 0$
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ, σ
Poisson	$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$	$\lambda > 0$

Additional parametric distributions can be found in the Tables of Exam C.

Now, when multiplying a random variable by a positive constant and the resulting distribution belongs to the same family of distributions of the original random variable then we call the distribution **scale distribution.**

Example 15.1
Show that the Pareto distribution is a scale distribution.

Solution.
The cdf of the Pareto distribution is

$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha.$$

Let $Y = cX$. Then

$$\begin{aligned} F_Y(y) =& \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right) \\ =& 1 - \left(\frac{c\theta}{y + c\theta}\right)^\alpha. \end{aligned}$$

This is a Pareto distribution with parameters α and $c\theta$ ■

Example 15.2
Show that the Weibull distribution with parameters θ and τ is a scale distribution.

Solution.
The cdf of the Weibull distribution is

$$F_X(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^\tau}.$$

Let $Y = cX$. Then

$$\begin{aligned} F_Y(y) =& \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right) \\ =& 1 - e^{-\left(\frac{x}{c\theta}\right)^\tau}. \end{aligned}$$

This is a Weibull distribution with parameters $c\theta$ and τ ■

A parameter θ in a scale distribution X is called a **scale parameter** if $c\theta$ is a parameter of cX and θ is the only changed parameter.

Example 15.3
Show that the parameter θ in the Pareto distribution is a scale parameter.

Solution.
This follos from Example 15.1 ■

Example 15.4
Find the scale parameter of the Weibull distribution with parameters θ and τ.

Solution.
According to Example 15.2, the scale parameter is θ ■

Example 15.5
The amount of money in dollars that Clark received in 2010 from his investment in Simplicity futures follows a Pareto distribution with parameters $\alpha = 3$ and θ. Annual inflation in the United States from 2010 to 2011 is i%. The 80$^{\text{th}}$ percentile of the earning size in 2010 equals the mean earning size in 2011. If Clark's investment income keeps up with inflation but is otherwise unaffected, determine i.

Solution.
Let X be the earning size in 2010 and Y that in 2011. Then Y is a Pareto distribution with parameters $\alpha = 3$ and $(1+i)\theta$. We are told that

$$\pi_{0.80} = E(Y) = (1+i)E(X) = \frac{(1+i)\theta}{3-1} = \frac{(1+i)\theta}{2}.$$

Thus,

$$0.8 = \Pr\left(X < \frac{(1+i)\theta}{2}\right) = F_X\left(\frac{(1+i)\theta}{2}\right) = 1 - \left(\frac{\theta}{\frac{(1+i)\theta}{2} + \theta}\right)^3.$$

Solving the above equation for i, we find $i = 0.42$ ■

By assigning all possible numerical values to the parameters of a particular parametric distribution we obtain a family of distributions that we call a **parametric distribution family**.

Example 15.6
Show that exponential distributions belong to the Weibull distribution family with parameters θ and τ.

Solution.
Weibull distributions with parameters θ and τ have pdf $f_X(x) = \frac{\tau\left(\frac{x}{\theta}\right)^\tau e^{-\left(\frac{x}{\theta}\right)^\tau}}{x}$. Letting $\tau = 1$, the pdf reduces to $f_X(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}$ which is the pdf of an exponential distribution ■

Practice Problems

Problem 15.1
Show that the exponential distribution is a scale distribution.

Problem 15.2
Let X be a random variable with pdf $f_X(x) = 2xe^{-x^2}$ for $x > 0$ and 0 otherwise. Let $Y = cX$ for $c > 0$.

Find $F_Y(y)$.

Problem 15.3
Let X be a uniform random variable on the interval $(0, \theta)$. Let $Y = cX$ for $c > 0$.

Find $F_Y(y)$.

Problem 15.4
Show that the Fréchet distribution with cdf $F_X(x) = e^{-\left(\frac{x}{\theta}\right)^{-\alpha}}$ and parameters θ and α is a scale distribution.

Problem 15.5
Show that the three-parameter Burr distribution with cdf $F_X(x) = 1 - \frac{1}{\left[1+\left(\frac{x}{\theta}\right)^{\gamma}\right]^{\alpha}}$ is a scale distribution.

Problem 15.6
Find the scale parameter of the following distributions:

(a) The exponential distribution with parameter θ.
(b) The uniform distribution on $(0, \theta)$.
(c) The Fréchet distribution with parameters θ and α.
(d) The Burr distribution with parameters α, θ, and γ.

Problem 15.7
Claim severities are modeled using a continuous distribution and inflation impacts claims uniformly at an annual rate of r.

Which of the following are true statements regarding the distribution of claim severities after the effect of inflation?

(1) An exponential distribution will have a scale parameter of $(1 + r)\theta$.
(2) A Pareto distribution will have scale parameters $(1 + r)\alpha$ and $(1 + r)\theta$
(3) A Burr distribution will have scale parameters $\alpha, (1 + r)\theta, \gamma$.

Problem 15.8
Let X be the lognormal distribution with parameters μ and σ and cdf $F_X(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$.

Show that X has a scale distribution with no scale parameters.

Problem 15.9
Show that the Gamma distribution is a scale distribution. Is there a scale parameter?

Problem 15.10
Earnings during 2012 follow a Gamma distribution with variance 2,500. For the 2013, earnings are expected to be subject to $P\%$ inflation and the expected variance for the 2013 year is 10,000.

Determine the value of P.

Problem 15.11
The Gamma distribution with parameters α and θ has the pdf $f_X(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\theta}}}{\Gamma(\alpha)}$.

Show that the exponential distributions belong to this family.

Problem 15.12
Hardy Auto Insurance claims X are represented by a Weibull distribution with parameters $\alpha = 2$ and $\theta = 400$. It is found that the claim sizes are inflated by 30% uniformly.

Calculate the probability that a claim will be at least 90 counting inflation.

16 Discrete Mixture Distributions

In probability and statistics, a **mixture distribution** is the probability distribution of a random variable whose values can be interpreted as being derived from an underlying set of other random variables. For example, a dental claim may be from a check-up, cleaning, filling cavity, a surgical procedure, etc.

A random variable X is a **k-point mixture** of the random variables $X_1, \cdots, X_k$ if its cumulative distribution function (cdf) is given by

$$F_X(x) = a_1 F_{X_1}(x) + a_2 F_{X_2}(x) + \cdots + a_k F_{X_k}(x)$$

where each **mixing weight** $a_i > 0$ and $a_1 + a_2 + \cdots a_k = 1$. The mixture X is defined in terms of its pdf or cdf and is not the sum of the random variables $a_1X_1, a_2X_2, \cdots, a_nX_n$.

The mixing weights are discrete probabilities. To see this, let Θ be the discrete random variable with support $\{1, 2, \cdots, k\}$ and pmf $\Pr(\Theta = i) = a_i$. We can think of the distribution of Θ as a conditioning distribution where $X = X_i$ is conditioned on $\Theta = i$, or equivalently, $F_{X|\Theta}(x|\Theta = i) = F_{X_i}(x)$. In this case, X is the unconditional distribution with cdf

$$F_X(x) = a_1 F_{X_1}(x) + a_2 F_{X_2}(x) + \cdots + a_k F_{X_k}(x) = \sum_{i=1}^{k} F_{X|\Theta}(x|\Theta = i)\Pr(\Theta = i).$$

In actuarial and insurance terms, discrete mixtures arise in situations where the risk class of a policyholder is uncertain, and the number of possible risk classes is discrete.

The continuous mixture of distributions will be discussed in Section 20 of this study guide.

Example 16.1
Let Y be a 2-point mixture of two random variables X_1 and X_2 with mixing weights 0.6 and 0.4 respectively. The random variable X_1 is a Pareto random variable with parameters $\alpha = 3$ and $\theta = 900$. The random variable X_2 is a Pareto random variable with parameters $\alpha = 5$ and $\theta = 1500$. Find the pdf of Y.

Solution.
We are given

$$f_{X_1} = \tfrac{3(900)^3}{(x+900)^4} \quad \text{and} \quad f_{X_2}(x) = \tfrac{5(1500)^5}{(x+1500)^6}.$$

Thus,

$$f_Y(x) = 0.6f_{X_1} + 0.4f_{X_2} = 0.6\left[\frac{3(900)^3}{(x+900)^4}\right] + 0.4\left[\frac{5(1500)^5}{(x+1500)^6}\right] \blacksquare$$

Example 16.2 ‡
The random variable N has a mixed distribution:
(i) With probability p, N has a binomial distribution with $q = 0.5$ and $m = 2$.
(ii) With probability $1-p$, N has a binomial distribution with $q = 0.5$ and $m = 4$.
Calculate $\Pr(N = 2)$.

Solution.
We have

$$\begin{aligned}
p_{N_1}(N_1 = 2) =& C(2,2)(0.5)^2 = 0.25\\
p_{N_2}(N_2 = 2) =& C(4,2)(0.5)^2(0.5)^2 = 0.375\\
p_N(N = 2) =& pp_{N_1}(N_1 = 2) + (1-p)p_{N_2}(N_1 = 2) = 0.375 - 0.125p \blacksquare
\end{aligned}$$

Example 16.3
Determine the mean and second moment of the two-point mixture distribution with the cdf

$$F_X(x) = 1 - \alpha\left(\frac{\theta_1}{x+\theta_1}\right)^\alpha - (1-\alpha)\left(\frac{\theta_2}{x+\theta_2}\right)^{\alpha+2}.$$

Solution.
The first part is the distribution of a Pareto random variable X_1 with parameters $\alpha_1 = \alpha$ and θ_1. The second part is the distribution of a Pareto random variable X_2 with parameters $\alpha_2 = \alpha + 2$ and θ_2. Thus,

$$\begin{aligned}
E(X_1) =& \frac{\theta_1}{\alpha_1 - 1} = \frac{\theta_1}{\alpha - 1}\\
E(X_2) =& \frac{\theta_2}{\alpha_2 - 1} = \frac{\theta_2}{\alpha + 1}\\
E(X) =& \alpha\frac{\theta_1}{\alpha} + (1-\alpha)\frac{\theta_2}{\alpha+1}
\end{aligned}$$

$$E(X_1^2) = \frac{\theta_1^2 2!}{(\alpha_1 - 1)(\alpha_1 - 2)} = \frac{2\theta_1^2}{(\alpha - 1)(\alpha - 2)}$$
$$E(X_2^2) = \frac{\theta_2^2 2!}{(\alpha_2 - 1)(\alpha_2 - 2)} = \frac{2\theta_2^2}{(\alpha)(\alpha + 1)}$$
$$E(X^2) = \alpha\left(\frac{2\theta_1^2}{(\alpha - 1)(\alpha - 2)}\right) + (1 - \alpha)\left(\frac{2\theta_2^2}{(\alpha)(\alpha + 1)}\right) \blacksquare$$

Next, we consider mixtures where the number of random variables in the mixture is unknown. A **variable-component mixture distribution** has a distribution function that can be written as

$$F_X(x) = a_1 F_{X_1}(x) + a_2 F_{X_2}(x) + \cdots + a_N F_{X_N}(x)$$

where each $a_j > 0$ and $\sum_{i=1}^N a_i = 1$ and $N \in \mathbb{N}$.

In a variable-component mixture distribution, each of the mixture weights associated with each individual $F_{X_j}(x)$ is a parameter. Also, there are $(N-1)$ parameters corresponding to the weights a_1 through a_{N-1}. The weight a_N is not itself a parameter, since the value of a_N is determined by the value of the constants a_1 through a_{N-1}.

Example 16.4
Determine the distribution, density, and hazard rate functions for the variable mixture of exponential distributions.

Solution.
The distribution function of the variable mixture is

$$F_X(x) = 1 - a_1 e^{-\frac{x}{\theta_1}} - a_2 e^{-\frac{x}{\theta_2}} - \cdots - a_N e^{-\frac{x}{\theta_N}}$$

where $a_j > 0$ and $\sum_{i=1}^N a_i = 1$.
The density function is

$$f_X(x) = \frac{a_1}{\theta_1} e^{-\frac{x}{\theta_1}} + \frac{a_2}{\theta_2} e^{-\frac{x}{\theta_2}} + \cdots + \frac{a_N}{\theta_N} e^{-\frac{x}{\theta_N}}$$

and the hazard rate function is

$$h_X(x) = \frac{\frac{a_1}{\theta_1} e^{-\frac{x}{\theta_1}} + \frac{a_2}{\theta_2} e^{-\frac{x}{\theta_2}} + \cdots + \frac{a_N}{\theta_N} e^{-\frac{x}{\theta_N}}}{a_1 e^{-\frac{x}{\theta_1}} + a_2 e^{-\frac{x}{\theta_2}} + \cdots + a_N e^{-\frac{x}{\theta_N}}}.$$

Note that the parameters corresponding to $N = 3$ are $(a_1, a_2, \theta_1, \theta_2, \theta_3)$ and those corresponding to $N = 5$ are $(a_1, a_2, a_3, a_4, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ $\blacksquare$

Example 16.5 ‡
You are given claim count data for which the sample mean is roughly equal to the sample variance. Thus you would like to use a claim count model that has its mean equal to its variance. An obvious choice is the Poisson distribution.
Determine which of the following models may also be appropriate.
(A) A mixture of two binomial distributions with different means
(B) A mixture of two Poisson distributions with different means
(C) A mixture of two negative binomial distributions with different means
(D) None of (A), (B) or (C)
(E) All of (A), (B) and (C).

Solution.
Let X be a 2-point mixture of the random variables X_1 and X_2 with mixing weights α and $1-\alpha$. Let Θ be the discrete random variable such that $\Pr(\Theta = 1) = \alpha$ and $\Pr(\Theta = 2) = 1-\alpha$. Thus, we have

$$\begin{aligned}
E(X) =&\alpha E(X_1) + (1-\alpha)E(X_2)\\
\mathrm{Var}(X) =&E(X^2) - E(X)^2\\
=&\alpha E(X_1^2) + (1-\alpha)E(X_2^2) - [\alpha E(X_1) + (1-\alpha)E(X_2)]^2\\
=&\alpha \mathrm{Var}(X_1) + (1-\alpha)\mathrm{Var}(X_2) + \alpha(1-\alpha)[E(X_1) - E(X_2)]^2.
\end{aligned}$$

If X_1 and X_2 are Poisson with means λ_1 and λ_2 respectively with $\lambda_1 \neq \lambda_2$, then

$$\begin{aligned}
\mathrm{Var}(X) =&\alpha\lambda_1 + (1-\alpha)\lambda_2 + \alpha(1-\alpha)(\lambda_1 - \lambda_2)^2\\
>&\alpha\lambda_1 + (1-\alpha)\lambda_2 = E(X).
\end{aligned}$$

If X_1 and X_2 are negative binomial with parameters (r_1, β_1) and (r_2, β_2) respectively with $r_1\beta_1 \neq r_2\beta_2$, then

$$\begin{aligned}
\mathrm{Var}(X) =&\alpha r_1\beta_1(1+\beta_1) + (1-\alpha)r_2\beta_2(1+\beta_2) + \alpha(1-\alpha)(r_1\beta_1 - r_2\beta_2)^2\\
>&\alpha r_1\beta_1 + (1-\alpha)r_2\beta_2 = E(X).
\end{aligned}$$

If X_1 and X_2 are binomial with parameters (m_1, q_1) and (m_2, q_2) respectively with $m_1q_1 \neq m_2q_2$, then

$$\begin{aligned}
\mathrm{Var}(X) =&\alpha m_1q_1(1-q_1) + (1-\alpha)m_2q_2(1-q_2) + \alpha(1-\alpha)(m_1q_1 - m_2q_2)^2\\
=&E(X) + \alpha(1-\alpha)(m_1q_1 - m_2q_2)^2 - \alpha m_1q_1^2 - (1-\alpha)m_2q_2^2.
\end{aligned}$$

The expression $\alpha(1-\alpha)(m_1q_1 - m_2q_2)^2 - \alpha m_1q_1^2 - (1-\alpha)m_2q_2^2$ can be positive, negative, or zero. Thus, a mixture of two binomial distributions with different means may result in the variance being equal to the mean ■

Example 16.6 ‡
Losses come from an equally weighted mixture of an exponential distribution with mean m_1 , and an exponential distribution with mean m_2. Determine the least upper bound for the coefficient of variation of this distribution.

Solution.
Let X be the random variable with pdf

$$f(x) = \frac{1}{2}\left(\frac{1}{m_1}e^{-\frac{x}{m_1}} + \frac{1}{m_2}e^{-\frac{x}{m_2}}\right).$$

We have

$$\begin{aligned}
E(X) =& \frac{1}{2}(m_1 + m_2)\\
E(X^2) =& \frac{1}{2}(2m_1^2 + 2m_2^2)\\
\mathrm{Var}(X) =& \frac{1}{2}(2m_1^2 + 2m_2^2) - \left[\frac{1}{2}(m_1 + m_2)\right]^2.
\end{aligned}$$

The square of coefficient of variation of X is

$$\begin{aligned}
CV^2 =& \frac{\frac{1}{2}(2m_1^2 + 2m_2^2) - \left[\frac{1}{2}(m_1 + m_2)\right]^2}{\left[\frac{1}{2}(m_1 + m_2)\right]^2}\\
=& \frac{3m_1^2 - 2m_1m_2 + 3m_2^2}{(m_1 + m_2)^2}\\
=& 3 - \frac{8m_1m_2}{(m_1 + m_2)^2}.
\end{aligned}$$

Let $r = \frac{m_1}{m_2}$. Then

$$CV^2 = 3 - \frac{8r}{(1 + r)^2}.$$

Thaking the derivative and setting it to 0, we find

$$-8(1 + r)^{-2} + 16r(1 + r)^{-3} = 0 \Longrightarrow r = 1.$$

Moreover, $r''(1) = 1 > 0,\ r(0) = r(\infty) = 3$ so that $r = 1$ is a global minimum. Hence, the least upper bound of the coefficient of variation is $\sqrt{3}$ ■

Practice Problems

Problem 16.1
The distribution of a loss, X, is a 2-point mixture:
(i) With probability 0.6, X_1 is a Pareto distribution with parameters $\alpha = 3$ and $\theta = 900$.
(ii) With probability 0.4, X_2 is a Pareto distribution with parameters $\alpha = 5$ and $\theta = 1500$.

Determine $\Pr(X > 1000)$.

Problem 16.2
The distribution of a loss, X, is a 2-point mixture:
(i) With probability 0.5, X_1 is a Burr distribution with parameters $\alpha = 1, \gamma = 2$ and $\theta = 1000^{0.5}$.
(ii) With probability 0.5, X_2 is a Pareto distribution with parameters $\alpha = 1$ and $\theta = 1000$.

Determine the median of X.

Problem 16.3
You are given:
• X is a 2-point mixture of two exponential random variables X_1 and X_2 with parameters $\theta_1 = 1$ and $\theta_2 = 3$ and mixing weights $\frac{1}{2}$ and $\frac{1}{6}$ respectively.
• $Y = 2X$ and Y is a mixture of two exponential random variables Y_1 and Y_2.

Find $E(Y_1)$ and $E(Y_2)$.

Problem 16.4
The severity distribution function for losses from your renters insurance is the following:

$$F_X(x) = 1 - 0.3\left(\frac{1000}{1000+x}\right)^5 - 0.7\left(\frac{3500}{3500+x}\right)^3.$$

Calculate the mean and the variance of the loss size.

Problem 16.5
Seventy-five percent of claims have a normal distribution with a mean of 3,000 and a variance of 1,000,000. The remaining 25% have a normal distribution with a mean of 4,000 and a variance of 1,000,000.

Determine the probability that a randomly selected claim exceeds 5,000.

Problem 16.6
How many parameters are there in a variable component mixture consisting of 9 Burr distributions?

Problem 16.7
Determine the distribution, density, and hazard rate functions for the variable mixture of two-parameter Pareto distribution.

Problem 16.8
A Weibull distribution has two parameters: θ and τ. An actuary is creating variable-component mixture distribution consisting of K Weibull distributions. If the actuary chooses to use 17 Weibull distributions instead of 12, how many more parameters will the variable-component mixture distribution have as a result?

Problem 16.9
Let X be a 2-point mixture with underlying random variables X_1 and X_2. The distribution of X_1 is a Pareto distribution with parmaters $\alpha_1 = 3$ and θ. The distribution of X_2 is a Gamma distribution with parameters $\alpha_2 = 2$ and $\theta_2 = 2000$.

Given that $a_1 = 0.7, a_2 = 0.3$, and $E(X) = 1340$, determine the value of θ.

Problem 16.10
Let X be a 3-point mixture of three variables X_1, X_2, X_3. You are given the following information:

R.V.	Weight	Mean	Standard Deviation
X_1	0.2	0.10	0.15
X_2	0.5	0.25	0.45
X_3	0.3	0.17	0.35

Determine Var(X).

Problem 16.11 ‡
The distribution of a loss, X, is a 2-point mixture:
(i) With probability 0.8, X_1 is a Pareto distribution with parameters $\alpha = 2$ and $\theta = 100$.
(ii) With probability 0.2, X_2 is a Pareto distribution with parameters $\alpha = 4$ and $\theta = 3000$.

Determine $\Pr(X \leq 200)$.

17 Data-dependent Distributions

In Section 15, we discussed parametric distributions. In Section 16, we introduced the $k-$point mixture distributions that are also known as **semi-parametric distributions.** In this section, we look at **non-parametric distributions**.

According to [1], *a* **data-dependent distribution** *is at least as complex as the data or knowledge that produced it, and the number of "parameters" increases as the number of data points or the amount of knowledge increases.*

We consider two-types of data-dependent distributions:

• **The empirical distribution** is obtained by assigning a probability of $\frac{1}{n}$ to each data point in a sample with n data points.

Example 17.1
Below are the losses suffered by policyholders of an insurance company:

$$49, 50, 50, 50, 60, 75, 80, 120, 230.$$

Let X be the random variable representing the losses incurred by the policyholders. Find the pmf and the cdf of X.

Solution.
The pmf is given by the table below.

x	49	50	60	75	80	120	130
$p(x)$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

The cdf is defined by

$$F_X(x) = \frac{1}{9}\text{number of elements in the sample that are} \leq x.$$

Thus, for example,

$$F_X(73) = \frac{5}{9} \blacksquare$$

• **A kernel smoothed distribution**
Given an empirical distribution, we wish to create a continuous distribution whose pdf will be a good estimation of the (discrete) empirical distribution. The density function is given by

$$f_X(x) = \sum_{i=1}^{n} p_n(x)k_i(x)$$

where $p_n(x) = \frac{1}{n}$ and $k_i(x)$ is the **kernel smoothed density function**. We illustrate these concepts next.

Example 17.2
Below are the losses suffered by policyholders of an insurance company:

$$49, 50, 50, 50, 60, 75, 80, 120, 230.$$

Develop a kernel-smoothed distribution associated with this set, such that each point x is associated with a uniform distribution that has positive probability over the interval $(x-5, x+5)$. As your answer, write the probability density function (pdf) of the kernel smoothed distribution.

Solution.
For $i = 1, 2, \cdots, 9$, we have

$$k_i(x) = \begin{cases} \frac{1}{10}, & x_i - 5 \le x \le x_i + 5 \\ 0, & \text{otherwise.} \end{cases}$$

We refer to $k_i(x)$ as the **uniform kernel with bandwith** 5.Thus,

$$f_X(x) = \sum_{i=1}^{9} \left(\frac{1}{9}\right) k_i(x) \blacksquare$$

A futher discussion of kernel density models will be covered in Section 56 of this book.

Practice Problems

Problem 17.1
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

$$94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.$$

Let X be the random variable representing the losses incurred by the policyholders.

Find the mean, the variance and the mode of X.

Problem 17.2
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

$$94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.$$

Let X be the random variable representing the losses incurred by the policyholders. The insurance company issued a policy with an ordinary deductible of 105.

Calculate $E(X \wedge 105)$ and the cost per payment $e_X(105)$.

Problem 17.3
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

$$94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.$$

Let X be the random variable representing the losses incurred by the policyholders.

Find the value of β in the standard deviation principle $\mu + \beta\sigma$ so that the standard deviation principle is equal to $\mathrm{VaR}_{0.8}(X)$.

Problem 17.4
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.

Let X be the random variable representing the losses incurred by the policyholders.

Find the empirical pmf and cdf of X.

Problem 17.5
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.

Let X be the random variable representing the losses incurred by the policyholders.

Find the coefficient of variation of X.

Problem 17.6
You are given the following empirical distribution of losses suffered by policyholders Prevent Dental Insurance Company:

94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.

Let X be the random variable representing the losses incurred by the policyholders.

Find the coefficient of skewness of X.

Problem 17.7
You are given the following the distribution of losses suffered by policyholders Prevent Dental Insurance Company:

94, 104, 104, 104, 134, 134, 180, 180, 180, 180, 210, 350, 524.

Let X be the random variable representing the losses incurred by the policyholders.

Calculate $f_X(x)$ using smoothed kernel setting with uniform kernel of bandwith 4.

Generating New Distributions

In this chapter a collection of continuous models that are commonly used for most actuarial modeling situations will be developed. Processes for creating new distributions from existing ones will be introduced. We point out here that most of the distributions that are used in actuarial modeling have nonnegative support so that $F_X(0) = 0$.

18 Scalar Multiplication of Random Variables

The first process that we discuss for creating new distributions from existing ones is the multiplication of a random variable by a constant. The following theorem sheds information about this process.

Theorem 18.1
Let X be a continuous random variable and c a positive constant. Let $Y = cX$. Then $f_Y(y) = \frac{1}{c} f_X\left(\frac{y}{c}\right)$ and $F_Y(y) = F_X\left(\frac{y}{c}\right)$. Thus, c is a scale parameter for Y.

Proof.
We start by finding the cdf of Y and then we generate the pdf by differentiation. We have

$$F_Y(y) = \Pr(Y \leq y) = \Pr\left(X \leq \frac{y}{c}\right) = F_X\left(\frac{y}{c}\right).$$

Now, differentiating $F_Y(y)$ and using the chain rule we find

$$f_Y(y) = \frac{1}{c} f_X\left(\frac{y}{c}\right) \blacksquare$$

Example 18.1
Suppose that random losses are exponentially distributed with parameter θ. Find the pdf and the cdf of the random variable $Y = cX, c > 0$.

Solution.
We have

$$F_Y(y) = \Pr(Y \leq y) = \Pr\left(X \leq \frac{y}{c}\right) = 1 - e^{-\frac{y}{c\theta}}.$$

Thus,

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{c} e^{-\frac{y}{c\theta}} \blacksquare$$

Scalar Multiples of random variables are useful in actuarial modeling when annual losses are subject to future uniform inflation. For example, if X is the random variable representing this year's losses and uniform losses are known to be $i\%$ for the next year then the next year's losses can be modeled with the random variable $Y = (1 + 0.01i)X$.

Example 18.2
You are given:
(i) In 2011, losses follow a Pareto distribution with parameters $\alpha = 2$ and

$\theta = 100$.
(ii) Inflation of 3.5% impacts all losses uniformly from 2011 to 2012.
What is the probability that the losses will exceed 350 in 2012?

Solution.
Let X and Y be the random variables representing the losses in 2011 and 2012 respectively. Then $Y = 1.035X$. We want to find $\Pr(Y > 350)$. Recall that the cdf of the Pareto distribution is

$$F_X(x) = 1 - \left(\frac{100}{x+100}\right)^2.$$

Thus,

$$\Pr(Y > 350) = \Pr\left(X > \frac{350}{1.035}\right) = S_X\left(X > \frac{350}{1.035}\right) = \left(\frac{100}{\frac{350}{1.035} + 100}\right)^2 = 0.0521 \blacksquare$$

Practice Problems

Problem 18.1
You are given:
(i) X has a Pareto distribution with parameters $\alpha = 3$ and $\theta = 2000$.
(ii) $Y = cX,\ c > 0$.
(iii) $\sigma_Y = 1500$.

Determine the value of c.

Problem 18.2
Losses in 2011 are represented by a random variable X with pdf $f_X(x) = \frac{3}{8}x^2$ for $0 < x < 2$ and 0 otherwise. Let Y be the random variable of losses in 2012. It is expected that losses in 2012 will go down by 50% than the current year.

Find $f_Y(y), F_Y(y)$, and $S_Y(y)$.

Problem 18.3
Losses from auto accidents are modeled by a Pareto distribution with parameters $\alpha = 3$ and $\theta = 2000$. The insurance policy pays only 75% of any auto accident claim.

Find the mean and the standard deviation of the claims for this policy.

Problem 18.4
Let X have cdf $F_X(x) = 1 - (1 + x)^{-\alpha}$ where $x > 0$ and $\alpha > 0$. Determine the pdf and the cdf of $Y = \theta X$.

Problem 18.5
Let Y have the lognormal distribution with parameters μ and σ. Let $Z = \theta Y$.

Show that Z also have the lognormal distribution.

Problem 18.6
Losses in 1993 follow the density function $f_X(x) = 3x^{-4}, x > 1$ where x is the loss in millions of dollars. Inflation of 10% impacts all claims uniformly from 1993 to 1994.

Determine the cdf of losses for 1994 and use it to determine the probability that a 1994 loss exceeds 2,200,000.

Problem 18.7

You are given:

(i) X is a loglogistic random variable with parameters $\gamma = 2$ and $\theta = 10\sqrt{10}$.

(ii) Y is a Pareto distribution with parameters $\alpha = 1$ and $\theta = 1000$.

(iii) Z is a 2-point mixture of X and Y with equal mixing weights.

(iv) $W = (1+r)Z$ where $r > 0$.

Show that W is an equal mixture of a loglogistic and a Pareto distribution and determine the parameters of W.

19 Powers and Exponentiation of Random Variables

A new distribution is obtained by raising a random variable to a certain power such as

$$Y = X^{\frac{1}{\tau}} \quad \text{or} \quad Y = X^{-\frac{1}{\tau}}$$

where $\tau > 0$. In the first case, Y is called **transformed**. In the second case, assuming $\tau \neq 1$, Y is called **inverse transformed**. If $\tau = 1$, we call Y the **inverse** of X.

The pdf and the cdf of the new distribution are provided by the next theorem.

Theorem 19.1
Let X be a continuous random variable with pdf and cdf $f_X(x)$ and $F_X(x)$ with $F_X(0) = 0$. Let $\tau > 0$. We have
(a) In the transformed case, $Y = X^{\frac{1}{\tau}}$

$$F_Y(y) = F_X(y^\tau) \quad \text{and} \quad f_Y(y) = \tau y^{\tau-1} f_X(y^\tau).$$

(b) In the inverse transformed case, $Y = X^{-\frac{1}{\tau}}$

$$F_Y(y) = 1 - F_X(y^{-\tau}) \quad \text{and} \quad f_Y(y) = \tau y^{-\tau-1} f_X(y^{-\tau}).$$

(c) In the inverse case, $Y = X^{-1}$

$$F_Y(y) = 1 - F_X(y^{-1}) \quad \text{and} \quad f_Y(y) = y^{-2} f_X(y^{-1}).$$

Proof.
(a) We have

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^\tau) = F_X(y^\tau).$$

Differentiating this function with respect to y and using the chain rule, we find

$$f_Y(y) = \tau y^{\tau-1} f_X(y^\tau).$$

(b) We have

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \geq y^\tau) = 1 - F_X(y^{-\tau}).$$

Differentiating this function with respect to y and using the chain rule, we find

$$f_Y(y) = \tau y^{-\tau-1} f_X(y^{-\tau}).$$

(c) We have

$$F_Y(y) = \Pr(Y \le y) = \Pr(X \ge y^{-1}) = 1 - F_X(y^{-1}).$$

Differentiating this function with respect to y and using the chain rule, we find

$$f_Y(y) = y^{-2} f_X(y^{-1}) \blacksquare$$

Example 19.1
Let X be a random variable with pdf $f_X(x) = x$ for $0 < x < \sqrt{2}$ and 0 otherwise. Let $Y = X^{\frac{1}{4}}$. Find the pdf and the cdf of Y.

Solution.
The cdf of X is $F_X(x) = x^2$ for $0 \le x \le \sqrt{2}$ and 1 for $x > \sqrt{2}$. Thus, $F_Y(y) = F_X(y^4) = y^8$ for $0 \le y \le 2^{\frac{1}{8}}$ and 1 for $y \ge 2^{\frac{1}{8}}$. The pdf of Y is $f_Y(y) = 4y^3 f_X(y^4) = 8y^7$ for $0 < y < 2^{\frac{1}{8}}$ and 0 otherwise $\blacksquare$

Example 19.2
Let X have the beta distribution with pdf

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \ 0 < x < 1$$

and 0 otherwise. Find the pdf of $Y = X^{\frac{1}{\tau}}$ where $\tau > 0$.

Solution.
We have

$$f_Y(y) = -\tau y^{\tau-1} f_X(y^\tau) = -\tau y^{\tau-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\tau(\alpha-1)}(1-y^\tau)^{\beta-1}, \ 0 < y < 1$$

and 0 otherwise $\blacksquare$

Let X be a continuous random variable and $Y = e^X$. That is, Y has a lognormal distribution.

Theorem 19.2
Let X be a continuous random variable with pdf $f_X(x)$ and cdf $F_X(x)$ such that $f_X(x) > 0$ for all $x \in \mathbb{R}$. Let $Y = e^X$. Then, for $y > 0$, we have

$$F_Y(y) = F_X(\ln y) \quad \text{and} \quad f_Y(y) = \tfrac{1}{y} f_X(\ln y).$$

Proof.
We have

$$F_Y(y) = \Pr(Y \le y) = \Pr(e^X \le y) = \Pr(X \le \ln y) = F_X(\ln y).$$

Differentiating with respect to y, we find

$$f_Y(y) = \frac{1}{y} f_X(\ln y) \blacksquare$$

Example 19.3
Let X be a normal distribution with parameters $\mu = 1$ and $\sigma^2 = 4$. Define the random variable $Y = e^X$.
(a) Find $E(Y)$.
(b) The 95^{th} percentile of the standard normal distribution is 1.645. Find the 95^{th} percentile of Y.

Solution.
(a) We have

$$\begin{aligned} E(Y) =& \int_{-\infty}^{\infty} e^x \frac{1}{2\sqrt{2\pi}} e^{-\left(\frac{(x-1)^2}{8}\right)} dx \\ =& e^3 \int_{-\infty}^{\infty} \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(x-5)^2} dx \\ =& e^3 \cdot 1 = 20.086. \end{aligned}$$

Note that the last integral is the integral of the density function of the normal distribution with parameters $\mu = 5$ and $\sigma^2 = 4$.
(b) Let $\pi_{0.95}$ be the 95^{th} percentile of Y. Then $\Pr(Y \le \pi_{0.95}) = 0.95$. Thus,

$$0.95 = \Pr(Y \le \pi_{0.95}) = \Pr(X \le \ln 0.95) = \Pr\left(\frac{X-1}{2} \le \frac{\ln \pi_{0.95} - 1}{2}\right).$$

Hence,

$$\frac{\ln \pi_{0.95} - 1}{2} = 1.645 \Longrightarrow \pi_{0.95} = 72.97 \blacksquare$$

Practice Problems

Problem 19.1
Let X be the exponential distribution with parameter θ. Determine the pdf and the cdf of the transformed, inverse transformed, and inverse exponential distribution.

Problem 19.2
Find the cdf of the inverse of a Pareto distribution with parameters α and θ. What's the name of the new distribution and its parameter(s)?

Problem 19.3
Let X be a random variable with pdf $f_X(x) = 2x$ for $0 < x < 1$ and 0 otherwise.

Find the pdf of $Y = X^{-1}$.

Problem 19.4
Find the pdf of the inverse of the Gamma distribution with parameters α and $\theta = 1$.

Problem 19.5
Let X have a uniform distribution in $(0, b)$. Find the pdf of $Y = X^{\frac{1}{\tau}}$, with $\tau > 0$.

Problem 19.6
Let X have a Pareto distribution with parameters α and θ. Let $Y = \ln\left(1 + \frac{X}{\theta}\right)$. Determine the name of the distribution of Y and its parameter(s).

Problem 19.7
Let X have the normal distribution with parameters μ and σ^2. Find the pdf and cdf of $Y = e^X$.

Problem 19.8
Let X have a uniform distribution in $(0, b)$. Find the pdf of $Y = e^X$.

Problem 19.9
Let X have an exponential distribution with parameter θ. Find the pdf of $Y = e^X$.

Problem 19.10 ‡
You are given:
(i) X has a Pareto distribution with parameters $\alpha = 2$ and $\theta = 100$.
(ii) $Y = \ln\left(1 + \frac{x}{\theta}\right)$.

Calculate the variance of Y.

20 Continuous Mixing of Distributions

In Section 16 we introduced the concept of discrete mixture of distributions. In this section we will define the continuous version of mixing where the discrete probabilities are replaced by the pdf of a continuous random variable. In actuarial terms, continuous mixtures arise when a risk parameter from the loss distribution is uncertain and the uncertain parameter is continuous.

Suppose that Λ is a continuous random variable with pdf $f_\Lambda(\lambda)$. Let X be a continuous random variable that depends on a parameter λ. We say that X is a **mixture of distributions** if its pdf is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda.$$

The distribution $f_X(x)$ is a marginal density function: Let $f(x, \lambda)$ be the joint distribution of X and Λ. Then $f_X(x)$ and $f_\Lambda(\lambda)$ are referred to as **marginal densities**. Then the product $f_{X|\Lambda}(x|\lambda)f_\Lambda(\lambda)$ is the **joint density function** and the marginal density function can be recovered by integrating the joint density function(See Section 33 of [2].)

Theorem 20.1
For the mixture distribution X as defined above, we have
(a) $F_X(x) = \int_{-\infty}^{\infty} F_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda$.
(b) $E(X^k) = E[E(X^k|\Lambda)]$.
(c) $\text{Var}(X) = E[\text{Var}(X|\Lambda)] + \text{Var}[E(X|\Lambda)]$.

Proof.
(a) Assuming that the order of integration can be reversed, we have

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X|\Lambda}(t|\lambda) f_\Lambda(\lambda) d\lambda dt = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X|\Lambda}(t|\lambda) f_\Lambda(\lambda) dt d\lambda \\ &= \int_{-\infty}^{\infty} F_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda. \end{aligned}$$

(b) We have

$$\begin{aligned} E(X^k) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda dx \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x^k f_{X|\Lambda}(x|\lambda) dx \right] f_\Lambda(\lambda) d\lambda \\ &= \int_{-\infty}^{\infty} E(X^k|\Lambda) f_\Lambda(\lambda) d\lambda = E[E(X^k|\Lambda)]. \end{aligned}$$

(c) We have

$$\begin{aligned}\text{Var}(X) =&E(X^2) - (E(X))^2\\ =&E[E(X^2|\Lambda)] - \{E[E(X|\Lambda)]\}^2\\ =&E\{\text{Var}(X|\Lambda) + [E(X|\Lambda)]^2\} - \{E[E(X|\Lambda)]\}^2\\ =&E[\text{Var}(X|\Lambda)] + E\{[E(X|\Lambda)]^2\} - \{E[E(X|\Lambda)]\}^2\\ =&E[\text{Var}(X|\Lambda)] + \text{Var}[E(X|\Lambda)] \blacksquare\end{aligned}$$

Mixture distributions tend to be heavy-tailed as seen in the next example.

Example 20.1
The distribution of $X|\Lambda$ is exponential with parameter $\frac{1}{\Lambda}$. The distribution of Λ is Gamma with parameters α and θ. Find $f_X(x)$.

Solution.
We have

$$\begin{aligned}f_X(x) =& \int_0^\infty \lambda e^{-\lambda x}\frac{\theta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\lambda\theta}d\lambda\\ =&\frac{\theta^\alpha}{\Gamma(\alpha)}\int_0^\infty \lambda^\alpha e^{-\lambda(x+\theta)}d\lambda\\ =&\frac{\theta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{(x+\theta)^{\alpha+1}} = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}.\end{aligned}$$

This is the distribution of a Pareto random variable and we know that this distribution is heavy-tailed $\blacksquare$

Example 20.2
Suppose that $X|\Lambda$ has a normal distribution with parameters λ and σ_1. That is,

$$f_{X|\Lambda}(x|\lambda) = \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-\lambda)^2}{2\sigma_1^2}}, \ -\infty < x < \infty.$$

Suppose that Λ has a normal distribution with parameters μ and σ_2. That is,

$$f_\Lambda(\lambda) = \frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(\lambda-\mu)^2}{2\sigma_2^2}}, \ -\infty < x < \infty.$$

Determine the unconditional pdf of X.

Solution.
We first establish the following identity:

$$\left(\frac{x-\lambda}{\sigma_1}\right)^2+\left(\frac{\lambda-\mu}{\sigma_2}\right)^2=\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)\left(\lambda-\frac{\sigma_2^2x+\mu\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{(x-\mu)^2}{\sigma_1^2+\sigma_2^2}.$$

By completing the square, we find

$$\begin{aligned}\left(\frac{x-\lambda}{\sigma_1}\right)^2+\left(\frac{\lambda-\mu}{\sigma_2}\right)^2=&\frac{x^2-2\lambda x+\lambda^2}{\sigma_1^2}+\frac{\lambda^2-2\mu\lambda+\mu^2}{\sigma_2^2}\\
=&\lambda^2\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)-2\lambda\left(\frac{\sigma_2^2x+\sigma_1^2\mu}{\sigma_1^2\sigma_2^2}\right)+\frac{\sigma_2^2x^2+\sigma_1^2\mu^2}{\sigma_1^2\sigma_2^2}\\
=&\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)\left(\lambda-\frac{\sigma_2^2x+\mu\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2-\left(\frac{(\sigma_2^2x+\mu\sigma_1^2)^2}{\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)}\right)+\frac{\sigma_2^2x^2+\sigma_1^2\mu^2}{\sigma_1^2\sigma_2^2}\\
=&\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)\left(\lambda-\frac{\sigma_2^2x+\mu\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{1}{\sigma_1^2\sigma_2^2}\left[\frac{\sigma_1^2\sigma_2^2x^2-2\mu x\sigma_1^2\sigma_2^2+\sigma_1^2\sigma_2^2\mu^2}{\sigma_1^2+\sigma_2^2}\right]\\
=&\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)\left(\lambda-\frac{\sigma_2^2x+\mu\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{(x-\mu)^2}{\sigma_1^2+\sigma_2^2}.\end{aligned}$$

Now, the marginal density function of X is

$$\begin{aligned}f_X(x)=&\int_{-\infty}^{\infty}\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-\lambda)^2}{2\sigma_1^2}}\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(\lambda-\mu)^2}{2\sigma_2^2}}d\lambda\\
=&\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left[\left(\frac{x-\lambda}{\sigma_1}\right)^2+\left(\frac{\lambda-\mu}{\sigma_2}\right)^2\right]}d\lambda\\
=&\frac{e^{-\frac{(x-\mu)^2}{2(\sigma_1^2+\sigma_2^2)}}}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\int_{-\infty}^{\infty}\sqrt{\frac{\sigma_1^2+\sigma_2^2}{2\pi\sigma_1^2\sigma_2^2}}\exp\left[-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(\lambda-\frac{\sigma_2^2x+\mu\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2\right]d\lambda.\end{aligned}$$

The integrand in the last integral is the pdf of a normal distribution with parameters $\frac{\sigma_1^2x+\sigma_2^2\mu}{\sigma_1^2+\sigma_2^2}$ and $\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ so that the integral is 1. Hence,

$$f_X(x)=\frac{e^{-\frac{(x-\mu)^2}{2(\sigma_1^2+\sigma_2^2)}}}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}$$

which is the pdf of the normal distribution with parameters μ and $\sigma_1^2+\sigma_2^2$ ■

Example 20.3 ‡
The scores on the final exam in Ms. B's Latin class have a normal distribution with mean θ and standard deviation equal to 8. θ is a random variable

with a normal distribution with mean equal to 75 and standard deviation equal to 6.
Each year, Ms. B chooses a student at random and pays the student 1 times the student's score. However, if the student fails the exam (score ≤ 65), then there is no payment.
Calculate the conditional probability that the payment is less than 90, given that there is a payment.

Solution.
Let S denote the scores. Since $S|\Theta$ and Θ are normal, S is also normally distributed with mean

$$E(S) = E[E(S|\Theta)] = E(\Theta) = 75$$

and variance

$$\mathrm{Var}(S) = E[\mathrm{Var}(S|\Theta)] + \mathrm{Var}[E(S|\Theta)] = 64 + 36 = 100.$$

We want

$$\begin{aligned}
\Pr(S < 90|S > 65) =& \frac{\Pr(65 < S < 90)}{\Pr(S > 65)} = \frac{F_S(90) - F_S(65)}{1 - F_S(65)} \\
=& \frac{\Phi\left(\frac{90-75}{10}\right) - \Phi\left(\frac{65-75}{10}\right)}{1 - \Phi\left(\frac{65-75}{10}\right)} \\
=& \frac{\Phi(1.5) - \Phi(-1.0)}{1 - \Phi(-1.0)} = \frac{\Phi(1.5) - [1 - \Phi(1)]}{\Phi(1)} \\
=& \frac{0.9332 - (1 - 0.8413)}{0.8413} = 0.9206 \blacksquare
\end{aligned}$$

Example 20.4
Let $X|\Lambda$ have a Poisson distribution with parameter λ. Let Λ have a Gamma distribution with parameters α and β. That is,

$$f_\Lambda(\lambda) = \frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}}{\beta^\alpha\Gamma(\alpha)}.$$

Find the unconditional probability $\Pr(X = 1)$.

Solution.
We have[8]

$$\begin{aligned}
\Pr(X=1)=&\int_0^\infty \Pr(X=1|\Lambda)f_\Lambda(\lambda)d\lambda\\
=&\int_0^\infty \lambda e^{-\lambda}\frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}}{\beta^\alpha\Gamma(\alpha)}d\lambda\\
=&\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty \lambda^\alpha e^{-\lambda\left(1+\frac{1}{\beta}\right)}d\lambda\\
=&\frac{\left(1+\frac{1}{\beta}\right)^{-(\alpha+1)}\Gamma(\alpha+1)}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty \frac{\lambda^{(\alpha+1)-1}e^{-\frac{\lambda}{\left(1+\frac{1}{\beta}\right)^{-1}}}}{\left(1+\frac{1}{\beta}\right)^{-(\alpha+1)}\Gamma(\alpha+1)}d\lambda.
\end{aligned}$$

The integrand is the pdf of a Gamma distribution with parameters $\alpha+1$ and $\left(1+\frac{1}{\beta}\right)^{-1}$ so that the integral is 1. Hence,

$$\Pr(X=1)=\frac{\left(1+\frac{1}{\beta}\right)^{-(\alpha+1)}\Gamma(\alpha+1)}{\beta^\alpha\Gamma(\alpha)}=\frac{\alpha\beta}{(1+\beta)^{\alpha+1}} \blacksquare$$

Example 20.5 ‡
Bob is a carnival operator of a game in which a player receives a prize worth $W=2^N$ if the player has N successes, $N=0,1,2,\cdots$, Bob models the probability of success for a player as follows:
(i) N has a Poisson distribution with mean Λ.
(ii) Λ has a uniform distribution on the interval $(0,4)$.
Calculate $E[W]$.

Solution.
We know that $P_N(z)=e^{\lambda(z-1)}=E(Z^N)$. In particular, $E(W|\Lambda)=P_N(2)=e^\lambda$. Thus,

$$E(W)=\int_0^4 E(W|\Lambda)f_\Lambda(\lambda)d\lambda=\frac{1}{4}\int_0^4 e^\lambda d\lambda=13.4 \blacksquare$$

[8]See P. 380 of [2]

Practice Problems

Problem 20.1
Let X be a loss random variable having a Pareto distribution with parameters α and Θ. The parameter Θ is uniformly distributed in $(0, b)$.

Find the unconditional variance of X.

Problem 20.2
Let $X|\Theta$ be the inverse exponential random variable with parameter Θ. Its pdf is

$$f_{X|\Theta}(x|\theta) = \frac{1}{x^2}\theta e^{-\theta x},\ x > 0$$

and 0 otherwise. Let Θ have the exponential distribution with mean 4.

Determine the unconditional distribution of X.

Problem 20.3
Let $X|\Lambda$ have the pdf

$$f_{X|\Lambda}(x|\lambda) = \lambda e^{-x} + (1-\lambda)xe^{-x},\ x > 0$$

and 0 otherwise.

Find the unconditional variance of X given that $E(\Lambda) = 0.45$.

Problem 20.4
Let $X|\Lambda$ have a Poisson distribution with parameter Λ. Let Λ have a Gamma distribution with parameters α and β. That is,

$$f_\Lambda(\lambda) = \frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}}{\beta^\alpha\Gamma(\alpha)}.$$

Determine the expectation of Λ and then the unconditional expectation of X.

Problem 20.5
Suppose that $X|\Lambda$ has a Weibull distribution with cdf $F_{X|\Lambda}(x|\lambda) = 1 - e^{-\lambda x^\gamma}$ for $x \geq 0$. Suppose that Λ is exponentially distributed with mean θ.

Find the unconditional cdf of X.

Problem 20.6
Let N have a Poisson distribution with mean Λ. Let Λ have a uniform distribution on the interval (0,5).

Determine the unconditional probability that $N \geq 2$.

Problem 20.7
Let $N|\Lambda$ have a negative binomial distribution with $r = 1$ and Λ. Let Λ have a Gamma distribution with $\alpha = 1$ and $\theta = 2$.

Find the unconditional variance of N.

Problem 20.8 ‡
A claim count distribution can be expressed as a mixed Poisson distribution. The mean of the Poisson distribution is uniformly distributed over the interval [0, 5].

Calculate the probability that there are 2 or more claims.

Problem 20.9 ‡
The length of time T, in years, that a person will remember an actuarial statistic is modeled by an exponential distribution with mean $\frac{1}{Y}$. In a certain population, Y has a gamma distribution with $\alpha = \theta = 2$.

Calculate the probability that a person drawn at random from this population will remember an actuarial statistic less than $\frac{1}{2}$ year.

21 Frailty (Mixing) Models

A continuous mixture model that arises within the context of survival analysis is the **frailty model**. We will discuss mainly the mathematical components of this model so that the focus remains on the importance of generating new distributions by mixing.

A frailty model is defined as follows: Let $X|\Lambda$ be a random variable with **conditional hazard function** given by

$$h_{X|\Lambda}(x|\lambda) = \lambda a(x)$$

where $a(x)$ is some specified and known function of x. The **frailty random variable** Λ is supposed to have positive support.

Theorem 21.1
Let $A(x) = \int_0^x a(z)dz$. We have:
(a) The conditional survival function of X given Λ is

$$S_{X|\Lambda}(x|\lambda) = e^{-\lambda A(x)}.$$

(b) The unconditional survival function of X is

$$S_X(x) = M_\Lambda(-A(x)).$$

Proof.
(a) Recall that a survival function can be recovered from the hazard rate function so that we have

$$S_{X|\Lambda}(x|\lambda) = e^{-\int_0^x h_{X|\Lambda}(z|\lambda)dz} = e^{-\int_0^x \lambda a(z)dz} = e^{-\lambda A(x)}.$$

(b) The unconditional survival function is

$$S_X(x) = \int_\lambda S_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda)d\lambda = E[S_{X|\Lambda}(x|\lambda)] = E[e^{-\lambda A(x)}] = M_\Lambda(-A(x)) \blacksquare$$

Remark 21.1
If $X|\Lambda$ has an exponential distribution in the frailty model, $a(x)$ will be 1, and $A(x)$ will be x. When $X|\Lambda$ has Weibull distribution in the frailty model, $a(x)$ will be $\gamma x^{\gamma-1}$, and $A(x)$ will be x^γ.

Example 21.1
Let $X|\Lambda$ have a Weibull distribution with conditional survival function $S_{X|\Lambda}(x|\lambda) = e^{-\lambda x^\gamma}$. Let Λ have a Gamma distribution with parameters α and θ. Find the unconditional or marginal survival function of X.

Solution.
We first find the moment generating function of Λ. We have

$$\begin{aligned}E(e^{tX}) =&\frac{1}{\theta^\alpha\Gamma(\alpha)}\int_0^\infty e^{tx}x^{\alpha-1}e^{-\frac{x}{\theta}}dx\\ =&\frac{1}{\theta^\alpha\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-x\left(-t+\frac{1}{\theta}\right)}dx\\ =&\int_0^\infty \frac{y^{\alpha-1}(-t+\frac{1}{\theta})^{-\alpha}e^{-y}}{\theta^\alpha\Gamma(\alpha)}dy\\ =&\frac{\left(-t+\frac{1}{\theta}\right)^{-\alpha}\Gamma(\alpha)}{\theta^\alpha\Gamma(\alpha)}=(1-\theta t)^{-\alpha},\ t<\frac{1}{\theta}.\end{aligned}$$

Also, we know that $A(x)=x^\gamma$. Hence,

$$S_X(x)=M_\Lambda(-A(x))=(1+\theta x^\gamma)^{-\alpha}.$$

This is the survival function of a Burr distribution with parameters α and $\theta^{-\frac{1}{\gamma}}$ ■

Example 21.2
A continuous mixture is used in a frailty model with frailty random variable Λ, such that $a(x)=\frac{1}{x+1},\ x>0$. Find the conditional survival function of X.

Solution.
We first find $A(x)$:

$$A(x)=\int_0^x\frac{dt}{1+t}=\ln{(1+x)}.$$

Thus,

$$S_{X|\Lambda}(x|\lambda)=e^{-\lambda A(x)}=e^{-\lambda\ln{(1+x)}}=\frac{1}{(1+x)^\lambda}\ ■$$

Example 21.3
Given that the marginal survival function in a frailty model is $S_X(x)=x^{-2},\ x\geq 1$. Let the frailty random variable have an exponential distribution. Determine $M_\Lambda(x)$.

Solution.
Since Λ have an exponential distribution, $A(x)=x$. Thus, $x^{-2}=S_X(x)=M_\Lambda(-A(x))=M_\Lambda(-x)$. Hence, $M_\Lambda(x)=x^{-2},\ x\geq 1$ ■

Example 21.4
The marginal survival function in a frailty model is given to be $S_X(x) = 2^{\frac{1}{\alpha}} M_\Lambda(x)$. The frailty random variable Λ is a Gamma random variable with parameters α and θ. Determine $A(x)$.

Solution.
The moment generating function of Λ is $M_\lambda(x) = (1-\theta x)^{-\alpha}$. Thus, $2^{\frac{1}{\alpha}}(1-\theta x)^{-\alpha} = S_X(x) = M_\Lambda(-A(x)) = (1+\theta A(x))^{-\alpha}$. Hence, $2(1-\theta x) = 1 + \theta A(x)$. Solving for $A(x)$, we find $A(x) = \frac{2-\theta x}{\theta}$ ■

Example 21.5
Consider a frailty model where $X|\Theta$ has an exponential distribution with conditional hazard rate function $h_{X|\Theta}(x|\theta) = \theta$. The frailty random variable Θ has a uniform distribution in $(1, 11)$. Find the conditional survival function of X.

Solution.
We have

$$S_{X|\Theta}(x|\theta) = e^{-\int_0^x h_{X|\theta}(z|\theta)dz} = e^{-\int_0^x \theta dz} = e^{-\theta x}, \ x > 0 \ ■$$

Example 21.6
Consider the exponential-inverse Gaussian frailty model with

$$a(x) = \frac{\theta}{2\sqrt{1+\theta x}}, \ \theta > 0.$$

Determine $A(x)$.

Solution.
We have

$$A(x) = \int_0^x a(t)dt = \int_0^x \frac{\theta}{2\sqrt{1+\theta t}}dt = \sqrt{1+\theta t}\Big|_0^x = \sqrt{1+\theta x} - 1 \ ■$$

Practice Problems

Problem 21.1
Let $X|\Lambda$ have an exponential distribution with conditional survival function $S_{X|\Lambda}(x|\lambda) = e^{-\lambda x}$. Let Λ have a Gamma distribution with parameters α and θ.

Find the unconditional or marginal survival function of X.

Problem 21.2
A continuous mixture is used in a frailty model with frailty random variable Λ, such that $a(x) = \frac{1}{x+1},\ x > 0$. The frailty random variable has a uniform distribution in $(0, 1)$.

Find the marginal survival function of X.

Problem 21.3
Given that the marginal survival function in a frailty model is $S_X(x) = e^{-\frac{x}{\theta}},\ x \geq 0$. Let the frailty random variable have an exponential distribution.

Determine $M_\Lambda(x)$.

Problem 21.4
The marginal survival function in a frailty model is given to be $S_X(x) = 2^{\frac{1}{\alpha}} M_\Lambda(x)$. The frailty random variable Λ is a Gamma random variable with parameters α and θ.

Determine $a(x)$.

Problem 21.5
The probability generating function of the frailty random variable is $P_\Lambda(x) = e^{x-1}$. Suppose $X|\Lambda$ has an exponential distribution.

Find the marginal survival function of X.

Problem 21.6
Consider a frailty model where $X|\Theta$ has an exponential distribution with conditional hazard rate function $h_{X|\Theta}(x|\theta) = \theta$. The frailty random variable Θ has a uniform distribution in $(1, 11)$.

Find the marginal survival function of X.

Problem 21.7
Suppose that $X|\Lambda$ has the Weibull distribution with conditional survival function $S_{X|\Lambda}(x|\lambda) = e^{-\lambda x^{\gamma}},\ x \geq 0$. The frailty random variable Λ has an exponential distribution with mean θ.

Find the marginal survival function of X.

Problem 21.8
Consider the exponential-inverse Gaussian frailty model with

$$a(x) = \frac{\theta}{2\sqrt{1+\theta x}},\ \theta > 0.$$

Determine the conditional survival function $S_{X|\Lambda}(x|\lambda)$.

Problem 21.9
Consider the exponential-inverse Gaussian frailty model with

$$a(x) = \frac{\theta}{2\sqrt{1+\theta x}},\ \theta > 0.$$

Suppose Λ has a Gamma distribution with parameters 2α and $\theta = 1$.

Determine the marginal survival function of X.

Problem 21.10
Determine the unconditional probability density function of frailty distribution.

Problem 21.11
Determine the unconditional hazard rate function of frailty distribution.

22 Spliced Distributions

A **spliced distribution** consists of different distributions one for each part of the domain of the random variable. For example, a **n-component spliced distribution** has the following pdf

$$f_X(x) = \begin{cases} \alpha_1 f_1(x), & c_1 < x < c_2 \\ \alpha_2 f_2(x), & c_2 < x < c_3 \\ \vdots & \\ \alpha_n f_n(x), & c_{n-1} < x < c_n \end{cases}$$

where each $\alpha_j > 0, \alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$, and each $f_i(x)$ is a legitimate pdf with support (c_i, c_{i+1}).

An interpretation in insurance claims is that the distributions vary by size of claims.

Example 22.1
Show that the pdf below is a spliced distribution of some random variable X and then find $E(X)$.

$$f(x) = \begin{cases} 0.01, & 0 \le x < 50 \\ 0.02, & 50 \le x \le 75 \end{cases}$$

and 0 otherwise.

Solution.
For the given pdf, we have $f_1(x) = 0.02$ for $0 \le x < 50$; $f_2(x) = 0.04$ for $50 \le x \le 75$; $\alpha_1 = \alpha_2 = 0.5$. Note that $f_1(x)$ and $f_2(x)$ are uniform densities. The mean of X is

$$E(X) = \int_0^{75} xf(x)dx = \int_0^{50} 0.01xdx + \int_{50}^{75} 0.02xdx = 43.75 \blacksquare$$

Example 22.2
Let $g_1(x)$ and $g_2(x)$ be two nonnegative functions defined on the intervals (c_1, c_2) and $[c_2, c_3)$ respectively. Create a pdf of a 2-component spliced distribution with support (c_1, c_3).

Solution.
Let $f_1(x) = \frac{g_1(x)}{\int_{c_1}^{c_2} g_1(x)dx}, f_2(x) = \frac{g_2(x)}{\int_{c_1}^{c_2} g_2(x)dx}, \alpha_1 = \alpha > 0, \alpha_2 = 1 - \alpha$ where α is some positive constant. Then

$$f(x) = \begin{cases} \alpha f_1(x), & c_1 < x < c_2 \\ (1-\alpha) f_2(x), & c_2 \le x < c_3 \end{cases} \blacksquare$$

Example 22.3 ‡
An actuary for a medical device manufacturer initially models the failure time for a particular device with an exponential distribution with mean 4 years. This distribution is replaced with a spliced model whose density function:
(i) is Uniform over [0,3]
(ii) is proportional to the initial modeled density function after 3 years
(iii) is continuous.
Calculate the probability of failure in the first 3 years under the revised distribution.

Solution.
The two-spliced pdf is

$$f(x)=\begin{cases}\frac{\alpha}{3}, & 0<x<3\\ (1-\alpha)e^{0.75}[0.25e^{-0.25x}], & x\geq 3\end{cases}$$

Using continuity, we have

$$\lim_{x\to 3^-}\frac{\alpha_1}{3}=\lim_{x\to 3^+}0.25(1-\alpha)e^{0.75}e^{-0.25x}\Longrightarrow\frac{\alpha}{3}=0.25(1-\alpha).$$

Solving, we find $\alpha=\frac{3}{7}$. Finally,

$$\Pr(X\leq 3)=\int_0^3\frac{3}{7}\frac{dx}{3}=\frac{3}{7}\ \blacksquare$$

Practice Problems

Problem 22.1
Find the density function of a random variable that is uniform on $(0, c)$ and exponential thereafter.

Problem 22.2 ‡
Suppose a loss distribution is a two-component spliced model with:
(i) a Weibull distribution having parameters $\theta_1 = 1500$ and $\tau = 1$ for losses up to \$4,000; and
(ii) a Pareto distribution having parameters $\theta_2 = 12000$ and $\alpha = 2$ for losses \$4,000 and up.
The probability that losses are less than \$4,000 is 0.60.

Calculate the probability that losses are less than \$25,000.

Problem 22.3
A random variable X follows a continuous two-component spliced distribution that is uniform in $(0, 3]$ and exponential (with mean 1) thereafter. Find the 95^{th} percentile of X.

Problem 22.4
Using the results of the previous problem, find $E(X|X > \pi_{0.95})$.

Problem 22.5
Write the density function for a 2-component spliced model in which the density function is proportional to a uniform density over the interval from 0 to 1000 and is proportional to an exponential density function from 1000 to ∞. Ensure that the resulting density function is continuous.

Problem 22.6
The pdf of two-component spliced distribution is given below.

$$f(x) = \begin{cases} 0.01, & 0 \le x < 50 \\ 0.02, & 50 \le x \le 75 \end{cases}$$

and 0 otherwise. Find the variance of X.

23 Limiting Distributions

In addition to the methods described in the previous sections, we can obtain new distributions as limiting cases of other ones. This is accomplished by letting the parameters go to either infinity or zero.

Example 23.1
For a Pareto distribution with parameters α and θ, let both α and θ go to infinity with the ratio $\frac{\alpha}{\theta} \to \xi$ held constant. Show that the result is an exponential distribution.

Solution.
Let $\xi = \frac{\alpha}{\theta}$ so that $\alpha = \xi\theta$. Substituting into the cdf of the Pareto distribution, we find

$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\xi\theta}.$$

Let $w = \left(\frac{\theta}{x+\theta}\right)^{\xi\theta}$. We have

$$\begin{aligned}
\lim_{\theta\to\infty} \ln w = & \lim_{\theta\to\infty} \xi\theta[\ln\theta - \ln(x+\theta)]\\
=&\xi \lim_{\theta\to\infty} \frac{\ln\theta - \ln(x+\theta)}{\theta^{-1}}\\
=&\xi \lim_{\theta\to\infty} \frac{\frac{d}{d\theta}[\ln\theta - \ln(x+\theta)]}{\frac{d}{d\theta}(\theta^{-1})}\\
=&\xi \lim_{\theta\to\infty} \frac{\theta^{-1} - (x+\theta)^{-1}}{-\theta^{-2}}\\
=&-\xi \lim_{\theta\to\infty} \frac{x\theta}{x+\theta} = -\xi x.
\end{aligned}$$

It follows that $\lim_{\theta\to\infty} w = e^{-\xi x}$ and $\lim_{\theta\to\infty} F_X(x) = 1 - e^{-\xi x}$ which is the cdf of an exponential distribution with mean $\frac{1}{\xi}$ ■

Example 23.2
For a transformed beta distribution with parameters α, γ and θ, let both α and θ go to infinity with the ratio $\frac{\theta}{\alpha^{\frac{1}{\gamma}}} \to \xi$. Show that the result is a transformed gamma distribution.

Solution.
For large α, Stirling's formula gives

$$\Gamma(\alpha) \approx e^{-\alpha}\alpha^{\alpha-\frac{1}{2}}(2\pi)^{\frac{1}{2}}.$$

Also, we let $\xi = \frac{\theta}{\alpha^{\frac{1}{\gamma}}}$ so that $\theta = \xi\alpha^{\frac{1}{\gamma}}$.
Using this and Stirling's formula in the pdf of a transformed beta distribution, we find

$$\begin{aligned}f_X(x) =&\frac{\Gamma(\alpha+\tau)\gamma x^{\gamma\tau-1}}{\Gamma(\alpha)\Gamma(\tau)\theta^{\gamma\tau}(1+x^\gamma\theta^{-\gamma})^{\gamma+\tau}}\\ \approx&\frac{e^{-\alpha-\tau}(\alpha+\tau)^{\alpha+\tau-\frac{1}{2}}(2\pi)^{\frac{1}{2}}\gamma x^{\gamma\tau-1}}{e^{-\alpha}(\alpha)^{\alpha-\frac{1}{2}}(2\pi)^{\frac{1}{2}}\Gamma(\tau)(\alpha^{\frac{1}{\gamma}})^{\gamma\tau}(1+x^\gamma\xi^{-\gamma}\alpha^{-1})^{\gamma+\tau}}\\ =&\frac{e^{-\tau}[(\alpha+\tau)/\alpha]^{\alpha+\tau-\frac{1}{2}}\gamma x^{\gamma\tau-1}}{\Gamma(\tau)\xi^{\gamma\tau}[1+(x/\xi)^\gamma/\alpha]^{\alpha+\tau}}.\end{aligned}$$

Now, let

$$w_1 = \left(1+\frac{\tau}{\alpha}\right)^{\alpha+\tau-\frac{1}{2}}.$$

We have

$$\begin{aligned}\lim_{\alpha\to\infty}\ln w_1 =&\lim_{\alpha\to\infty}(\alpha+\tau-\frac{1}{2})\ln\left(1+\frac{\tau}{\alpha}\right)\\ =&\lim_{\alpha\to\infty}\frac{\ln\left(1+\frac{\tau}{\alpha}\right)}{(\alpha+\tau-\frac{1}{2})^{-1}}\\ =&\lim_{\alpha\to\infty}\frac{-\tau\alpha^{-2}\left(1+\frac{\tau}{\alpha}\right)^{-1}}{-(\alpha+\tau-\frac{1}{2})^{-2}}\\ =&\tau\lim_{\alpha\to\infty}\left(1+\frac{\tau}{\alpha}\right)^{-1}\left(1+\frac{\tau}{\alpha}-\frac{1}{1\alpha}\right)^2\\ =&\tau.\end{aligned}$$

Thus, $\lim_{\alpha\to\infty} w_1 = e^\tau$. Now, let

$$w_2 = \left[1+\frac{(x/\xi)^\gamma}{\alpha}\right]^{\alpha+\tau}.$$

We have

$$\begin{aligned}
\lim_{\alpha\to\infty} \ln w_2 =& \lim_{\alpha\to\infty} (\alpha+\tau)\ln\left[1+\frac{(x/\xi)^\gamma}{\alpha}\right] \\
=& \lim_{\alpha\to\infty} \frac{\ln\left[1+\frac{(x/\xi)^\gamma}{\alpha}\right]}{(\alpha+\tau)^{-1}} \\
=& \lim_{\alpha\to\infty} \frac{\left[1+\frac{(x/\xi)^\gamma}{\alpha}\right]^{-1}\left(-(x/\xi)^\gamma\alpha^{-2}\right)}{-(\alpha+\gamma)^{-2}} \\
=&(x/\xi)^\gamma \lim_{\alpha\to\infty}\left(1+\frac{\tau}{\alpha}\right)^2\left[1+\frac{(x/\xi)^\gamma}{\alpha}\right]^{-1} \\
=&(x/\xi)^\gamma.
\end{aligned}$$

Hence,

$$\lim_{\alpha\to\infty} w_2 = e^{(x/\xi)^\gamma}.$$

Finally,

$$\lim_{\alpha\to\infty} f_X(x) = \frac{\gamma x^{\gamma\tau-1}e^{-\left(\frac{x}{\xi}\right)^\gamma}}{\Gamma(\tau)\xi^{\gamma\tau}}$$

which is the pdf of the transformed gamma distribution ■

Practice Problems

Problem 23.1
Show:

$$\lim_{\tau\to\infty}\left(1+\frac{\alpha}{\tau}\right)^{\alpha+\tau-\frac{1}{2}}=e^{\alpha} \quad \text{and} \quad \lim_{\tau\to\infty}\left[1+\frac{(\xi/x)^{\gamma}}{\tau}\right]^{\alpha+\tau}=e^{\left(\frac{\xi}{x}\right)^{\gamma}}.$$

Problem 23.2
For a transformed beta distribution with parameters α, γ and θ, let θ go to infinity with the ratio $\theta\tau^{\frac{1}{\gamma}} \to \xi$.

Show that the result is the inverse transformed Gamma distribution.

Problem 23.3
For an inverse Pareto distribution with parameters τ and θ, let $\theta \to 0, \tau \to \infty$ and $\tau\theta \to \xi$.

Show that the result is the inverse exponential distribution.

24 The Linear Exponential Family of Distributions

A random variable X is said to belong to the **linear exponential family** (LEF) if its pdf can be reformulated in terms of a parameter θ as

$$f(x,\theta)=\frac{p(x)}{q(\theta)}e^{r(\theta)x}$$

where $p(x)$ depends solely of x. The support of X must not depend on θ. Other parameters of the distribution (parameters that are not θ) may occur in the expressions $p(x), q(\theta)$, and $r(\theta)$. But they have no role in determining whether the distribution belongs to the linear exponential family.

Example 24.1
Show that a normal distribution with parameters μ and σ^2 belongs to the linear exponential family with $\theta=\mu$.

Solution.
The pdf of the normal distribution function can be written as

$$\begin{aligned}
f(x,\mu)=&\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\\
=&\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x^2-2\mu x+\mu^2}{\sigma^2}\right)}\\
=&\frac{1}{\sigma\sqrt{2\pi}e^{\frac{\mu^2}{2\sigma^2}}}e^{-\frac{x^2}{2\sigma^2}}e^{\frac{\mu}{\sigma^2}x}.
\end{aligned}$$

Thus, X belongs to the linear exponential family with $p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$, $q(\mu)=e^{\frac{\mu^2}{2\sigma^2}}$, and $r(\mu)=\frac{\mu}{\sigma^2}$ ■

Example 24.2
Let X belong in the linear exponential family. Find an expression of $E(X)$.

Solution.
Taking the logarithm of $f(x,\theta)$ we find

$$\ln f(x,\theta)=\ln p(x)+r(\theta)x-\ln q(\theta).$$

Now, differentiating both sides with respect to θ, we find

$$\frac{1}{f(x,\theta)}\frac{\partial f}{\partial\theta}(x,\theta)=xr'(\theta)-\frac{q'(\theta)}{q(\theta)}\Longrightarrow\frac{\partial f}{\partial\theta}(x,\theta)=r'(\theta)f(x,\theta)x-\frac{q'(\theta)}{q(\theta)}f(x,\theta).$$

Now we integrate both sides with respect to x to obtain

$$\int \frac{\partial f}{\partial \theta}(x,\theta)dx = \int xr'(\theta)f(x,\theta)dx - \int \frac{q'(\theta)}{q(\theta)}f(x,\theta)dx.$$

By the definition of the family, the support of X is independent of θ and so is the range of x. Thus, we can write

$$\begin{aligned}\frac{\partial}{\partial \theta}\int f(x,\theta)dx =& r'(\theta)\int xf(x,\theta)dx - \frac{q'(\theta)}{q(\theta)}\int f(x,\theta)dx\\ \frac{\partial}{\partial \theta}(1) =& r'(\theta)E(X) - \frac{q'(\theta)}{q(\theta)}.\end{aligned}$$

That is,

$$r'(\theta)E(X) - \frac{q'(\theta)}{q(\theta)} = 0.$$

Solving for $E[X]$ yields

$$E(X) = \frac{q'(\theta)}{r'(\theta)q(\theta)} = \mu(\theta) \blacksquare$$

Example 24.3

Let X belong in the linear exponential family. Find the variance of X.

Solution.

From the previous example, we have

$$\frac{\partial f}{\partial \theta}(x,\theta) = [x-\mu(\theta)]r'(\theta)f(x,\theta).$$

Differentiating with respect to θ and using the already obtained first derivative of $f(x,\theta)$ yield

$$\begin{aligned}\frac{\partial^2 f}{\partial \theta^2}(x,\theta) =& r''(\theta)[x-\mu(\theta)]f(x,\theta) - \mu'(\theta)r'(\theta)f(x,\theta) + [x-\mu(\theta)]r'(\theta)\frac{\partial f}{\partial \theta}(x,\theta)\\ =& r''(\theta)[x-\mu(\theta)]f(x,\theta) - \mu'(\theta)r'(\theta)f(x,\theta) + [x-\mu(\theta)]^2[r'(\theta)]^2 f(x,\theta).\end{aligned}$$

Now we integrate both sides with respect to x to obtain

$$\int \frac{\partial^2 f}{\partial \theta^2}(x,\theta)dx = [r'(\theta)]^2\text{Var}(X) - r'(\theta)\mu'(\theta).$$

By the definition of the family, the support of X is independent of θ and so is the range of x. Thus, we can write

$$\frac{\partial^2}{\partial \theta^2}\int v(x,\theta)dx = [r'(\theta)]^2\text{Var}(X) - r'(\theta)\mu'(\theta).$$

That is,

$$[r'(\theta)]^2\mathrm{Var}(X) - r'(\theta)\mu'(\theta) = 0.$$

Solving for $\mathrm{Var}(X)$ yields

$$\mathrm{Var}(X) = \frac{\mu'(\theta)}{r'(\theta)} \blacksquare$$

Example 24.4
Let X be the normal distribution with $\theta = \mu = 24$ and $\sigma = 3$. Use the formulas of this section to verify that $E(X) = 24$ and $\mathrm{Var}(X) = 9$.

Solution.
For the given distribution, we have $p(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$, $q(\theta) = e^{\frac{\theta^2}{2\sigma^2}}$, and $r(\theta) = \frac{\theta}{\sigma^2}$. Thus,

$$E(X) = \frac{q'(\theta)}{r'(\theta)q(\theta)} = \frac{\frac{\theta}{\sigma^2}e^{\frac{\theta^2}{2\sigma^2}}}{\frac{1}{\sigma^2}e^{\frac{\theta^2}{2\sigma^2}}} = 24.$$

Likewise, we have

$$\mathrm{Var}(X) = \frac{\mu'(\theta)}{r'(\theta)} = \frac{1}{\frac{1}{9}} = 9 \blacksquare$$

Practice Problems

Problem 24.1
Show that the Gamma distributions belongs to the linear exponential family.

Problem 24.2
Show that the Poisson distribution with parameter λ belongs to the linear exponential family.

Problem 24.3
Show that the binomial distribution with m trials and parameter p belongs to the linear exponential family.

Problem 24.4
Use the formulas of this section to find the mean and the variance of the Poisson distribution.

Problem 24.5
Use the formulas of this section to find the mean and the variance of the binomial distribution.

Discrete Distributions

The distributions and the families of distributions that we have been discussing so far are mainly used to describe the amount of risks. Next, we turn our attention to distributions that describe the number of risks or claims. In this chapter, we introduce classes of counting distributions. By a **counting distribution** we mean a discrete distribution with support a subset of $\mathbb{N} \cup \{0\}$. We will adopt the following notation: If N is the random variable representing the number of events (or claims) then the probability mass function or the **probability function** $\Pr(N = k)$ will be denoted by p_k.

25 The Poisson Distribution

The first counting distribution that we consider is the Poisson distribution.

A random variable X is said to be a **Poisson** random variable with parameter λ if its probability mass function has the form

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!}, \ k = 0, 1, 2, \cdots .$$

Note that

$$\sum_{k=0}^{\infty} p_k = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1.$$

The Poisson random variable is most commonly used to model the number of random occurrences of some phenomenon in a specified unit of space or time. For example, the number of phone calls received by a telephone operator in a 10-minute period or the number of typos per page made by a secretary.

The probability generating function of N is

$$P_N(z) = E(z^N) = \sum_{k=0}^{\infty} p_k z^k = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(z\lambda)^k}{k!} = e^{\lambda(z-1)}$$

and the moment generating function is

$$M_N(t) = P_N(e^t) = e^{\lambda(e^t-1)}.$$

Using the pgf, we can find the mean and the variance:

$$\begin{aligned}
E(N) =& P_N'(1) = \lambda\\
E[N(N-1)] =& P_N''(1) = \lambda^2\\
\mathrm{Var}(N) =& E[N(N-1)] + E(N) - [E(N)]^2\\
=& \lambda^2 + \lambda - \lambda^2 = \lambda.
\end{aligned}$$

We next discuss two useful properties of Poisson distributions. The first concerns the sum of independent Poisson random variables.

Theorem 25.1
Let $N_1, N_2, \cdots, N_n$ be n independent Poisson random variables with parameterd $\lambda_1, \lambda_2, \cdots, \lambda_n$ respectively. Then the random variable $S = N_1 + N_2 + \cdots + N_n$ is also a Poisson distribution with parameters $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Proof.
Since $N_1, N_2, \cdots, N_n$ are independent so are $e^{tN_1}, e^{tN_2}, \cdots, e^{tN_m}$. Using the fact that the expectation of a product of independent random variables is the product of the individual expectations, we have

$$\begin{aligned} M_S(t) =& E(e^{t(N_1+N_2+\cdots+N_n)} = E[e^{tN_1}e^{tN_2}\cdots e^{tN_n}] \\ =& E[e^{tN_1}]E[e^{tN_2}]\cdots E[e^{tN_n}] \\ =& e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)}\cdots e^{\lambda_n(e^t-1)} \\ =& e^{\lambda(e^t-1)} \end{aligned}$$

where $\lambda = \lambda_1+\lambda_2+\cdots+\lambda_n$. Since the moment generating function uniquely determines the distribution, S has a Poisson random variable with parameter λ ■

The second result is very useful when modeling insurance losses or claims where claims have classification. This result is known as the **decomposition property** of Poisson distribution.

Theorem 25.2
Let the total number of events N be a Poisson random variable with mean λ. Suppose that the events can be classified into independent types such as Type 1, Type 2,$\cdots$, Type m with probabilities $p_k,\ k = 1, \cdots, m$. Let N_k be the random variable representing the number of events of Type $k,\ k = 1, \cdots, m$. Then $N_1, N_2, \cdots, N_m$ are mutually independent Poisson distributions with means $\lambda p_1, \lambda p_2, , \cdots, \lambda p_m$ respectively.

Proof.
Given $n = n_1+n_2+\cdots+n_m$, the conditional joint distribution of $N_1, N_2, \cdots, N_m$ is a multinomial distribution with parameters $n, p_1, p_2, \cdots, p_m$. Its conditional pmf is

$$\Pr(N_1 = n_1, N_2, n_2, \cdots, N_m|N = n) = \frac{n!}{n_1!n_2!\cdots n_m!}p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m}.$$

(See p.155 of [2]).
Now since N has a Poisson distribution with mean λ, we have

$$\Pr(N = n) = \frac{e^{-\lambda}\lambda^n}{n!}.$$

The unconditional joint distribution of $N_1, N_2, \cdots, N_m$ is

$$\begin{gathered}\Pr(N_1 = n_1, N_2 = n_2, \cdots, N_m = n_m) = \\ \Pr(N_1 = n_1, N_2 = n_2, \cdots, N_m = n_m | N = n) \times \Pr(N = n) = \\ \frac{n!}{n_1! n_2! \cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} \frac{e^{-\lambda} \lambda^n}{n!}.\end{gathered}$$

With some algebra, we find the following

$$\begin{aligned}\Pr(N_1 = n_1, N_2 = n_2, \cdots, N_m = n_m) =& \\ \frac{p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} e^{-\lambda(p_1 + p_2 + \cdots + p_m)} \lambda^{n_1 + n_2 + \cdots + n_m}}{n_1! n_2! \cdots n_m!} =& \\ \frac{(\lambda p_1)^{n_1} (\lambda p_2)^{n_2} \cdots (\lambda p_m)^{n_m} e^{-\lambda n_1} e^{-\lambda n_2} \cdots e^{-\lambda n_m}}{n_1! n_2! \cdots n_m!} =& \\ \prod_{i=1}^{m} \frac{e^{\lambda p_i} (\lambda p_i)^{n_i}}{n_i!}.&\end{aligned}$$

Next, the unconditional marginal pmf of $N_i, \ i = 1, 2, \cdots, m$ is

$$\begin{aligned}\Pr(N_i = n_i) =& \sum_{n=n_i}^{\infty} \Pr(N_i = n_i | N = n) \Pr(N = n) \\ =& \sum_{n=n_i}^{\infty} C(n, n_i) p^{n_i} (1 - p)^{n - n_i} \frac{e^{-\lambda} \lambda^n}{n!} \\ =& \frac{e^{-\lambda} (\lambda p_i)^{n_i}}{n_i!} \sum_{n=n_i}^{\infty} \frac{[\lambda(1 - p_i)^{n - n_i}}{(n - n_i)!} \\ =& \frac{e^{-\lambda} (\lambda p_i)^{n_i}}{n_i!} e^{\lambda(1 - p_i)} \\ =& \frac{e^{-\lambda} (\lambda p_i)^{n_i}}{n_i!}.\end{aligned}$$

Thus, N_i has a Poisson distribution with mean λp_i. Since the joint distribution is the product of the marginal distributions, $N_1, N_2, \cdots, N_m$ are mutually independent ■

Example 25.1
The number of patients that arrive at an emergency room during an hour has a Poisson distribution with mean 6. Of every 50 patients that arrive to the emergency, one has an appendicitis. Calculate the probability that the number of patients with appendicitis that arrive to the emergency room in a 24-hour period will exceed the expected number of appendicitis.

Solution.
Let N be the number of patients with appendicitis that arrive to the emergency room in the 24 hour period. Then

$$E(N) = 6 \times 24 \times \frac{1}{50} = 2.88.$$

Thus,

$$\Pr(N > 2.88) = 1 - p_0 - p_1 - p_2 = 1 - e^{-2.88} - 2.88e^{-2.88} - 2.88^2\frac{e^{-2.88}}{2} = 0.5494 \blacksquare$$

Example 25.2
In a portfolio of insurance, a claim can be classified as Type A, Type B, or Type C with probabilities $0.2, 0.3$, and 0.5 respectively. Suppose that the total number of claims is a Poisson random variable with mean 10. Each type has a Poisson distribution and these random variables are supposed to be independent. What is the probability that out of 5 claims, 2 are of Type A?

Solution.
We have

$$\begin{aligned}
\Pr(N_A = 2|N = 5) =& \frac{\Pr(N_A = 2, N_B + N_C = 3)}{\Pr(N = 5)} \\
=& \frac{\frac{e^{-2}2^2}{2!}\frac{e^{-8}8^3}{3!}}{\frac{e^{-10}10^5}{5!}} = 0.2048 \blacksquare
\end{aligned}$$

Practice Problems

Problem 25.1
The number of students in College Algebra that go to Math Help room has a Poisson distribution with mean 2 per hour. The number of students in Trigonometry that go to Math Help room has a Poisson distribution with mean 1 per hour. The number of students in Pre-calculus that go to Math Help room has a Poisson distribution with mean 0.5 per hour.

Calculate the probability that more than 3 students (in any class) go to Math Help room between 2:00om and 4:pm.

Problem 25.2
Let N_1 and N_2 be two independent Poisson random variables with mean 2 and 3 respectively.

Calculate $\Pr(N_1 + N_2 = 1)$.

Problem 25.3
The number of monthly car wrecks in Russellville follows a Poisson distribution with mean 10. There are three possible mutually exclusive causes of accidents. The probability of a wreck due to icy road is 0.40. The probability of a wreck due to break malfunction is 0.45 and the probability of a wreck due to driver's error is 0.15.

What is the probability that exactly 5 wrecks will occur this month due to driver's error?

Problem 25.4
Suppose that the number of cars arriving for service at a service facility in one week has a Poisson distribution with a mean of 20. Suppose that each car is classified as either domestic or foreign. Suppose also that each time a car arrives for service there is a 75% chance that it domestic and a 25% chance that it is foreign.

What is the weekly expected number of foreign cars arriving at the service facility?

26 The Negative Binomial Distribution

The Poisson distribution is a one-parameter counting distribution. The negative binomial distribution is a two-parameter counting distribution.

A random variable X is a **negative binomial distribution** if its pdf is of the form

$$p_n = \Pr(N = n) = \binom{n+r-1}{n} \left(\frac{1}{1+\beta)}\right)^r \left(\frac{\beta}{1+\beta)}\right)^n$$

where $n = 0, 1, 2, \cdots$ and $r > 0, \beta > 0$. Note that r is a positive real number not just a positive integer.

Example 26.1
Find the probability generating function of the negative binomial distribution.

Solution.
We have

$$\begin{aligned}
P_N(z) =& E(z^N) = \sum_{n=0}^{\infty} z^n \binom{n+r-1}{n} \left(\frac{1}{1+\beta)}\right)^r \left(\frac{\beta}{1+\beta}\right)^n \\
=& \sum_{n=0}^{\infty} \binom{n+r-1}{n} \left(\frac{1}{1+\beta}\right)^r \left(\frac{z\beta}{1+\beta}\right)^n \\
=& \left(\frac{1}{1+\beta}\right)^r \left(1 - \frac{z\beta}{1+\beta}\right)^{-r} \sum_{n=0}^{\infty} \binom{n+r-1}{n} (p^*)^r (1-p^*)^n \\
=& \left(\frac{1}{1+\beta}\right)^r \left(1 - \frac{z\beta}{1+\beta}\right)^{-r} = [1 - \beta(z-1)]^{-r}
\end{aligned}$$

where $p^* = 1 - \frac{z\beta}{1+\beta}$ ■

Example 26.2
Find the mean and the variance of a negative binomial distribution N.

Solution.
We have

$$\begin{aligned}
E(N) =& P_N'(1) = r\beta \\
E[N(N-1)] =& P_N''(1) = r(r+1)\beta^2 \\
\mathrm{Var}(N) =& E[N(N-1)] + E(N) - (E(N))^2 \\
=& r\beta(1+\beta).
\end{aligned}$$

Note that since $\beta > 0$, we have $\mathrm{Var}(N) > E(N)$ ■

Example 26.3
Show that a negative binomial distribution is a mixture of a Poisson distribution with a random parameter distributed Gamma.

Solution.
Suppose that $N|\Lambda$ has a Poisson distribution with parameter λ. Suppose that Λ has a Gamma distribution with parameters α and θ. We wish to find the distribution of N. We have

$$\begin{aligned}
p_n =&\Pr(N=n) = \int_0^\infty \Pr(N=n|\Lambda=\lambda)\Pr(\Lambda=\lambda)d\lambda\\
=&\int_0^\infty \frac{e^{-\lambda}\lambda^n}{n!}\frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\theta}}}{\theta^\alpha\Gamma(\alpha)}d\lambda = \frac{1}{n!}\frac{1}{\theta^\alpha\Gamma(\alpha)}\int_0^\infty e^{-\lambda\left(1+\frac{1}{\theta}\right)}\lambda^{n+\alpha-1}d\lambda\\
=&\binom{n+\alpha-1}{n}\left(\frac{1}{1+\theta)}\right)^\alpha\left(\frac{\theta}{1+\theta)}\right)^n\int_0^\infty \frac{e^{-\frac{\lambda}{(1+1/\theta)^{-1}}}\lambda^{\alpha+n-1}}{(1+1/\theta)^{-(\alpha+n)}\Gamma(\alpha+n)}d\lambda\\
=&\binom{n+\alpha-1}{n}\left(\frac{1}{1+\theta}\right)^\alpha\left(\frac{\theta}{1+\theta}\right)^n.
\end{aligned}$$

This shows that the mixed Poisson, with a Gamma mixing distribution is the same as the negative binomial distribution ■

Example 26.4 ‡
Glen is practicing his simulation skills. He generates 1000 values of the random variable X as follows:
(i) He generates the observed value λ from the gamma distribution with $\alpha = 2$ and $\theta = 1$ (hence with mean 2 and variance 2).
(ii) He then generates x from the Poisson distribution with mean λ.
(iii) He repeats the process 999 more times: first generating a value λ, then generating x from the Poisson distribution with mean λ.
(iv) The repetitions are mutually independent.
Calculate the expected number of times that his simulated value of X is 3.

Solution.
By the previous result, X is a negative binomial distribution with $r = \alpha = 2$ and $\beta = \theta = 1$. From Table C, we have

$$p_3 = \frac{r(r+1)(r+2)\beta^3}{3!(1+\beta)^{r+3}} = \frac{(2)(3)(4)1^3}{3!(2)^5} = 0.125.$$

Thus we expect $1000p_3 = 125$ out of 1000 simulated values to be 3 ■

Example 26.5
Show that a negative binomial distribution with $r \to \infty, n \to 0$, and $rn \to \xi$ results in a Poisson distribution.

Solution.
Replace β in the pgf of N by $\frac{\xi}{r}$. We have

$$\begin{aligned}
\lim_{r\to\infty} P_N(z) =& \lim_{r\to\infty}\left[1-\frac{\xi}{r}(z-1)\right]\\
=&\exp\left\{\lim_{r\to\infty} -r\ln\left[1-\frac{\xi}{r}(z-1)\right]\right\}\\
=&\exp\left\{\lim_{r\to\infty} -\frac{\ln\left[1-\xi(z-1)r^{-1}\right]}{r^{-1}}\right\}\\
=&\exp\left\{\lim_{r\to\infty} \frac{[1-\xi(z-1)r^{-1}]\xi(z-1)r^{-2}}{r^{-2}}\right\}\\
=&\exp\left[\lim_{r\to\infty}\frac{r\xi(z-1)}{r-\xi(z-1)}\right] = e^{\xi(z-1)}.
\end{aligned}$$

This is the pgf of a Poisson distribution with parameter ξ ■

Example 26.6 ‡
Actuaries have modeled auto windshield claim frequencies and have concluded that the number of windshield claims filed per year per driver follows the Poisson distribution with parameter Λ, where Λ follows the Gamma distribution with mean 3 and variance 3. Calculate the probability that a driver selected at random will file no more than 1 windshield claim next year.

Solution.
We are given that $\alpha\theta = 3$ and $\alpha\theta^2 = 3$. Thus, $\alpha = 3$ and $\theta = 1$. On the other hand, the number of windshield is a negative binomial distribution with $r = \alpha = 3$ and $\beta = \theta = 1$. Hence,

$$\Pr(N \le 0) = p_0 + p_1 = \frac{1}{8} + \frac{3}{16} = \frac{5}{16} = 0.3125 \ ■$$

For the special case $r = 1$, the negative binomial random variable is called the **geometric** random variable. The geometric distribution, like the exponential distribution, has the **memoryless property**: the distribution of a variable that is known to be in excess of some value d does not depend on d.

Example 26.7
The number of students having the flu in a particular college follows a geometric distribution with $\beta = 0.4$. What is the difference between the following two values:
A. The expected number of students having the flu in excess of 3 if it is known that the number of students having the flu is greater than 6.
B. The expected number of students having the flu in excess of 2 if it is known that the number of students having the flu is greater than 2.

Solution.
They are equal by the memoryless property of the geometric distribution ■

Practice Problems

Problem 26.1
You are modeling the frequency of events, and you need to select a distribution to use. You observe that the variance of the number of events is less than the mean number of events.

Which of the following distributions should you use?

(a) Poisson
(b) Negative binomial
(c) Geometric
(d) None of the above

Problem 26.2
Assume that a certain type of claims in one month follows a geometric distribution with $\beta = 3$.

What is the probability that there is at least 1 claim in a month?

Problem 26.3
Suppose that the number of claims N in one month follows a geometric distribution with $\beta = 3$.

Calculate $E(N)$ and $\mathrm{Var}(N)$.

Problem 26.4
Suppose that the number of claims N in one month folloes a negative binomial distribution with $r = 3$ and $\beta = 2$.

Calculate $E(N)$ and $\mathrm{Var}(N)$.

Problem 26.5
Let $N|\Lambda$ have a negative binomial distribution with $r = 2$ and λ. Let Λ have a Gamma distribution with $\alpha = 2$ and $\theta = 3$.

Find $\mathrm{Var}[N]$.

Problem 26.6
Let N be a negative binomial random variable with mean 8 and variance 40.

Find the parameters r and β.

Problem 26.7
Find the probability generating function of a geometric random variable N.

Problem 26.8
Find the coefficient of variation of a negative binomial random variable with parameters r and β.

27 The Bernoulli and Binomial Distributions

Binomial experiments are problems that consist of a fixed number of trials n, with each trial having exactly two possible outcomes: Success and failure. The probability of a success is denoted by q and that of a failure by $1-q$. Also, we assume that the trials are independent, that is what happens to one trial does not affect the probability of a success in any other trial. The prefix "bi" in binomial experiment refers to the fact that there are two possible outcomes (e.g., head or tail, true or false, working or defective) to each trial in the binomial experiment.

Let N represent the number of successes that occur in m trials. Then N is said to be a **binomial random variable** with parameters (m, q) If $m = 1$ then N is said to be a **Bernoulli** random variable. The central question of a binomial experiment is to find the probability of r successes out of m trials. In the next paragraph we'll see how to compute such a probability. Now, anytime we make selections from a population without replacement, we do not have independent trials. For example, selecting a ball from an urn that contain balls of two different colors.

The probability of r successes in a sequence out of m independent trials is given by $q^r(1-q)^{m-r}$. Since the binomial coefficient $C(m, r)$ counts all the number of outcomes that have r successes and $m-r$ failures, the probability of having r successes in any order is given by the binomial mass function

$$p_r = \Pr(N = r) = C(m, r)q^r(1-q)^{m-r}$$

where q denotes the probability of a success. Note that

$$\sum_{k=0}^{m} p_k = \sum_{k=0}^{m} C(m, k)q^k(1-q)^{m-k} = (q + 1 - q)^m = 1.$$

Example 27.1
Suppose that in a particular sheet of 100 postage stamps, 3 are defective. The inspection policy is to look at 5 randomly chosen stamps on a sheet and to release the sheet into circulation if none of those five is defective. Write down the random variable, the corresponding probability distribution and then determine the probability that the sheet described here will be allowed to go into circulation.

Solution.
Let N be the number of defective stamps in the sheet. Then N is a binomial

random variable with probability distribution

$$\Pr(N = k) = C(5,k)(0.03)^k(0.97)^{5-k},\ \ k = 0,1,2,3,4,5.$$

Now,

$$\Pr(\text{sheet goes into circulation}) = \Pr(N = 0) = (0.97)^5 = 0.859 \ \blacksquare$$

Example 27.2
Suppose 40% of the student body at a large university are in favor of a ban on drinking in dormitories. Suppose 5 students are to be randomly sampled. Find the probability that
(a) 2 favor the ban.
(b) less than 4 favor the ban.
(c) at least 1 favor the ban.

Solution.
(a) $\Pr(N = 2) = C(5,2)(.4)^2(0.6)^3 \approx 0.3456.$
(b) $\Pr(N < 4) = \Pr(0) + \Pr(1) + \Pr(2) + \Pr(3) = C(5,0)(0.4)^0(0.6)^5 + C(5,1)(0.4)^1(0.6)^4 + C(5,2)(0.4)^2(0.6)^3 + C(5,3)(0.4)^3(0.6)^2 \approx 0.913.$
(c) $\Pr(N \geq 1) = 1 - \Pr(N < 1) = 1 - C(5,0)(0.4)^0(0.6)^5 \approx 0.922 \ \blacksquare$

Example 27.3
A student has no knowledge of the material to be tested on a true-false exam with 10 questions. So, the student flips a fair coin in order to determine the response to each question.
(a) What is the probability that the student answers at least six questions correctly?
(b) What is the probability that the student answers at most two questions correctly?

Solution.
(a) Let N be the number of correct responses. Then N is a binomial random variable with parameters $m = 10$ and $q = \frac{1}{2}$. So, the desired probability is

$$\begin{aligned}\Pr(N \geq 6) =& \Pr(N = 6) + \Pr(N = 7) + \Pr(N = 8) + \Pr(N = 9) + \Pr(N = 10)\\ =& \sum_{k=6}^{10} C(10,k)(0.5)^k(0.5)^{10-k} \approx 0.3769.\end{aligned}$$

(b) We have

$$\Pr(N \leq 2) = \sum_{k=0}^{2} C(10,k)(0.5)^k(0.5)^{10-k} \approx 0.0547 \ \blacksquare$$

Theorem 27.1
Let N be binomial distribution with parameters (m, q). Then
(a) $P_N(z) = (zq + 1 - q)^m$.
(b) $E(N) = mq >$ and $\text{Var}(N) = mq(1-q)$. Note that $\text{Var}(N) < E(N)$.

Proof.
(a) The probability generating function is found by using the binomial formula $(a+b)^n = \sum_{k=0}^n a^k b^{n-k}$:

$$P_N(z) = \sum_{k=0}^{m} C(m,k) z^k q^k (1-q)^{m-k} = \sum_{k=0}^{m} (zq)^k (1-q)^{m-k} = (zq + 1 - q)^m.$$

(b) We have:

$$\begin{aligned}
P_N'(z) =& mq(zq + 1 - q)^{m-1} \\
E(N) =& P_N'(1) = mq \\
P_N''(z) =& m(m-1)q^2(zq + 1 - q)^{m-2} \\
E[N(N-1)] =& P_N''(1) = m(m-1)q^2 \\
\text{Var}(N) =& E[N(N-1)] + E(N) - (E(N))^2 \\
=& m(m-1)q^2 + mq - m^2q^2 = mq(1-q) \blacksquare
\end{aligned}$$

Example 27.4
Let N be a binomial random variable with parameters $(12, 0.5)$. Find the variance and the standard deviation of N.

Solution.
We have $m = 12$ and $q = 0.5$. Thus, $\text{Var}(N) = mq(1-q) = 6(1-0.5) = 3$. The standard deviation is $SD(N) = \sqrt{3}$ $\blacksquare$

Example 27.5
An exam consists of 25 multiple choice questions in which there are five choices for each question. Suppose that you randomly pick an answer for each question. Let N denote the total number of correctly answered questions. Write an expression that represents each of the following probabilities.
(a) The probability that you get exactly 16, or 17, or 18 of the questions correct.
(b) The probability that you get at least one of the questions correct.

Solution.
(a) We have

$$\begin{aligned}\Pr(N = 16 \textit{ or } N = 17 \textit{ or } N = 18) =& C(25,16)(0.2)^{16}(0.8)^9 + C(25,17)(0.2)^{17}(0.8)^8\\ &+C(25,18)(0.2)^{18}(0.8)^7.\end{aligned}$$

(b) $\Pr(N \geq 1) = 1 - P(N = 0) = 1 - C(25,0)(0.8)^{25}$ ■

A useful fact about the binomial distribution is a recursion for calculating the probability mass function.

Theorem 27.2
Let N be a binomial random variable with parameters (m, q). Then for $k = 1, 2, 3, \cdots, n$

$$p(k) = \frac{q}{1-q}\frac{m-k+1}{k}p(k-1)$$

Proof.
We have

$$\begin{aligned}\frac{p(k)}{p(k-1)} =& \frac{C(m,k)q^k(1-q)^{m-k}}{C(m,k-1)q^{k-1}(1-q)^{m-k+1}}\\ =& \frac{\frac{m!}{k!(m-k)!}q^k(1-q)^{m-k}}{\frac{m!}{(k-1)!(m-k+1)!}q^{k-1}(1-q)^{m-k+1}}\\ =& \frac{(m-k+1)q}{k(1-q)} = \frac{q}{1-q}\frac{m-k+1}{k}\end{aligned}$$ ■

Practice Problems

Problem 27.1
You are again modeling the frequency of events, and you need to select a distribution to use. You observe that the variance of the number of events is less than the mean number of events.

Which of the following distributions should you use?

(a) Binomial
(b) Poisson
(c) Negative binomial
(d) Geometric
(e) None of the above are correct.

Problem 27.2
Let N be a random variable which follows a binomial distribution with parameters $m = 20$ and $q = 0.2$

Calculate $E(2^N)$.

Problem 27.3
Suppose that the number of claims N has a binomial distribution with $m = 2$ and $q = 0.7$.

Calculate $E(N)$ and $\text{Var}(N)$.

Problem 27.4
An insurance company insures 15 risks, each with a 2.5% probability of loss. The probabilities of loss are independent.

What is the probability of 3 or more losses in the same year?

Problem 27.5
Suppose that $N|\Lambda$ has a binomial distribution with parameters Λ and $q = 0.4$. Suppose that Λ has a probability function defined by $p(1) = p(2) = p(3) = p(4) = 0.25$.

Calculate the unconditional variance of N.

Problem 27.6
An actuary has determined that the number of claims follows a binomial distribution with mean 6 and variance 3.

Calculate the probability that the number of claims is at least 3 but less than 5.

Problem 27.7
Let N be a binomial random variable with $m = 10$ and $q = 0.2$. Let $F(m)$ denote the cdf of N. Complete the following table.

m	0	1	2	3	4
p_m					
$F(m)$					

Problem 27.8
Let N_1 and N_2 be two independent binomial random variables with respective parameters (m_1, q) and (m_2, q).

Compute the pmf of $N_1 + N_2$.

28 The $(a, b, 0)$ Class of Discrete Distributions

The Poisson, negative binomial, geometric, and binomial distributions all satisfy the recursive equation

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} \tag{28.1}$$

for some constants a and b and for $k \in \mathbb{N}$. We will denote the collection of these discrete distributions by $C(a, b, 0)$. The table below list the parameters a and b for each distribution together with the probability function at 0.

Distributions	a	b	p_0
Poisson	0	λ	$e^{-\lambda}$
Negative binomial	$\frac{\beta}{1+\beta}$	$(r-1)\frac{\beta}{1+\beta}$	$(1+\beta)^{-r}$
Geometric	$\frac{\beta}{1+\beta}$	0	$(1+\beta)^{-1}$
Binomial	$-\frac{q}{1-q}$	$(m+1)\frac{q}{1-q}$	$(1-q)^m$

Note that (28.1) can be written as

$$k\frac{p_k}{p_{k-1}} = ak + b \tag{28.2}$$

so that the right-side is a linear function of k with slope a. Thus, if we graph $k\frac{p_k}{p_{k-1}}$ against k, we get a straight line that has a positive slope for the negative binomial or geometric distributions, negative slope for the binomial distribution, and 0 slope for the Poisson distribution. This characterization of the sign of a, helps in the selection of the distribution when (28.2) is given. We illustrate this idea in the next example.

Example 28.1
Let N be a member of $C(a, b, 0)$ satisfying the recursive probabilities

$$k\frac{p_k}{p_{k-1}} = \frac{3}{4}k + 3.$$

Identify the distribution N.

Solution.
Since $a > 0$ and $b > 0$, N is the negative binomial distribution. We have $\frac{\beta}{1+\beta} = \frac{3}{4}$ which implies $\beta = 3$. Also, we have $(r-1)\frac{\beta}{1+\beta} = 3$ which yields $r = 5$ ■

Example 28.2 ‡
The distribution of accidents for 84 randomly selected policies is as follows:

# of Accidents	# of Policies
0	32
1	26
2	12
3	7
4	4
5	2
6	1

Identify the frequency model that best represents these data.

Solution.
We have

# of Accidents	# of Policies	$k\frac{p_k}{p_{k-1}}$
0	32	NA
1	26	0.8125
2	12	0.9231
3	7	1.75
4	4	2.2857
5	2	2.5
6	1	3

Plotting the points $\left(k, k\frac{p_k}{p_{k-1}}\right)$ we find that the negative binomial distribution is the best model from $C(a,b,0)$ to use (slope of line is positive) ■

Example 28.3
The number of dental claims in a year follows a Poisson distribution with a mean of λ. The probability of exactly six claims during a year is 40% of the probability that there will be 5 claims. Determine the probability that there will be 4 claims.

Solution.
Let N be the number of dental claims in a year. Since N is a member of $C(a,b,0)$ we can write

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} = \frac{\lambda}{k},\ k = 1, 2, \cdots.$$

We are told that

$$0.4 = \frac{p_6}{p_5} = \frac{\lambda}{6} \Longrightarrow \lambda = 2.4.$$

Thus,

$$p_4 = \frac{e^{-2.4}(2.4)^4}{4!} = 0.1254 \blacksquare$$

Example 28.4 ‡
X is a discrete random variable with a probability function which is a member of the $C(a, b, 0)$ class of distributions.
You are given:
(i) $\Pr(X = 0) = \Pr(X = 1) = 0.25$
(ii)$\Pr(X = 2) = 0.1875$
Calculate $\Pr(X = 3)$.

Solution.
Let N denote the distribution under consideration. Since N is a member of the $C(a, b, 0)$ class, we have the recursive relation

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1}, \ k = 1, 2, \cdots .$$

From (i), we obtain $0.25 = (a + b)(0.25)$ which implies $a + b = 1$. From (ii), we have $0.1875 = \left(1 - \frac{b}{2}\right)(0.25)$ which implies $b = 0.5$. Hence, $a = 0.5$ and

$$\Pr(X = 3) = \left(0.5 + \frac{0.5}{3}\right)(0.1875) = 0.125 \blacksquare$$

Example 28.5 ‡
A discrete probability distribution has the following properties:
(i) $p_k = c\left(1 + \frac{1}{k}\right) p_{k-1}, \ k = 1, 2, \cdots$
(ii) $p_0 = 0.5$.
Determine the value of c.

Solution.
This is a class $C(a, b, 0)$ distribution with $a = b = c$. We will go through each distribution in the class and see which one fits the condition of the problem.
• If the distribution is Poisson then $a = b = c = 0$. In this case, $p_0 = 0.5$ and $p_k = 0$ for $k = 1, 2, \cdots$. Since $\sum_{k=1}^{\infty} p_k = 0.5 \neq 1$, this distribution can not be the answer.

• If the distribution is geometric then $a = b = c = 0$ and as in the previous case, the distribution can not be geometric.
• The distribution can not be binomial since $a = b$. For binomial a and b are of opposite signs.
• The only possibility remaining is that the distribution is negative binomial. In this case, $a = \frac{\beta}{1+\beta}$ and $b = (r-1)\frac{\beta}{1+\beta}$. We have

$$\frac{a}{b} = 1 = \frac{1}{r-1} \Longrightarrow r = 2.$$

Also,

$$p_0 = 0.5 = (1+\beta)^{-r} \Longrightarrow \beta = 0.414.$$

Finally,

$$c = \frac{\beta}{1+\beta} = \frac{0.414}{1.414} = 0.29 \blacksquare$$

Practice Problems

Problem 28.1
Let N be a member of $C(a, b, 0)$ satisfying the recursive probabilities

$$\frac{p_k}{p_{k-1}} = \frac{4}{k} - \frac{1}{3}.$$

Identify the distribution N.

Problem 28.2
Let N be a member of $C(a, b, 0)$ satisfying the recursive probabilities

$$\frac{p_k}{p_{k-1}} = \frac{4}{k} - \frac{1}{3}.$$

Find $E(N)$ and $\text{Var}(N)$.

Problem 28.3 ‡
The number of claims is being modeled with an $C(a, b, 0))$ class of distributions. You are given:
- $p_0 = p_1 = 0.25$
- $p_2 = 0.1875$.

Find the value of p_3.

Problem 28.4
Let N be a counting distribution in $C(a, b, 0)$ satisfying:
- $p_0 = \frac{1}{4^5}$
- $\frac{p_k}{p_{k-1}} = c\left(0.25 + \frac{1}{k}\right), \ k = 1, 2, 3, \cdots$.

Determine the value of c.

Problem 28.5
Suppose that the number of claims N has a Poisson distribution with mean $\lambda = 3$.

Calculate $\frac{p_{255}}{p_{254}}$.

Problem 28.6
Let N be a negative binomial random variable with $r = 2.5$ and $\beta = 5$.

Find the smallest value of k such that $\frac{p_k}{p_{k-1}} < 1$.

Problem 28.7
Let N be a member of $C(a, b, 0)$ such that $p_0 = p_1$ and $p_2 = 0.6p_1$.

Determine the values of a and b.

Problem 28.8
For N in $C(a, b, 0)$ you are given the following:
- $p_0 = p_1$.
- $p_2 = 0.6p_1$.

Based on this information, which of the following are true statements?

(I) N has a Poisson distribution.
(II) N has a negative binomial distribution.
(III) N has a binomial distribution.
(IV) None of the above.

Problem 28.9
Let N be a member of $C(a, b, 0)$. You are given:
- $p_5 = 0.00144$
- $p_4 = 0.006$
- $p_3 = 0.02$

Determine the value of p_0.

Problem 28.10
Let N be a member of $C(a, b, 0)$. You are given:
- $p_2 = 0.1536$
- $p_1 = 0.4096$
- $p_0 = 0.4096$

Determine $E(N)$.

Problem 28.11 ‡
For a discrete probability distribution, you are given the recursion relation

$$p(k) = \frac{2}{k}p(k-1),\ \ k = 1, 2, \cdots.$$

Determine $p(4)$.

29 The Class $C(a, b, 1)$ of Discrete Distributions

All the four members of $C(a, b, 0)$ have a fixed positive probability at 0. For some models, an adjustment at 0 is sometimes required. We would like to be able to assign any value in the interval $[0, 1)$ to p_0 and not be bounded by the fixed values provided by the four distributions of $C(a, b, 0)$.

Let N be a discrete non-negative random variable with pmf $p_k,\ k = 0, 1, 2, \cdots$. We can create a new random variable N^* with pmf p_k^* such that p_0^* has a preassigned value in the interval $[0, 1)$. The pmf of N^* is given by

$$p_k^* = \frac{1 - p_0^*}{1 - p_0} p_k,\ k = 1, 2, \cdots.$$

Note that

$$\sum_{k=0}^{\infty} p_k^* = p_0^* + \sum_{k=1}^{\infty} p_k^* = p_0^* + \frac{1 - p_0^*}{1 - p_0} \sum_{k=1}^{\infty} p_k = p_0^* + \frac{1 - p_0^*}{1 - p_0}(1 - p_0) = 1.$$

We say that N^* has a **zero-modified** distribution. In the spacial case when $p_0^* = 0$, we say that N^* has a **zero-truncated** distribution. Note that the zero-truncated distribution is a special case of the zero-modified distribution. We will therefore develop properties of the zero-modified distributions.

We will consider only the distributions of $C(a, b, 0)$. The collection of all members of $C(a, b, 0)$ together with the associated zero-modified distributions belong to a class denoted by $C(a, b, 1)$. Note that, for $k = 1, 2, \cdots$, we have

$$\frac{p_k^*}{p_{k-1}^*} = \frac{p_k}{p_{k-1}} = a + \frac{b}{k}.$$

Note that both members of $C(a, b, 0)$ and $C(a, b, 1)$ satisfy

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1}$$

but for the $C(a, b, 0)$ class the k starts from 1 whereas for the $C(a, b, 1)$ class the k starts from 2.

The probabilities of a zero-modified distribution are denoted by p_k^M where

$$p_k^M = \frac{1 - p_0^M}{1 - p_0} p_k,\ k = 1, 2, \cdots$$

and p_0^M is an arbitrary number in $[0, 1)$. The probabilities of a zero-truncated distribution are denoted by p_k^T where

$$p_k^T = \frac{1}{1-p_0}p_k, \ \ k = 1, 2, \cdots$$

and $p_0^T = 0$.

Theorem 29.1
Let N be in $C(a, b, 0)$ with corresponding moment generating function $M_N(t)$. Then the moment generating function of N^M is

$$M_N^M(t) = \frac{p_0^M - p_0}{1-p_0} + \left(\frac{1-p_0^M}{1-p_0}\right) M_N(t).$$

Proof.
We have

$$\begin{aligned}
M_N^M(t) =& E(e^{tN}) = \sum_{n=0}^{\infty} e^{tn} p_n^M \\
=& p_0^M + \left(\frac{1-p_0^M}{1-p_0}\right) \sum_{n=1}^{\infty} e^{tn} p_n \\
=& p_0^M + \left(\frac{1-p_0^M}{1-p_0}\right) \left[\sum_{n=0}^{\infty} e^{tn} p_n - p_0\right] \\
=& p_0^M + \left(\frac{1-p_0^M}{1-p_0}\right) [M_N(t) - p_0] \\
=& p_0^M - p_0 \left(\frac{1-p_0^M}{1-p_0}\right) + \left(\frac{1-p_0^M}{1-p_0}\right) M_N(t) \\
=& \frac{p_0^M - p_0}{1-p_0} + \left(\frac{1-p_0^M}{1-p_0}\right) M_N(t) \ \blacksquare
\end{aligned}$$

Corollary 29.1
Let N be in $C(a, b, 0)$ with corresponding probability generating function $P_N(t)$. Then the probability generating function of N^M is

$$P_N^M(t) = \frac{p_0^M - p_0}{1-p_0} + \left(\frac{1-p_0^M}{1-p_0}\right) P_N(t).$$

Proof.
We have

$$\begin{aligned}P_N^M(t) =&M_N^M(\ln t) = \frac{p_0^M - p_0}{1-p_0} + \left(\frac{1-p_0^M}{1-p_0}\right) M_N(\ln t)\\ =&\frac{p_0^M - p_0}{1-p_0} + \left(\frac{1-p_0^M}{1-p_0}\right) P_N(t) \blacksquare\end{aligned}$$

Remark 29.1
Note that

$$P_N^M(t) = \left(1 - \frac{1-p_0^M}{1-p_0}\right) \cdot 1 + \left(\frac{1-p_0^M}{1-p_0}\right) P_N(t).$$

Thus, N^M is a mixture of a degenerate distribution (i.e. a distribution with support consisting of a single point) with mixing weight $\left(\frac{p_0^M - p_0}{1-p_0}\right)$ and the corresponding member of $C(a, b, 0)$ with mixing weight $\left(\frac{1-p_0^M}{1-p_0}\right)$.

Example 29.1
Let N be the Poisson distribution with parameter λ. Find the probability functions of
(a) the zero-truncated distribution
(b) the zero-modified distribution with preassigned $p_0^M = 0.10$.

Solution.
(a) We have $p_0 = e^{-\lambda}$ and $p_k = \frac{e^{-\lambda}\lambda^k}{k!}$. Hence,

$$p_k^T = \frac{1}{1-p_0} p_k = \frac{1}{1-e^{-\lambda}} \frac{e^{-\lambda}\lambda^k}{k!}, \ k = 1, 2, \cdots.$$

(b) We have, with $p_0^M = 0.10$,

$$p_k^M = \frac{1-p_0^M}{1-p_0} p_k = \frac{0.90}{1-e^{-\lambda}} \frac{e^{-\lambda}\lambda^k}{k!}, \ k = 1, 2, \cdots \blacksquare$$

Example 29.2
Let N have the negative binomial distribution with $r = 2.5$ and $\beta = 0.5$.
(a) Determine $p_k, \ k = 0, 1, 2, 3$.
(b) Determine p_1^T, p_2^T, p_3^T.
(c) Determine p_1^M, p_2^M, p_3^M given that $p_0^M = 0.6$.

Solution.
(a) From the table in the previous section, we find

$$\begin{aligned} p_0 =& (1+\beta)^{-2.5} = 0.362887 \\ a =& \frac{\beta}{1+\beta} = \frac{0.5}{1+0.5} = \frac{1}{3} \\ b =& (r-1)\frac{\beta}{1+\beta} = \frac{(2.5-1)}{3} = \frac{1}{2}. \end{aligned}$$

Hence,

$$\begin{aligned} p_1 =& p_0\left(\frac{1}{3}+\frac{1}{2}\frac{1}{1}\right) = 0.302406 \\ p_2 =& p_1\left(\frac{1}{3}+\frac{1}{2}\frac{1}{2}\right) = 0.176404 \\ p_3 =& p_2\left(\frac{1}{3}+\frac{1}{2}\frac{1}{3}\right) = 0.088202. \end{aligned}$$

(b) We have

$$\begin{aligned} p_0^T =& 0 \\ p_1^T =& \frac{1}{1-p_0}p_1 = 0.474651 \\ p_2^T =& \frac{1}{1-p_0}p_2 = 0.276880 \\ p_3^T =& \frac{1}{1-p_0}p_3 = 0.138440. \end{aligned}$$

(c) We have

$$\begin{aligned} p_0^M =& 0.6 \\ p_1^M =& \frac{1-p_0^M}{1-p_0}p_1 = 0.189860 \\ p_2^M =& \frac{1-p_0^M}{1-p_0}p_2 = 0.110752 \\ p_3^M =& \frac{1-p_0^M}{1-p_0}p_3 = 0.055376 \ \blacksquare \end{aligned}$$

Practice Problems

Problem 29.1
Show that

$$p_k^M = (1 - p_0^M)p_k^T,\ k = 1, 2, \cdots$$

Problem 29.2
Show that

$$E(N^M) = \left(\frac{1 - p_0^M}{1 - p_0}\right) E(N).$$

Problem 29.3
Let N^M be the zero-modified distribution associated to N.

Find $E[N^M(N^M - 1)]$ and $\mathrm{Var}(N^M)$.

Problem 29.4
Let N have a Poisson distribution with mean 1.

(a) Determine p_0, p_1, and p_2.
(b) Find the mean and the variance of $E(N^T)$.

Problem 29.5
Consider the zero-modified geometric distribution:

$$\begin{aligned} p_0^M =& \frac{1}{2} \\ p_k^M =& \frac{1}{6}\left(\frac{2}{3}\right)^{k-1},\ k = 1, 2, 3, \cdots . \end{aligned}$$

(a) Find the moment generating function of N^M.
(b) Find the mean and the variance of N^M.

Problem 29.6
You are given: $p_1^M = \frac{1}{6},\ p_2^M = \frac{1}{9}$, and $p_3^M = \frac{2}{27}$. Find p_0^M.

30 The Extended Truncated Negative Binomial Model

The $C(a, b, 1)$ class not only contains the members of $C(a, b, 0)$ but also two additional distributions the so-called the **extended truncated negative binomial distribution**(ETNB) and the **logarithmic distribution**.

The extended truncated negative binomial distribution has the pobability function defined recursively as

$$\begin{aligned} p_0 =& 0 \\ \frac{p_k}{p_{k-1}} =& a + \frac{b}{k}, \ k = 2, 3, \cdots \end{aligned}$$

where

$$a = \tfrac{\beta}{1+\beta}, \ \beta > 0 \quad \text{and} \quad b = (r-1)\tfrac{\beta}{1+\beta}, \ r > -1, \ r \neq 0.$$

Example 30.1
Show that

$$p_k = p_1 \left(\frac{\beta}{\beta+1}\right)^{k-1} \frac{r+1}{2} \cdot \frac{r+2}{3} \cdots \frac{r+k-1}{k}, \ k = 2, 3, \cdots .$$

Solution.
The proof is by induction on $k = 2, 3, \cdots$. For $k = 2$, we have

$$p_2 = p_1 \left(\frac{\beta}{1+\beta} + \frac{r-1}{2} \frac{\beta}{1+\beta}\right) = p_1 \left(\frac{\beta}{1+\beta}\right) \frac{r+1}{2}.$$

Suppose that up to k, we have

$$p_k = p_1 \left(\frac{\beta}{\beta+1}\right)^{k-1} \frac{r+1}{2} \cdot \frac{r+2}{3} \cdots \frac{r+k-1}{k}.$$

Then,

$$p_{k+1} = p_1 \left(\frac{\beta}{1+\beta} + \frac{r-1}{k+1} \frac{\beta}{1+\beta}\right) = p_1 \left(\frac{\beta}{\beta+1}\right)^k \frac{r+1}{2} \cdot \frac{r+2}{3} \cdots \frac{r+k-1}{k} \frac{r+k}{k+1} \blacksquare$$

Example 30.2
(a) Show that if $p_1 > 0$ then $p_k > 0$ for all $k = 2, 3, \cdots$.
(b) Show that $\sum_{k=1}^{\infty} p_k < \infty$.

Solution.
(a) Since $r > -1, \beta > 0$ and $k = 2, 3, \cdots$, from the previous example we conclude that $p_k > 0$ for all $k = 2, 3, \cdots$.
(b) We have

$$\sum_{k=1}^{\infty} p_k = p_1 \sum_{k=2}^{\infty} \left(\frac{\beta}{\beta+1}\right)^{k-1} \frac{r+1}{2} \cdot \frac{r+2}{3} \cdots \frac{r+k-1}{k}.$$

Let

$$a_k = \left(\frac{\beta}{\beta+1}\right)^{k-1} \frac{r+1}{2} \cdot \frac{r+2}{3} \cdots \frac{r+k-1}{k}.$$

Then

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{\beta}{1+\beta} \frac{r+k}{k+1} = \frac{\beta}{1+\beta} < 1.$$

Hence, by the ratio series test, the given series is convergent ■

Now, letting $r \to 0$ we have

$$\begin{aligned} 1 = \sum_{k=1}^{\infty} p_k = & p_1 \sum_{k=2}^{\infty} \left(\frac{\beta}{1+\beta}\right)^{k-1} \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{k-1}{k} \\ = & p_1 \sum_{k=2}^{\infty} \left(\frac{\beta}{1+\beta}\right)^{k-1} \frac{1}{k} \\ = & p_1 \left(\frac{1+\beta}{\beta}\right) \sum_{k=2}^{\infty} \left(\frac{\beta}{1+\beta}\right)^{k} \frac{1}{k} \\ = & p_1 \left[-\ln\left(1 - \frac{1}{1+\beta}\right)\right] \end{aligned}$$

where we used the Taylor expansion of $\ln(1-x)$. Hence, $p_0 = 0$ and

$$p_k = \left(\frac{\beta}{1+\beta}\right)^k \frac{1}{k\ln(1+\beta)}, \ k = 1, 2, \cdots.$$

But this is the probability function of the **logarithmic distribution** with parameter β.

Example 30.3
Find the probability generating function of the logarithmic distribution.

Solution.
We have

$$\begin{aligned}
P(z) = \sum_{n=0}^{\infty} p_n z^n =& \frac{1}{\ln(1+\beta)} \sum_{n=1}^{\infty} \left(\frac{\beta}{1+\beta}\right)^n \frac{1}{n} z^n \\
=& \frac{1}{\ln(1+\beta)} \left[-\ln\left(1 - \frac{z\beta}{1+\beta}\right)\right] \\
=& \frac{1}{\ln(1+\beta)} \ln\left(\frac{1+\beta}{1-\beta(z-1)}\right) \\
=& 1 - \frac{\ln[1-\beta(z-1)]}{\ln(1+\beta)} \blacksquare
\end{aligned}$$

Example 30.4
Consider the extended zero-truncated negative binomial distribution with $r = -0.5$ and $\beta = 1$. Calculate p_2^T, and p_3^T given that $p_1^T = 0.853553$.

Solution.
We have

$$\begin{aligned}
a =& \frac{\beta}{1+\beta} = \frac{1}{1+1} = 0.5 \\
b =& (r-1)\frac{\beta}{1+\beta} = -0.75 \\
p_2^T =& p_1^T \left(0.5 - \frac{0.75}{2}\right) = 0.106694 \\
p_3^T =& p_2^T \left(0.5 - \frac{0.75}{3}\right) = 0.026674 \blacksquare
\end{aligned}$$

Remark 30.1
The $C(a, b, 1)$ class consists of the following distributions:
• Poisson, Negative binomial, Geometric, Binomial, and Logarithmic.
• Zero-truncated: Poisson, Binomial, Negative Binomial (ETNB), and Geometric.
• Zero-modified: Poisson, Binomial, ETNB, Geometric, and Logarithmic.

Practice Problems

Problem 30.1
Consider the extended zero-modified negative binomial distribution with $r = -0.5$ and $\beta = 1$.

Calculate p_1^M, p_2^M, and p_3^M given that $p_1^T = 0.853553$ and $p_0^M = 0.6$.

Problem 30.2
Let N denote the logarithmic distribution introduced in this section. Find $E(N)$.

Problem 30.3
Let N denote the logarithmic distribution introduced in this section. Find $E[N(N-1)]$.

Problem 30.4
Let N denote the logarithmic distribution introduced in this section. Find $\mathrm{Var}(N)$.

Problem 30.5
Let N^T denote the zero-truncated distribution corresponding to the distribution N. Let $P_N^T(z)$ be the probability generating function of N^T and $P_N(z)$ be the probability generating function of N. Show that

$$p_N^T(z) = \frac{P_N(z) - p_0}{1 - p_0}.$$

Problem 30.6
Find the probability generating function for the extended truncated negative binomial distribution.

Problem 30.7
Find the mean of the extended truncated negative binomial.

Problem 30.8
Let N^T denote the ETNB. Find $E[N^T(N^T-1)]$.

Problem 30.9
Let N^T denote the ETNB. Find $\mathrm{Var}(N^T)$.

Modifications of the Loss Random Variable

The purpose of this chapter is to relate the random variables introduced earlier in this text to insurance applications. In this chapter, all the random vairables have a support that is a subset of $[0, \infty)$. For purposes of notation, we shall refer to X as the (ground-up) loss amount random variable−before any modifications. We shall denote the modified loss amount to be Y and will be referred to as the **claim amount** paid by the insurer. We will use the notation Y^L to denote the loss-variable and the notation Y^P to denote the payment-variable. In this chapter, we will discuss three types of coverage modifications: deductibles, policy limits, and coinsurance factors.

31 Ordinary Policy Deductibles

For an insurance policy, in order for a claim to be paid, a threshold d must be exceeded. That is, the ground-up loss X must exceed d. In this case, the insurance pays the policyholder the amount $X-d$. For loss amount less than d the insurance pays 0. This is referred to as **ordinary deductible**. The amount paid by the insurance is called the **cost-per loss**[9] or left censored and shifted variable (see Section 5). It is denoted by

$$Y^L = \max\{X-d, 0\} = (X-d)_+ = \begin{cases} 0 & X \le d \\ X-d & X > d. \end{cases}$$

The cost-per loss Y^L is an example of a mixed random variable: For $X \le d$, Y^L is a discrete random variable with probability at 0 given by

$$p_{Y^L}(0) = \Pr(Y^L = 0) = \Pr(X \le d) = F_X(d).$$

The continuous part of Y^L is given by

$$f_Y(y) = f_X(y+d),\ \ y > 0.$$

The cdf of Y^L is given by

$$F_{Y^L}(y) = \Pr(Y^L \le y) = \Pr(X-d \le y) = F_X(y+d),\ \ y \ge 0.$$

From this, it follows that

$$S_{Y^L}(y) = S_X(y+d),\ \ y \ge 0.$$

The kth moments of Y^L are given by

$$E[(Y^L)^k] = \sum_{x>d} (x-d)^k p_X(x)$$

if X is discrete and

$$E[(Y^L)^k] = \int_d^\infty (x-d)^k f_X(x)dx$$

if X is continuous. Moreover, for the continuous case, if $k=1$ we have

$$E(Y^L) = \int_d^\infty (x-d) f_X(x)dx = \int_d^\infty [1-F_X(x)]dx.$$

[9]The random variable Y^L is also referred to as the **claim amount paid per-loss event.**

It follows that $E(Y^L) \leq E(X)$ since

$$\int_d^\infty [1 - F_X(x)]dx \leq \int_0^\infty [1 - F_X(x)]dx.$$

Note that with the presence of deductibles, the number of payments is fewer than the losses since losses with amount less than or equal to the deductible will result in no payments.

Example 31.1
Determine the pdf, cdf, and the sdf for Y^L if the gound-up loss amount function has an exponential distribution with mean $\frac{1}{\theta}$ and an ordinary deductible of d.

Solution.
Recall that

$$f_X(x) = \theta e^{-\theta x}, \ x > 0.$$

Thus,

$$f_{Y^L}(y) = \begin{cases} 1 - e^{-\theta d} & y = 0 \\ \theta e^{-\theta(y+d)} & y > 0. \end{cases}$$

$$F_{Y^L}(y) = \begin{cases} 1 - e^{-\theta d} & y = 0 \\ 1 - e^{-\theta(y+d)} & y > 0. \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} e^{-\theta d} & y = 0 \\ \theta e^{-\theta(y+d)} & y > 0 \ \blacksquare \end{cases}$$

In the cost per loss situation, all losses below or at the deductible level are recorded as 0. We next examine the situation where all losses below or at the deductible level are completely ignored and not recorded in any way. This sitution is represented by the random variable

$$Y^P = (Y^L | Y^L > 0) = (Y^L | X > d) = \begin{cases} \text{undefined} & X \leq d \\ X - d & X > d. \end{cases}$$

We call Y^P, the **excess loss variable**[10], the **cost per-payment**, or the **left truncated and shifted variable** as defined in Section 5. Also, recall

[10] We refer to Y^P the **amount paid per-payment event**

the following results from Section 5:

$$\begin{aligned}
f_{Y^P}(y) =& \frac{f_X(y+d)}{S_X(d)}, \ y > 0\\
F_{Y^P}(y) =& \frac{F_X(y+d) - F_X(d)}{S_X(d)}, \ y > 0\\
S_{Y^P}(y) =& \frac{S_X(y+d)}{S_X(d)}, \ y > 0\\
h_{Y^P}(y) =& \frac{f_X(y+d)}{S_X(y+d)} = h_X(y+d), \ y > 0.
\end{aligned}$$

Note that the excess loss random variable places no probability at 0. Also, recall from Section 5,

$$E(Y^P) = \frac{E(Y^L)}{1 - F_X(d)} > E(Y^L).$$

That is, the mean excess loss is larger than the expected amount paid per-loss event.

Example 31.2
Determine the pdf, cdf, sdf, and the hazard rate function for Y^P if the gound-up loss amount function has an exponential distribution with mean $\frac{1}{\theta}$ and an ordinary deductible of d.

Solution.
Recall that

$$f_X(x) = \theta e^{-\theta x}, \ x > 0.$$

Thus,

$$\begin{aligned}
f_{Y^P}(y) =& \theta^{-\theta y}\\
F_{Y^P}(y) =& 1 - e^{-\theta y}\\
S_{Y^P}(y) =& e^{-\theta y}\\
h_{Y^P}(y) =& \theta \ \blacksquare
\end{aligned}$$

The following theorem deals with calculating the expectation of Y^L and Y^P under ordinary deductibles.

Theorem 31.1
For an ordinary deductible d, we have

$$E(Y^L) = E(X) - E(X \wedge d)$$

and

$$E(Y^P) = \frac{E(X) - E(X \wedge d)}{1 - F_X(d)}$$

where $X \wedge d$ is the limited loss variable (see Section 5) defined by

$$X \wedge d = \min(X, d) = \begin{cases} X, & X < d \\ d, & X \geq d \end{cases}$$

and

$$E(X \wedge d) = \int_0^d S_X(x)dx.$$

Proof.
Buying one policy with a deductible d and another one with a limit d is equivalent to purchasing full cover. That is,

$$Y^L + X \wedge d = X.$$

Hence,

$$E(Y^L) = E(X) - E(X \wedge d).$$

Also,

$$E(Y^P) = \frac{E(Y^L)}{1 - F_X(d)} = \frac{E(X) - E(X \wedge d)}{1 - F_X(d)} \blacksquare$$

Example 31.3
Losses are distributed exponentially with parameter θ. Policies are subject to ordinary deductible d. Find $E(Y^L)$ and $E(Y^P)$.

Solution.
We have

$$E(X \wedge d) = \int_0^d e^{-\theta x} dx = \frac{1}{\theta}(1 - e^{-\theta d})$$

and

$$E(X) = \int_0^\infty x\theta e^{-\theta x} dx = \frac{1}{\theta}.$$

Hence,

$$E(Y^L) = \frac{1}{\theta} - \frac{1}{\theta}(1 - e^{-\theta d}) = \frac{e^{-\theta d}}{\theta}$$

and

$$E(Y^P) = \frac{\frac{e^{-\theta d}}{\theta}}{e^{-\theta d}} = \frac{1}{\theta} \blacksquare$$

Example 31.4 ‡
The annual number of doctor visits for each individual in a family of 4 has a geometric distribution with mean 1.5. The annual numbers of visits for the family members are mutually independent. An insurance pays 100 per doctor visit beginning with the 4^{th} visit per family.
Calculate the expected payments per year for this family.

Solution.
Let X_i be the annual number of visits by member i of the family, where $i = 1, 2, 3, 4$. Let $Y = X_1 + X_2 + X_3 + X_4$ be the annual number of visits by the whole family. For each X_i, the probability generating function is

$$P(z) = [1 - 1.5(z - 1)]^{-1}.$$

Using independence,

$$PY(z) = [1 - 1.5(z - 1)]^{-4}.$$

Thus, Y is has a negative binomial distribution with $r = 4$ and $\beta = 1.5$. The number of visits resulting in insurance payments is

$$(Y - 3)_+ = Y - Y \wedge 3$$

so that

$$E[(Y - 3)_+] = E(Y) - E(Y \wedge 3).$$

From Table C, the pmf of the negative binomial with $r = 4$ and $\beta = 1.5$ is

$$\Pr(Y = k) = \frac{r(r+1)\cdots(r+k-1)\beta^k}{k!(1+\beta)^{r+k}}.$$

Now, we have the following

$$\begin{aligned}
E(Y) =& 4E(X) = 4(1.5) = 6\\
E(Y \wedge 3) =& 1 \cdot \Pr(Y = 1) + 2 \cdot \Pr(Y = 2) + 3 \cdot \Pr(Y = 3) + 3[1 - \Pr(Y \leq 3)\\
=& 1 \cdot \Pr(Y = 1) + 2 \cdot \Pr(Y = 2) + +3\Pr(Y \geq 3)
\end{aligned}$$

$$\begin{aligned}
\Pr(Y = 0) =& \frac{1}{(2.5)^4} = 0.0256\\
\Pr(Y = 1) =& \frac{4(1.5)}{(2.5)^{4+1}} = 0.06144\\
\Pr(Y = 2) =& \frac{4(5)(1.5)^2}{2(2.5)^{4+2}} = 0.09216\\
\Pr(Y \geq 3) =& 1 - 0.0256 - 0.06144 - 0.09216 = 0.8208.
\end{aligned}$$

Thus,

$$E(Y \wedge 3) = 0.06144 + 2(0.09216) + 3(0.8208) = 2.71$$

and the expected number of visits resulting in insurance payments is

$$E(Y) = 6 - 2.71 = 3.29.$$

The insurance pays 100 per visit so that the total expected insurance payment for the year is $3.29 \times 100 = 329$ ■

Practice Problems

Problem 31.1
The cdf of a loss amount distribution is expressed as:

$$F_X(x) = 1 - e^{-\left(\frac{x}{100}\right)^2},\ x > 0.$$

The ordinary deductible for an insurance policy is 50. Find the pdf, cdf, and the survival function of the cost per loss Y^L.

Problem 31.2
The cdf of a loss amount distribution is expressed as:

$$F_X(x) = 1 - e^{-\left(\frac{x}{100}\right)^2},\ x > 0.$$

The ordinary deductible for an insurance policy is 50. Find the pdf, cdf, survival function, and the hazard rate function of the cost per payment Y^P.

Problem 31.3
Loss amounts are exponentially distributed with parameter $\theta = 1000$. An insurance policy is subject to deductible of 500.

Find $E(Y^L)$.

Problem 31.4
Loss amounts are exponentially distributed with parameter $\theta = 1000$. An insurance policy is subject to deductible of 500.

Find $E(Y^P)$.

Problem 31.5 ‡
Losses follow an exponential distribution with parameter θ. For an ordinary deductible of 100, the expected payment per loss is 2000.

What is the expected payment per loss in terms of θ for an ordinary deductible of 500?

Problem 31.6
Loss amounts are distributed as a single Pareto with parameters $\alpha = 4$ and $\theta = 90$. An insurance policy is subject to an ordinary deductible of 100.

Determine $\mathrm{Var}(Y^L)$.

Problem 31.7
Losses are uniformly distributed in $(0, b)$. An insurance policy is subject to an ordinary deductible of d.

Calculate $\text{Var}(Y^P)$.

Problem 31.8
Loss amounts have a discrete distribution with the following probabilities:

Loss Amount	Probability
150	0.30
300	0.25
425	0.15
750	0.30

An insurance coverage for these losses has an ordinary deductible of 250.

(a) Calculate the expected insurance payment per loss.
(b) Calculate the expected insurance payment per payment.

Problem 31.9 ‡
Risk 1 has a Pareto distribution with parameters $\alpha > 2$ and θ. Risk 2 has a Pareto distribution with parameters 0.8α and θ. Each risk is covered by a separate policy, each with an ordinary deductible of k.

(a) Determine the expected cost per loss for risk 1.
(b) Determine the limit as k goes to infinity of the ratio of the expected cost per loss for risk 2 to the expected cost per loss for risk 1.

Problem 31.10
Loss amounts are uniformly distributted in $(0, 10)$. An ordinary policy deductible of d is applied.

Find d if $E(Y^L) = 20$.

32 Franchise Policy Deductibles

The second type of deductibles that we consider is the **franchise deductibles**. Under this deductible policy, the loss is fully paid when the loss exceeds the deductible.

The cost per-loss random variable is defined by

$$Y^L = \begin{cases} 0, & X \leq d \\ X, & X > d \end{cases}$$

and the cost per-payment is

$$Y^P = (X|X > d) = \begin{cases} \text{undefined}, & X \leq d \\ X, & X > d. \end{cases}$$

The key distributional functions are

$$f_{Y^L}(y) = \begin{cases} F_X(d), & y = 0 \\ f_X(y), & y > d \end{cases}$$

$$F_{Y^L}(y) = \begin{cases} F_X(d), & 0 \leq y \leq d \\ F_X(y), & y > d \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} S_X(d), & 0 \leq y \leq d \\ S_X(y), & y > d \end{cases}$$

$$h_{Y^L}(y) = \begin{cases} 0, & 0 < y < d \\ h_X(y), & y > d \end{cases}$$

for the cost per-loss and

$$f_{Y^P}(y) = \frac{f_X(y)}{S_X(d)}, \ y > d$$

$$F_{Y^P}(y) = \begin{cases} 0, & 0 \leq y \leq d \\ \frac{F_X(y)-F_X(d)}{S_X(d)}, & y > d \end{cases}$$

$$S_{Y^P}(y) = \begin{cases} 1, & 0 \leq y \leq d \\ \frac{S_X(y)}{S_X(d)}, & y > d \end{cases}$$

$$h_{Y^P}(y) = \begin{cases} 0, & 0 < y < d \\ h_X(y), & y > d \end{cases}$$

for the cost per-payment.

Example 32.1
Loss amounts are distributed exponentially with mean $\frac{1}{\theta}$. Insurance policies are subject to a franchise deductible of d. Find the pdf, cdf, sdf, and the hazard rate function for the cost per-loss Y^L.

Solution.
The pdf of X is

$$f_X(x) = \theta e^{-\theta x}.\ x > 0.$$

We have

$$f_{Y^L}(y) = \begin{cases} 1 - e^{-\theta d}, & y = 0 \\ \theta e^{-\theta y}, & y > d. \end{cases}$$

$$F_{Y^L}(y) = \begin{cases} 1 - e^{-\theta d}, & 0 \leq y \leq d \\ 1 - e^{-\theta y}, & y > d. \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} e^{-\theta d}, & 0 \leq y \leq d \\ e^{-\theta y}, & y > d. \end{cases}$$

$$h_{Y^L}(y) = \begin{cases} 0, & 0 < y < d \\ \theta, & y > d \ \blacksquare \end{cases}$$

Example 32.2
Loss amounts are distributed exponentially with mean $\frac{1}{\theta}$. Insurance policies are subject to a franchise deductible of d. Find the pdf, cdf, sdf, and the hazard rate function for the cost per-payment Y^P.

Solution.
We have

$$f_{Y^P}(y) = \theta e^{-\theta(y-d)},\ y > d.$$

$$F_{Y^P}(y) = \begin{cases} 0, & 0 \leq y \leq d \\ 1 - e^{-\theta(y-d)}, & y > d. \end{cases}$$

$$S_{Y^P}(y) = \begin{cases} 1, & 0 \leq y \leq d \\ e^{-\theta(y-d)}, & y > d. \end{cases}$$

$$h_{Y^P}(y) = \begin{cases} 0, & 0 < y < d \\ \theta, & y > d \ \blacksquare \end{cases}$$

The following theorem deals with calculating the expectation of Y^L and Y^P under franchise deductibles.

Theorem 32.1
For a franchise deductible d, we have

$$E(Y^L) = E(X) - E(X \wedge d) + d[1 - F_X(d)]$$

and

$$E(Y^P) = \frac{E(X) - E(X \wedge d)}{1 - F_X(d)} + d$$

where $X \wedge d$ is the limited loss variable (see Section 5) defined by

$$X \wedge d = \min(X, d) = \begin{cases} X, & X < d \\ d, & X \geq d \end{cases}$$

and

$$E(X \wedge d) = \int_0^d S_X(x)dx.$$

Proof.
We will prove the results for the continuous case. We have

$$\begin{aligned} E(Y^L) =& \int_d^\infty x f_X(x)dx = \int_d^\infty (x-d)f_X(x)dx + d\int_d^\infty f_X(x)dx \\ =& E[(X-d)_+] + d[1 - F_X(d)] \\ =& E(X) - E(X \wedge d) + d[1 - F_X(d)] \\ E(Y^P) =& \frac{E(Y^L)}{1 - F_X(d)]} = \frac{E(X) - E(X \wedge d)}{1 - F_X(d)} + d \blacksquare \end{aligned}$$

Example 32.3
Losses are distributed exponentially with parameter θ. Policies are subject to franchise deductible d. Find $E(Y^L)$ and $E(Y^P)$.

Solution.
Using Example 31.3, we have

$$E(Y^L) = \frac{e^{-\theta d}}{\theta} + de^{-\theta d}$$

and

$$E(Y^P) = \frac{e^{\theta d}}{\theta} + d \blacksquare$$

Example 32.4 ‡
Insurance agent Hunt N. Quotum will receive no annual bonus if the ratio of incurred losses to earned premiums for his book of business is 60% or more for the year. If the ratio is less than 60%, Hunt's bonus will be a percentage of his earned premium equal to 15% of the difference between his ratio and 60%. Hunt's annual earned premium is 800,000.
Incurred losses are distributed according to the Pareto distribution, with $\theta = 500,000$ and $\alpha = 2$.
Calculate the expected value of Hunt's bonus.

Solution.
Let L be the incurred losses and B Hunt's bonuses. We are told that if $\frac{L}{800,000} < 0.6$, that is, $L \le 480,000$, then $B = 0.15\left(0.6 - \frac{L}{800,000}\right)(800,000) = 0.15(480,000 - L)$. This can be written as

$$B = 0.15\begin{cases} 480,000 - L, & L < 480,000 \\ 0, & L \ge 480,000 \end{cases} = 0.15[480,000 - X \wedge 480,000)].$$

Its expected value is

$$E(B) = 0.15[480,000 - E(X \wedge 480,000)].$$

From Table C, we have

$$E(X \wedge 480,000) = \frac{500,000}{1}\left[1 - \left(\frac{500,000}{500,000 + 480,000}\right)^1\right] = 244,898.$$

Hence,

$$E(B) = 0.15(480,000 - 244,898) = 35,265 \blacksquare$$

Practice Problems

Problem 32.1
Loss amounts are uniformly distributed on $(0,\theta)$. For a franchise deductible $d<\theta$, find $f_{Y^L}(y), F_{Y^L}(y), S_{Y^L}(y)$, and $h_{Y^L}(y)$.

Problem 32.2
Loss amounts are uniformly distributed on $(0,\theta)$. For a franchise deductible $d<\theta$, find $f_{Y^P}(y), F_{Y^P}(y), S_{Y^P}(y)$, and $h_{Y^P}(y)$.

Problem 32.3
Loss amounts are uniformly distributed on $(0,\theta)$. For a franchise deductible $d<\theta$, find $E(Y^L)$ and $E(Y^P)$.

Problem 32.4
Claim amounts X have the following Pareto distribution

$$F_X(x)=1-\left(\frac{800}{x+800}\right)^3.$$

An insurance policy has a franchise deductibe of 300. Find the expected cost per-loss.

Problem 32.5 ‡
Auto liability losses for a group of insureds (Group R) follow a Pareto distribution with $\alpha=2$ and $\theta=2000$. Losses from a second group (Group S) follow a Pareto distribution with $\alpha=2$ and $\theta=3000$. Group R has an ordinary deductible of 500, Group S has a franchise deductible of 200.

Calculate the amount that the expected cost per payment for Group S exceeds that for Group R.

Problem 32.6
Loss amounts are exponentially distributed with parameter θ. For a franchise deductible d, it is given that $E(Y^L)=0.40E(Y^P)$. Express d in terms of θ.

Problem 32.7
Loss amounts have a discrete distribution with the following probabilities:

Loss amounts	Probability
100	20%
300	70%
1000	10%

An insurance coverage for these losses has a franchise deductible of 250.

(a) Calculate the expected insurance payment per loss.
(b) Calculate the expected insurance payment per payment.

Problem 32.8
Losses in 2011 are distributed as a Pareto distribution with $\alpha = 2$ and $\theta = 2000$. An insurance company sells a policy that covers these losses with a franchise deductible of 500 during 2011. Losses in 2012 increase by 20%. During 2012, the insurance company will sell a policy covering the losses. However, instead of the franchise deductible used in 2010, the company will implement an ordinary deductible of d. The expected value of per-loss is the same for both years.

Find the value of d.

Problem 32.9
Losses are distributed as a Pareto distribution with $\alpha = 5$ and $\theta = 1000$. Losses are subject to a franchise deductible of d. The expected value per payment after the deductible is 820.

Calculate d.

Problem 32.10
Loss amounts follow a Pareto distribution with parameters $\alpha = 2$ and θ. The expected value per-payment for an ordinary deductible of 10000 is 20000.

Calculate the expected value per-loss when losses are subject to a franchise deductible of 15000.

33 The Loss Elimination Ratio and Inflation Effects for Ordinary Deductibles

Loss amounts X subject to an ordinary deductible d satisfy the relation

$$X = (X-d)_+ + X \wedge d = \begin{cases} 0, & X \le d \\ X-d, & X > d \end{cases} + \begin{cases} X, & X \le d \\ d, & X > d. \end{cases}$$

In words, this relation says that for a policy with a deductible d, losses below d are not covered and therefore can be covered by another policy with a limit of d. Rewriting the above relation as

$$X \wedge d = X - (X-d)_+$$

we see that $X \wedge d$ is a decrease of the overall losses and hence can be considered as savings to the policyholder in the presence of deductibles. The expected savings (due to the deductible) expressed as a percentage of the loss (no deductible at all) is called the **Loss Elimination Ratio**:

$$LER = \frac{E(X \wedge d)}{E(X)}.$$

For a continuous loss amounts X we have

$$LER = \frac{\int_0^d [1 - F_X(x)]dx}{\int_0^\infty [1 - F_X(x)]dx}.$$

Example 33.1
Loss amounts X follow an exponential distribution with mean $\theta = 1000$. Suppose that insurance policies are subject to an ordinary deductible of 500. Calculate the loss elimination ratio.

Solution.
We have

$$LER = \frac{\int_0^d [1 - F_X(x)]dx}{\int_0^\infty [1 - F_X(x)]dx} = \frac{\int_0^{500} e^{-\frac{x}{1000}} dx}{\int_0^\infty e^{-\frac{x}{1000}} dx} = 1 - e^{-0.5} \blacksquare$$

Example 33.2 ‡
You are given:
(i) Losses follow an exponential distribution with the same mean in all years.
(ii) The loss elimination ratio this year is 70%.
(iii) The ordinary deductible for the coming year is 4/3 of the current deductible.
Compute the loss elimination ratio for the coming year.

Solution.
We have

$$LER = \frac{\int_0^d e^{-\frac{x}{\theta}}dx}{\int_0^\infty e^{-\frac{x}{\theta}}dx} = 1 - e^{-\frac{d}{\theta}} \Longrightarrow d = \theta \ln(1 - LER).$$

Thus,

$$d_{\text{Last year}} = \theta \ln 0.30$$

and

$$d_{\text{Last year}} = \theta \ln(1 - LER_{\text{Next year}}).$$

We are told that $d_{\text{Next year}} = \frac{4}{3}d_{\text{Last year}}$. Thus,

$$\theta \ln(1 - LER_{\text{Next year}}) = \frac{4}{3}\theta \ln 0.30 \Longrightarrow LER_{\text{Next year}} = 0.80 \blacksquare$$

Example 33.3
An insurance company offers two types of policies: Type A and Type B. The distribution of each type is presented below.

Type A		Type B	
Loss Amount	Probability	Loss Amount	Probability
100	0.65	300	0.70
200	0.35	400	0.30

55% of the policies are of Type A and the rest are of Type B. For an ordinary deductible of 125, calculate the loss elimination ratio and interpret its value.

Solution.
The expected losses without deductibles is

$$\begin{aligned}E(X) =& E(X|A)\Pr(A) + E(X|B)\Pr(B)\\ =& [100(0.65) + 200(0.35](0.55) + [300(0.70) + 400(0.30)](0.45)\\ =& 222.75.\end{aligned}$$

The expected savings (due to the deductible) is

$$\begin{aligned}E(X \wedge 125) =& E(X \wedge 125|A)\Pr(A) + E(X \wedge 125|B)\Pr(B)\\ =& [100(0.65) + 125(1 - 0.65)](0.55) + 125(0.45) = 116.0625.\end{aligned}$$

Hence,

$$LER = \frac{116.0625}{222.75} = 0.521 = 52.1\%$$

This is the percentage of savings in claim payments due to the presence of the deductible 125 ■

The effect of inflation

To see the effect of inflation, consider the following situation: Suppose an event formerly produced a loss of 475. With a deductible of 500, the insurer has no payments to make. Inflation of 12% will increase the loss to 532 and thus results of insurer's payment of 32.

Theorem 33.1

Let loss amounts be X and let Y be the loss amounts after uniform inflation of r. That is, $Y = (1+r)X$. For an ordinary deductible of d, the expected cost per-loss is

$$E(Y^L) = (1+r)\left[E(X) - E\left(X \wedge \frac{d}{1+r}\right)\right].$$

The expected cost per-payment is

$$E(Y^P) = \frac{E(Y) - E(Y \wedge d)}{1 - F_Y(d)} = \left[1 - F_X\left(\frac{d}{1+r}\right)\right]^{-1} (1+r)\left[E(X) - E\left(X \wedge \frac{d}{1+r}\right)\right].$$

Proof.

We have

$$\begin{aligned}
F_Y(y) =& \Pr(Y \le y) = \Pr\left(X \le \frac{y}{1+r}\right) = F_X\left(\frac{y}{1+r}\right)\\
f_Y(y) =& \frac{1}{1+r} f_X\left(\frac{y}{1+r}\right)\\
E(Y) =& (1+r)E(X).
\end{aligned}$$

$$\begin{aligned}
E(Y \wedge d) =& \int_0^d y f_Y(y)dy + d[1 - F_Y(d)]\\
=& \int_0^d \frac{y f_X\left(\frac{y}{1+r}\right)}{1+r} dy + d\left[1 - F_X\left(\frac{d}{1+r}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
=&\int_0^{\frac{d}{1+r}}(1+r)xf_X(x)dx+d\left[1-F_X\left(\frac{d}{1+r}\right)\right]\\
=&(1+r)\left\{\int_0^{\frac{d}{1+r}}xf_X(x)dx+\frac{d}{1+r}\left[1-F_X\left(\frac{d}{1+r}\right)\right]\right\}\\
=&(1+r)E\left(X\wedge\frac{d}{1+r}\right)
\end{aligned}
$$

Hence, the expected cost per loss is

$$E(Y^L)=E(Y)-E(Y\wedge d)=(1+r)\left[E(X)-E\left(X\wedge\frac{d}{1+r}\right)\right].$$

For the cost per-payment, see Theorem 31.1 ■

Example 33.4
Determine the effect of inflation at 10% on an ordinary deductible of 500 applied to an exponential distribution with mean 1000.

Solution.
Before the inflation, We have

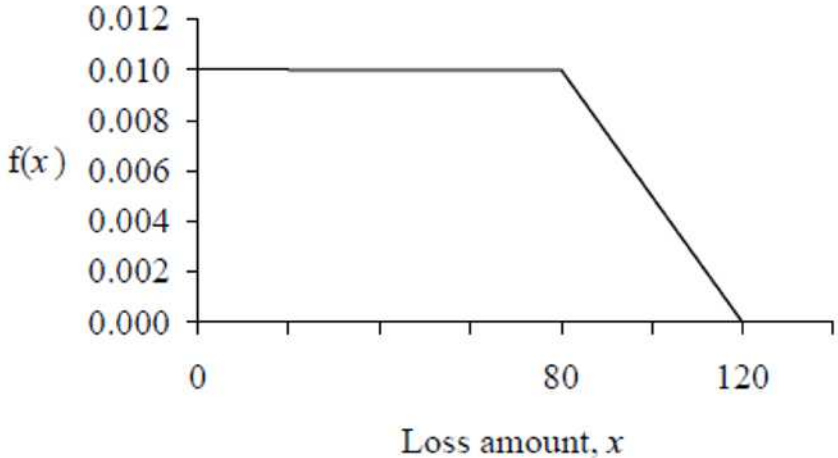
$$
\begin{aligned}
E(X)=&1000\\
E(X\wedge 500)=&\int_0^{500}e^{-\frac{x}{1000}}dx=1000(1-e^{-0.5})\\
E(Y^L_{BI})=&1000e^{-0.5}\\
E(Y^P_{BI})=&\frac{1000e^{-0.5}}{e^{-0.5}}=1000.
\end{aligned}
$$

After the inflation, we have

$$
\begin{aligned}
E\left(X\wedge\frac{500}{1.1}\right)=&\int_0^{\frac{500}{1.1}}e^{-\frac{x}{1000}}dx=1000(1-e^{-\frac{1}{2.2}})\\
E(Y^L_{AI})=&(1.1)[1000-1000(1-e^{-\frac{1}{2.2}})=1100e^{-\frac{1}{2.2}}\\
E(Y^P_{AI})=&\frac{1100e^{-\frac{1}{2.2}}}{e^{-\frac{1}{2.2}}}=1100.
\end{aligned}
$$

Thus, the expected cost per loss increased from $1000e^{-0.5}$ to $1100e^{-\frac{1}{2.2}}$, an increase of $\frac{1100e^{-\frac{1}{2.2}}-1000e^{-0.5}}{1000e^{-0.5}}=15.11\%$. The cost per-pay increased from 1000 to 1100, an increase of 10% ■

Example 33.5 ‡
The graph of the density function for losses is:



Calculate the loss elimination ratio for an ordinary deductible of 20.

Solution.
We will use the formula

$$LER = 1 - \frac{E[(X-20)_+}{E(X)}.$$

We have

$$\begin{aligned}
E(X) =& \int_0^{80} 0.01x dx + \int_{80}^{120} (0.03x - 0.00025x^2) dx \\
=& 50.66667 \\
E[(X-20)_+] =& E(X) - \int_0^{20} x f(x) dx - 20[1 - \int_0^{20} f(x) dx] \\
=& E(X) - \int_0^{20} 0.01x^2 dx - 20[1 - \int_0^{20} 0.01x dx] \\
=& 50.6667 - 2 - 20(0.8) = 32.6667 \\
LER =& 1 - \frac{32.6667}{50.6667} = 0.3553 \blacksquare
\end{aligned}$$

Practice Problems

Problem 33.1
Loss amounts are being modeled with a distribution function expressed below:

$$F_X(x) = \frac{x^2}{2500}, \ 0 \le x \le 50.$$

An insurance policy comes with an ordinary deductible of 30.

Calculate the loss elimination ration and interpret its value.

Problem 33.2
Loss amounts are being exponentially distributed with the same mean for all years. Suppose that with an ordinary deductible of d, $LER(2011) = 0.75$. In 2012, the deductible is expected to increase by 45%.

Calculate $LER(2012)$.

Problem 33.3 ‡
Losses have an exponential distribution with a mean of 1000. There is an ordinary deductible of 500. The insurer wants to double loss elimination ratio.

Determine the new deductible that achieves this.

Problem 33.4
Losses follow an exponential distribution with mean of $\theta = 1000$. An insurance company applies an ordinary policy deductible d which results in a Loss Elimination Ratio of $1 - e^{-0.5}$.

Calculate d.

Problem 33.5 ‡
Losses follow a distribution prior to the application of any deductible with a mean of 2000. The loss elimination ratio at a deductible of 1000 is 0.3. The probability of a loss being greater than 1000 is 0.4.

Determine the average size of a loss given it is less than or equal to the deductible of 1000, that is, find $E(X|X \le 1000)$.

Problem 33.6
Loss amounts are being modeled with a distribution function expressed below

$$F_X(x) = 1 - e^{-\frac{x}{100}}.$$

An insurance policy comes with a deductible of 50.

Calculate the difference in the loss elimination ratio before and after a uniform inflation of 30%

Problem 33.7 ‡
Losses have a Pareto distribution with $\alpha = 2$ and $\theta = k$. There is an ordinary deductible of $2k$.

Determine the loss elimination ration before and after 100% inflation.

Problem 33.8 ‡
Claim sizes this year are described by a 2-parameter Pareto distribution with parameters $\alpha = 4$ and $\theta = 1500$.

What is the expected claim size per loss next year after 20% inflation and the introduction of a $100 ordinary deductible?

Problem 33.9 ‡
Losses in 2003 follow a two-parameter Pareto distribution with $\alpha = 2$ and $\theta = 5$. Losses in 2004 are uniformly 20% higher than in 2003. An insurance covers each loss subject to an ordinary deductible of 10.

Calculate the Loss Elimination Ratio in 2004.

34 Policy Limits

If a policy has a limit u, then the insurer will pay the full loss as long as the losses are less than or equal to u, otherwise, the insurer pays only the amount u. Thus, the insurer is subject to pay a maximum covered loss of u. Let Y denote the claim amount random variable for policies with limit u. Then

$$Y = \min\{X, u\} = X \wedge u = \begin{cases} X, & X \le u \\ u, & X > u. \end{cases}$$

We call Y the **limited loss random variable**(See Section 5).

Just as in the case of a deductible, Y has a mixed distribution, with the discrete part

$$p_Y(u) = \Pr(X > u) = 1 - F_X(u)$$

and a continuous part given by

$$f_Y(y) = f_X(y), \quad y < u.$$

Moreover, $f_Y(y) = 0$ for $y > u$. The cdf of Y is

$$F_Y(y) = \Pr(Y \le y) = \begin{cases} F_X(y), & y < u \\ 1, & y \ge u. \end{cases}$$

Example 34.1
Show that: $E(Y) = \int_0^u S_X(x)dx$.

Solution.
First note that

$$Y = u + Z$$

where

$$Z = \begin{cases} X - u, & X \le u \\ 0, & X > u. \end{cases}$$

Thus,

$$\begin{aligned} E(Y) =& u + \int_0^u (x-u)f_X(x)dx \\ =& u + [(x-u)F_X(x)]_0^u - \int_0^u F_X(x)dx \\ =& \int_0^u [1 - F_X(x)]dx \blacksquare \end{aligned}$$

Theorem 34.1
Let $E(Y)$ be the expected cost before inflation. Suppose that the same policy limit applies after an inflation at rate r. Then the after inflation expected cost is given by

$$E((1+r)X \wedge u) = (1+r)E\left(X \wedge \frac{u}{1+r}\right).$$

Proof.
See the proof of Theorem 33.1 ■

Example 34.2
Losses follow a Pareto distribution with parameters $\alpha = 2$ and $\theta = 1000$. For a coverage with policy limit of 2000, find the pdf and the cdf of the limited loss random variable.

Solution.
Recall that

$$f_X(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}} = \frac{2(1000)^2}{(1000+x)^3}$$

and

$$F_X(x) = 1 - \left(\frac{\theta}{\theta+x}\right)^\alpha = 1 - \left(\frac{1000}{1000+x}\right)^2.$$

Thus,

$$f_Y(y) = \begin{cases} \frac{1}{9}, & y = 2000 \\ \frac{2(1000)^2}{(1000+y)^3}, & y < 2000 \\ 0, & y > 2000 \end{cases}$$

and

$$F_Y(y) = \begin{cases} 1 - \left(\frac{1000}{1000+x}\right)^2, & y < 2000 \\ 1 & y \ge 2000 \end{cases}$$ ■

Example 34.3
Losses follow a Pareto distribution with parameters $\alpha = 2$ and $\theta = 1000$. Calculate the expected cost for a coverage with policy limit of 2000.

Solution.
Recall that

$$S(x) = \left(\frac{\theta}{\theta+x}\right)^\alpha.$$

Thus,

$$E(X \wedge 2000) = \int_0^{2000} S(x)dx = \int_0^{2000} \left(\frac{1000}{1000+x}\right)^2 dx = 666.67 \blacksquare$$

Example 34.4
Losses follow a Pareto distribution with parameters $\alpha = 2$ and $\theta = 1000$. For a coverage with policy limit 2000 and after an inflation rate of 30%, calculate the after inflation expected cost.

Solution.
We have

$$E(1.3X \wedge 2000) = 1.3E\left(X \wedge \frac{2000}{1.3}\right) = 1.3\frac{2000}{2-1}\left[1 - \left(\frac{1000}{1000 + \frac{2000}{1.3}}\right)\right] = 1575.76 \blacksquare$$

Example 34.5 ‡
A jewelry store has obtained two separate insurance policies that together provide full coverage. You are given:
(i) The average ground-up loss is 11,100.
(ii) Policy A has an ordinary deductible of 5,000 with no policy limit.
(iii) Under policy A, the expected amount paid per loss is 6,500.
(iv) Under policy A, the expected amount paid per payment is 10,000.
(v) Policy B has no deductible and a policy limit of 5,000.
Given that a loss has occurred, determine the probability that the payment under policy B is 5,000.

Solution.
Let X denote the ground-up loss random variable. By (i), $E(X) = 11,100$. By (ii) and (iii), we have

$$E(Y^L) = E(X) - E(X \wedge 5000) = 6,500.$$

By (iv), we habe

$$E(Y^P) = \frac{E(X) - E(X \wedge 5000)}{1 - F_X(500)} = 10,000.$$

Thus,

$$\frac{6500}{1 - F_X(5000)} = 10,000 \Longrightarrow F_X(5000) = 0.35.$$

For Policy B, a payment of 5000 will occur if $X \geq 5000$. Hence,

$$\Pr(X \geq 5000) = 1 - F_X(5000) = 1 - 0.35 = 0.65 \blacksquare$$

Practice Problems

Problem 34.1
Losses follow an exponential distribution with mean $\frac{1}{\theta}$. For a coverage with policy limit u, find $f_Y(y)$ and $F_Y(y)$.

Problem 34.2
Losses follow an exponential distribution with mean $\frac{1}{\theta}$. For a coverage with policy limit u, find $E(X \wedge u)$.

Problem 34.3
Losses follow an exponential distribution with mean $\frac{1}{\theta}$. For a coverage with policy limit u and with an inflation at rate r, find the expected cost.

Problem 34.4
Losses are distributed uniformly between 0 and 100. An insurance policy which covers the losses has an upper limit of 80. Find the expected cost.

Problem 34.5 ‡
An insurance company offers two types of policies: Type Q and Type R. Type Q has no deductible, but has a policy limit of 3000. Type R has no limit, but has an ordinary deductible of d. Losses follow a Pareto distribution with $\alpha = 3$ and $\theta = 2000$.

Calculate the deductible d such that both policies have the same expected cost per loss.

Problem 34.6
Suppose that the ground-up losses for 2010 follow an exponential distribution with a mean of 1000. In 2011, all losses are subject to uniform inflation of 25%. The policy in 2011 has limit u.

Determine the value of u if the expected cost in 2011 is equal to the expected loss in 2010.

Problem 34.7
Suppose that the ground-up losses for 2010 follow a Pareto distribution with parameters $\alpha = 3$ and $\theta = 9800$. In 2011, all losses are subject to uniform inflation of 6%. The policy limit in 2011 is 170,000.

Calculate the expected cost in 2011.

35 Combinations of Coinsurance, Deductibles, Limits, and Inflations

Another coverage modification is the use of coinsurance factor. In a policy with a **coinsurance** factor $0 < \alpha < 1$, the insurer portion of the loss is αX and the insured portion is $(1-\alpha)X$. The claim amount random variable is $Y = \alpha X$ and its distribution function is

$$F_Y(y) = \Pr(Y \le y) = \Pr\left(X \le \frac{y}{\alpha}\right) = F_X\left(\frac{y}{\alpha}\right).$$

Its density function is given by

$$f_Y(y) = \frac{1}{\alpha} f_X\left(\frac{y}{\alpha}\right).$$

Now, suppose that a policy is subject to a coinsurance factor α, an ordinary deductible d, and a policy limit u_*. By convention, the coinsurance factor is applied after the application of any deductible or limit. We define the **maximum covered loss** to be $u = u_* + d$. Thus, in the absence of a deductible the maximum covered loss is just the policy limit. The claim amount per loss random variable is given by

$$Y^L = \begin{cases} 0, & X \le d \\ \alpha(X-d), & d < X \le u \\ \alpha(u-d), & X > u. \end{cases}$$

In compact form, we can write

$$Y^L = \alpha[X \wedge u - X \wedge d].$$

Note that $\alpha(u-d)$ is the maximum payment per loss.

Example 35.1
A health insurance policy has a deductible of 200, a policy limit of 5000 and a coinsurance factor of 80%. Calculate the expected claim amount per loss event and the expected claim amount per payment event for this policy if losses follow an exponential distribution with mean 1000.

Solution.
Note that $u = 5000 + 200 = 5200$. The expected claim amount per loss is

$$\begin{aligned} E(Y^L) =& \alpha[E(X \wedge u) - E(X \wedge d)] \\ =& \alpha\left[\frac{1}{\theta}(1-e^{-\frac{u}{\theta}}) - \frac{1}{\theta}(1-e^{-\frac{d}{\theta}})\right] \\ =& 0.8\left[1000(1-e^{-5.2}) - 1000(1-e^{-0.2})\right] \\ =& 650.57 \end{aligned}$$

The expected claim amount per payment is

$$E(Y^P) = \frac{E(Y^L)}{1 - F_X(200)} = \frac{650.57}{e^{-0.2}} = 794.61 \blacksquare$$

For a policy subject to a coinsurance factor α, an ordinary deductible d, policy limit u_*, and uniform inflation at rate r applied only to losses, the claim amount per loss is given

$$Y^L = \begin{cases} 0, & X \leq \frac{d}{1+r} \\ \alpha[(1+r)X - d], & \frac{d}{1+r} < X \leq \frac{u}{1+r} \\ \alpha(u-d), & X > \frac{u}{1+r} \end{cases}$$

where $u = u_* + d$. In compact form, we can write

$$Y^L = \alpha[(1+r)X \wedge u - (1+r)X \wedge d].$$

Theorem 35.1
The expected value of the per-loss random variable is

$$E(Y^L) = \alpha(1+r)\left[E\left(X \wedge \frac{u}{1+r}\right) - E\left(X \wedge \frac{d}{1+r}\right)\right].$$

The expected value of the per-payment random variable is

$$E(Y^P) = \frac{E(Y^L)}{1 - F_X\left(\frac{d}{1+r}\right)}.$$

Proof.
We have

$$\begin{aligned} E(Y^L) =& \alpha\{E[(1+r)X \wedge u] - E[(1+r)X \wedge d]\} \\ =& \alpha\left[(1+r)E\left(X \wedge \frac{u}{1+r}\right) - (1+r)E\left(X \wedge \frac{d}{1+r}\right)\right] \\ =& \alpha(1+r)\left[E\left(X \wedge \frac{u}{1+r}\right) - E\left(X \wedge \frac{d}{1+r}\right)\right]. \end{aligned}$$

For the expected cost per payment divide the expression by $1 - F_X\left(\frac{d}{1+r}\right)$ $\blacksquare$

To find the variance of Y^L, the second moment is required which is provided by the following theorem.

Theorem 35.2
The second moment for the per-loss random variable is

$$E[(Y^L)^2] = [\alpha(1+r)]^2 \left\{E[(X \wedge u^*)^2] - E[(X \wedge d^*)^2] - 2d^*[E(X \wedge u^*) - E(X \wedge d^*)]\right\}.$$

Proof.
Using

$$Y^L = \alpha[(1+r)X \wedge u - (1+r)X \wedge d] = \alpha(1+r)[X \wedge u^* - X \wedge d^*]$$

we can write

$$\begin{aligned}\left[\frac{Y^L}{\alpha(1+r)}\right]^2 =&[X \wedge u^* - X \wedge d^*]^2\\ =&(X \wedge u^*)^2 - 2(X \wedge u^*)(X \wedge d^*) + (X \wedge d^*)^2\\ =&(X \wedge u^*)^2 - (X \wedge d^*)^2 - 2(X \wedge d^*)[X \wedge u^* - X \wedge d^*].\end{aligned}$$

But

$$(X\wedge d^*)[X\wedge u^* - X\wedge d^*] = \begin{cases} 0, & X \le \frac{d}{1+r}\\ d^*(X - d^*), & \frac{d}{1+r} < X \le \frac{u}{1+r}\\ d^*(u^* - d^*), & X > \frac{u}{1+r}\end{cases} = d^*[X\wedge u^* - X\wedge d^*].$$

Thus,

$$\left[\frac{Y^L}{\alpha(1+r)}\right]^2 = (X \wedge u^*)^2 - (X \wedge d^*)^2 - 2d^*[X \wedge u^* - X \wedge d^*].$$

Now, the result of the theorem follows by taking the expectation of both sides ■

Example 35.2
Determine the mean and the standard deviation per loss for an exponential distribution with mean 1000 and with a deductible of 500 and a policy limit of 2500.

Solution.
Recall that $E(X \wedge x) = \theta(1 - e^{-\frac{x}{\theta}})$. Thus,

$$E(Y^L) = E(X\wedge 3000) - E(X\wedge 500) = 1000(1-e^{-\frac{3000}{1000}}) - 1000(1-e^{-\frac{500}{1000}}) = 556.74.$$

For the second moment of Y^L we have

$$\begin{aligned}
E[(X\wedge 3000)^2] =& \int_0^{3000} x^2(0.001)e^{-\frac{x}{1000}}dx + 3000^2e^{-3} = 1,601,703.453\\
E[(X\wedge 500)^2] =& \int_0^{500} x^2(0.001)e^{-\frac{x}{1000}}dx + 500^2e^{-0.5} = 180,408.0209\\
E[(Y^L)^2] =& 1,601,703.453 - 180,408.0209 - 2(500)(556.74)\\
=& 864,555.4321.
\end{aligned}$$

The variance of Y^L is

$$\text{Var}(Y^L) = 864555.4321 - 556.74^2 = 554,596.0045 \blacksquare$$

Example 35.3 ‡
A group dental policy has a negative binomial claim count distribution with mean 300 and variance 800.
Ground-up severity is given by the following table:

Severity	Probability
40	0.25
80	0.25
120	0.25
200	0.25

You expect severity to increase 50% with no change in frequency. You decide to impose a per claim deductible of 100.
Calculate the expected total claim payment after these changes.

Solution.
After imposing the 50% increase, the severity values are

Severity (X)	Probability
60	0.25
120	0.25
180	0.25
300	0.25

Let N be the claim frequency. Then the expected total claim payment is the expected number of losses times the expected payment per loss. That is,

$$E(N)[E(X) - E(X \wedge 100)].$$

We have

$$\begin{aligned}E(N) =&300\\E(X) =&(60+120+180+300)(0.25)=165\\E(X\wedge 100) =&60(0.25)+100(1-0.25)=90.\end{aligned}$$

Thus, the expected total claim payment is $300(165-90)=22,500$ ■

Example 35.4 ‡
An insurer has excess-of-loss reinsurance on auto insurance. You are given:
(i) Total expected losses in the year 2001 are 10,000,000.
(ii) In the year 2001 individual losses have a Pareto distribution with

$$F(x)=1-\left(\frac{2000}{x+2000}\right)^2,\ x>0.$$

(iii) Reinsurance will pay the excess of each loss over 3000.
(iv) Each year, the reinsurer is paid a ceded premium, C_{year}, equal to 110% of the expected losses covered by the reinsurance.
(v) Individual losses increase 5% each year due to inflation.
(vi) The frequency distribution does not change.
(a) Calculate C_{2001}.
(b) Calculate $\frac{C_{2002}}{C_{2001}}$.

Solution.
(a) The reinsurance fraction per loss is given by

$$\frac{E(X)-E(X\wedge 3000)}{E(X)}$$

where

$$\begin{aligned}E(X) =&\frac{\theta}{\alpha-1}=\frac{2000}{2-1}=2000\\E(X\wedge 3000) =&\frac{\theta}{\alpha-1}\left[1-\left(\frac{2000}{3000+2000}\right)^{\alpha-1}\right]\\=&\frac{2000}{2-1}\left[1-\left(\frac{2000}{3000+2000}\right)^{2-1}\right]=1200.\end{aligned}$$

Thus,

$$\frac{E(X)-E(X\wedge 3000)}{E(X)}=1-\frac{1200}{2000}=0.40.$$

Finally,

$$C_{2001} = 1.10(0.40)(10,000,000) = 4,400,000.$$

(b) Due to inflation, the amount per loss in 2002 is $X_{2002} = 1.05X_{2001}$. Thus, $E(X_{2002}) = 1.05E(X_{2001}) = 1.05(2000) = 2100$. Now, X_{2002} is a Pareto distribution with parameters $\alpha = 2$ and $\theta = 2100$) (see Section 18). Thus,

$$E(X_{2002} \wedge 3000) = \frac{2100}{2-1}\left[1 - \left(\frac{2100}{3000+2100}\right)^{2-1}\right] = 1235.$$

Hence,

$$C_{2002} = 1.10\left(\frac{2100-1235}{2100}\right)(10,000,000)(1.05) = 4,758,600$$

and

$$\frac{C_{2002}}{C_{2001}} = \frac{4,758,600}{4,400,000} = 1.08 \blacksquare$$

Example 35.5 ‡

Annual prescription drug costs are modeled by a two-parameter Pareto distribution with $\theta = 2000$ and $\alpha = 2$.

A prescription drug plan pays annual drug costs for an insured member subject to the following provisions:

(i) The insured pays 100% of costs up to the ordinary annual deductible of 250.

(ii) The insured then pays 25% of the costs between 250 and 2250.

(iii) The insured pays 100% of the costs above 2250 until the insured has paid 3600 in total.

(iv) The insured then pays 5% of the remaining costs.

Determine the expected annual plan payment.

Solution.

Let X denote the annual drug cost and Y the insurer's payment. What is the first value of X where the insured total payment reaches 3600? From what is given, if $X = 2250$ the insured's payment is $250 + 0.25(2250 - 250) = 750$. After this point, the insured's will pay 100% of the costs above 2250 until the insured has paid 3600 in total. But this means that insured will pay $3600 - 750 = 2850$ past the 2250 mark. In other words, the insured reaches the total payment of 3600 when $X = 2250 + 2850 = 5100$. Having said that, Y can be expressed as follows

$$Y = \begin{cases} 0, & 0 \leq X \leq 250 \\ 0.75(X - 250), & 250 < X \leq 2250 \\ 0.75(2250 - 250), & 2250 < X \leq 5100 \\ 1500 + 0.95(X - 5100), & X > 5100. \end{cases}$$

Next, note the following,

$$\begin{cases} 0, & 0 \leq X \leq 250 \\ 0.75(X - 250), & 250 < X \leq 2250 \\ 0.75(2250 - 250), & 2250 < X \leq 5100 \end{cases} = 0.75 \begin{cases} X - X, & 0 \leq X \leq 250 \\ X - 250, & 250 < X \leq 2250 \\ 2250 - 250, & 2250 < X \leq 5100 \end{cases}$$

$$= 0.75(X \wedge 2250 - X \wedge 250).$$

Thus, we can express Y as

$$Y = 0.75(X \wedge 2250 - X \wedge 250) + 0.95(X - X \wedge 5100).$$

Hence,

$$\begin{aligned} E(Y) =& 0.75[E(X \wedge 2250) - E(X \wedge 250)] + 0.95[E(X) - E(X \wedge 5100)] \\ =& 0.75\left\{\left(\frac{2000}{2-1}\right)\left[1 - \frac{2000}{2000 + 2250}\right] - \left(\frac{2000}{2-1}\right)\left[1 - \frac{2000}{2000 + 250}\right]\right\} \\ &+ 0.95\left\{\left(\frac{2000}{2-1}\right) + \left(\frac{2000}{2-1}\right)\left[1 - \frac{2000}{2000 + 5100}\right]\right\} \\ =& 1163 \blacksquare \end{aligned}$$

Example 35.6 ‡
Loss amounts have the distribution function

$$F(x) = \begin{cases} \left(\frac{x}{100}\right)^2, & 0 \leq x \leq 100 \\ 1, & x > 100 \end{cases}$$

An insurance pays 80% of the amount of the loss in excess of an ordinary deductible of 20, subject to a maximum payment of 60 per loss.
Calculate the conditional expected claim payment, given that a payment has been made.

Solution.
The maximum covered loss is

$$u = 20 + 0.8(60) = 95.$$

Thus,

$$\begin{aligned}
E(Y^L) =&0.8[E(X\wedge 95) - E(X\wedge 20)]\\
=&0.8\left[\int_0^{95} S(x)dx - \int_0^{20} S(x)dx\right]\\
=&0.8\left[\int_0^{95}\left[1-\left(\frac{x}{100}\right)^2\right]dx - \int_0^{20}\left[1-\left(\frac{x}{100}\right)^2\right]dx\right]\\
=&0.8\int_{20}^{95}\left[1-\left(\frac{x}{100}\right)^2\right]dx = 37.35.
\end{aligned}$$

The expected cost per payment is

$$E(Y^P) = \frac{E(Y^L)}{1-F(20)} = \frac{37.35}{1-0.04} = 38.91 \blacksquare$$

Example 35.7 ‡

For a special investment product, you are given:
(i) All deposits are credited with 75% of the annual equity index return, subject to a minimum guaranteed crediting rate of 3%.
(ii) The annual equity index return is normally distributed with a mean of 8% and a standard deviation of 16%.
(iii) For a random variable X which has a normal distribution with mean μ and standard deviation σ, you are given the following limited expected values:

$E(X \wedge 3\%)$		
	$\mu = 6\%$	$\mu = 8\%$
$\sigma = 12\%$	-0.43%	0.31%
$\sigma = 16\%$	-1.99%	-1.19%

$E(X \wedge 4\%)$		
	$\mu = 6\%$	$\mu = 8\%$
$\sigma = 12\%$	0.15%	0.95%
$\sigma = 16\%$	-1.43%	-0.58%

Calculate the expected annual crediting rate.

Solution.

Let$Y$ denote the annual credit rating and X the annual equity index return. Then we have

$$Y = \begin{cases} 3, & 0.75X \le 3 \\ 0.75X, & 0.75X > 3 \end{cases} = \begin{cases} 3, & X \le 4 \\ 0.75X, & X > 4 \end{cases} = 3 + \begin{cases} 0, & X \le 4 \\ 0.75X - 3, & X > 4 \end{cases}$$

That is,

$$Y = 3 + (0.75X - 3)_+ = 3 + 0.75X - (0.75X \wedge 3).$$

Hence,

$$E(Y) = 3 + 0.75E(X) - 0.75E(X \wedge 4).$$

By (ii), $E(X) = 8$. From the second given table, we see that $E(X \wedge 4) = -0.58$. Hence,

$$E(Y) = 3 + 6 - 0.75(-0.58) = 9.435\% \blacksquare$$

Example 35.8 ‡
A risk has a loss amount which has a Poisson distribution with mean 3.
An insurance covers the risk with an ordinary deductible of 2. An alternative insurance replaces the deductible with coinsurance α, which is the proportion of the loss paid by the insurance, so that the expected insurance cost remains the same.
Calculate α.

Solution.
The expected cost per loss with a deductible of 2 is

$$\begin{aligned}
E[(X-2)_+] =& E(X) - E(X \wedge 2) = 3 - [\Pr(X=1) + 2\Pr(X=2) + 2(1 - \Pr(X \le 2))] \\
=& 3 - \Pr(X=1) + 2[1 - \Pr(X=0) - \Pr(X=1)] \\
=& 3 - 3e^{-3} - 2[1 - (e^{-3} + 3e^{-3})] = 1.249.
\end{aligned}$$

The expected cost per loss with a coinsurance of α is $\alpha E(X) = 3\alpha$. We are told that $3\alpha = 1.249$ so that $\alpha = \frac{1.249}{3} = 0.42$ $\blacksquare$

Example 35.9 ‡
Michael is a professional stuntman who performs dangerous motorcycle jumps at extreme sports events around the world.
The annual cost of repairs to his motorcycle is modeled by a two parameter Pareto distribution with $\theta = 5000$ and $\alpha = 2$.
An insurance reimburses Michael's motorcycle repair costs subject to the following provisions:
(i) Michael pays an annual ordinary deductible of 1000 each year.
(ii) Michael pays 20% of repair costs between 1000 and 6000 each year.
(iii) Michael pays 100% of the annual repair costs above 6000 until Michael has paid 10,000 in out-of-pocket repair costs each year.
(iv) Michael pays 10% of the remaining repair costs each year.
Calculate the expected annual insurance reimbursement.

Solution.
Let X denote the annual repair cost. For $1000 \leq X \leq 6000$, the insurance pays $0.80(X-1000)$. Thus, if cost is 6000, Michael's share is 2000 and the insurance share is 4000. For the next 8000, Michael's share is the whole amount. That is, the insurance pays nothing. But in this case, Michael's out-of-pocket reaches 10,000. It follows that, the insurance pays 90% of cost over 14,000 which is $0.90(X-14000)$.
The annual insurance reimbursement is

$$\begin{aligned}R =&0.8[(X-1000)_+ - (X-6000)_+] + 0.9(X-14000)_+\\ =&0.8[X - X\wedge 1000 - (X - X\wedge 6000)] + 0.9(X - X\wedge 14000)\\ =&0.8(X\wedge 6000 - X\wedge 1000) + 0.9(X - X\wedge 14000).\end{aligned}$$

The expected annual insurance reimbursement is

$$E(R) = 0.8[E(X\wedge 6000) - E(X\wedge 1000)] + 0.9[E(X) - E(X\wedge 14000)].$$

From Table C,

$$E(X\wedge x) = \theta\left(1 - \frac{\theta}{\theta+x}\right) = \frac{5000x}{x+5000}$$

and

$$E(X) = \frac{\theta}{\alpha-1} = 5000.$$

Hence,

$$E(R) = 0.8\left(\frac{5000(6000)}{6000+5000} - \frac{5000(1000)}{1000+5000}\right) + 0.9\left(5000 - \frac{5000}{14000+5000}\right) = 2699.36 \blacksquare$$

Practice Problems

Problem 35.1 ‡
Losses this year have a distribution such that $E(X \wedge x) = -0.025x^2 + 1.475x - 2.25$ for $x = 11, 12, \cdots, 26$. Next year, losses will be uniformly higher by 10%. An insurance policy reimburses 100% of losses subject to a deductible of 11 up to a maximum reimbursement of 11.

Determine the ratio of next year 's reimbursements to this year's reimbursement.

Problem 35.2 ‡
Losses have an exponential distribution with mean 1000. An insurance company will pay the amount of each claim in excess of a deductible of 100.

Calculate the variance of the amount paid by the insurance company for one claim, including the possibility that the amount paid is 0.

Problem 35.3 ‡
Losses follow a Poisson distribution with mean $\lambda = 3$. Consider two insurance contracts. One has an ordinary deductible of 2. The second has no deductible and coinsurance in which the insurance company pays α of the loss.

Determine α so that the expected cost of the two contracts is the same.

Problem 35.4 ‡
You are given that $e(0) = 25$ and $S(x) = 1 - \frac{x}{w},\ 0 \le x \le w$, and Y^P is the excess loss variable for $d = 10$.

Determine the variance of Y^P.

Problem 35.5 ‡
Total hospital claims for a health plan were previously modeled by a two-parameter Pareto distribution with $\alpha = 2$ and $\theta = 500$. The health plan begins to provide financial incentives to physicians by paying a bonus of 50% of the amount by which total hospital claims are less than 500. No bonus is paid if total claims exceed 500. Total hospital claims for the health plan are now modeled by a new Pareto distribution with $\alpha = 2$ and $\theta = K$. The expected claims plus the expected bonus under the revised model equals expected claims under the previous model.

Calculate K.

Problem 35.6
The amount of a loss has a Pareto distribution with $\alpha = 2$ and $\theta = 5000$. An insurance policy on this loss has an ordinary deductible of 1,000, a policy limit of 10,000, and a coinsurance of 80%.

With a uniform inflation of 2%, calculate the expected claim amount per payment on this policy.

Problem 35.7
Claim amounts follow a Pareto distribution with parameters $\alpha_X = 3.$ and $\theta_X = 2000$. A policy is subject to a coinsurance rate of α. The standard deviation of the claims for this policy is 1472.24.

Determine the value of α.

Problem 35.8
The loss size distribution is exponential with mean 50. An insurance policy pays the following for each loss. There is no insurance payment for the first 20. The policy has a coinsurance in which the insurance company pays 75% of the loss. The maximum covered loss is 100.

Calculate $E(Y^P)$.

Problem 35.9
The loss size distribution is a Pareto $\alpha = 2$ and $\theta = 100$. An insurance policy pays the following for each loss. There is no insurance payment for the first 20. The policy has a coinsurance in which the insurance company pays 75% of the loss. The policy has a limit of u_*. The maximum covered loss is 100.

Given that $E(Y^P) = 34.2857$, find the maximum payment per loss for this policy.

Problem 35.10 ‡
In 2005 a risk has a two-parameter Pareto distribution with $\alpha = 2$ and $\theta = 3000$. In 2006 losses inflate by 20%.
An insurance on the risk has a deductible of 600 in each year. P_i, the premium in year i, equals 1.2 times the expected claims.
The risk is reinsured with a deductible that stays the same in each year. R_i,

the reinsurance premium in year i, equals 1.1 times the expected reinsured claims.
Suppose $R_{2005}/P_{2005} = 0.55$. Calculate R_{2006}/P_{2006}.

36 The Impact of Deductibles on the Number of Payments

Deductibles always affect the number of payments. For example, when an imposed deductible is increased the number of payments per period is decreased whereas decreasing deductibles results in an increasing number of payments per period.

Let X_j denote the jth ground-up loss and assume that there are no coverage modifications. Let N^L be the total number of losses. Now, suppose that a deductible is imposed. Let $v = \Pr(X > d)$ be the probability that a loss will result in a payment. We will let I_j be the indicator random variable whose value is 1 if the jth loss occur (and thus results in a payment) and is 0 otherwise. Then I_j is a Bernoulli random variable such that $\Pr(I_j = 1) = v$ and $\Pr(I_j = 0) = 1 - v$. The corresponding pgf is $P_{I_j}(z) = vz + 1 - v$.

Now, let N^P be the total number of payments. Then $N^P = I_1 + I_2 + \cdots + I_{N^L}$. We call N^P a **compound frequency model** or **aggregate claims model**. If $I_1, I_2, \cdots, I_j, N_L$ are mutually independent and the individual claims I_j are identically and independently distributed (iid) then N^P is a **compound distribution** with **primary distribution** N^L and **secondary distribution** a Bernoulli distribution. For such a distribution, we have

$$\begin{aligned}
P_{N^P}(z) =&\sum_{k=0}^{\infty}\Pr(N^P = k)z^k = \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\Pr(N^P = k|N^L = n)\Pr(N^L = n)z^k \\
=&\sum_{n=0}^{\infty}\Pr(N^L = n)\sum_{k=0}^{\infty}\Pr(I_1 + I_2 + \cdots + I_n = k|N^L = n)z^k \\
=&\sum_{n=0}^{\infty}\Pr(N^L = n)[P_{I_j}(z)]^n \\
=&P_{N^L}[P_{I_j}(z)] = P_{N^L}[1 + v(z - 1)]
\end{aligned}$$

where we used the fact that the pgf of a sum of independent random variables is the product of pgf of the individual random variables.

Example 36.1
Suppose that losses follow an exponential distribution all with the same mean 100. Suppose that insurance policies are subject to an ordinary deductible of 20 and that N^L follows a Poisson distribution with $\lambda = 3$.

(a) Calculate the probability that a loss will result in a payment.
(b) Find the pgf of an indicator random variable I_j.
(c) Find the pgf of the total number of losses N^L.
(d) Find the pgf of the total number of payments N^P.

Solution.
(a) The probability of a loss that results in a payment is

$$v = \int_{20}^{\infty} e^{-\frac{x}{100}} dx = 100e^{-0.2}.$$

(b) $P_{I_j}(z) = 1 - 100e^{-0.2} + 100e^{-0.2}z$.
(c) $P_{N^L}(z) = e^{3(z-1)}$.
(d) $P_{N^P}(z) = e^{3(100e^{-0.2}z - 100e^{-0.2})}$ ■

Remark 36.1
In general, if K is a compound distribution with primary distribution N and secondary distribution M then

$$P_K(z) = P_N[P_M(z)].$$

Example 36.2 ‡
An actuary has created a compound claims frequency model with the following properties:
(i) The primary distribution is the negative binomial with probability generating function

$$P(z) = [1 - 3(z-1)]^{-2}$$

(ii) The secondary distribution is the Poisson with probability generating function

$$P(z) = e^{\lambda}(z-1).$$

(iii) The probability of no claims equals 0.067.
Calculate λ.

Solution.
From the above remark, we have

$$P_K(z) = P_N[P_M(z)] = P_N[e^{\lambda(z-1)}] = [1 - 3(e^{\lambda(z-1)} - 1)]^{-2}.$$

From Section 8, we know that $P_K(0) = \Pr(K = 0)$. Hence

$$0.067 = [1 - 3(e^{-\lambda} - 1]^{-2} \Longrightarrow \lambda = 3.1$$ ■

Now, suppose that N^L depends on a parameter θ and that its pgf $P_{N^L}(z;\theta)$ satisfies the equation

$$P_{N^L}(z;\theta) = B[\theta(z-1)]$$

where $B(z)$ is independent of θ and both z and θ only appear in the pgf as $\theta(z-1)$. Then we have

$$\begin{aligned} P_{N^P}(z) =& P_{N^P}(1-v+vz) = B[\theta(1-v+vz-1)] \\ =& B[v\theta(z-1)] = P_{N^L}(z, v\theta). \end{aligned}$$

This shows that N^L and N^P are both from the same parametric family and only the parameter θ need to be changed to $v\theta$.

Example 36.3
Suppose that N^L follows a negative binomial distribution with parameters r and β. Find $P_{N^L}(z)$ and $B(z)$.

Solution.
We have $P_{N^L}(z) = [1-\beta(z-1)]^{-r}$ so that $B(z) = (1-z)^{-r}$. Note that β takes on the role of θ in the above result ■

Example 36.4
Losses follow a Pareto distribution with $\alpha = 3$ and $\theta = 1000$. Assume that N^L follow a negative binomial distribution with $r = 2$ and $\beta = 3$. Find $P_{N^P}(z)$ when a deductible of 250 is imposed on N^L.

Solution.
N^P has a negative binomial distribution with $r^* = 2$ and $\beta^* = \beta v$. Since

$$v = 1 - F_X(250) = \left(\frac{1000}{1000+250}\right)^3 = 0.512$$

we have

$$P_{N^P}(z) = B[v\beta(z-1)] = [1-0.512(3)(z-1)]^{-2} = [1-1.536(z-1)]^{-2} \blacksquare$$

Now, suppose that N^L depends on two parameters θ and α with the pgf satisfying the equation

$$P_{N^L}(z) = P_{N^L}(z;\theta;\alpha) = \alpha + (1-\alpha)\frac{B[\theta(z-1)] - B(-\theta)}{1-B(-\theta)}.$$

Note that $\alpha = P_{N^L}(0) = \Pr(N^L = 0)$ so that N^L is a zero-trancated distribution. With this at hand, by letting

$$\alpha^* = P_{N^L}(1-v;\theta;\alpha) = \alpha + (1-\alpha)\frac{B(-v\theta) - B(-\theta)}{1-B(-\theta)}$$

we have

$$\begin{aligned}
P_{N^P}(z) =& P_{N^L}[1-v+vz] \\
=& \alpha + (1-\alpha)\frac{B[v\theta(z-1)] - B(-\theta)}{1-B(-\theta)} \\
=& \left[\alpha + (1-\alpha)\frac{B[-v\theta] - B(-\theta)}{1-B(-\theta)}\right] \\
&+ \left\{1 - \left[\alpha + (1-\alpha)\frac{B[-v\theta] - B(-\theta)}{1-B(-\theta)}\right]\right\}\frac{B[v\theta(z-1)] - B(-v\theta)}{1-B(-v\theta)} \\
=& \alpha^* + (1-\alpha^*)\frac{B[v\theta(z-1)] - B(-v\theta)}{1-B(-v\theta)} \\
=& P_{N^L}(z; v\theta; \alpha^*)
\end{aligned}$$

where

$$\alpha^* = P_{N^P}(0) = \Pr(N^P = 0).$$

Hence, if N^L is zero-modified then N^P is zero-modified.

Example 36.5
Losses follow a Pareto distribution with $\alpha = 3$ and $\theta = 1000$. Assume that N^L follow a zero-modified negative binomial distribution with $r = 2,\ \beta = 3$ and $p_0^M = 0.4$ Find $P_{N^P}(z)$ when a deductible of 250 is imposed on N^L.

solution.
The pgf of N^L is (See Problem 30.6)

$$P_{N^L}(z) = p_0^M + (1-p_0^M)\frac{[1-\beta(z-1)]^{-r} - (1+\beta)^{-r}}{1-(1+\beta)^{-r}}$$

where $\alpha = p_0^M$ and $B(z) = (1-z)^{-r}$. Hence, N^P is a zero-modified negative binomial distribution with $r^* = r2,\ \beta^* = v\beta = 1.536$ and

$$\alpha^* = p_0^M * = 0.4 + (1-0.4)\frac{(1+1.536)^{-2} - (1+3)^{-2}}{1-(1+3)^{-2}} = 0.4595.$$

Hence,

$$P_{N^P}(z) = 0.4595 + 0.5405\frac{[1-1.536(z-1)]^{-2} - 2.536^{-2}}{1-2.536^{-2}} \blacksquare$$

Now, we can obtain the pgf of N^L given the pgf of N^P :

$$P_{N^P}(z) = P_{N^L}[P_{I_j}(z)] \Longrightarrow P_{N^L}(z) = P_{N^P}[P_{I_j}^{-1}(z)] = P_{N^P}(1 - v^{-1} + v^{-1}z).$$

If N^P is a function of some parameter θ then we have

$$P_{N^L}(z;\theta) = P_{N^P}\left(z; \frac{\theta}{v}\right).$$

Example 36.6
Suppose payments on a policy with a deductible of 250 have the zero-modified negative binomial distribution with $r^* = 2$, $\beta^* = 1.536$ and $p_0^M * = 0.4595$. Losses have the Pareto distribution with $\alpha = 3$ and $\beta = 1000$. Determine the distribution of payments when the deductible is removed.

solution.
When the deductibles are removed, the number of payments follow a zero-modified negative binomial with $r = 2$, $\beta = \frac{\beta^*}{v} = \frac{1.536}{0.512} = 3$ and

$$p_0^M = \frac{p_0^M * -(1+\beta^*)^{-r} + (1+\frac{\beta^*}{v})^{-r} - p_0^M * (1+\frac{\beta^*}{v})^{-r}}{1-(1+\beta^*)^{-r}}$$

$$p_0^M = \frac{0.4595 - 2.536^{-2} + 4^{-2} - 0.4595(4)^{-2}}{1-2.536^{-2}} = 0.4 \blacksquare$$

Practice Problems

Problem 36.1 ‡
The frequency distribution for the number of losses when there is no deductible is negative binomial with $r = 2$ and $\beta = 5$. Loss amounts have a Weibull distribution with $r = 0.3$ and $\theta = 1000$.

Determine the expected number of payments when a deductible of 200 is applied.

Problem 36.2
Individual losses have an exponential distribution with mean 340. With a deductible of 200, the frequency distribution for the number of payments is Poisson with mean $\lambda = 0.5$.

Determine the probability that a loss will result in a payment.

Problem 36.3
Individual losses have an exponential distribution with mean 340. With a deductible of 200, the frequency distribution for the number of payments is Poisson with mean $\lambda = 0.5$.

Find the pgf of the number of payments.

Problem 36.4
Individual losses have an exponential distribution with mean 340. With a deductible of 200, the frequency distribution for the number of payments is Poisson with mean $\lambda = 0.5$.

What fraction of policyholders would be expected to have at least one claim paid to them?

Problem 36.5
Loss amounts follow a Pareto distribution with $\alpha = 3$ and $\theta = 1000$. With a deductible of 500, the frequency distribution for the number of payments is geometric with $\beta = 0.4$.

Find $P_{N^P}(z)$.

Problem 36.6
Individual losses have an exponential distribution with mean 1000. With a

deductible of 200, the frequency distribution for the number of payments is zero-modified Poisson with mean $\lambda = 0.5$.

Determine $p_0^M *$ if $p_0^M = 0.4$.

Problem 36.7
Suppose payments on a policy with deductible 200 have a zero-modified Poisson distribution with $\lambda^* = 0.27756$ and $p_0^M * = 0.6304$. Losses have an exponential distribution with mean $\theta = 1000$.

Find p_0^M if the deductible is removed.

Aggregate Loss Models

We continue our discussion of modeling losses. Keep in mind that losses depend on the **loss frequency**(i.e., the number of losses) and on the **loss severity**(i.e., the size or the amount of the loss). In this chapter, we will concentrate our attention on aggregate loss models. An **aggregate loss** refers to the total amount of losses in one period of time, which is often encountered in the analysis of a portfolio of risks such as a group insurance policy.

37 Individual Risk and Collective Risk Models

Consider a set of n insurance policies. Let $X_i,\ i = 1, 2, \cdots, n$, denote the loss amount of each policy. We assume that the $X_i's$ are independent but are not necessarily identically distributed. The **individual risk model** represents the aggregate loss as a sum

$$S = X_1 + X_2 + \cdots + X_n.$$

The $X_i's$ usually have mixed distributions with a probability mass at zero corresponding to the probability of no loss or payments.

This type of models is used in modeling group life or health insurance policy of a group of n individuals where each individual can have different coverage and different level of loss probabilities.

If the $X_i's$ are identically distributed then the individual risk model becomes a special case of the so-called collective risk model which we define next.

Let N be a random variable representing the number of claims which we refer to as the **claim count random variable** or **frequency** and whose distribution is called the **claim count distribution** or **frequency distribution**. Let $X_1, X_2, \cdots, X_N$ be the loss amounts. We refer to each X_i as a **severity** and to its distribution as the **severity distribution**. We assume that the $X_i's$ are independent and identically distributed. We further assume that the $X_i's$ are independent of N. The **collective risk model** represents the aggregate loss as the sum

$$S = X_1 + X_2 + \cdots + X_N.$$

S is an example of a **compound distribution** with **primary distribution** the distribution of N and **secondary distribution** the distribution of the losses. Also, we mention here that when $S = 0$ then $N = 0$.

There are advantages in modeling the claim frequency and claim severity separately, and then combine them to obtain the aggregate loss distribution. For example, expansion of insurance business may have impacts on the claim frequency but not the claim severity. In contrast, a cost increae may a affect the claim severity with no effects on the claim frequency.

Example 37.1
An insurance company insures 500 individuals against accidental death. Which of the following risk models is better suited for evaluating the risk to the insurance company?
(a) A collective risk model.
(b) An individual risk model.

Solution.
Since the number of policies is fixed, the model that is most suited to the insurance company is the individual risk model ■

Example 37.2
Which of the following statements are true?
(a) In collective risk model, the number of summands in the aggregate loss is fixed.
(b) In individual risk model, the number of summands is a random variable.
(c) In individual risk models, the summands are independent but not necessarily identically distributed.

Solution.
(a) This is false. In collective risk model, the number of summands in the aggregate loss S is a random variable−which we refer to as the frequency random variable.
(b) This is false. In individual risk model, the number of summands is a fixed number.
(c) This is true ■

Example 37.3 ‡
You are given:

# of Claims	Probability	Claim Size	Probability
0	$\frac{1}{5}$		
1	$\frac{3}{5}$	25 150	$\frac{1}{3}$ $\frac{2}{3}$
2	$\frac{1}{5}$	50 200	$\frac{2}{3}$ $\frac{1}{3}$

Claim sizes are independent. Determine the variance of the aggregate loss.

Solution.
For $N = 0$, we have $S = 0$. For $N = 1$ either $S = 25$ or $S = 150$. For $N = 2$,

either $S = 50+50 = 100$, $S = 200+200 = 400$, or $S = 50+200 = 250$. The probability distribution of the aggregate loss is as follows:

$$\begin{aligned}
\Pr(S=0) =& \frac{1}{5} \\
\Pr(S=25) =& \Pr(S=25|N=1)\Pr(N=1) \\
=& \Pr(X=25|N=1)\Pr(N=1) = \left(\frac{1}{3}\right)\left(\frac{3}{5}\right) = \frac{1}{5} \\
\Pr(S=100) =& \Pr(S=100|N=2)\Pr(N=2) \\
=& \Pr[(X_1=50)\text{and}(X_2=50)]\Pr(N=2) = \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{5}\right) = \frac{4}{45} \\
\Pr(S=150) =& \Pr(S=150|N=1)\Pr(N=1) \\
=& \Pr(X=150|N=1)\Pr(N=1) = \left(\frac{2}{3}\right)\left(\frac{3}{5}\right) = \frac{2}{5} \\
\Pr(S=250) =& \Pr(S=250|N=2)\Pr(N=2) \\
=& [\Pr[(X_1=50)\text{and}(X_2=200)] + \Pr[(X_1=200)\text{and}(X_2=50)]\Pr(N=2) \\
=& 2\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) = \frac{4}{45} \\
\Pr(S=400) =& \Pr(S=400|N=2)\Pr(N=2) \\
=& \Pr[(X_1=200)\text{and}(X_2=200)]\Pr(N=2) = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) = \frac{1}{45}.
\end{aligned}$$

Thus,

$$\begin{aligned}
E(S) =& 0\left(\frac{1}{5}\right) + 25\left(\frac{1}{5}\right) + 100\left(\frac{4}{45}\right) + 150\left(\frac{2}{5}\right) + 250\left(\frac{4}{45}\right) + 400\left(\frac{1}{45}\right) \\
=& 105 \\
E(S^2) =& 0^2\left(\frac{1}{5}\right) + 25^2\left(\frac{1}{5}\right) + 100^2\left(\frac{4}{45}\right) + 150^2\left(\frac{2}{5}\right) + 250^2\left(\frac{4}{45}\right) + 400^2\left(\frac{1}{45}\right) \\
=& 19125 \\
\text{Var}(S) =& E(S^2) - E(S)^2 = 19125 - 105^2 = 8100 \ \blacksquare
\end{aligned}$$

In [1], it is suggested that the frequency distribution N is best to have a distribution with the following probability generating function:

$$P_N(z;\alpha) = Q(z)^\alpha,$$

where $Q(z)$ is some function of z independent of α. This implies that if the volume of business of the insurance company increases by 100r%, expected claims will increase in a manner proportional to $(1+r)\alpha$.

Example 37.4
Show that the Poisson distribution with parameter λ has a probability generating function of the form $P(z) = Q(z)^\alpha$.

Solution.
The probability generating function is

$$P(z) = e^{\lambda(z-1)} = [e^{(z-1)}]^\lambda.$$

Thus, $Q(z) = e^{z-1}$ and $\alpha = \lambda$ ■

Example 37.5 ‡
In order to simplify an actuarial analysis Actuary A uses an aggregate distribution $S = X_1 + X_2 + \cdots + X_N$, where N has a Poisson distribution with mean 10 and $X_i = 1.5$ for all i.
Actuary A's work is criticized because the actual severity distribution is given by

$$\Pr(Y_i = 1) = \Pr(Y_i = 2) = 0.5, \text{ for all} i,$$

where $Y_i's$ are independent.
Actuary A counters this criticism by claiming that the correlation coefficient between S and $S^* = Y_1 + Y_2 + \cdots + Y_N$ is high.
Calculate the correlation coefficient between S and S^*.

Solution.
The coefficient of corrolation between S and S^* is given by

$$\rho = \frac{\text{Cov}(S, S^*)}{\sqrt{\text{Var}(S)}\sqrt{\text{Var}(S^8)}}.$$

We have

$$\begin{aligned}
E(S) =& E(N)E(X) = 10(1.5) = 15 \\
\text{Var}(S) =& E(N)\text{Var}(X) + \text{Var}(N)E(X)^2 \\
=& 10(0) + 10(1.5)^2 = 22.5 \\
E(S^*) =& E(N)E(Y) = 10(1.5) = 15 \\
\text{Var}(S^*) =& E(N)\text{Var}(Y) + \text{Var}(N)E(Y)^2 \\
=& 10(0.25) + 10(1.5)^2 = 25
\end{aligned}$$

$$\begin{aligned}
E(SS^*) =& E[E(SS^*|N)] = E\{E[(X_1 + X_2 + \cdots + X_N)(Y_1 + Y_2 + \cdots + Y_N)|N)\} \\
=& E\{E(1.5N(Y_1 + Y_2 + \cdots + Y_N)|N)\} = E[1.5NE(NY)] \\
=& E[1.5N^2E(Y)] = E[1.5N^2(1.5)] = 2.25E(N^2) \\
=& 2.25[\mathrm{Var}(N) + E(N)^2] = 2.25(10 + 100) = 247.5 \\
\mathrm{Cov}(S, S^*) =& E(SS^*) - E(S)E(S^*) \\
=& 247.5 - 15^2 = 22.5 \\
\rho =& \frac{22.5}{\sqrt{22.5}\sqrt{25}} = 0.949 \blacksquare
\end{aligned}$$

Practice Problems

Problem 37.1
Which of the following statements are true?

(a) In a collective risk model one policy can be modeled using a Pareto distribution and another can be modeled with an exponential distribution.
(b) In an individual risk model, the loss amounts can have different probabilities at zero.
(c) In the collective risk model, the frequency and severity of payments are modeled separately.
(d) In the collective risk model, the number of payments affects the size of each individual payment.

Problem 37.2
Consider the aggregate loss sum in an individual loss model:

$$S = X_1 + X_2 + \cdots + X_n.$$

Assume that the loss amounts are identically distributed. Find the limit of the coefficient of variation of S as n goes to infinity.

Problem 37.3
Consider a portfolio of two policies. One policy follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 100$ and the other policy follows an exponential distribution with parameter 0.05. Assume that the two policies are independent.

Calculate the mean and the variance of the aggregate loss S.

Problem 37.4
Determine whether the following model is individual or collective: The number of claims per day N has a geometric distribution with mean 2. The size of each claim has an exponential distribution with mean 1000. The number of losses and loss sizes are mutually independent.

Problem 37.5
Let N, the claim count random variable, follow a negative binomial distribution with parameters r and β. Show that $P_N(z;\alpha) = [Q(z)]^\alpha$.

38 Aggregate Loss Distributions via Convolutions

In this section, we look at distributional quantities of the aggregate loss random variable for the collective model. Let

$$S = X_1 + X_2 + \cdots + X_N$$

where N is the counting random variable or the frequency random variable and the $X_i's$ are independent and identically distributed. Moreover, the $X_i's$ are independent of N i.e., the claim counts are independent of the claim sizes. Thus, S is a compound distribution with primary distribution N and secondary distribution the common distribution of the $X_i's$. Analytical methods for the calculation of compound distributions can be accomplished using methods of convolutions. This section presents these methods.

The distribution of S can be accomplished via convolutions as:

$$F_S(x) = \Pr(S \le x) = \sum_{n=0}^{\infty} \Pr(S \le x | N = n)\Pr(N = n) = \sum_{n=0}^{\infty} \Pr(N = n) F_X^{*n}(x)$$

where $F_X^{*n}(x) = \Pr(X_1 + X_2 + \cdots + X_n \le x)$ is the $n-$th fold convolution of $F_X(x)$ defined recursively by

$$F_X^{*n}(x) = \int_0^x F_X^{*(n-1)}(x-y) f_X(y) dy$$

and

$$F_X^{*0}(x) = \begin{cases} 1, x \ge 0 \\ 0, x < 0. \end{cases}$$

The random variable X is the common distribution of the $X_i's$. Note that $F_X^{*1} = F_X(x)$. By differentiation, and assuming that differentiation and integration can be reversed, we find the pdf

$$f_X^{*n}(x) = \int_0^x f_X^{*(n-1)}(x-y) f_X(y) dy, \ n = 2, 3, \cdots.$$

In words, $f_X^{*n}(x)$ is the probability that, given exactly n claims occur, that the aggregate amount is x. Now, for a continuous X, the pdf of S is given by

$$f_S(x) = \sum_{n=1}^{\infty} \Pr(N = n) f_X^{*n}(x)$$

and a discrete mass point at $x=0$ given by $\Pr(S=0)=\Pr(N=0)$.

If X has a discrete counting distribution with probabilities at $0,1,2,\cdots$ then

$$F_X^{*n}(x)=\sum_{y=0}^{x}F_X^{*(n-1)}(x-y)f_X(y),\ x=0,1,2,\cdots;n=2,3,\cdots.$$

The corresponding pdf is

$$f_X^{*n}(x)=\sum_{y=0}^{x}f_X^{*(n-1)}(x-y)f_X(y),\ x=0,1,2,\cdots;n=2,3,\cdots.$$

By defining $f_X^{*0}(0)=1$ and $f_X^{*0}(x)=0$ for $x\neq 0$, we can write

$$f_S(x)=\Pr(S=x)=\sum_{n=0}^{\infty}\Pr(N=n)f_X^{*n}(x),\ x=0,1,2,\cdots.$$

Though the obtained formula is analytic, its direct calculation is difficult because, in general, the convolution powers are not available in closed-form. Panjer recursion , discussed in Section 41 is a very efficient numerical method to calculate these convolutions.

Example 38.1
An insurance portfolio produces N claims, where

n	$\Pr(N=n)$
0	0.5
1	0.2
2	0.2
3	0.1

Individual claim amounts have the following distribution:

x	$f_X(x)$
1	0.9
2	0.1

Individual claim amounts and N are mutually indepedent. Complete the following table:

x	$f_X^{*0}(x)$	$f_X^{*1}(x)$	$f_X^{*2}(x)$	$f_X^{*3}(x)$	$f_S(x)$
0					
1					
2					
3					
4					
$\Pr(N=n)$					

Solution.

x	$f_X^{*0}(x)$	$f_X^{*1}(x)$	$f_X^{*2}(x)$	$f_X^{*3}(x)$	$f_S(x)$
0	1	0	0	0	0.5
1	0	0.9	0	0	0.18
2	0	0.1	0.81	0	0.182
3	0	0	0.09	0.729	0.0909
4	0	0	0.01	0.162	0.0182
$\Pr(N=n)$	0.5	0.2	0.2	0.1	

To find $f_X^{*n}(x)$ pick two columns whose superscripts sum to n. Then add all combinations of products from these columns where the arguments sum to x. For example, $f_X^{*3}(4) = f_X^{*1}(1)f_X^{*2}(3)+f_X^{*1}(3)f_X^{*2}(1)+f_X^{*1}(2)f_X^{*2}(2)$. To obtain, $f_S(x)$, each row of the matrix of convolutions of $f_X(x)$ is multiplied by the probabilities from the row below the table and the products are summed ■

Example 38.2 ‡
The number of claims in a period has a geometric distribution with mean 4. The amount of each claim X follows $\Pr(X=x) = 0.25,\ x = 1,2,3,4$. The number of claims and the claim amounts are independent. S is the aggregate claim amount in the period.
Calculate $F_S(3)$.

Solution.
Note first that S is discrete so that

$$F_S(3) = f_S(0) + f_S(1) + f_S(2) + f_S(3)$$

where

$$f_S(x) = \sum_{n=0}^{\infty} \Pr(N=n) f_X^{*n}(x),\ x = 0,1,2,\cdots.$$

With four claims or more, i.e., $n \geq 4$ the aggregate sum is at least 4. Since we are interested in $S \leq 3$, we see that $f_X^{*n}(x) = 0$ for $n \geq 4$. Next, we create the chart

x	$f_X^{*0}(x)$	$f_X^{*1}(x)$	$f_X^{*2}(x)$	$f_X^{*3}(x)$	$f_S(x)$
0	1	0	0	0	0.2
1	0	0.25	0	0	0.04
2	0	0.25	0.0625	0	0.048
3	0	0.25	0.125	0.0156	0.0576
$\Pr(N=n)$	0.2	0.16	0.128	0.1024	

where

$$\Pr(N=n) = \frac{4^n}{5^{n+1}}, \ n = 1, 2, \cdots.$$

Hence,

$$F_S(3) = 0.2 + 0.04 + 0.048 + 0.0576 = 0.3456 \blacksquare$$

Example 38.3

Severities have a uniform distribution on $[0, 100]$. The frequency distribution is given by

n	Probability
0	0.60
1	0.30
2	0.10

Find the 2-fold convolution of $F_X(x)$.

Solution.

We know that $F_X^{*0}(x) = 1$ for $x \geq 0$. Thus,

$$F_X^{*1}(x) = \int_0^x \frac{dt}{100} = \frac{x}{100}.$$

The 2-fold convolution is

$$F_X^{*2}(x) = \int_0^x \frac{t}{100}\frac{dt}{100} = \frac{x^2}{20000} \blacksquare$$

Example 38.4

Severities have a uniform distribution on $[0, 100]$. The frequency distribution is given by

n	Probability
0	0.60
1	0.30
2	0.10

Find $F_S(x)$.

Solution.
We have

$$F_S(x) = 0.60 + 0.30\frac{x}{100} + 0.10\frac{x^2}{20000} \blacksquare$$

The pgf of S is found as follows:

$$\begin{aligned}
P_S(x) =& E(z^S) = E[z^0]\Pr(N=0) + \sum_{n=1}^{\infty} E[z^{X_1+X_2+\cdots+X_n}|N=n]\Pr(N=n) \\
=& \Pr(N=0) + \sum_{n=1}^{\infty} E\left[\Pi_{j=1}^{n} z^{X_j}\right]\Pr(N=n) = \sum_{n=1}^{\infty} \Pr(N=n)[P_X(z)]^n \\
=& P_N[P_X(z)].
\end{aligned}$$

The moment generating function of S is

$$M_S(z) = P_S(e^z) = P_N[P_X(e^z)] = P_N[M_X(z)].$$

From this, we find the following:

$$\begin{aligned}
E(S) =& M_S'(0) = P_N'[M_X(0)]M_X'(0) = P_N'(1)M_X'(0) = E(N)E(X) \\
E(S^2) =& M_S''(0) = P_N''[M_X(0)][M_X'(0)]^2 + P_N'[M_X(0)]M_X''(0) \\
=& P_N''(1)[M_X'(0)]^2 + P_N'(1)M_X''(0) \\
=& E[N(N-1)][E(X)]^2 + E(N)E(X^2) \\
\mathrm{Var}(S) =& E[N(N-1)][E(X)]^2 + E(N)E(X^2) - E(N)^2E(X)^2 \\
=& E(N)[E(X^2) - E(X)^2] + [E(N^2) - E(N)^2]E(X)^2 \\
=& E(N)\mathrm{Var}(X) + \mathrm{Var}(N)E(X)^2.
\end{aligned}$$

Example 38.5
Loss frequency N follows a Poisson distribution with parameter λ. Loss severity X follows an exponential distribution with mean θ. Find the expected value and the variance of the aggregate loss random variable.

Solution.
We have

$$E(S) = E(N)E(X) = \lambda\theta$$

and

$$E(S^2) = E[N(N-1)][E(X)]^2 + E(N)E(X^2) = \lambda^2\theta^2 + \lambda(2\theta^2).$$

Thus,

$$\text{Var}(S) = \lambda^2\theta^2 + \lambda(2\theta^2) - \lambda^2\theta^2 = 2\lambda\theta^2 \blacksquare$$

Example 38.6
Loss frequency N follows a Poisson distribution with parameter λ. Loss severity X follows an exponential distribution with mean θ. Find an expression of the probability that the aggregate loss random variable will exceed a value of n
(a) if S can be approximated with the standard normal distribution
(b) if S can be approximated with the lognormal distribution.

Solution.
(a) We have

$$\begin{aligned}
\Pr(S > n) =& \Pr\left(\frac{S - E(S)}{\sqrt{\text{Var}(S)}} > \frac{n - E(S)}{\sqrt{\text{Var}(S)}}\right)\\
=& \Pr\left(Z > \frac{n - E(S)}{\sqrt{\text{Var}(S)}}\right)\\
=& 1 - \Phi\left(\frac{n - E(S)}{\sqrt{\text{Var}(S)}}\right).
\end{aligned}$$

(b) If S is approximated with a lognormal distribution then we must have $E(S) = e^{\mu + \frac{1}{2}\sigma^2}$ and $E(S^2) = e^{2(\mu+\sigma^2)}$. Thus, μ and θ^2 are the solutions to the equations $\mu + \frac{1}{2}\sigma^2 = \ln(\lambda\theta)$ and $\mu + \theta^2 = \frac{\ln(\lambda^2\theta^2 + \lambda(2\theta^2))}{2}$. Hence,

$$\begin{aligned}
\Pr(S > n) =& \Pr\left(\frac{\ln S - \mu}{\sigma} > \frac{\ln n - \mu}{\sigma}\right)\\
=& \Pr\left(Z > \frac{\ln n - \mu}{\sigma}\right)\\
=& 1 - \Phi\left(\frac{\ln n - \mu}{\sigma}\right) \blacksquare
\end{aligned}$$

Example 38.7 ‡
You own a fancy light bulb factory. Your workforce is a bit clumsy they keep dropping boxes of light bulbs. The boxes have varying numbers of light bulbs in them, and when dropped, the entire box is destroyed.
You are given:
(i) Expected number of boxes dropped per month : 50
(ii) Variance of the number of boxes dropped per month: 100

(iii) Expected value per box: 200
(iv)Variance of the value per box: 400
You pay your employees a bonus if the value of light bulbs destroyed in a month is less than 8000.
Assuming independence and using the normal approximation, calculate the probability that you will pay your employees a bonus next month.

Solution.
Let S denote the total value of boxes destroyed in a month. Then

$$S = X_1 + X_2 + \cdots + X_N$$

where N is the number of boxes destroyed in a month and X_i is the value of the i^{th} box destroyed. S is a compound distribution so that

$$\begin{aligned} E(S) =& E(N)E(X) = 50(200) = 10,000 \\ Var(S) =& E(N)\text{Var}(X) + \text{Var}(N)E(X)^2 \\ =& 50(400) + 100(200)^2 = 4,020,000. \end{aligned}$$

Using normal approximation, we have

$$\Pr(S < 8,000) = \Pr\left(\frac{S - 10,000}{\sqrt{4,020,000}} < \frac{8,000 - 10,000}{\sqrt{4,020,000}}\right) = \quad P(Z < -0.9975) = 0.16 \blacksquare$$

Example 38.8 ‡
The number of claims, N, made on an insurance portfolio follows the following distribution:

n	$\Pr(N = n)$
0	0.7
2	0.2
3	0.1

If a claim occurs, the benefit is 0 or 10 with probability 0.8 and 0.2, respectively. The number of claims and the benefit for each claim are independent. Calculate the probability that aggregate benefits will exceed expected benefits by more than 2 standard deviations.

Solution.
We have

$$\begin{aligned}
E(N) =&0(0.7) + 2(0.2) + 3(0.1) = 0.7\\
E(N^2) =&0^2(0.7) + 2^2(0.2) + 3^2(0.1) = 1.7\\
\text{Var}(N) =&1.7 - 0.7^2 = 2\\
E(X) =&0(0.8) + 10(0.2) = 2\\
E(X^2) =&0^2(0.8) + 10^2(0.2) = 20\\
\text{Var}(X) =&20 - 2^2 = 16\\
E(S) =&E(N)E(X) = 0.7(0.2) = 1.4\\
\text{Var}(S) =&E(N)\text{Var}(X) + \text{Var}(N)E(X)^2\\
=&0.7(16) + 2(2)^2 = 16.04\\
STD =&\sqrt{16.04} = 4\\
\Pr(S > 1.4 + 2(4)) =&\Pr(S > 9.4)\\
=&1 - \Pr(S = 0).
\end{aligned}$$

But,

$$\begin{aligned}
\Pr(S = 0) =&\Pr[(S = 0) \cap (N = 0)] + \Pr[(S = 0) \cap (N = 2)] + \Pr[(S = 0) \cap (N = 3)]\\
=&\Pr(S = 0|N = 0)\Pr(N = 0) + \Pr(S = 0|N = 2)\Pr(N = 2)\\
&+\Pr(S = 0|N = 3)\Pr(N = 3)\\
=&(1)(0.7) + (0.8)^2(0.2) + (0.8)^3(0.1) = 0.8792.
\end{aligned}$$

Finally,

$$\Pr(S > 9.4) = 1 - 0.8792 = 0.1208 \blacksquare$$

Example 38.9 ‡
For a collective risk model the number of losses, N, has a Poisson distribution with $\lambda = 20$.
The common distribution of the individual losses has the following characteristics:
(i) $E(X) = 70$
(ii) $E(X \wedge 30) = 25$
(iii) $\Pr(X > 30) = 0.75$
(iv) $E(X^2|X > 30) = 9000$
An insurance covers aggregate losses subject to an ordinary deductible of 30 per loss.
Calculate the variance of the aggregate payments of the insurance.

Solution.
Let S denote the aggregate payments. Then S is a compound distribution with primary distribution N and secondary distribution the payment per loss $(X-30)_+$. We are asked to find

$$\text{Var}(S) = E(N)E[(X-30)_+^2] = 20E[(X-30)_+^2].$$

We have

$$\begin{aligned}
E[(X-30)_+^2] =& E[(X-30)_+^2|X>30](1-F_X(30)) = 0.75E[(X-30)_+^2|X>30]\\
=& 0.75E[(X-30)^2|X>30] = 0.75E(X^2-60X+900|X>30)\\
=& 0.75[E(X^2|X>30) - 60E(X|X>30)+900]\\
=& 0.75[E(X^2|X>30) - 60E(X-30|X>30) - 1800 + 900]\\
=& 0.75[E(X^2|X>30) - 60(E(X)-E(X\wedge 30))(1-F_X(30))^{-1} - 900]\\
=& 0.75[9000 - 60[70-25)(1-0.75)^{-1} - 900] = 3375\\
\text{Var}(S) =& 20(3375) = 67500 \blacksquare
\end{aligned}$$

Example 38.10 ‡
The repair costs for boats in a marina have the following characteristics:

Boat type	Number of boats	Probability that repair is needed	Mean of repair cost given a repair	Variance of repair cost given a repair
Power boats	100	0.3	300	10,000
Sailboats	300	0.1	1000	400,000
Luxury yachts	50	0.6	5000	2,000,000

At most one repair is required per boat each year. Repair costs are independent.
The marina budgets an amount, Y, equal to the aggregate mean repair costs plus the standard deviation of the aggregate repair costs.
Calculate Y.

Solution.
Let N be the number of boats that need repair. Then N is a binomial distribution with mean mq and variance $mq(1-q)$. Let X denote the repair cost. Let S be the aggregate cost of a type of boats. Then

$$\begin{aligned}
E(S) =& E(N)E(X)\\
\text{Var}(S) =& E(N)\text{Var}(X) + \text{Var}(N)E(X)^2.
\end{aligned}$$

We have

Boat type	$E(N)$	Var(N)	$E(X)$	Var(X)	$E(S)$	Var(S)
Power boats	30	21	300	10,000	9,000	2,190,000
Sailboats	30	27	1000	400,000	30,000	39,000,000
Luxury yachts	30	12	5000	2,000,000	150,000	360,000,000

Thus,

$$E(Y) = 9,000 + 30,000 + 150,000 + \sqrt{(2.19 + 39 + 360) \times 10^6} \approx 209,000 \blacksquare$$

Example 38.11 ‡
For an insurance:
(i) The number of losses per year has a Poisson distribution with $\lambda = 10$.
(ii) Loss amounts are uniformly distributed on $(0, 10)$.
(iii) Loss amounts and the number of losses are mutually independent.
(iv) There is an ordinary deductible of 4 per loss.
Calculate the variance of aggregate payments in a year.

Solution.
Let S be the aggregate claims. Then S is a compound distribution with primary distribution N, the number of claims, and secondary distribution $Y = (X - 4)_+$, the amount paid per loss, where X is the loss amount. We want

$$\text{Var}(S) = E(N)\text{Var}(Y) + \text{Var}(N)E(Y)^2 = \lambda E(Y^2)$$

where

$$E(Y^2) = \int_4^{10} (x-4)^2 f_X(x)dx = \int_4^{10} (x-4)^2(0.1)dx = 7.2.$$

Hence,

$$\text{Var}(S) = 10(7.2) = 72 \blacksquare$$

Example 38.12 ‡
For an insurance portfolio:
(i) The number of claims has the probability distribution

n	p_n
0	0.1
1	0.4
2	0.3
3	0.2

(ii) Each claim amount has a Poisson distribution with mean 3.
(iii) The number of claims and claim amounts are mutually independent.
Calculate the variance of aggregate claims.

Solution.
Let S be the aggreagte claims, N the number of claims, and X the amount of claim. Then S is a compound distribution with primary distribution N and a secondary distribution X. Thus,

$$\text{Var}(S) = E(N)\text{var}(X) + \text{Var}(N)E(X)^2$$

where

$$\begin{aligned}
E(X) =& \lambda = 3 \\
\text{Var}(X) =& \lambda = 3 \\
E(N) =& 0(0.1) + 1(0.4) + 2(0.3) + 3(0.2) = 1.6 \\
E(N^2) =& 0^2(0.1) + 1^2(0.4) + 2^2(0.3) + 3^2(0.2) = 3.4 \\
\text{Var}(N) =& 3.4 - 1.6^2 = 0.84.
\end{aligned}$$

Hence,

$$\text{Var}(S) = 1.6(3) + 0.84(3^2) = 12.36 \blacksquare$$

Example 38.13 ‡
Aggregate losses are modeled as follows:
(i) The number of losses has a Poisson distribution with $\lambda = 3$.
(ii) The amount of each loss has a Burr (Burr Type XII, Singh-Maddala) distribution with $\alpha = 3, \theta = 2$, and $\gamma = 1$.
(iii) The number of losses and the amounts of the losses are mutually independent.
Calculate the variance of aggregate losses.

Solution.
Let N denote the number of losses and X be the severity random variable. Then

$$\text{Var}(S) = E(N)\text{Var}(X) + \text{Var}(N)E(X)^2 = \lambda E(X^2).$$

From Table C, we find

$$E(X^2) = \frac{2^2\Gamma(3)\Gamma(1)}{\Gamma(3)} = 4.$$

Hence,

$$\text{Var}(S) = 3(4) = 12 \blacksquare$$

Example 38.14 ‡
You are the producer for the television show Actuarial Idol. Each year, 1000 actuarial clubs audition for the show. The probability of a club being accepted is 0.20.
The number of members of an accepted club has a distribution with mean 20 and variance 20. Club acceptances and the numbers of club members are mutually independent.
Your annual budget for persons appearing on the show equals 10 times the expected number of persons plus 10 times the standard deviation of the number of persons.
Calculate your annual budget for persons appearing on the show.

Solution.
Let S be the number of people appearing on the show. Then S is a compound distribution with frequency N (number of clubs being accepted) and severity X (number of members of an accepted club).
The frequency N has a binomial distribution with parameters $m = 1000$ and $q = 0.20$. Thus,

$$\begin{aligned}E(N) =&mq = 1000(0.20) = 200\\ \mathrm{Var}(N) =&mq(1-q)200(1-0.2) = 160.\end{aligned}$$

The severity X has mean 20 and variance 20. Hence,

$$\begin{aligned}E(S) =&E(N)E(X) = 200(20) = 4000\\ \mathrm{Var}(S) =&E(N)\mathrm{Var}(X) + \mathrm{Var}(N)E(X)^2\\ =&200(20) + 160(20)^2 = 68000.\end{aligned}$$

The annual budget for persons appearing on the show is

$$10(4000) + 10\sqrt{68000} \approx 42610 \blacksquare$$

Example 38.15 ‡
For an aggregate loss distribution S :
(i) The number of claims has a negative binomial distribution with $r = 16$ and $\beta = 6$.
(ii) The claim amounts are uniformly distributed on the interval $(0, 8)$.
(iii) The number of claims and claim amounts are mutually independent.
Using the normal approximation for aggregate losses, calculate the premium such that the probability that aggregate losses will exceed the premium is 5%.

Solution.
The aggregate losses S has a compound distribution with primary function the frequency N and secondary function the severity X. We have

$$\begin{aligned}
E(N) =& r\beta = 16(6) = 96\\
\text{Var}(N) =& r\beta(1+\beta) = 16(6)(7) = 672\\
E(X) =& \frac{0+8}{2} = 4\\
\text{Var}(X) =& \frac{(8-0)^2}{12} = 5.333\\
E(S) =& E(N)E(X) = 96(4) = 384\\
\text{Var}(S) =& E(N)\text{Var}(X) + \text{Var}(N)E(X)^2\\
=& 96(5.333) + 672(4^2) = 11264.
\end{aligned}$$

Let P denote the premium such that $\Pr(S > P) = 0.05$. Using normal approximation, we have

$$\begin{aligned}
0.05 =& \Pr(S > P) = \Pr\left(\frac{S-384}{\sqrt{11264}} > \frac{P-384}{\sqrt{11264}}\right)\\
=& \Pr\left(Z > \frac{P-384}{\sqrt{11264}}\right)\\
=& 1 - \Pr\left(Z \le \frac{P-384}{\sqrt{11264}}\right)\\
=& 1 - \Phi\left(\frac{P-384}{\sqrt{11264}}\right).
\end{aligned}$$

Hence,

$$\Phi\left(\frac{P-384}{\sqrt{11264}}\right) = 0.95 \Longrightarrow \frac{P-384}{\sqrt{11264}} = 1.645.$$

Solving this last equation, we find $P = 558.59$ ■

Practice Problems

Problem 38.1
Find the third raw moment of the aggregate loss random variable $S : E[S^3]$.

Problem 38.2
Find the third central moment of the aggregate loss random variable $S : E[(S - E(S))^3]$.

Problem 38.3
Let $S = X_1 + X_2 + \cdots X_N$ where the $X_i's$ are independent with common distribution the lognormal distribution with parameters μ and σ. The frequency random variable N has a Poisson distribution with parameter λ.

Write the integral representing $F_X^{*2}(x)$.

Problem 38.4
Find the first and the second raw moments of S in Example 38.1.

Problem 38.5
Find the variance of S in Example 38.1.

Problem 38.6 ‡
For an insured, Y is the random variable representing the total amount of time spent in a hospital each year. The distribution of the number of hospital admissions in a year is:

# of Admissions	Probability
0	0.60
1	0.30
2	0.10

The distribution of the length of each stay for an admission follows a Gamma distribution with $\alpha = 1$ and $\theta = 5$.

Determine $E(Y)$ and $\mathrm{Var}(Y)$.

Problem 38.7
For an insurance company, each loss has a mean of 100 and a variance of 100. The number of losses follows a Poisson distribution with a mean of 500. Each loss and the number of losses are mutually independent.

Using the normal approximation, calculate the probability that the aggregate losses will exceed 52,250.

Problem 38.8
For an insurance company, each loss has a mean of 100 and a variance of 100. The number of losses follows a Poisson distribution with a mean of 500. Each loss and the number of losses are mutually independent.

Using the lognormal approximation, calculate the probability that the aggregate losses will exceed 52,250.

Problem 38.9
An insurance company offers car insurance to a group of 1000 employees. The frequency claim has negative binomial distribution with $r = 1$ and $\beta = 1.5$. Severity claims are exponentially distributed with a mean of 5000. Assume that the number of claims and the size of the claim are independent and identically distributed.

Using the normal distribution as an approximating distribution of aggregate losses, calculate the probability that losses will exceed 8 million.

Problem 38.10 ‡
When an individual is admitted to the hospital, the hospital charges have the following characteristics:

Charges	Mean	Std deviation
Room	1000	500
Other	500	300

The covariance between an individual's room charges and other charges is 100,000. An insurer issues apolicy that reimburses 100% for room charges and 80% for other charges. The number of hospital admissions has Poisson distribution with parameter 4.

Determine the mean and the standard deviation of the insurer's payout for the policy.

Problem 38.11 ‡
Computer maintenance costs for a department are modeled as follows:
(i) The distribution of the number of maintenance calls each machine will need in a year is Poisson with mean 3.
(ii) The cost for a maintenance call has mean 80 and standard deviation 200.
(iii) The number of maintenance calls and the costs of the maintenance calls

are all mutually independent.
The department must buy a maintenance contract to cover repairs if there is at least a 10% probability that aggregate maintenance costs in a given year will exceed 120% of the expected costs.

Using the normal approximation for the distribution of the aggregate maintenance costs, calculate the minimum number of computers needed to avoid purchasing a maintenance contract.

Problem 38.12 ‡
A towing company provides all towing services to members of the City Automobile Club. You are given:

Towing Distance (in miles)	Towing Cost	Frequency
0-9.99	80	50%
10-29.99	100	40%
30^+	160	10%

(i) The automobile owner must pay 10% of the cost and the remainder is paid by the City Automobile Club.
(ii) The number of towings has a Poisson distribution with mean of 1000 per year.
(iii) The number of towings and the costs of individual towings are all mutually independent.

Using the normal approximation for the distribution of aggregate towing costs, calculate the probability that the City Automobile Club pays more than 90,000 in any given year.

Problem 38.13 ‡
The number of auto vandalism claims reported per month at Sunny Daze Insurance Company (SDIC) has mean 110 and variance 750. Individual losses have mean 1101 and standard deviation 70. The number of claims and the amounts of individual losses are independent.

Using the normal approximation, calculate the probability that SDIC's aggregate auto vandalism losses reported for a month will be less than 100,000.

Problem 38.14 ‡
At the beginning of each round of a game of chance the player pays 12.5.

The player then rolls one die with outcome N. The player then rolls N dice and wins an amount equal to the total of the numbers showing on the N dice. All dice have 6 sides and are fair.

Using the normal approximation, calculate the probability that a player starting with 15,000 will have at least 15,000 after 1000 rounds. Caution: Use continuity correction since a discrete distribution is being approximated by a continuous one.

Problem 38.15 ‡
A dam is proposed for a river which is currently used for salmon breeding. You have modeled:
(i) For each hour the dam is opened the number of salmon that will pass through and reach the breeding grounds has a distribution with mean 100 and variance 900.
(ii) The number of eggs released by each salmon has a distribution with mean of 5 and variance of 5.
(iii) The number of salmon going through the dam each hour it is open and the numbers of eggs released by the salmon are independent.

Using the normal approximation for the aggregate number of eggs released, determine the least number of whole hours the dam should be left open so the probability that 10,000 eggs will be released is greater than 95%.

Problem 38.16 ‡
You are the producer of a television quiz show that gives cash prizes. The number of prizes, N, and prize amounts, X, have the following distributions:

n	$\Pr(N=n)$
1	0.8
2	0.2

x	$\Pr(X=x)$
0	0.2
100	0.7
1000	0.1

Your budget for prizes equals the expected prizes plus the standard deviation of prizes.

Calculate your budget.

Problem 38.17 ‡
The number of accidents follows a Poisson distribution with mean 12. Each

accident generates 1, 2, or 3 claimants with probabilities $\frac{1}{2}, \frac{1}{3}, \frac{1}{6}$, respectively.

Calculate the variance of the total number of claimants.

Problem 38.18 ‡
In a clinic, physicians volunteer their time on a daily basis to provide care to those who are not eligible to obtain care otherwise. The number of physicians who volunteer in any day is uniformly distributed on the integers 1 through 5. The number of patients that can be served by a given physician has a Poisson distribution with mean 30.

Determine the probability that 120 or more patients can be served in a day at the clinic, using the normal approximation with continuity correction.

Problem 38.19 ‡
For an individual over 65:
(i) The number of pharmacy claims is a Poisson random variable with mean 25.
(ii) The amount of each pharmacy claim is uniformly distributed between 5 and 95.
(iii) The amounts of the claims and the number of claims are mutually independent.

Determine the probability that aggregate claims for this individual will exceed 2000 using the normal approximation.

Problem 38.20 ‡
Two types of insurance claims are made to an insurance company. For each type, the number of claims follows a Poisson distribution and the amount of each claim is uniformly distributed as follows:

Type of Claim	Poisson Parameter λ for Number of Claims	Range of Each Claim Amount
I	12	$(0,1)$
II	4	$(0,5)$

The numbers of claims of the two types are independent and the claim amounts and claim numbers are independent.

Calculate the normal approximation to the probability that the total of claim amounts exceeds 18.

Problem 38.21 ‡

For aggregate losses, S :

(i) The number of losses has a negative binomial distribution with mean 3 and variance 3.6.

(ii) The common distribution of the independent individual loss amounts is uniform from 0 to 20.

Calculate the 95^{th} percentile of the distribution of S as approximated by the normal distribution.

Problem 38.22 ‡

In a CCRC, residents start each month in one of the following three states: Independent Living (State #1), Temporarily in a Health Center (State #2) or Permanently in a Health Center (State #3). Transitions between states occur at the end of the month.

If a resident receives physical therapy, the number of sessions that the resident receives in a month has a geometric distribution with a mean which depends on the state in which the resident begins the month. The numbers of sessions received are independent. The number in each state at the beginning of a given month, the probability of needing physical therapy in the month, and the mean number of sessions received for residents receiving therapy are displayed in the following table:

State #	Number in state	Probability of needing therapy	Mean number of visits
1	400	0.2	2
2	300	0.5	15
3	200	0.3	9

Using the normal approximation for the aggregate distribution, calculate the probability that more than 3000 physical therapy sessions will be required for the given month.

Problem 38.23 ‡

You are given:

(i) Aggregate losses follow a compound model.

(ii) The claim count random variable has mean 100 and standard deviation 25.

(iii) The single-loss random variable has mean 20,000 and standard deviation 5000.

Determine the normal approximation to the probability that aggregate claims exceed 150% of expected costs.

39 Stop Loss Insurance

When a deductible d is applied to the aggregate loss S over a definite period, then the insurance payment will be

$$(S-d)_+ = \max\{S-d, 0\} = S - S\wedge d = \begin{cases} 0, & S \le d \\ S-d, & S > d. \end{cases}$$

We refer to this payment as the **stop-loss insurance**. Its expected cost is called the **net stop-loss premium** and can be computed as

$$E[(S-d)_+] = \int_d^\infty S(x)dx = \int_d^\infty [1-F_S(x)]dx = \int_d^\infty (x-d)f_S(x)dx$$

in the continuous case and

$$E[(S-d)_+] = \sum_{x>d}(x-d)f_S(x)$$

in the discrete case. Note that this is identical to the discussion of ordinary deductible of Section 31.

Example 39.1
The distribution of aggregate losses covered under a policy of stop-loss insurance is given by $F_S(x) = 1 - \frac{1}{x^2},\ x > 1$. Calculate $E[(S-3)_+]$.

Solution.
We have

$$E[(S-3)_+] = \int_3^\infty \frac{dx}{x^2} = -\frac{1}{x}\Big|_3^\infty = \frac{1}{3} \blacksquare$$

The following provides a simple calculation of the net stop-loss premium in a special case.

Theorem 39.1
Suppose that $\Pr(a < S < b) = 0$. Then, for $a \le d \le b$, we have

$$E[(S-d)_+] = \frac{b-d}{b-a}E[(S-a)_+] + \frac{d-a}{b-a}E[(S-b)_+].$$

Proof.
Let $a \le x \le b$. Since $a \le x$ and $F_S(x)$ is nondecreasing, we have $F_S(a) \le F_S(x)$. On the other hand,

$$F_S(x) = \int_0^x f_S(y)dy = \int_0^a f_S(y)dy + \int_a^x f_S(y)dy \le F_S(a) + \int_a^b f_S(y)dy = F_S(a).$$

Hence, $F_S(x) = F_S(a)$ for all $a \le x \le b$. Next, we have

$$\begin{aligned}
E[(S-d)_+] =& \int_d^\infty [1-F_S(x)]dx\\
=& \int_a^\infty [1-F_S(x)]dx - \int_a^d [1-F_S(x)]dx\\
=&E[(S-a)_+] - \int_a^d [1-F_S(a)]dx\\
=&E[(S-a)_+] - (d-a)[1-F_S(a)].
\end{aligned}$$

The above is true for all $a \le d \le b$. In particular, if $d = b$, we have

$$E[(S-b)_+] = E[(S-a)_+] - (b-a)[1-F_S(a)] \Longrightarrow 1-F_S(a) = \frac{E[(S-a)_+ - E[(S-b)_+]}{b-a}.$$

Hence,

$$\begin{aligned}
E[(S-d)_+] =&E[(S-a)_+] - \frac{d-a}{b-a}\left\{E[(S-a)_+ - E[(S-b)_+]\right\}\\
=&\frac{b-d}{b-a}E[(S-a)_+] + \frac{d-a}{b-a}E[(S-b)_+] \blacksquare
\end{aligned}$$

Example 39.2
A reinsurer pays aggregate claim amounts in excess of d, and in return it receives a stop-loss premium $E[(S-d)_+]$. You are given $E[(S-100)_+] = 15$, $E[(S-120)_+] = 10$, and the probability that the aggregate claim amounts are greater than 100 and less than to 120 is 0. Calculate $E[(S-105)_+]$.

Solution.
We have

$$E[(S-105)_+] = \left(\frac{120-105}{120-100}\right)(15) + \left(\frac{105-100}{120-100}\right)(10) = 13.75 \blacksquare$$

More simplification result follows.

Theorem 39.2
Suppose S is discrete and $\Pr(S = kh) \ge 0$ for some fixed h and $k = 0, 1, 2, \cdots$. Also, $\Pr(S = x) = 0$ for all $x \ne kh$. Then, for any nonnegative integer j, we have

$$E[(S-jh)_+] = h\sum_{n=0}^\infty \{1 - F_S[(n+1)j]\}.$$

In particular,

$$E[(S-(j+1)h)_+]-E[(S-jh)_+]=h[F_S(jh)-1].$$

Proof.
We have

$$\begin{aligned}E[(S-jh)_+]=&\sum_{x>jh}(x-jh)f_S(x)\\=&\sum_{k=j}^{\infty}(kh-jh)\Pr(S=kh)=h\sum_{k=j}^{\infty}(k-j)\Pr(S=kh)\\=&h\sum_{k=j}^{\infty}\sum_{n=0}^{k-j-1}\Pr(S=kh)=h\sum_{n=0}^{\infty}\sum_{k=n+j+1}^{k-j-1}\Pr(S=kh)\\=&h\sum_{n=0}^{\infty}\left\{1-F_S[(n+1)j]\right\}.\end{aligned}$$

Finally, we have

$$\begin{aligned}E[(S-(j+1)h)_+]-E[(S-jh)_+]=&h\sum_{n=0}^{\infty}\left\{1-F_S[(n+1)j+n+1]\right\}\\&-h\sum_{n=0}^{\infty}\left\{1-F_S[(n+1)j]\right\}\\=&h\sum_{n=1}^{\infty}\left\{1-F_S[(n+1)j]\right\}\\&-h\sum_{n=0}^{\infty}\left\{1-F_S[(n+1)j]\right\}\\=&h[F_S(jh)-1]\ \blacksquare\end{aligned}$$

Example 39.3
Given the following information about the distribution of a discrete aggregate loss random variable:

x	0	25	50	75
$F_S(x)$	0.05	0.065	0.08838	0.12306

Calculate $E[(S-25)_+], E[(S-50)_+], E[(S-75)_+]$, and $E[(S-100)_+]$ given that $E(S)=314.50$.

Solution.
We have

$$\begin{aligned}E[(S-25)_+] =&314.50-25(1-0.05)=290.75\\E[(S-50)_+] =&290.75-25(1-0.065)=267.375\\E[(S-75)_+] =&267.375-25(1-0.08838)=244.5845\\E[(S-100)_+] =&244.5845-25(1-0.12306)=222.661 \blacksquare\end{aligned}$$

Example 39.4 ‡
Prescription drug losses, S, are modeled assuming the number of claims has a geometric distribution with mean 4, and the amount of each prescription is 40. Calculate $E[(S-100)_+]$.

Solution.
Let N denote the number of prescriptions and S the aggregate losses. Then $S=40N$. We have

$$\begin{aligned}E[(S-100)_+] =&E(S)-E(S\wedge 100)\\=&40E(N)-40E(N\wedge 2.5)\\=&40E(N)-40[f_N(1)+2f_N(2)+2.5(1-F_N(2))]\\=&40(4)-40[0.16+2(0.1280)+2.5(0.5120)]\\=&92.16\end{aligned}$$

where

$$f_N(n)=\frac{4^n}{5^{n+1}} \blacksquare$$

Example 39.5 ‡
WidgetsRUs owns two factories. It buys insurance to protect itself against major repair costs. Profit equals revenues, less the sum of insurance premiums, retained major repair costs, and all other expenses. WidgetsRUs will pay a dividend equal to the profit, if it is positive.
You are given:
(i) Combined revenue for the two factories is 3.
(ii) Major repair costs at the factories are independent.
(iii) The distribution of major repair costs (k) for each factory is

k	Probability(k)
0	0.4
1	0.3
2	0.2
3	0.1

(iv) At each factory, the insurance policy pays the major repair costs in excess of that factorys ordinary deductible of 1. The insurance premium is 110% of the expected claims.
(v) All other expenses are 15% of revenues.
Calculate the expected dividend.

Solution.
Let R denote the retained major repair cost for both factories. That is, this is the portion of the major repair costs to be covered by the ownership of the factory. Note that R is an aggregate random variable and $R = 0, 1, 2$. Let D stand for the dividend. The expected claims per factory for the insurance company is

$$E[(X-1)_+] = (2-1)(0.2) + (3-1)(0.1) = 0.4.$$

Thus, the insurance company charges WidgetRUs a premium of $(1.10)(2)(0.4) = 0.88$. Hence, the formula of the dividend can be expressed as

$$D = 3 - 0.88 - R - 0.15(3) = 1.67 - R.$$

The distribution of R is found as follows:
For $R = 0$, both factories have no major repair costs and in this case $\Pr(R = 0) = (0.4)(0.4) = 0.16$. For $R = 1$, one factory has a major repair cost and the other does not. In this case, $\Pr(R = 1) = 2(0.4)(0.6) = 0.48$. Hence, $\Pr(R = 2) = 1 - \Pr(R = 0) - \Pr(R = 1) = 0.36$. The distribution of D is

D	Probability(D)
1.67	0.16
0.67	0.48
0	0.36

Finally,
$$E(D) = 1.67(0.16) + 0.67(0.48) = 0.5888 \blacksquare$$

Example 39.6 ‡
For a collective risk model:
(i) The number of losses has a Poisson distribution with $\lambda = 2$.
(ii) The common distribution of the individual losses is:

x	$f(x)$
1	0.6
2	0.4

An insurance covers aggregate losses subject to a deductible of 3. Calculate the expected aggregate payments of the insurance.

Solution.
We have

$$\begin{aligned}
E[(S-3)_+] =& E[(S-2)_+] + F_S(2) - 1 \\
=& E[(S-1)_+] + F_S(2) + F_S(1) - 2 \\
=& E(S) + F_S(0) + F_S(1) + F_S(2) - 3 \\
=& E(S) + 3f_S(0) + 2f_S(1) + f_S(2) - 3 \\
=& 2(1.4) + 3e^{-2} + 2e^{-2}(0.6) + 2e^{-2}(0.4) + \frac{e^{-2}2^2}{2!}(0.6)^2 - 3 \\
=& 0.7365 \blacksquare
\end{aligned}$$

Example 39.7 ‡
A compound Poisson claim distribution has $\lambda = 5$ and individual claim amounts distributed as follows:

x	$f(x)$
5	0.6
k	0.4

where $k > 5$. The expected cost of an aggregate stop-loss insurance subject to a deductible of 5 is 28.03.
Calculate k.

Solution.
The stop-loss insurance with deductible 5 pays

$$(S-5)_+ = S - S \wedge 5.$$

Thus,

$$E[(S-5)_+] = E(S) - E(S \wedge 5).$$

We have

$$E(S) = E(N)E(X) = 5[5(0.6) + k(0.4)] = 15 + 2k$$

and

$$\begin{aligned}
E(S \wedge 5) =& 5[1 - F_S(5)] \\
=& 5[1 - \Pr(S = 0)] \\
=& 5[1 - \Pr(N = 0)] \\
=& 5(1 - e^{-5}) = 4.9663.
\end{aligned}$$

Thus,

$$28.03 = E[(S-5)_+] = 15 + 2k - 4.9663 \Longrightarrow k = 9 \blacksquare$$

Practice Problems

Problem 39.1
You are given: $E[(S-15)_+ = 0.34$ and $E[(S-30)_+] = 0.55$. Calculate $F_S(15)$.

Problem 39.2 ‡
An aggregate claim distribution has the following characteristics: $\Pr(S = i) = \frac{1}{6}$ for $i = 1, 2, 3, 4, 5, 6$. A stop-loss insurance with deductible amount d has an expected insurance payment of 1.5 .

Find d.

Problem 39.3 ‡
You are given:
(i) S takes only positive integer values.
(ii) $E(S) = \frac{5}{3}$.
(iii) $E[(S-2)_+] = \frac{1}{6}$.
(iv) $E[(S-3)_+] = 0$.

Find the value of $f_S(1)$.

Problem 39.4 ‡
For a stop-loss insurance on a three person group:
(i) Loss amounts are independent.
(ii) The distribution of loss amount for each person is:

Loss Amount	Probability
0	0.4
1	0.3
2	0.2
3	0.1

iii) The stop-loss insurance has a deductible of 1 for the group.

Calculate the net stop-loss premium.

Problem 39.5
Suppose that the aggregate loss random variable is discrete satisfying $\Pr(S = 50k) = \frac{1}{2^{k+1}}$ for $k = 0, 1, 2, \cdots$ and $\Pr(S = x) = 0$ for all other x.

Calculate $E[(S-150)_+]$.

Problem 39.6 ‡

For a certain company, losses follow a Poisson frequency distribution with mean 2 per year, and the amount of a loss is 1, 2, or 3, each with probability 1/3. Loss amounts are independent of the number of losses, and of each other.

An insurance policy covers all losses in a year, subject to an annual aggregate deductible of 2.

Calculate the expected claim payments for this insurance policy.

Problem 39.7 ‡

For a stop-loss insurance on a three person group:

(i) Loss amounts are independent.

(ii) The distribution of loss amount for each person is:

Loss Amount (X)	Probability (X)
0	0.4
1	0.3
2	0.2
3	0.1

(iii) The stop-loss insurance has a deductible of 1 for the group.

Calculate the net stop-loss premium.

Problem 39.8 ‡

The number of annual losses has a Poisson distribution with a mean of 5. The size of each loss has a two-parameter Pareto distribution with $\theta = 10$ and $\alpha = 2.5$. An insurance for the losses has an ordinary deductible of 5 per loss.

Calculate the expected value of the aggregate annual payments for this insurance.

Problem 39.9 ‡

In a given week, the number of projects that require you to work overtime has a geometric distribution with $\beta = 2$. For each project, the distribution of the number of overtime hours in the week is the following:

x	$f(x)$
5	0.2
10	0.3
20	0.5

The number of projects and number of overtime hours are independent. You will get paid for overtime hours in excess of 15 hours in the week.

Calculate the expected number of overtime hours for which you will get paid in the week.

40 Closed Form of Aggregate Distributions

Finding the distribution of an aggregate loss random variable via the method of convolutions is practically very difficult. Mostly, this is done through numerical methods which we will discuss in the coming sections. However, it is possible to find a closed form for the distribution of some special families of severities which is the topic of this section.

Let S be an aggregate loss random variable such that the severities are all independent and identically distributed with common distribution the exponential distribution with mean θ. We assume that the severities and the frequency distributions are independent. By independence, we have

$$M_S(z) = [M_X(z)]^n = (1-\theta z)^{-n}.$$

Thus, S is equivalent to a Gamma distribution with parameters θ and n and with cdf

$$\Pr(X_1 + X_2 + \cdots + X_N \le x) = F_X^{*n}(x) = \Gamma\left(n; \frac{x}{\theta}\right)$$

where $\Gamma(\alpha; x)$ is the **incomplete Gamma** function defined by

$$\Gamma(\alpha; x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt, \ x > 0, \ \alpha > 0$$

and where

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

If $\alpha = n$ is a positive integer then we can find a closed form of $\Gamma(n : x)$.

Theorem 40.1
Let n be a positive integer. Then

$$\Gamma(n; x) = 1 - \sum_{j=0}^{n-1} \frac{x^j e^{-x}}{j!}.$$

Proof.
The proof is by induction on n. For $n = 1$, we have

$$\Gamma(1; x) = \int_0^x e^{-t} dt = 1 - e^{-x} = 1 - \sum_{j=0}^{1-1} \frac{x^j e^{-x}}{j!}.$$

So suppose the formula is true for $1, 2, \cdots, n$. Using integration by parts, we have

$$\begin{aligned}
\Gamma(n+1;x) =& \frac{1}{\Gamma(n+1)}\int_0^x t^n e^{-t}dt \\
=& \frac{1}{n!}\left(-t^n e^{-t}\Big|_0^x + \int_0^x nt^{n-1}e^{-t}dt\right) \\
=& \frac{1}{n!}(-x^n e^{-x}) + \Gamma(n;x) \\
=& \frac{-x^n e^{-x}}{n!} + 1 - \sum_{j=0}^{n-1}\frac{x^j e^{-x}}{j!} \\
=& 1 - \sum_{j=0}^{n}\frac{x^j e^{-x}}{j!} \blacksquare
\end{aligned}$$

Now, from Section 38, we have

$$\begin{aligned}
F_S(x) =& \sum_{n=0}^{\infty}\Pr(N=n)F_X^{*n}(x) = \Pr(N=0) + \sum_{n=1}^{\infty}\Pr(N=n)\Gamma\left(n;\frac{x}{\theta}\right) \\
=& \Pr(N=0) + \sum_{n=1}^{\infty}\Pr(N=n)\left[1 - \sum_{j=0}^{n-1}\frac{(x/\theta)^j e^{-x/\theta}}{j!}\right] \\
=& 1 - \sum_{n=1}^{\infty}\Pr(N=n)\sum_{j=0}^{n-1}\frac{(x/\theta)^j e^{-x/\theta}}{j!} \\
=& 1 - e^{-\frac{x}{\theta}}\sum_{j=0}^{\infty}\frac{(x/\theta)^j}{j!}\sum_{n=j+1}^{\infty}\Pr(N=n) \\
=& 1 - e^{-\frac{x}{\theta}}\sum_{j=0}^{\infty}\overline{P}_j\frac{(x/\theta)^j}{j!}, \ x \ge 0
\end{aligned}$$

where

$$\overline{P}_j = \sum_{n=j+1}^{\infty}\Pr(N=n), \ j = 0, 1, \cdots.$$

The pdf of S can be found upon differentiating $F_S(x)$.

Example 40.1
Find $F_S(x)$ if the frequency N has a binomial distribution with parameters n and m.

Solution.
We have $\overline{P}_j = 0$ for $j = m, m+1, \cdots$. Hence,

$$F_S(x) = 1 - \sum_{n=1}^{m} \binom{m}{n} q^n (1-q)^{m-n} \sum_{j=0}^{n-1} \frac{(x/\theta)^j}{j!} \blacksquare$$

When the subscripted independent random variables $X_i's$ belong to a same family of distributions and the resulting sum S belongs to the same family, we say that the family is **closed under convolution.**

Example 40.2
Suppose that $X_1, X_2, \cdots, X_N$ are independent normal random variables with parameters $(\mu_1, \sigma_1^2), (\mu_2, \sigma_2^2), \cdots, (\mu_N, \sigma_N^2)$. Show that $S = X_1 + X_2 + \cdots + X_N$ is also a normal random variable.

Solution.
By independence, we have

$$\begin{aligned} M_S(z) =& M_{X_1}(z) M_{X_2}(z) \cdots M_{X_N}(z) \\ =& e^{\mu_1 + \sigma_1^2 t} e^{\mu_2 + \sigma_2^2 t} \cdots e^{\mu_N + \sigma_N^2 t} \\ =& e^{(\mu_1 + \mu_2 + \cdots + \mu_N) + (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_N^2) t}. \end{aligned}$$

Hence, S has a normal distribution with parameters $(\mu_1 + \mu_2 + \cdots + \mu_N, \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_N^2)$ $\blacksquare$

A **compound Poisson distribution** is an aggregate distribution S where the frequency random variable is Poisson.

Theorem 40.2
Suppose that $S_1, S_2, \cdots, S_n$ are compound Poisson distributions with parameters $\lambda_1, \lambda_2, \cdots, \lambda_n$ and severity distributions with cdf's $F_1(x), F_2(x), \cdots, F_n(x)$. Suppose that the $S_i's$ are independent. Then $S = S_1 + \cdots + S_n$ has a compound Poisson distribution with parameter $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ and severity distribution with cdf

$$F(x) = \sum_{j=1}^{n} \frac{\lambda_i}{\lambda} F_j(x).$$

Proof.
Let $M_j(t)$ be the mgf of $F_j(x)$, Then

$$M_{S_j}(t) = P_{S_j}[M_j(t)] = e^{\lambda_j [M_j(t) - 1]}.$$

Hence, by independence, we have

$$M_S(t) = \prod_{j=1}^{n} M_{S_j}(t) = \prod_{j=1}^{n} e^{\lambda_j[M_j(t)-1]}$$
$$= e^{(\lambda_1 M_1(t)+\lambda_2 M_2(t)+\cdots+\lambda_n M_n(t))-\lambda}$$
$$= e^{\lambda\left(\frac{\lambda_1}{\lambda}M_1(t)+\frac{\lambda_2}{\lambda}M_2(t)+\cdots+\frac{\lambda_n}{\lambda}M_n(t)\right)-\lambda}.$$

Thus, S is a compound Poisson distribution with parameter $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ and a severity distribution with cdf

$$F(x) = \sum_{j=1}^{n} \frac{\lambda_i}{\lambda} F_j(x) \blacksquare$$

Example 40.3
Let S_1, S_2, S_3, and S_4 be independent compound Poisson with parameters $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 4$ and severities with cdf $F_j(x) = x^2$ for $0 \le x \le 1$. Let $S = S_1 + S_2 + S_3 + S_4$. Find the mgf of S.

Solution.
The moment generating function of $F_j(x)$ is

$$M(t) = M_j(t) = \int_0^1 e^{tx}(2x)dx = \frac{e^t}{t} - \frac{e^t}{t^2} + \frac{1}{t^2}.$$

Thus,

$$\lambda\left\{\left[\sum_{j=1}^{4} \frac{\lambda_j}{\lambda} M_j(t)\right] - 1\right\} = 10[M(t)-1].$$

Hence,

$$M_S(t) = e^{10[M(t)-1]} = e^{10\left(\frac{e^t}{t} - \frac{e^t}{t^2} + \frac{1}{t^2} - 1\right)} \blacksquare$$

Practice Problems

Problem 40.1
Suppose that the frequency distribution N of S has a negative binomial with integer value r and parameter β. Suppose also that the severities are identically distributed with common distribution the exponential distribution with mean θ. Suppose that the severities are independent among each other and with N.

Find the moment generating function. Show that the given model is equivalent to the binomial-exponential model.

Problem 40.2
Show that the cdf of the model of Exercise 40.1 is given by

$$F_S(x) = 1 - \sum_{n=1}^{r} \binom{r}{n} \left(\frac{\beta}{1+\beta}\right)^n \left(\frac{1}{1+\beta}\right)^{r-n} \times P_n(x)$$

where

$$P_n(x) = \sum_{j=0}^{n-1} \frac{[x\theta^{-1}(1+\beta)^{-1}]^j e^{-x\theta^{-1}(1+\beta)^{-1}}}{j!}.$$

Problem 40.3
Suppose N has a geometric distribution with parameter β. The severities are all exponential with mean θ. The severities and frequency are independent.

Find a closed form of $F_S(x)$.

Problem 40.4
Suppose N has a geometric distribution with parameter β. The severities are all exponential with mean θ. The severities and frequency are independent.

Find a closed form of $f_S(x)$.

Problem 40.5
Show that the family of Poisson distributions is closed under convolution.

Problem 40.6
Show that the family of binomial distributions is closed under convolution.

Problem 40.7
Show that the family of negative binomial distributions with the same parameter β but different $r's$ is closed under convolution.

Problem 40.8
Show that the family of Gamma distributions with common paramter θ is closed under convolution.

Problem 40.9
Let S be an aggregate loss random variable with a discrete frequency distribution N defined by the table below.

n	0	1	2
$\Pr(N = n)$	0.5	0.4	0.1

The severity claim has an exponential distribution with mean 2. Find $F_S(x)$.

Problem 40.10
Let $S_1, S_2, \cdots, S_5$ be i.i.d. compound Poisson random variables, each with parameter $\lambda_j = 7$. The pdf of the severity distribution for each S_j is $F_j(x) = x^2$ for $0 < x < 1$. For $S = S_1 + S_2 + \cdots + S_5$, find $M_S(2)$.

41 Distribution of S via the Recursive Method

As pointed out in Section 38, the $n-$fold convolutions are often extremely difficult to compute in practice and therefore one encounters difficulties dealing with the probability distribution of S. An alternative approach is to use various approximation techniques. In this section, we consider a technique known as the **Panjer recursive formula**.

The recursion method is used to find the distribution of S when the frequency distribution belongs to either the $(a,b,0)$ class or the $(a,b,1)$ class and the severity is integer valued and nonnegative. So let N be in the $(a,b,1)$ class and denote its distribution by $p_n = \Pr(N=n)$. Then p_n satisfies the equation

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, \quad n = 2, 3, \cdots . \tag{41.1}$$

We have the following theorem.

Theorem 41.1
For a frequency distribution in $(a,b,1)$, the pdf of S is recursively defined by

$$f_S(n) = \frac{[p_1 - (a+b)p_0]f_X(n) + \sum_{j=1}^{n}\left(a + \frac{b}{n}j\right) f_X(j) f_S(n-j)}{1 - af_X(0)} \tag{41.2}$$

where $f_S(n) = \Pr(S=n)$.

Proof.
Rewrite (41.1) in the form

$$np_n = (n-1)ap_{n-1} + (a+b)p_{n-1}, \ n = 2, 3, \cdots$$

Multiplying each side by $[P_X(z)]^{n-1}P'_X(z)$, summing over n and reindexing yields

$$\sum_{n=1}^{\infty} np_n[P_X(z)]^{n-1}P'_X(z) - p_1P'_X(z) = a\sum_{n=2}^{\infty}(n-1)p_{n-1}[P_X(z)]^{n-1}P'_X(z)$$

$$+(a+b)\sum_{n=2}^{\infty} p_{n-1}[P_X(z)]^{n-1}P'_X(z)$$

$$
\begin{aligned}
&=a\sum_{n=1}^{\infty} np_n[P_X(z)]^nP_X'(z)+(a+b)\sum_{n=1}^{\infty}p_n[P_X(z)]^nP_X'(z)\\
&=a\sum_{n=1}^{\infty} np_n[P_X(z)]^nP_X'(z)+(a+b)\sum_{n=0}^{\infty}p_n[P_X(z)]^nP_X'(z)-p_0(a+b)P_X'(z).
\end{aligned}
$$

Because $P_S(z)=\sum_{n=0}^{\infty}p_n[P_X(z)]^n$ (see similar argument in Section 36), the previous calculation yields

$$P_S'(z)-p_1P_X'(z)=aP_S'(z)P_X(z)+(a+b)P_S(z)P_X'(z)-p_0(a+b)P_X'(z)$$

or

$$P_S'(z)=[p_1-(a+b)p_0]P_X'(z)+aP_S'(z)P_X(z)+(a+b)P_S(z)P_X'(z).$$

Each side can be expanded in powers of z. The coefficient of z^{n-1} in such an expansion must be the same on both sides of the equation. Hence, obtaining

$$
\begin{aligned}
nf_S(n)=&[p_1-(a+b)p_0]nf_X(n)+a\sum_{j=0}^{n}(n-j)f_X(j)f_S(n-j)\\
&+(a+b)\sum_{j=0}^{n}jf_X(j)f_S(n-j)=[p_1-(a+b)p_0]nf_X(n)+anf_X(0)f_S(n)\\
&+a\sum_{j=1}^{n}(n-j)f_X(j)f_S(n-j)+(a+b)\sum_{j=1}^{n}jf_X(j)f_S(n-j)\\
&=[p_1-(a+b)p_0]nf_X(n)+anf_X(0)f_S(n)\\
&+an\sum_{j=1}^{n}f_X(j)f_S(n-j)+b\sum_{j=1}^{n}jf_X(j)f_S(n-j).
\end{aligned}
$$

Hence

$$[1-af_X(0)]f_S(n)=[p_1-(a+b)p_0]f_X(n)+\sum_{j=1}^{n}\left(a+\frac{bj}{n}\right)f_X(j)f_S(n-j).$$

Finally, the result follows by dividing both sides of the last equation by $1-af_X(0)$ ■

Corollary 41.2
If N is in the $(a,b,0)$ class then

$$f_S(x)=\Pr(S=x)=\frac{\sum_{j=1}^{n}\left(a+\frac{b}{x}j\right)f_X(j)f_S(x-j)}{1-af_X(0)} \qquad (41.3)$$

Proof.
If N is in the $(a,b,0)$ class then from $\frac{p_1}{p_0} = a+b$ we find $p_1 - (a+b)p_0 = 0$ so that (41.2) reduces to (41.3) ■

The recursive method requires an initial starting value $f_S(0)$ which can be found as follows:

$$f_S(0) = \Pr(S=0) = \Pr(N=0) = P_N[P_X(0)] = P_N[\Pr(X=0)].$$

Example 41.1
Develop the recursive formula for the case of compound Poisson distribution with parameter λ.

Solution.
The initial value is

$$f_S(0) = P_N[P_X(0)] = e^{\lambda(f_X(0)-1)}.$$

For the Poisson distribution, $a=0$ and $b=\lambda$. The recursive method reduces to

$$f_S(n) = \sum_{j=1}^{n} \frac{\lambda j}{n} f_X(j) f_S(n-j) \ ■$$

Example 41.2 ‡
For a tyrannosaur with a taste for scientists:
(i) The number of scientists eaten has a binomial distribution with $q=0.6$ and $m=8$.
(ii) The number of calories of a scientist is uniformly distributed on $(7000, 9000)$.
(iii) The numbers of calories of scientists eaten are independent, and are independent of the number of scientists eaten.
Calculate the probability that two or more scientists are eaten and exactly two of those eaten have at least 8000 calories each.

Solution.
If X denotes the number of calories of a scientist and N the number of scientists eaten. Then

$$\Pr(8000 \le X \le 9000) = F_X(8000) - F_X(9000) = \frac{9000-8000}{9000-7000} = \frac{1}{2}.$$

Thus, half the scientists are considered heavy. Let Y denote a heavy scientist S the total number of heavy scientist. We are asked to find $\Pr(S=2) =$

$f_S(2)$. Note that S is a compound distribution with primary distribution the number of scientist (a binomial distribution) and a secondary distribution standing for a heavy scientist (a Bernoulli distribution with a probability of a success equals to 0.5). Using Panjer algorithm, we have

$$\begin{aligned}
a =& -\frac{q}{1-q} = -\frac{-0.6}{0.4} = -1.5 \\
b =& (m+1)\frac{q}{1-q} = 9(1.5) = 13.5 \\
f_S(0) =& P_N[\Pr(X=0)] = P_N(0.5) \\
=& [1+0.6(0.5-1)^8] = 0.057648 \\
f_S(1) =& \frac{1}{1+1.5(0.5)}\left(-1.5+\frac{13.5}{1}\right)f_X(1)f_S(0) \\
=& \frac{1}{1.75}(12)(0.5)(0.057648) = 0.197650 \\
f_S(2) =& \frac{1}{1+1.5(0.5)}\left[\left(-1.5+\frac{13.5}{2}\right)f_X(1)f_S(0) + (-1.5+13.5)f_X(2)f_S(1)\right] \\
=& \frac{1}{1.75}[(5.25)(0.5)(0.057648) + 12(0)(0.197650)] = 0.29647 \blacksquare
\end{aligned}$$

Example 41.3 ‡
Let S have a Poisson frequency distribution with parameter $\lambda = 5$. The individual claim amount has the following distribution:

x	$f_X(x)$
100	0.8
500	0.16
1000	0.04

Calculate the probability that aggregate claims will be exactly 600.

Solution.
In order to have exactly 600 in aggregate claims, one of following must happen:
• $N = 2,\ X_1 = 100$, and $X_2 = 500$. In this case, the probability is

$$\frac{e^{-5}5^2}{2!}(0.8)(0.16) = 0.010781.$$

• $N = 2,\ X_1 = 500$, and $X_2 = 100$. In this case, the probability is

$$\frac{e^{-5}5^2}{2!}(0.16)(0.8) = 0.010781.$$

- $N = 6,\ X_1 = \cdots = X_6 = 100$. In this case, the probability is

$$\frac{e^{-5}5^6}{6!}(0.8)^6 = 0.038331.$$

Summing up, we find a total probability of about 0.06 ∎

Practice Problems

Problem 41.1
Let S have a Poisson frequency distribution with parameter $\lambda = 0.04$. The individual claim amount has the following distribution:

x	$f_X(x)$
1	0.5
2	0.4
3	0.1

Find $f_S(x),\ x = 1, 2, 3, 4$ using Panjer recursion formula.

Problem 41.2
Let S be a compound Poisson distribution with parameter $\lambda = 0.04$ and individual claim distribution given by

x	$f_X(x)$
1	0.5
2	0.4
3	0.1

Show that: $f_S(n) = \frac{1}{n}[0.02f_S(n-1) + 0.032f_S(n-2) + 0.012f_S(n-3)]$.

Problem 41.3 ‡
You are given:
• S has a compound Poisson distribution with $\lambda = 2$.
• Individual claim amounts are distributed as follows: $f_X(1) = 0.4$ and $f_X(2) = 0.6$.

Determine $f_S(4)$.

Problem 41.4 ‡
Aggregate claims S has a compound Poisson distribution with parameter λ and with discrete individual claim amount distributions of $f_X(1) = \frac{1}{3}$ and $f_X(3) = \frac{2}{3}$. Also, $f_S(4) = f_S(3) + 6f_S(1)$.

Determine the value of λ.

Problem 41.5 ‡
Aggregate claims S has a compound Poisson distribution with parameter λ and with discrete individual claim amount distributions of $f_X(1) = \frac{1}{3}$ and

$f_X(3) = \frac{2}{3}$. Also, $f_S(4) = f_S(3) + 6f_S(1)$.

Determine $\mathrm{Var}(S)$.

Problem 41.6 ‡
For aggregate claim S, you are given:

$$f_S(x) = \sum_{n=0}^{\infty} f_X^{*n}(x)\frac{e^{-50}(50^n)}{n!}.$$

Losses are distributed as follows: $f_X(1) = 0.4$ $f_X(2) = 0.5$, and $f_X(3) = 0.1$.

Calculate $\mathrm{Var}(S)$.

Problem 41.7
Let S have a Poisson frequency distribution with parameter $\lambda = 0.04$. The individual claim amount has the following distribution:

x	$f_X(x)$
1	0.5
2	0.4
3	0.1

Find $F_S(4)$.

Problem 41.8
The frequency distribution of an aggregate loss S follows a binomial distribution with $m = 4$ and $q = 0.3$. Loss amount has the distribution: $f_X(0) = 0.2,\ f_X(1) = 0.7,\ f_X(2) = 0.1$

Find the starting value $f_S(0)$ in the recursion formula.

Problem 41.9
The frequency distribution of an aggregate loss S follows a binomial distribution with $m = 4$ and $q = 0.3$. Loss amount has the distribution: $f_X(0) = 0.2,\ f_X(1) = 0.7,\ f_X(2) = 0.1$

Calculate $f_S(1)$ and $f_S(2)$.

Problem 41.10
The frequency distribution of an aggregate loss S follows a binomial distribution with $m = 4$ and $q = 0.3$. Loss amount has the distribution:

$f_X(0) = 0.2,\ f_X(1) = 0.7,\ f_X(2) = 0.1$

Calculate $\Pr(S \geq 3)$.

Problem 41.11

Annual aggregate losses for a dental policy follow the compound Poisson distribution with $\lambda = 3$. The distribution of individual losses is:

Loss	Probability
1	0.4
2	0.3
3	0.2
4	0.1

Calculate the probability that aggregate losses in one year do not exceed 3.

42 Discretization of Continuous Severities

The severity distribution in the recursive method of Section 41 was assumed to be discrete. Most severity distributions are of continuous type. An analog of Panjer recursive method with a continuous severity is

$$f_S(x) = (a+b)\Pr(S=0)f_X(x) + \int_0^x \left(a + \frac{by}{x}\right) f_X(y)f_S(x-y)dy$$

where the probability of the severity distribution is taken on the positive real line. This is a Volterra integral equation of the second kind. This type of equations is usually very difficult to solve analytically. Instead, numerical methods are used. The two methods discussed in this section are based on the discretization of the continuous severiy distribution. Once a continuous severity is discretized, the recursive method can be applied.

The Method of Rounding
An **arithmetic distribution** is a distribution with support $\mathbb{N} = \{1, 2, \cdots\}$. An equispaced arithmetic distribution is defined on positive multiples of a unit of measurement h which we call the **span**. Transforming a continuous distribution to an arithmetic distribution is referred to as **discretizing** or **arithmetizing** the distribution.
Let X denote the random variable to be arithmetized and $h > 0$ denote the span. The **method of rounding** (also known as the method of **mass dispersal**) concentrates the probability one-half span either side of jh and places it at jh :

$$\begin{aligned} f_0 =& \Pr\left(X < \frac{h}{2}\right) = F_X\left(\frac{h}{2}\right) \\ f_j =& \Pr\left(jh - \frac{h}{2} \leq X < jh + \frac{h}{2}\right) \\ =& F_X\left(jh + \frac{h}{2}\right) - F_X\left(jh - \frac{h}{2}\right), \ j = 1, 2, \cdots \end{aligned}$$

At some point, it is reasonable to halt the discretization process at some point, say mh, once most of the probability has been accounted for. At this terminal point, we have $f_m = 1 - F_X\left(mh - \frac{h}{2}\right)$. With this method, all the f_j sum up to 1, ensuring that the discretization leads to a legitimate probability distribution.

Example 42.1
Let S be an aggregate random variable with a frequency distribution that

has a Poisson distribution with $\lambda = 2$ and a severity distribution that has a Pareto distribution with parameters $\alpha = 2$ and $\theta = 1000$.
(a) Use the method of rounding to approximate f_0, f_1, f_2 and f_3 in the arithmetize distribution such that f_3 is the last positive probability. Use the span $h = 6$.
(b) Find $f_S(12)$.

Solution.
(a) The cdf of the severity distribution is

$$F_X(x) = 1 - \left(\frac{1000}{x+1000}\right)^2.$$

Thus,

$$\begin{aligned}
f_0 =& F_X(3) = 1 - \left(\frac{1000}{3+1000}\right)^2 = 0.006\\
f_1 =& F_X(9) - F_X(3) = \left(\frac{1000}{3+1000}\right)^2 - \left(\frac{1000}{9+1000}\right)^2 = 0.0118\\
f_2 =& F_X(15) - F_X(9) = \left(\frac{1000}{9+1000}\right)^2 - \left(\frac{1000}{16+1000}\right)^2 = 0.0116\\
f_3 =& 1 - F_X(15) = \left(\frac{1000}{15+1000}\right)^2 = 0.9707.
\end{aligned}$$

(b) Let A denote the aggregate random variable with frequency distribution the Poisson distribution with $\lambda = 2$ and severity distribution the arithmetize distribution. We want to find $f_S(12) = f_A(2)$. Using the recursive method, we find

$$\begin{aligned}
f_A(0) =& e^{\lambda(f_0-1)} = 0.137\\
f_A(1) =& 2[f_1 f_A(0)] = 0.0032332\\
f_A(2) =& [f_1 f_A(1) + 2f_2 f_A(0)] = 0.00322 \blacksquare
\end{aligned}$$

The Method of Local Moment Matching
The rounding method has the disadvantage that the approximate distribution does not preserve any moment of the exact distribution. Gerber proposed a method that matched moments locally: The **local moment matching method**. This method constructs discrete equispaced distributions that matches some moments of the exact distribution. The idea

is to replace the probability of the severity distribution X over intervals $[x_k, x_k + ph)$ by point masses located at $x_i^k = x_k + ih$ where $i = 0, 1, \cdots, p$ and where the integer p represents the number of moments to be matched. The method then tries to calculate the $p+1$ point masses $m_0^k, m_1^k, \cdots, m_p^k$ located at $x_0^k, x_1^k, \cdots, x_p^k$ which are solutions to the $(p+1)$ equations

$$\sum_{j=0}^{p} (x_k + jh)^r m_j^k = \int_{x_k}^{x_k+ph} x^r dF_X(x), \quad r = 0, 1, 2, \cdots, p. \tag{42.1}$$

In practice, the above system is solved in each of the intervals

$$[0, ph), [ph, 2ph), [2ph, 3ph), \cdots,$$

that is, $x_{k+1} = x_k + ph$ with $x_0 = 0$ and therefore $x_k = kph$. The final probabilities of the discretized distribution are the point masses obtained as solutions of the previous system in each interval, summed for the endpoints of each interval:

$$\begin{array}{llll} f_0 = m_0^0 & f_1 = m_1^0 & f_2 = m_2^0 & \cdots \\ f_p = m_p^0 + m_0^1 & f_{p+1} = m_1^1 & f_{p+2} = m_2^1 & \cdots \end{array}$$

The unique solution to the system (42.1) is provided by the following theorem.

Theorem 42.1
The solution to (42.1) is

$$m_j^k = \int_{x_k}^{x_k+ph} \left(\prod_{i \neq j} \frac{x - x_k - ih}{(j-i)h} \right) dF_X(x), \;\; j = 0, 1, \cdots, p.$$

Proof.
Let $f(x)$ be a polynomial and consider the set of data points

$$(x_0, f(x_0)), (x_1, f(x_1)), \cdots, (x_p, f(x_p)).$$

Then Lagrange interpolation formula allows us to write

$$f(x) = \sum_{j=0}^{p} f(x_j) \left(\prod_{i \neq j} \frac{x - x_i}{x_j - x_i} \right).$$

In particular, by letting $f(x) = x^r$ and $x_i = x_k + ih$ $(i = 0, 1, \cdots, p)$ we find

$$x^r = \sum_{j=0}^{p} (x_k + jh)^r \left(\prod_{i \neq j} \frac{x - x_k - ih}{(j-i)h} \right), \; r = 0, 1, 2, \cdots$$

Integrate both sides with respect to $F_X(x)$ over the interval $[x_k, x_k + ph)$ we find

$$\sum_{j=0}^{p} (x_k + jh)^r m_j^k = \sum_{j=0}^{p} (x_k + jh)^r \int_{x_k}^{x_k+ph} \left(\prod_{i \neq j} \frac{x - x_k - ih}{(j-i)h} \right) dF_X(x).$$

By uniqueness of the solution to the system the desired result follows ∎

Example 42.2
Suppose X has the exponential distribution with pdf $f_X(x) = 0.1e^{-0.1x}$. Use the method of local moment mathing with $p = 1$ and a span $h = 2$
(a) to find the equation corresponding to $r = 0$ in the resulting system of equations. Assume that $x_0 = 0$;
(b) to find m_0^k and m_1^k using Theorem 42.1;
(c) to find f_0 and f_n.

Solution.
(a) We have

$$\sum_{j=0}^{1} (x_k + jh)^0 m_j^k = \int_{2k}^{2k+2} f_X(x)dx = 0.1 \int_{2k}^{2k+2} e^{-0.1x}dx = e^{-0.2k} - e^{-0.2k-0.2}$$

which implies $m_0^k + m_1^k = e^{-0.2k} - e^{-0.2k-0.2}$.
(b) Using Theorem 42.1, we find

$$m_0^k = \int_{2k}^{2k+2} \frac{x - 2k - 2}{-2} [0.1e^{-0.1x}]dx = 5e^{-0.2k-0.2} - 4e^{-0.2k}$$

and

$$m_1^k = \int_{2k}^{2k+2} \frac{x - 2k}{2} [0.1e^{-0.1x}]dx = -6e^{-0.2k-0.2} + 5e^{-0.2k}.$$

(c) We have

$$\begin{aligned} f_0 =& m_0^0 = 5e^{-0.2} - 4 = 0.09365 \\ f_n =& m_1^{n-1} + m_0^n = 5e^{-0.2n+0.2} - 10e^{-0.2n} + 5e^{-0.2n-0.2} \end{aligned}$$ ∎

Example 42.3
Loss amounts follow a Pareto distribution with parameters $\alpha = 3$ and $\theta = 5$. Use the Method of Local Moment Matching with $h = 3$ and $p = 1$ to find the system of two equations in the unknowns m_0^0 and m_1^0. Solve this system. Assume $x_0 = 0$.

Solution.
We have

$$m_0^0 + m_1^0 = \Pr(0 \le X < 3) = F_X(3) - F_X(0) = 1 - \left(\frac{5}{3+5}\right)^3 = 0.7559.$$

On the other hand,

$$x_0 m_0^0 + x_1 m_1^0 = \int_0^3 x f_X(x) dx = \int_0^3 \frac{3(5)^3 x}{(x+5)^4} dx = 0.79102.$$

This implies

$$3m_1^0 = 0.79102 \Longrightarrow m_1^0 = 0.2637$$

and

$$m_0^0 = 0.7559 - 0.2637 = 0.4922 \blacksquare$$

Practice Problems

Problem 42.1
Let S be an aggregate random variable with a frequency distribution that has a Poisson distribution with $\lambda = 2$ and a severity distribution that has a uniform distribution in $(0, 50)$.
(a) Use the method of rounding to approximate f_0, f_1, f_2 and f_3 in the arithmetize distribution such that f_3 is the last positive probability. Use the span $h = 10$.
(b) Find $f_S(30)$.

Problem 42.2
Medical claims have an exponential distribution with mean $\theta = 500$. Using the method of rounding with a span $h = 50$ estimate the probability that the claim amount is 500.

Problem 42.3
Loss amounts follow a Weibull distribution with parameters $\alpha = 2$ and $\theta = 1000$. Using the method of rounding with a span $h = 500$ estimate the probability that the loss amount is 2000.

Problem 42.4
Loss amounts follow a Pareto distribution with parameters $\alpha = 3$ and $\theta = 5$. Use the Method of Local Moment Matching with $h = 3$ and $p = 1$ to find the system in two equations in the unknowns m_0^1 and m_1^1. Solve this system. Assume $x_0 = 0$.

Problem 42.5
Loss amounts follow a Pareto distribution with parameters $\alpha = 3$ and $\theta = 5$. Use the Method of Local Moment Matching with $h = 3$ and $p = 1$ to find f_0 and f_3.

Problem 42.6
Loss amounts follow an exponential distribution $\theta = 500$. A discrete distribution is created using the Method of Local Moment Matching such that $p = 1$ and $h = 50$. Calculate the probability to a loss amount of 500 using Theorem 42.1.

Problem 42.7
Loss amounts follow a Pareto distribution $\alpha = 4$ and $\theta = 10000$. A discrete distribution is created using the Method of Local Moment Matching such that $p = 1$ and $h = 300$. Calculate the probability to a loss amount of 3000 using Theorem 42.1.

43 Individual Policy Modifications Impact on Aggregate Losses

Any individual policy modifications such as deductibles, policy limits, or coinsurance will have an impact on the aggregate loss/payment random variable. In this section, it will be assumed that the ground-up distribution of the individual loss amount X is unaffected by policy modifications, and only the payments themselves are affected.

We next start by reminding the reader of the following notations: The number of losses will be denoted by N^L; the number of payments by N^P, the probability of a loss resulting in a payment by v; the amount of payment on a per-loss basis is Y^L with $Y^L = 0$ if a loss results in no payment; and the amount of payment on a per-payment basis is Y^P where only losses that result on a nonzero payment are considered and the losses that result in no payment are completely ignored. Note that $\Pr(Y^P = 0) = 0$ and $Y^P = Y^L | Y^L > 0$.

The cdfs of Y^L and Y^P are related as follows:

$$\begin{aligned}F_{Y^L}(y) =&\Pr(Y^L \le y | Y^L = 0)\Pr(Y^L = 0) + \Pr(Y^L \le y | Y^L > 0)\Pr(Y^L > 0)\\ =&1 - v + vF_{Y^P}(y).\end{aligned}$$

The mgfs of Y^L and Y^P are related as follows:

$$\begin{aligned}M_{Y^L}(t) =&E(e^{tY^L} | Y^L = 0)\Pr(Y^L = 0) + E(e^{tY^L} | Y^L > 0)\Pr(Y^L > 0)\\ =&(1 - v) + vM_{Y^P}(t).\end{aligned}$$

The pgfs of N^P and N^L are related as follows (see Section 36):

$$P_{N^P}(z) = P_N(1 - v + vz).$$

Now back to the aggregate payments. On a per-loss basis, the total payments may be expressed as

$$S = Y_1^L + Y_2^L + \cdots + Y_{N^L}^L$$

with $S = 0$ if $N^L = 0$ and where Y_j^L is the payment amount on the jth loss. Alternatively, ignoring all losses that do not result in a nonzero-payment, the aggregate S can be expressed as

$$S = Y_1^P + Y_2^P + \cdots + Y_{N^P}^P$$

with $S = 0$ if $N^P = 0$ and where Y_j^P is the payment amount on the jth loss which results in a nonzero payment. On a per-loss basis, S is a compound distribution with primary distribution N^L and secondary distribution Y^L so that

$$M_S(t) = P_{N^L}[M_{Y^L}(t)].$$

Likewise, on a per-payment basis, we have

$$M_S(t) = P_{N^P}[M_{Y^P}(t)].$$

Note that

$$P_{N^L}[M_{Y^L}(t)] = P_{N^L}[1 - v + vM_{Y^P}(t)] = P_{N^P}[M_{Y^P}(t)].$$

Note that individual policy modifications factor either in the expression of Y^L or Y^P and thus have an impact on the aggregate payment.

Example 43.1
A ground-up model of individual losses follows a Pareto distribution with $\alpha = 4$ and $\theta = 10$. The number of losses is a Poisson distribution with $\lambda = 3$. There is an ordinary deductible of 6, a policy limit of 18− applied before the policy limit and the coinsurance, and coinsurance of 75%. Find the expected value and the variance of S, the aggregate payments on a per-loss basis.

Solution.
We have

$$\begin{aligned}
E(N^L) =&3\\
E(X \wedge 24) =&\frac{\theta}{\alpha - 1}\left[1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha-1}\right]\\
=&\frac{10}{3}\left[1 - \left(\frac{10}{24+10}\right)^{4-1}\right] = 3.2485\\
E(X \wedge 6) =&\frac{10}{3}\left[1 - \left(\frac{10}{6+10}\right)^{4-1}\right] = 2.5195\\
E(Y^L) =&\alpha[E(X \wedge u) - E(X \wedge d)] = 0.75[E(X \wedge 24) - E(X \wedge 6)]\\
=&0.75(3.2485 - 2.5195) = 0.54675.
\end{aligned}$$

$$\begin{aligned}
E(S) =& E(N^L)E(Y^L) = 3(0.54675) = 1.64\\
E[(X \wedge 24)^2] =& \int_0^{24} x^2 \left[\frac{4(10)^4}{(x+10)^5}\right] dx + 24^2\left(\frac{10}{24+10}\right)^4 = 26.379\\
E[(X \wedge 6)^2] =& \int_0^{6} x^2 \left[\frac{4(10)^4}{(x+10)^5}\right] dx + 6^2\left(\frac{10}{6+10}\right)^4 = 10.5469\\
E[(Y^L)^2] =& (0.75)^2\left\{E[(X\wedge 24)^2] - E[(X\wedge 6)^2]\right]\\
&-2(6)E(X\wedge 24) + 2(d)E(X\wedge 6)\ \} = 3.98481\\
\mathrm{Var}(Y^L) =& 3.98481 - 0.54675^2 = 3.685874438\\
\mathrm{Var}(S) =& E(N^L)\mathrm{Var}(Y^L) + E(Y^L)^2\mathrm{Var}(N^L)\\
=& 3(3.685874438) + 0.54675^2(3) = 11.95443 \ \blacksquare
\end{aligned}$$

Example 43.2
A ground-up model of individual losses follows a Pareto distribution with $\alpha = 4$ and $\theta = 10$. The number of losses is a Poisson distribution with $\lambda = 3$. There is an ordinary deductible of 6, a policy limit of 18− applied before the policy limit and the coinsurance, and coinsurance of 75%. Find the distribution of S, the aggregate payments on a per-payment basis.

Solution.
Since we are treating S on a per-payment basis, we can look at S as an aggregate distribution with frequency distribution N^P and severity distribution Y^P. The probability that a loss will result in a payment is

$$v = \Pr(X > 6) = S_X(6) = \left(\frac{10}{10+6}\right)^4 = 0.15259.$$

We also have (see Section 36)

$$E(N^P) = E(N^L)E(I_j) = \lambda v = 3(0.15259) = 0.45776.$$

Let $Z = X - 6|X > 6$ denote the individual payment random variable with only a deductible of 6. Then

$$\Pr(Z > z) = \frac{\Pr(X > z+6)}{\Pr(X > 6)}.$$

Now with the 75% coinsurance, $Y^P = 0.75Z$ and the maximum payment is $0.75(24-6) = 13.5$ so that for $y < 13.5$ we have

$$F_{Y^P}(y) = 1 - \Pr(0.75Z > y) = \frac{\Pr(X > 6) - \Pr\left(X > 6 + \frac{y}{0.75}\right)}{\Pr(X > 6)}$$

and $F_{Y^P}(y) = 1$ for $y \geq 13.5$.
Now since Y^P is continuous, we will use the method of rounding to discretize it. We will use a span of 2.25 to obtain

$$\begin{aligned}
f_0 =& F_{Y^P}(2.25/2) = 1 - \frac{\Pr(X > 6 + \frac{1.125}{0.75})}{0.15259} = 0.30124\\
f_1 =& F_{Y^P}(3.375) - F_{Y^P}(1.125) = 0.32768\\
\vdots =& \vdots\\
f_6 =& F_{Y^P}(14.625) - F_{Y^P}(12.375) = 0.05874\\
f_n =& 1 - 1 = 0,\ n = 7, 8, \cdots .
\end{aligned}$$

Now, S can be computed using the compound Poisson recursive formula

$$f_S(n) = \frac{0.45776}{n}\sum_{j=1}^{n} y f_n f_S(n-j),\ n = 1, 2, 3, \cdots$$

with $f_S(0) = e^{0.45776(1-0.30124)} = 0.72625$ ■

Example 43.3 ‡
Aggregate losses for a portfolio of policies are modeled as follows:
(i) The number of losses before any coverage modifications follows a Poisson distribution with mean λ.
(ii) The severity of each loss before any coverage modifications is uniformly distributed between 0 and b.
The insurer would like to model the impact of imposing an ordinary deductible, d $(0 < d < b)$, on each loss and reimbursing only a percentage, $c(0 < c \leq 1)$, of each loss in excess of the deductible.
It is assumed that the coverage modifications will not affect the loss distribution. The insurer models its claims with modified frequency and severity distributions. The modified claim amount is uniformly distributed on the interval $[0, c(b - d)]$.
Determine the mean of the modified frequency distribution.

Solution.
Imposing the deductible will limit payments to those losses that are greater than d. In this case, the proprotion of losses that result in a payment is $\Pr(X > d) = 1 - \frac{d}{b}$ (remember that loss amounts are uniform in $(0, b)$.) Thus, the mean of the modified frequency distribution, that is, the expected number of losses that will result in a payment being made is the product of

the total expected number of losses and the probability of a loss resulting in a payment. That is

$$\lambda\left(1-\frac{b}{d}\right) \blacksquare$$

Example 43.4 ‡
A company insures a fleet of vehicles. Aggregate losses have a compound Poisson distribution. The expected number of losses is 20. Loss amounts, regardless of vehicle type, have exponential distribution with $\theta = 200$.
In order to reduce the cost of the insurance, two modifications are to be made:
(i) a certain type of vehicle will not be insured. It is estimated that this will reduce loss frequency by 20%.
(ii) a deductible of 100 per loss will be imposed.
Calculate the expected aggregate amount paid by the insurer after the modifications.

Solution.
We want

$$E(S) = E(N^L)E(Y^L) = E(N^L)E[(X-100)_+] = E(N^L)[E(X)-E(X\wedge 100)].$$

From Table C,

$$E(X \wedge 100) = 200(1 - e^{-\frac{100}{200}}).$$

Thus,

$$E(S) = (20)(0.8)\left[200 - 200(1 - e^{-\frac{100}{200}})\right] \approx 1941 \blacksquare$$

Practice Problems

Problem 43.1
A ground-up model of individual losses follows a Pareto distribution with $\alpha = 4$ and $\theta = 10$. The number of losses is a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 6, a policy limit of 18− applied before the policy limit and the coinsurance, and coinsurance of 75%.

Find $F_{Y^L}(y)$.

Problem 43.2
An insurance company has a policy where the amount of each payment for losses follows an exponential distribution with mean $\theta = 100$. The number of losses follows a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 30, a maximum covered losses of 340− applied before the policy limit and the coinsurance, and coinsurance of 53%.

Find the expected value and variance of S, the aggregate payments on a per-loss basis.

Problem 43.3
An insurance company has a policy where the amount of each payment for losses follows an exponential distribution with mean $\theta = 100$. The number of losses follows a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 30, a maximum covered losses of 340− applied before the policy limit and the coinsurance, and coinsurance of 53%. Let S be the aggregate payments on a per-payment basis.

Find the distribution of Y^P.

Problem 43.4
An insurance company has a policy where the amount of each payment for losses follows an exponential distribution with mean $\theta = 100$. The number of losses follows a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 30, a maximum covered losses of 340− applied before the policy limit and the coinsurance, and coinsurance of 53%.

Find f_0 and f_n under the method of rounding using a span of 30.

Problem 43.5
An insurance company has a policy where the amount of each payment for

losses follows an exponential distribution with mean $\theta = 100$. The number of losses follows a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 30, a maximum covered losses of $340-$ applied before the policy limit and the coinsurance, and coinsurance of 53%.

Find $f_S(0)$ and $f_S(n)$.

Problem 43.6
An insurance company has a policy where the amount of each payment for losses on a per-payment basis follows an exponential distribution with mean $\theta = 100$. The probability that a loss will result in a payment is 0.74082.

Find $M_{Y^L}(t)$.

Problem 43.7
A ground-up model of individual losses follows a Pareto distribution with $\alpha = 4$ and $\theta = 10$. The number of losses is a Poisson distribution with $\lambda = 7$. There is an ordinary deductible of 6, a policy limit of $18-$ applied before the policy limit and the coinsurance, and coinsurance of 75%.

Find $P_{N^P}(z)$.

44 Aggregate Losses for the Individual Risk Model

Consider a portfolio of n insurance policies or risks. Denote the loss to the insurer, for a fixed period, for each policy i by X_i, for $i = 1, 2, \cdots, n$. We assume that the $X_i's$ are independent but not necessarily identically distributed. The **individual risk model** represents the aggregate losses for the fixed period as a fixed sum

$$S = X_1 + X_2 + \cdots + X_n.$$

Let q_i denote the probability of claim within a fixed period of the ith policyholder and let b_i denote the fixed benefit to the ith policyholder. Then the pdf of the loss to the insurer is

$$f_{X_j}(x) = \begin{cases} 1 - q_i, & x = 0 \\ q_i, & x = b_i. \end{cases}$$

On the other hand, the pgf of X_i is

$$P_{X_i}(z) = E(z^{X_i}) = 1 - q_i + q_i z^{b_i}.$$

By the independence of the $X_i's$, we can write

$$P_S(z) = \prod_{i=1}^{n} (1 - q_i + q_i z^{b_i}).$$

The claim amount variable X_i for each policy is usually presented as

$$X_i = I_i B_i, \ i = 1, 2, \cdots, n$$

where random variables $I_1, I_2, \cdots, I_n, B_1, B_2, \cdots, B_n$ are independent. The random variable I_i is an indicator random variable such that $\Pr(I_i = 1) = q_i$ and $\Pr(I_i = 0) = 1 - q_i$. Thus, I_i indicates whether the ith policy produced a payment. The positive random variable B_i can have an arbitrary distribution and represents the amount of the payment with respect to the ith policy given that a payment was made.
To find the mean of S, we proceed as follows:

$$\begin{aligned} E(X_i) =& E[E(X_i|I_i)] = E[E(I_iB_i|I_i)] = E[I_iE(B_i|I_i)] \\ =& E[I_iE(B_i)] = E(I_i)E(B_i) = q_iE(B_i). \end{aligned}$$

Hence, by the independence of the $X_i's$ we have

$$E(S) = \sum_{i=1}^{n} q_i E(B_i).$$

On the other hand,

$$\begin{aligned}E(X_i^2) =&E[E(X_i^2|I_i)] = E[E(I_i^2B_i^2|I_i)] = E[I_i^2E(B_i^2|I_i)]\\ =&E[I_i^2E(B_i^2)] = E(I_i)E(B_i^2) = q_iE(B_i^2).\end{aligned}$$

Hence,

$$\mathrm{Var}(X_i) = q_iE(B_i^2) - q_i^2E(B_i)^2 = q_i\mathrm{Var}(B_i) + q_i(1-q_i)E(B_i)^2$$

and

$$\mathrm{Var}(S) = \sum_{i=1}^n [q_i\mathrm{Var}(B_i) + q_i(1-q_i)E(B_i)^2].$$

The moment generating function of X_i is

$$\begin{aligned}M_{X_i}(t) =&E(e^{tI_iB_i}) = E[e^{tI_iB_i}|I_i=0]\Pr(I_i=0) + E[e^{tI_iB_i}|I_i=1]\Pr(I_i=1)\\ =&1-q_i+q_iM_{B_i}(t).\end{aligned}$$

By independence, the moment generating function of S is

$$M_S(t) = \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n [1-q_i+q_iM_{B_i}(t)].$$

Example 44.1
For a portfolio of 5 life insurance policies, you are given:

i	q_i	b_i
1	0.32	500
2	0.10	1000
3	0.54	250
4	0.23	375
5	0.14	650

Calculate $E(S)$ and $\mathrm{Var}(S)$ where $S = X_1 + X_2 + \cdots + X_5$.

Solution.
The aggregate mean is

$$E(S) = 0.23(500) + 0.10(1000) + 0.54(250) + 0.23(375) + 0.14(650) = 527.25.$$

The variance of S is

$$\begin{aligned}\mathrm{Var}(S) =&0.32(1-.32)(500)^2 + 0.10(1-0.10)(1000)^2\\ &+0.54(1-0.54)(250)^2 + 0.23(1-0.23)(375)^2 + 0.14(1-0.14)(650)^2\\ =&235698.6875 \blacksquare\end{aligned}$$

Example 44.2
For a portfolio of 3 life insurance policies, you are given:

i	q_i	b_i
1	0.32	500
2	0.10	1000
3	0.54	250

Calculate $P_S(z)$.

Solution.
We have

$$P_S(z) = (0.68 + 0.32z^{500})(0.90 + 0.10z^{1000})(0.36 + 0.54z^{250}) \blacksquare$$

Example 44.3
A life insurance portfolio consists of 100 policies where each policy has a probability of a claim of 0.20. When a claim occurs, the amount of the claim follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 1000$. Caculate the mean and the variance of the aggregate loss.

Solution.
For $i = 1, 2 \cdots, 100$, we have

$$E(B_i) = \frac{\theta}{\alpha - 1} = \frac{1000}{3-1} = 500.$$

Hence

$$E(S) = nqE(B) = 100(0.20)(500) = 10,000.$$

Next,

$$\text{Var}(B) = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)} = 750,000.$$

Therefore,

$$\text{Var}(S) = 100[0.20(750,000) + 0.20(0.80)(500)^2] = 19,000,000 \blacksquare$$

Example 44.4 ‡
Each life within a group medical expense policy has loss amounts which follow a compound Poisson process with $\lambda = 0.16$. Given a loss, the probability that it is for Disease 1 is $\frac{1}{16}$.
Loss amount distributions have the following parameters:

	Mean per loss	Standard Deviation per loss
Disease 1	5	50
Other Diseases	10	20

Premiums for a group of 100 independent lives are set at a level such that the probability (using the normal approximation to the distribution for aggregate losses) that aggregate losses for the group will exceed aggregate premiums for the group is 0.24.
A vaccine which will eliminate Disease 1 and costs 0.15 per person has been discovered.
Define:
A= the aggregate premium assuming that no one obtains the vaccine, and
B= the aggregate premium assuming that everyone obtains the vaccine and the cost of the vaccine is a covered loss.

Calculate A/B.

Solution.
Let X_1 denote the loss per person due to Disease 1 and X_2 the loss per person due to Other Diseases. Then X_1 and X_2 are Bernoulli with parameter $q = \frac{1}{16}$.
The expected loss per perosn is

$$E(L) = \frac{1}{16}(5) + \frac{15}{16}(10) = \frac{155}{16} = 9.6875$$

and the variance is

$$\text{Var}(L) = E(\text{Var}(X)) + \text{Var}[E(X)].$$

Now, $E(X)$ is a random variable with outcomes 5 (with probability $\frac{1}{16}$) and 10 (with probability $\frac{5}{16}$). Thus,

$$\text{Var}(E(X)) = 5^2\left(\frac{1}{16}\right) + 10^2\left(\frac{15}{16}\right) - 9.6875^2 = 1.4648.$$

Also,

$$E(\text{Var}(X)) = 50^2\left(\frac{1}{16}\right) + 20^2\left(\frac{15}{16}\right) = 531.25.$$

Hence,

$$\text{Var}(L) == 531.25 + 1.4648 = 532.7148.$$

Let S be the aggregate losses. Then S consists of 100 independent compound distributions where each compound distribution has a primary distribution N that is Poisson with parameter $\lambda = 0.16$ and secondary distribution L. Hence,

$$E(S) = 100(0.16)(9.6875) = 155$$

and

$$\text{Var}(S) = 100[E(N)\text{Var}(L) + \text{Var}(N)E(L)^2] = 100(0.16)(531.25 + 9.68875^2) = 10025.$$

Now, suppose that no one gets the vaccine. Then

$$\Pr(S > A) = \Pr\left(\frac{S - 155}{\sqrt{10025}} > \frac{A - 155}{\sqrt{10025}}\right) = \Pr\left(Z > \frac{A - 155}{\sqrt{10025}}\right) = 0.24.$$

Thus,

$$\frac{A - 155}{\sqrt{10025}} = 0.7 \Longrightarrow A = 225.09.$$

Next, if everyone gets the vaccine then the losses due to Disease 1 are eliminated. In this case,

$$E(S) = 100\left[0.16\frac{15}{16}(10) + 0.15\right] = 165$$

and

$$\text{Var}(S) = 100(0.16)\left[20^2\left(\frac{15}{16}\right) + 10^2\left(\frac{15}{16}\right)\right] = 7500.$$

In this case,

$$\Pr(S > B) = \Pr\left(\frac{S - 165}{\sqrt{7500}} > \frac{B - 165}{\sqrt{7500}}\right) = \Pr\left(Z > \frac{B - 165}{\sqrt{7500}}\right) = 0.24.$$

Thus,

$$\frac{B - 165}{\sqrt{7500}} = 0.7 \Longrightarrow B = 225.62.$$

Finally,

$$\frac{A}{B} = \frac{225.09}{225.62} = 0.998 \blacksquare$$

Practice Problems

Problem 44.1
A life insurance portfolio consists of 100 policies where each policy has a probability of a claim of 0.20. When a claim occurs, the amount of the claim follows an exponential distribution with mean $\theta = 1000$.

Caculate the mean and the variance of the aggregate loss.

Problem 44.2
A life insurance portfolio consists of 100 policies where each policy has a probability of a claim of 0.20. When a claim occurs, the amount of the claim follows an exponential distribution with mean $\theta = 1000$.

Find $M_S(t)$.

Problem 44.3
A life insurance portfolio consists of 100 policies where each policy has a probability of a claim of 0.20. When a claim occurs, the amount of the claim follows an exponential distribution with mean $\theta = 1000$.

Find $P_S(z)$.

Problem 44.4 ‡
A group life insurance contract covering independent lives is rated in the three age groupings as given in the table below.

Age group	Number in age group	Probability of claim per life	Mean of the exponential distribution of claim amounts
18-35	400	0.03	5
36-50	300	0.07	3
51-65	200	0.10	2

(a) Find the mean and the variance of the aggregate claim.
(b) Find the 95^{th} percentile of S.

Problem 44.5 ‡
The probability model for the distribution of annual claims per member in a health plan is shown below. Independence of costs and occurrences among services and members is assumed. Suppose that the plan consists of n members.

Service	Probability of claim	Mean of claim dist.	Variance of claim dist.
office visits	0.7	160	4,900
Surgery	0.2	600	20,000
Other Services	0.5	240	8,100

Find the mean and the variance of the aggregate claim S.

45 Approximating Probabilities in the Individual Risk Model

In this section, we approximate probabilities of aggregate losses in the individual risk model using either the normal distribution or the lognormal distribution.

For large n, the normal approximation uses the Central Limit Theorem as follows:

$$\begin{aligned}\Pr(S \le s) =& \Pr\left(\frac{S - E(S)}{\sqrt{\text{Var}(S)}} \le \frac{s - E(S)}{\sqrt{\text{Var}(S)}}\right)\\ \approx& \Pr\left(Z \le \frac{s - E(S)}{\sqrt{\text{Var}(S)}}\right)\\ =& \Phi\left(\frac{s - E(S)}{\sqrt{\text{Var}(S)}}\right)\end{aligned}$$

where Φ is the cdf of the standard normal distribution.

For the lognormal approximation, we solve the system of two equations $\mu + 0.5\sigma^2 = \ln[E(S)]$ and $2\mu + 2\sigma^2 = \ln[\text{Var}(S) + E(S)^2]$ for μ and σ^2. The probability of the aggregate losses is approximated as follows:

$$\begin{aligned}\Pr(S \le s) \approx& \Pr\left(\frac{\ln S - \mu}{\sigma} \le \frac{\ln s - \mu}{\sigma}\right)\\ =& \Phi\left(\frac{\ln s - \mu}{\sigma}\right).\end{aligned}$$

Example 45.1 ‡

A group life insurance contract covering independent lives is rated in the three age groupings as given in the table below. The insurer prices the contract so that the probability that the total claims will exceed the premium is 0.05.

Age group	Number in age group	Probability of claim per life	Mean of the exponential distribution of claim amounts
18-35	400	0.03	5
36-50	300	0.07	3
51-65	200	0.10	2

(a) Find the mean and the variance of the aggregate claim.
(b) Using the normal approximation, determine the premium that the insurer will charge.

Solution.
(a) The mean is given by

$$E(S) = 400(0.03)(5) + 300(0.07)(3) + 200(0.10)(2) = 163.$$

The variance is given by

$$\begin{aligned}\mathrm{Var}(S) =&400[(0.03)(5)^2 + (0.03)(0.97)5^2]\\ &+300[(0.07)(3)^2 + 0.07(0.93)(3)^2] + 200[0.10(2)^2 + 0.10(0.9)(2)^2]\\ =&1107.77.\end{aligned}$$

(b) We have

$$\begin{aligned}\Pr(S > P) =&0.05\\ 1 - \Pr(S > P) =&0.95\\ \Pr(S \le P) =&0.95\\ \Pr\left(Z \le \frac{P-163}{\sqrt{1107.77}}\right) =&0.95\\ \frac{P-163}{\sqrt{1107.77}} =&1.645.\end{aligned}$$

Solving the last equation, we find $P = 217.75$ ■

Example 45.2
Repeat the previous example by replacing the normal approximation with the lognormal approximation.

Solution.
Solving the system $\mu + 0.5\sigma^2 - 163$ and $2\mu + 2\sigma^2 = \ln 163$ and $2\mu + 2\sigma^2 = \ln(163^2 + 1107.77)$ we find $\mu = 5.073$ and $\sigma^2 = 0.041$. Thus,

$$\begin{aligned}\Pr(S > P) =&0.05\\ 1 - \Pr(S > P) =&0.95\\ \Pr(S \le P) =&0.95\\ \Pr\left(Z \le \frac{\ln P - 5.073}{\sqrt{0.041}}\right) =&0.95\\ \frac{\ln P - 5.073}{\sqrt{0.041}} =&1.645.\end{aligned}$$

Solving the last equation, we find $P = 222.76$ ■

Practice Problems

Problem 45.1 ‡
The probability model for the distribution of annual claims per member in a health plan is shown below. Independence of costs and occurrences among services and members is assumed.

Service	Probability of claim	Mean of claim dist.	Variance of claim dist.
office visits	0.7	160	4,900
Surgery	0.2	600	20,000
Other Services	0.5	240	8,100

Using the normal approximation, determine the minimum number of members that a plan must have such that the probability that actual charges will exceed 115% of the expected charges is less than 0.10.

Problem 45.2
An insurer has a portfolio consisting of 5 one-year life insurance policies grouped as follows:

i	q_i	b_i
1	0.32	500
2	0.10	1000
3	0.54	250
4	0.23	375
5	0.14	650

The insurer sets up an initial capital of consisting of 150% of its expected loss to cover its future obligations.

Use a normal approximation to find the probability that the company will not meet its obligation next year.

Problem 45.3
Repeat the previous problem by replacing the normal approximation with the lognormal approximation.

Problem 45.4 ‡
An insurance company is selling policies to individuals with independent future lifetimes and identical mortality profiles. For each individual, the probability of death by all causes is 0.10 and the probability of death due

to an accident is 0.01. Each insurance policy pays a benefit of 10 for an accidental death and 1 for non-accidental death. The company wishes to have at least 95% confidence that premiums with a relative security loading of 0.20 are adequate to cover claims. (In other words, the premium is $1.20E(S)$.)

Using the normal approximation, determine the minimum number of policies that must be sold.

Problem 45.5
Consider a portfolio of 100 independent life insurance policies. It is determined that the death benefit of each insured follows a Poisson distribution with paramaeter λ. The probability of a death is 0.1. Using the normal approximation, it has been estimated that the probability of the aggregate loss exceeding 25 is 0.05.

Determine the value of λ.

Review of Mathematical Statistics

In this chapter we review some concepts of mathematical statistics which is required in the sequel. Mainly, our attention will be focused on estimation and hypothesis testing.

46 Properties of Point Estimators

In model estimation, we are interested in estimating quantities related to a random variable X and determine properties of those estimates.

By an **estimate** we mean a single value obtained when an estimation procedure is applied to a set of values.

By an **estimator** or **statistic** we mean a formula or rule that produced the estimate. An example of an estimator is the sample mean of a random sample of values $\{x_1, x_2, \cdots, x_n\}$ given by

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

A **point estimation** is an estimation procedure that results of a single estimate which is served as the best estimate of an unknown parameter. The point estimator of a parameter θ will be denoted by $\hat{\theta}$ and is itself a random variable.

The first property of estimators is the property of unbiasedness. An estimator $\hat{\theta}$ is said to be **unbiased** if $E(\hat{\theta}) = \theta$ for all θ. That is, the mean of the estimator is just the quantity being estimated. Any estimator that is not unbiased is called **biased**. The bias associated to a parameter estimator is $\text{bias}_{\hat{\theta}}(\theta) = E(\hat{\theta}) - \theta$. That is, the difference between this estimator's expected value and the true value of the parameter being estimated. Note also that $E(\hat{\theta}) - \theta = E(\hat{\theta} - \theta)$ so that the bias is the expected value of the error.

Example 46.1
A population consists of the values 1,3,5, and 9. We want to estimate the mean of the population μ. A random sample of two values from this population is taken without replacement, and the mean of the sample $\hat{\mu}$ is used as an estimator of the population mean μ.
(a) Find the probability distribution of $\hat{\mu}$.
(b) Is $\hat{\mu}$ an unbiased estimator?

Solution.
(a) The various samples are $\{1, 3\}, \{1, 5\}, \{1, 9\}, \{3, 5\}, \{3, 9\}$, and $\{5, 9\}$ each occurring with probability of $\frac{1}{6}$. The following table provides the sample mean.

Sample	{1,3}	{1,5}	{1,9}	{3,5}	{3,9}	{5,9}
$\hat{\mu}$	2	3	5	4	6	7

(b) We have

$$\mu = \frac{1+3+5+9}{4} = 4.5$$

and

$$E(\hat{\mu}|\mu) = \frac{2+3+5+4+6+7}{6} = 4.5 = \mu.$$

Hence, the estimator is unbiased ■

Example 46.2
Let $X_1, X_2, \cdots, X_n$ be normally distributed random variables each with mean μ and variance σ^2. Show that the estimator of the variance

$$\hat{\sigma}^2 = \sum_{i=1}^n \frac{(X_i-\hat{\mu})^2}{n} \quad \text{where} \quad \hat{\mu} = \frac{X_1+X_2+\cdots+X_n}{n}$$

is biased.

Solution.
The estimator can be expressed as

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \hat{\mu}^2.$$

Now, we have the following calculation,

$$\begin{aligned}E(\hat{\sigma}^2) =&E\left[\frac{1}{n}\sum_{i=1}^n X_i^2 - \hat{\mu}^2\right] = \frac{1}{n}\sum_{i=1}^n(\sigma^2+\mu^2) - \left(\frac{\sigma^2}{n}+\mu^2\right)\\ =&\sigma^2+\mu^2-\frac{\sigma^2}{n}-\mu^2 = \frac{(n-1)}{n}\sigma^2 \neq \sigma^2.\end{aligned}$$

Hence, $\hat{\sigma}^2$ is biased ■

Now when an estimator exhibits a small amount of bias that vanishes as the sample size increases without bound then we say that the estimator is **asymptotically unbiased**. Symbolically, we have

$$\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$$

for all θ.

Example 46.3
Let $X_1, X_2, \cdots, X_n$ be uniform random variables on $(0, \theta)$. Show that the estimator

$$\hat{\theta}_n = \max\{X_1, X_2, \cdots, X_n\}$$

is asymptotically unbiased.

Solution.
The cdf of $\hat{\theta}_n$ is expressed as

$$F_{\hat{\theta}_n}(y) = \Pr(X_1 \le y, X_2 \le y, \cdots, X_n \le y) = \left(\frac{y}{\theta}\right)^n.$$

Hence, the pdf is

$$f_{\hat{\theta}_n}(y) = \frac{ny^{n-1}}{\theta^n},\ 0 < y < \theta$$

and the expected value of $\hat{\theta}_n$ is

$$E(\hat{\theta}_n) = \int_0^\theta ny^n\theta^{-n}dy = \frac{n}{n+1}\theta \to \theta$$

as $n \to \infty$. Thus, the estimator $\hat{\theta}_n$ is asymptotically unbiased ■

The next property of estimators that we consider is the property of consistency: An estimator is (weakly)**consistent** if the probability of the error being greater than a small amount is zero as the size of the sample goes to infinity. Symbolically,

$$\lim_{n\to\infty} \Pr(|\hat{\theta}_n - \theta| > \delta) = 0,\ \forall \delta > 0.$$

Theorem 46.1
If $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$ and $\lim_{n\to\infty} \mathrm{Var}(\hat{\theta}_n) = 0$ then $\hat{\theta}_n$ is consistent.

Proof.
Let $\delta > 0$. Then by Chebyshev's inequality (see p.405 of [2]) we can write

$$0 \le \Pr(|\hat{\theta}_n - E(\hat{\theta}_n| \ge 2\delta > \delta) \le \frac{\mathrm{Var}(\hat{\theta}_n)}{4\delta^2}.$$

Letting $n \to \infty$ and using the squeeze rule of limits the result follows ■

Example 46.4
Let $X_1, X_2, \cdots, X_n$ be uniform random variables on $(0, \theta)$. Show that the estimator

$$\hat{\theta}_n = \max\{X_1, X_2, \cdots, X_n\}$$

is weakly consistent.

Solution.
We have already shown that $\hat{\theta}_n$ is asymptotically unbiased. It remains to show that its variance goes to 0 as n goes to infinity. We have

$$\begin{aligned} E(\hat{\theta}_n^2) =& \int_0^\theta n y^{n+1} \theta^{-n} dy = \frac{n}{n+2}\theta^2 \\ \mathrm{Var}(\hat{\theta}_n) =& \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 \\ =& \frac{n\theta^2}{(n+2)(n+1)^2} \to 0 \end{aligned}$$

as $n \to \infty$. That, is $\hat{\theta}$ is consistent ■

The third property of estimators that we consider is the mean-squared error: The **mean-squared error (MSE)** is the second moment (about the origin) of the error. That is,

$$\mathrm{MSE}_{\hat{\theta}}(\theta) = E[(\hat{\theta} - \theta)^2].$$

That is, the mean-squared error is an average of the squares of the difference between the actual observations and those estimated. The mean-squared error is arguably the most important criterion used to evaluate the performance of an estimator. The MSE incorporates both the variance of the estimator and its bias. Indeed, we have

$$\begin{aligned} \mathrm{MSE}_{\hat{\theta}}(\theta) =& E[(\hat{\theta} - \theta)^2] = E(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2) \\ =& E(\hat{\theta}^2 - 2\theta E(\hat{\theta}) + \theta^2 \\ =& \mathrm{Var}(\hat{\theta}) + E(\hat{\theta})^2 - 2\theta E(\hat{\theta}) + \theta^2 \\ =& \mathrm{Var}(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 = \mathrm{Var}(\hat{\theta}) + [\mathrm{bias}_{\hat{\theta}}(\theta)]^2. \end{aligned}$$

Since the MSE decomposes into a sum of the bias and variance of the estimator, both quantities are important and need to be as small as possible to achieve good estimation performance. It is common to trade-off some increase in bias for a larger decrease in the variance and vice-versa.

Example 46.5
Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed randm variables with common distribution a normal distribution with parameters (μ, σ^2). Find $\text{MSE}_{\hat{\mu}}(\mu)$.

Solution.
We have

$$\begin{aligned}\text{MSE}_{\hat{\mu}}(\mu) =& \text{Var}(\hat{\mu}) + [\text{bias}_{\hat{\mu}}(\mu)]^2 = \text{Var}(\hat{\mu})\\ =& \text{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}(X_i) = \frac{\sigma^2}{n} \blacksquare\end{aligned}$$

Given two unbiased estimators $\hat{\theta}$ and $\hat{\theta}'$ of θ, we say that $\hat{\theta}$ is **more efficient** than $\hat{\theta}'$ if $\text{Var}(\hat{\theta}) < \text{Var}(\hat{\theta}')$. Note that for an unbiased estimator $\hat{\theta}$, we have $\text{MSE}_{\hat{\theta}}(\theta) = \text{Var}(\hat{\theta})$. If the estimator $\hat{\theta}$ satisfies the property $\text{Var}(\hat{\theta}) < \text{Var}(\hat{\theta}')$ where $\hat{\theta}'$ is another unbiased estimator of θ then we call $\hat{\theta}$ a **uniformly minimum variance unbiased estimator(UMVUE)**. That is, $\hat{\theta}$ is the most efficient unbiased estimator of θ.

Example 46.6
Let $X_1, X_2, \cdots, X_n$ be independent uniform random variables with parameters $(0, \theta)$. Consider the two unbiased estimators

$$\hat{\theta}_a = 2\hat{\mu} \text{ and } \hat{\theta}_b = \tfrac{n+1}{n}\max\{X_1, X_2, \cdots, X_n\}.$$

Determine which one is more efficient than the other.

Solution.
We have

$$\text{Var}(\hat{\theta}_a) = 4\text{Var}(\hat{\mu}) = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

and

$$\text{Var}(\hat{\theta}_b) = \left(\frac{n+1}{n}\right)^2 \text{Var}(\hat{\theta}_n) = \left(\frac{n+1}{n}\right)^2 \frac{n}{(n+1)^2(n+2)}\theta^2.$$

Thus,

$$\text{Var}(\hat{\theta}_a) - \text{Var}(\hat{\theta}_b) = \frac{n-1}{3n(n+2)} > 0, \; n > 1.$$

Thus, $\hat{\theta}_b$ is more efficient than $\hat{\theta}_a$ $\blacksquare$

Practice Problems

Problem 46.1 ‡
You ar given:
- $E(X) = \theta > 0$.
- $\mathrm{Var}(X) = \frac{\theta^2}{25}$.
- $\hat{\theta} = \frac{kX}{k+1}$.
- $\mathrm{MSE}_{\hat{\theta}}(\theta) = 2[\mathrm{bias}_{\hat{\theta}}(\theta)]^2$.

Determine the value of k.

Problem 46.2 ‡
You are given two independent estimates of an unknown quantity θ :
a. Estimator $A : E(\hat{\theta}_A) = 1000$ and $\sigma(\hat{\theta}_A) = 400$.
b. Estimator $B : E(\hat{\theta}_B) = 1200$ and $\sigma(\hat{\theta}_B) = 200$.
Estimator C is a weighted average of Estimator A and Estimator B such that

$$\hat{\theta}_C = w\hat{\theta}_A + (1-w)\hat{\theta}_B.$$

Determine the value of w that minimizes $\sigma(\hat{\theta}_C)$.

Problem 46.3 ‡
Claim sizes are uniformly distributed over the interval $[0, \theta]$. A sample of 10 claims, denoted by $X_1, , X_2, \cdots, X_{10}$ was observed and an estimate of θ was obtained using

$$\hat{\theta} = \max\{X_1, X_2, \cdots, X_{10}\}.$$

Recall that the probability density function for $\hat{\theta}$ is

$$f_{\hat{\theta}}(y) = \frac{10y^9}{\theta^{10}}.$$

Calculate the mean-squared error for $\hat{\theta}$ for $\theta = 100$.

Problem 46.4 ‡
A random sample, $X_1, X_2, \cdots, X_n$ is drawn from a distribution with a mean of $2/3$ and a variance of $1/18$. An estimator of the distribution mean is given by

$$\hat{\mu} = \frac{X_1 + X_2 + \cdots + X_n}{n-1}.$$

Find $\mathrm{MSE}_{\hat{\mu}}(\mu)$.

Problem 46.5
A random sample of 10 independent values were taken from a Pareto distribution with parameters $\alpha = 3$ and θ. The estimator used to estimate the mean of the distribution is given by

$$\hat{\mu} = \frac{X_1 + X_2 + \cdots + X_{10}}{10}.$$

It is given that $\text{MSE}_{\hat{\mu}}(\mu) = 300$. Determine the value of θ.

Problem 46.6
Let $X_1, X_2, \cdots, X_n$ be uniform random variables on $(0, \theta)$. The parameter θ is estimated using the estimator

$$\hat{\theta}_n = \frac{2n}{n-1}\mu$$

where

$$\mu = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Find the bias of $\hat{\theta}$.

Problem 46.7
Let $\hat{\theta}$ denote an estimator of a parameter θ. Suppose that $E(\hat{\theta}) = 3$ and $E(\hat{\theta}^2) = 17$.

Calculate $\text{MSE}_{\hat{\theta}}(5)$.

Problem 46.8 ‡
Which of the following statements is true?

(A) A uniformly minimum variance unbiased estimator is an estimator such that no other estimator has a smaller variance.
(B) An estimator is consistent whenever the variance of the estimator approaches zero as the sample size increases to infinity.
(C) A consistent estimator is also unbiased.
(D) For an unbiased estimator, the mean squared error is always equal to the variance.
(E) One computational advantage of using mean squared error is that it is not a function of the true value of the parameter.

47 Interval Estimation

The next estimation procedure that we consider is the **interval estimation**.

Point estimation provides a single value as an estimate to the true value. In contrast, interval estimation provides a range of values, any of which is likely to be the true value.

A specific type of interval estimator is the confidence interval: A level $100(1-\alpha)\%$, $0 < \alpha < 1$, **confidence interval** for a population parameter θ is an interval $[L, U]$ computed from a sample data such that $\Pr(L \leq \theta \leq U) \geq 1-\alpha$ for all θ. The number $1-\alpha$ is called the **level of confidence**. This says that $100(1-\alpha)\%$ of all random samples yield intervals containing the true value of the parameter.

Example 47.1
Let $X_1, X_2, \cdots, X_n$ be a sample from a normal distribution with a known standard deviation σ but unknown mean μ. Find $100(1-\alpha)\%$ confidence interval of μ.

Solution.
Let $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ by an estimator of the mean. Then the random variable $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ has the standard normal distribution. But the standard normal distribution is symmetric about 0 so we choose $L = -U$ so that

$$\Pr\left(-U \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq U\right) = 1-\alpha.$$

We choose U as $U = z_{\frac{\alpha}{2}}$ where $z_{\frac{\alpha}{2}}$ is the $100\left(1-\frac{\alpha}{2}\right)$th percentile of the standard normal distribution. That is,

$$\int_{-\infty}^{z_{\frac{\alpha}{2}}} \phi(x)dx = 1 - \frac{\alpha}{2}$$

where $\phi(x)$ is the pdf of the standard normal distribution. Note that $z_{\frac{\alpha}{2}}$ is the area defined as the value of the normal random variable Z such that the area to its right is $\frac{\alpha}{2}$. Hence, a confidence interval is given by

$$\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \blacksquare$$

Example 47.2
For a 95% confidence interval, what is $z_{\frac{\alpha}{2}}$?

Solution.
We have $1-\alpha = 0.95$ so that $\alpha = 0.05$. Thus, $z_{\frac{\alpha}{2}}$ satisfies

$$\Phi(z_{\frac{\alpha}{2}}) = \int_0^{z_{\frac{\alpha}{2}}} \phi(x)dx = 1 - 0.025 = 0.975.$$

Using the table for the standard normal distribution, we find $z_{0.025} = 1.96$ ■

In general, confidence intervals are constructed as follows: Suppose that $\hat{\theta}$ is a point estimator of the population parameter θ such that $E(\hat{\theta}) \approx \theta$ and $\mathrm{Var}(\hat{\theta}) \approx v(\theta)$ with $\hat{\theta}$ being approximated by a normal distribution. At a later section, we will show that for a confidence interval $100(1-\alpha)\%$, we have

$$1-\alpha \approx \Pr\left(-z_{\frac{\alpha}{2}} \le \frac{\hat{\theta}-\theta}{\sqrt{v(\theta)}} \le z_{\frac{\alpha}{2}}\right).$$

Example 47.3
Construct a 95% confidence interval for the mean of a Poisson distribution.

Solution.
Let $\theta = \lambda$, the mean of the Poisson distribution. Let the point estimator be $\hat{\theta} = \overline{X}$. By the Central Limit Theorem, $\overline{X}$ is normally distributed with $E(\hat{\theta}) = E(X) = \theta$ and $\mathrm{Var}(\hat{\theta}) = \frac{\mathrm{Var}(X)}{n} = \frac{\theta}{n}$. Thus, we want

$$0.95 \approx \Pr\left(-1.96 \le \frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}} \le 1.96\right).$$

Thus, we want

$$\left|\frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}}\right| \le 1.96$$

which is equivalent to

$$(\overline{X}-\theta)^2 \le \frac{3.84160\theta}{n}$$

or

$$\theta^2 - \left(2\overline{X} + \frac{3.8416}{n}\right)\theta + \overline{X}^2 \le 0.$$

Solving this quadratic inequality, we find the confidence interval with endpoints

$$\left[\overline{X} \pm \frac{1.9208}{n} - \frac{1}{2}\sqrt{\frac{15.3664\overline{X} + 3.8416^2/n}{n}}\right] \blacksquare$$

Practice Problems

Problem 47.1
For a 90% confidence interval, what is $z_{\frac{\alpha}{2}}$?

Problem 47.2
For a 99% confidence interval, what is $z_{\frac{\alpha}{2}}$?

Problem 47.3
You are given that $\Pr(L \leq \theta \leq U) \geq 0.80$. What is the probability that a confidence interval will not include the population parameter?

Problem 47.4
You are given:
• a population has a normal distribution with mean μ and $\sigma = 3$.
• $\overline{X} = 80$
• $n = 10$

Construct a 95% confidence interval.

Problem 47.5 ‡
A sample of 2000 policies had 1600 with no claims and 400 with one or more claims. Using the normal approximation, determine the symmetric 95% confidence interval for the probability that a single policy has one or more claims.

48 Hypothesis Testing

Within the context of this book, a **statistical hypothesis** is a claim or a statement regarding a population parameter or a probability distribution.

To test the validity of a hypothesis with absolute certainty requires the examination of the entire population and this is hard to do in general. Instead, a **hypothesis testing** is used in collecting evidence from a random sample and based on the evidence one can judge whether to support the claim or not.

Hypothesis testing is formulated based on two hypotheses:

• The **null hypothesis** denoted by H_0 is the claim that is assumed to be true. The conclusion of any hypothesis testing is either to reject the null hypothesis or fail to reject because one does not have enough evidence, based on a sample, to reject it. Strangely, one will never accept the null hypothesis. The null hypothesis typically includes a statement of equality such as "=", "$\geq$", or "$\leq$".

• The **alternative hypothesis** denoted by H_1 (or H_a) is the one to be tested and is the claim that is oppose it to the null hypothesis. The alternative hypothesis typically includes a statement of strict inequality or not equal: "$>$", "$<$", "$\neq$".

Example 48.1
An insurance company is reviewing its current policy rates. Originally, the insurance company believed that the average claim amount should be \$1,200. Currently, they are suspecting that the true mean is actually higher than this. What are the hypothesis for this problem?

Solution.
The null and alternative hypothesis for this problem are:

$$\begin{aligned} H_0 :& \mu \leq 1,200 \\ H_1 :& \mu > 1,200 \blacksquare \end{aligned}$$

The testing method used in making the decision to either reject or not to reject the null hypothesis involves two concepts: A test statistic and a rejection region.

A **test statistic** is a random variable represented by a numerical value obtained from the random sample and this value is used (see below) to determine whether to reject or fail to reject the null hypothesis. The random samples are assumed to have a large number of data so that the data values have a normal distribution thanks to the Central Limit Theorem. Thus, a common test statistic is the Z score.

A **rejection region** is the set of all values of the test statistic for which the null hypothesis will be rejected,i.e., values that provide strong evidence in favor of the alternative hypothesis.

The boundaries of the rejection region are called the **critical values**. In the case of an alternative hypothesis with $>$ sign the rejection region lies in the right-tail of the distribution of the test statistic, with $<$ the tail is left, and with $\neq$ sign the region is two-tailed.

Because we are making a decision based on a finite sample, there is a possibility that we will make mistakes. The possible outcomes are listed in the chart:

	H_0 is true	H_1 is true
Do not reject H_0	Correct decision	Type II error
reject H_0	Type I error	correct decision

It is possible that we reject a null hypothesis in a situation where it happens to be true. In this case, we make an error which is referred to as **Type I error**. The maximum probability of making a Type I error given that the null hypothesis is true is called the **level of confidence** and is denoted by α. The lower significance level α, the less likely we are to commit a type I error. The level of confidence is usually set in advance by the experimenter and is often between 1% and 10%. So for example, if $\alpha = 0.05$, there is a 5% chance that, when the null hypothesis is true, we will erroneously reject it.

Example 48.2
Find the rejection region for a level of significance $\alpha = 0.01$ when the alternative hypothesis has the $\neq$ sign.

Solution.
We need to put half of α in the left tail, and the other half of α in the right

tail. The rejection region is the two-tailed region of the $Z-$distribution: $|Z| > 2.58$. The critical values are -2.58 and 2.58 ■

Example 48.3
An insurance company is reviewing its current policy rates. When originally setting the rates they believed that the average claim amount was $1,800. They are concerned that the true mean is actually higher than this, because they could potentially lose a lot of money. They randomly select 40 claims, and calculate a sample mean of $1,950. Assuming that the standard deviation of claims is $500, and set $\alpha = 0.025$, test to see if the insurance company should be concerned.

Solution.
The null and alternative hypothesis for this problem are:

$$\begin{aligned}H_0 :&\mu \leq 1,800\\ H_1 :&\mu > 1,800.\end{aligned}$$

The test statistic of the given sample is

$$z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{1950 - 1800}{500/\sqrt{40}} = 1.897.$$

Since

$$\Pr(Z > 1.96|\mu \leq 1,800) = 0.025$$

the rejection region is $Z > 1.96$. Since $1.897 < 1.96$, we fail to reject the null hypothesis. We cannot conclude anything statistically significant from this test, and cannot tell the insurance company whether or not they should be concerned about their current policies ■

Failure to reject H_0 when H_1 is true is called a **Type II error**. The probability of committing a type II error is denoted by β.
α and β are not independent of each other−as one increases, the other decreases. However, increases in the sample size cause both to decrease, since sampling error is reduced.

Example 48.4
Consider the following hypotheses:

$$\begin{aligned}H_0 :&\mu \geq 3\\ H_1 :&\mu < 3.\end{aligned}$$

You throw a fair die. If the face shows a 3 or less then you reject H_0. If the face shows a number greater than 3 you fail to reject H_0 or H_1 is true.
(a) Calculate the level of significance α.
(b) Calculate the probability of Type II error.

Solution.
(a) We have

$$\begin{aligned}\alpha =&\Pr(H_0 \text{ is true but is rejected})\\ =&\Pr(X \le 3) = \frac{3}{6} = \frac{1}{2}.\end{aligned}$$

(b) We have

$$\begin{aligned}\beta =&\Pr(H_1 \text{ is true but fail to rejected } \mathrm{H}_0)\\ =&\Pr(X > 3) = \frac{3}{6} = \frac{1}{2} \blacksquare\end{aligned}$$

The rejection region is one way to test hypotheses where you compare the test statistic to a critical value. Another method for testing hypothesis which allows us to draw a conclusion without calculating critical values is the $p-$value method: A $p-$**value** or **probability value** is the probability of a more extreme departure from the null hypothesis than the observed data. The $p-$value is the area under the normal curve.

By comparing the $p-$value to the alpha level we can easily decide to reject or fail to reject: If $p-$value $> \alpha$, fail to reject H_0. If $p-$value $\le \alpha$, reject H_0.

How do we calculate the $p-$value? One method is as follows:

1. Determine the value of the test statistic z corresponding to the result of the sampling experiment.

2. (a) If the test is one-tailed, the $p-$value is equal to the tail area beyond z in the same direction as the alternative hypothesis. Thus, if the alternative hypothesis is of the form $>$, the $p-$value is the area to the right of, or above the observed z value. The same is true in the case of $<$.

(b) If the test is two-tailed, the $p-$value is equal to twice the tail area beyond the observed z value in the direction of the sign of z. That is, if z is positive,the $p-$value is twice the area to the right of, or above the observed z value. The same holds true in the case where z is negative.

Example 48.5
Find the $p-$value of Example 48.3.

Solution.
The $p-$value is $2\Pr(Z > 1.897)$. This is equivalent to twice the area under the standard normal curve to the right of $z = 1.897$. Looking in a standard normal table we find that $z = 1.897 \approx 1.90$ corresponds to 0.0287. Since $2(0.0287) = 0.0574 > \alpha = 0.05$, we fail to reject the null hypothesis ■

Example 48.6
You are performing a hypothesis test as follows:

$$H_0 : X \text{ follows an exponential distribution mean } 0.1$$
$$H_1 : X \text{ follows an exponential distribution mean } 0.2$$

You pick a random value of X. This is x, your test statistic. Your test statistic in this case is 2. What is the $p-$value of this test?

Solution.
Recall that the $p-$value is the probability that, if the null hypothesis is true, a higher value than the test statistic is observed.The sdf of X if H_0 is true is $S(x) = e^{-10x}$. Thus, the $p-$value is $S(4) = e^{-10(2)} = e^{-20}$ ■

Practice Problems

Problem 48.1
Suppose the current mean cost to treat a cancer patient for one month is \$18,000. Consider the following scenarios.
(a) A hospital treatment plan is implemented which hospital authorities feel will reduce the treatment costs.
(b) It is uncertain how a new treatment plan will affect costs.
(c) A treatment plan is implemented which hospital authorities feel will increase treatment costs.
Let μ represent the mean cost per patient per month after the new treatment plan is implemented.

Give the research hypothesis in symbolic form for each of the above cases

Problem 48.2
Classify each of the following as a lower-tailed, upper-tailed, or two-tailed rejection region:
(i) $H_0 : \mu = 12$ and $H_1 : \mu \neq 12$.
(ii) $H_0 : \mu \geq 12$ and $H_1 : \mu < 12$.
(iii) $H_0 : \mu = 12$ and $H_1 : \mu > 12$.

Problem 48.3
Consider a college campus where finding a parking space is not easy. The university claims that the average time spent in finding a parking space is at least 30 minutes. Suppose you suspect that it takes less than that. So in a sample of five, you found that the average time is 20 minutes. Assuming that the time it takes to find a parking spot is normal with standard deviation $\sigma = 6$ minutes, then perform a hypothesis test with level of significance $\alpha = 0.01$ to see if your claim is correct.

Problem 48.4
A hypothesis test has a p$-$value of 0.037. At which of these significance levels would you reject the null hypothesis?

(i) 0.025 (ii) 0.074 (iii) 0.05 (iv) 0.042.

Problem 48.5
A new restaurant has opened in town. A statistician claims that the amount spent per customer for dinner is more than \$20. To verify whether his claim is valid or not, he randomly selected a group of 49 custimers and found that

the average amount spent was \$22.60. Assume that the standard deviation is known to be \$2.50.

Using $\alpha = 2\%$, would he conclude the typical amount spent per customer is more than \$20.00?

Problem 48.6
Suppose a production line operates with a mean filling weight of 16 ounces per container. Since over$-$ or under$-$filling can be dangerous, a quality control inspector samples 30 items to determine whether or not the filling weight has to be adjusted. The sample revealed a mean of 16.32 ounces. From past data, the standard deviation is known to be 0.8 ounces.

Using a 0.10 level of significance, can it be concluded that the process is out of control (not equal to 16 ounces)?

Problem 48.7
A dietician is trying to test the claim that a new diet plan will cause a person to lose 10 lbs. over 4 weeks. To test her claim, she selects a random sample of 49 overweighted individulas and found that an average weight loss of 12.5 pounds over the four weeks, with $\sigma = 7\ lbs$.

Identify the critical value suitable for conducting a two-tail test of the hypothesis at the 2% level of significance.

Problem 48.8
A Type II error is committed when
(a) we reject a null hypothesis that is true.
(b) we don't reject a null hypothesis that is true.
(c) we reject a null hypothesis that is false.
(d) we don't reject a null hypothesis that is false.

The Empirical Distribution for Complete Data

The focus of this chapter is on estimating distributional quantities in data-dependent models. Recall from Section 17 the two types of data-dependent distributions: The empirical distribution and the kernel-smoothed distribution. In what follows, by a **complete data** we mean a set of outcomes with known exact values. When the exact value of an outcome in not known then the outcome belongs to an **incomplete data**.

49 The Empirical Distribution for Individual Data

In Section 17 we defined the empirical model as follows: **The empirical distribution** is obtained by assigning a probability of $\frac{1}{n}$ to each data point in a sample with n individual data points.

The **empirical distribution function**(edf) is

$$F_n(x) = \frac{\text{number of elements in the sample that are } \leq \text{x}}{n}.$$

Example 49.1
Below are the losses suffered by policyholders of an insurance company:

$$49, 50, 50, 50, 60, 75, 80, 120, 130.$$

Let X be the random variable representing the losses incurred by the policyholders. Find the empirical distribution probability function and the empirical distribution function of X.

Solution.
The pmf is given by the table below.

x	49	50	60	75	80	120	130
$p(x)$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

The edf is defined by

$$F_X(x) = \frac{\text{number of elements in the sample that are } \leq x}{9}.$$

Thus, for example,

$$F_X(73) = \frac{5}{9} \blacksquare$$

In order to define empirical estimates, we need to introduce some additional notation. Because of possible duplications of values, we re-define the sample by considering the k distinct values arranged in increasing order as follows

$$y_1 < y_2 < y_3 < \cdots < y_k, \ k \leq n.$$

Let s_i denote the number of times the value y_i appears in the sample. Clearly, $\sum_{i=1}^{k} s_i = n$. Next, for each $1 \leq j \leq k$, let $r_j = \sum_{i=j}^{k} s_i$. That is, r_j is the number of observations greater than or equal to y_j. The set of observations greater than or equal to y_j is called the **risk set**[11]. The

[11] When listing the elements of this set, repeated observations must be listed

convention is to call r_j also as the risk set. Using the notation of r_j, we can express the edf as follows

$$F_n(x) = \begin{cases} 0, & x < y_1 \\ 1 - \frac{r_j}{n}, & y_{j-1} \le x < y_j,\ j = 2, \cdots, k \\ 1, & x \ge y_k. \end{cases}$$

Example 49.2
Determine the edf of Example 49.1 using the previous paragraph.

Solution.
We have the following chart:

j	y_j	s_j	r_j
1	49	1	9
2	50	3	8
3	60	1	5
4	75	1	4
5	80	1	3
6	120	1	2
7	130	1	1

Using the above chart, we find

$$F_n(x) = \begin{cases} 0, & x < 49 \\ 1 - \frac{8}{9} = \frac{1}{9}, & 49 \le x < 50 \\ 1 - \frac{5}{9} = \frac{4}{9}, & 50 \le x < 60 \\ 1 - \frac{4}{9} = \frac{5}{9}, & 60 \le x < 75 \\ 1 - \frac{3}{9} = \frac{2}{3}, & 75 \le x < 80 \\ 1 - \frac{2}{9} = \frac{7}{9}, & 80 \le x < 120 \\ 1 - \frac{1}{9} = \frac{8}{9}, & 120 \le x < 130 \\ 1, & x \ge 130 \blacksquare \end{cases}$$

Since the empirical model is a discrete model, the derivative required to create the density and hazard rate functions cannot be taken. The best one can do is to estimate the **cumulative hazard rate function** defined by:

$$H(x) = -\ln S(x) = -\ln[1 - F(x)].$$

Note that once an estimate of $H(x)$ is found, we can find estimates for $F(x) = 1 - e^{-H(x)}$. An estimate of the cumulative hazard rate function is the **Nelson-Åalen estimate** given by:

$$\hat{H}(x) = \begin{cases} 0, & x < y_1 \\ \sum_{i=1}^{j-1} \frac{s_i}{r_i}, & y_{j-1} \le x < y_j,\ j = 2, 3, \cdots, k \\ \sum_{i=1}^{k} \frac{s_i}{r_i}, & x \ge y_k. \end{cases}$$

Example 49.3
Determine the **Nelson-Åalen** estimate for Example 49.1.

Solution.
We have

$$\hat{H}(x) = \begin{cases} 0, & x < 49 \\ \frac{1}{9}, & 49 \le x < 50 \\ \frac{1}{9} + \frac{3}{8} = \frac{35}{72}, & 49 \le x < 50 \\ \frac{35}{72} + \frac{1}{5} = \frac{247}{360}, & 60 \le x < 75 \\ \frac{247}{360} + \frac{1}{4} = \frac{337}{360}, & 75 \le x < 80 \\ \frac{337}{360} + \frac{1}{3} = \frac{457}{360}, & 80 \le x < 120 \\ \frac{457}{360} + \frac{1}{2} = \frac{637}{360}, & 120 \le x < 130 \\ \frac{637}{360} + \frac{1}{1} = \frac{997}{360} & x \ge 130 \blacksquare \end{cases}$$

Example 49.4 ‡
A portfolio of policies has produced the following claims:

100 100 100 200 300 300 300 400 500 600

Determine the empirical estimate of $H(300)$.

Solution.
We have

j	y_j	s_j	r_j
1	100	3	10
2	200	1	7
3	300	3	6
4	400	1	3
5	500	1	2
6	600	1	1

Using the above chart, we find

$$\hat{S}(300) = \frac{s_3}{n} = \frac{3}{10}.$$

Thus,

$$\hat{H}(300) = -\ln[\hat{S}(300)] = -\ln(0.3) = 1.204 \blacksquare$$

Practice Problems

Problem 49.1
Twelve policyholders were monitored from the starting date of the policy to the time of first claim. The observed data are as follows:

Time of first claim	1	2	3	4	5	6	7
Number of claims	2	1	2	2	1	2	2

Calculate $p_{12}(x)$ and $F_{12}(x)$.

Problem 49.2
Twelve policyholders were monitored from the starting date of the policy to the time of first claim. The observed data are as follows:

Time of first claim	1	2	3	4	5	6	7
Number of claims	2	1	2	2	1	2	2

Find the empirical mean and the empirical variance.

Problem 49.3
Twelve policyholders were monitored from the starting date of the policy to the time of first claim. The observed data are as follows:

Time of first claim	1	2	3	4	5	6	7
Number of claims	2	1	2	2	1	2	2

(a) Find the cumulative hazard function from the **Nelson-Åalen** estimate.
(b) Find the survival distribution function from the **Nelson-Åalen** estimate.

Problem 49.4
Below are the losses suffered by policyholders of an insurance company:

$$49, 50, 50, 50, 60, 75, 80, 120, 230.$$

Let X be the random variable representing the losses incurred by the policyholders. Find the empirical survival function.

Problem 49.5
Below are the losses suffered by policyholders of an insurance company:

$$49, 50, 50, 50, 60, 75, 80, 120, 130.$$

Let X be the random variable representing the losses incurred by the policyholders. For the observation 50, find the number of elements in the associated risk set.

Problem 49.6 ‡
You are given a random sample of 10 claims consisting of two claims of 400, seven claims of 800, and one claim of 1600. Determine the empirical skewness coefficient.

Problem 49.7 ‡
You are given the following about 100 insurance policies in a study of time to policy surrender:
(i) The study was designed in such a way that for every policy that was surrendered, a new policy was added, meaning that the risk set, r_j, is always equal to 100.
(ii) Policies are surrendered only at the end of a policy year.
(iii) The number of policies surrendered at the end of each policy year was observed to be:

$$\begin{gathered} 1 \text{ at the end of the 1st policy year} \\ 2 \text{ at the end of the 2nd policy year} \\ 3 \text{ at the end of the 3rd policy year} \\ \vdots \\ n \text{ at the end of the } n\text{th policy year.} \end{gathered}$$

(iv) The Nelson-Åalen empirical estimate of the cumulative distribution function at time n is $\hat{F}(n) = 0.542$.

What is the value of n?

50 Empirical Distribution of Grouped Data

Empirical distribution of grouped data is not possible. However, an estimate of the empirical distribution can be obtained. For that purpose, denote the grouped data by the intervals

$$(c_0, c_1], (c_1, c_2], \cdots, (c_{k-1}, c_k], (c_k, \infty)$$

where c_k is the maximum value of the range and $c_0 = 0$.

Now, let n_j denote the number of observations in the interval $(c_{j-1}, c_j]$. Clearly, $\sum_{j=1}^{n} n_j = n$, where n is the total number of observations in $(c_0, c_k]$. For such a grouped data, the empirical distribution is determined at the boundary points via the formula

$$F_n(c_j) = \frac{1}{n}\sum_{i=1}^{j} n_i, \; j = 1, \cdots, k$$

where $F_n(c_0) = 0$. With this notation, the empirical distribution can be obtained by connecting the group boundaries with straight lines via the formula

$$F_n(x) = \frac{c_j - x}{c_j - c_{j-1}} F_n(c_{j-1}) + \frac{x - c_{j-1}}{c_j - c_{j-1}} F_n(c_j), \; c_{j-1} \le x \le c_j.$$

The graph of the resulting distribution is called an **ogive**.

Example 50.1
Given the following grouped data.

Interval	Number of observations
$(0, 2]$	25
$(2, 10]$	10
$(10, 100]$	10
$(100, 1000]$	5

Find the empirical distribution function of the ogive corresponding to this data set.

Solution.
We have $c_0 = 0, \; c_1 = 2, \; c_2 = 10, \; c_3 = 100$, and $c_4 = 1000$. We first evaluate

the empirical distribution at the boundary points:

$$\begin{aligned}
F_{50}(0) =&0\\
F_{50}(2) =&\frac{25}{50}=0.5\\
F_{50}(10) =&\frac{35}{50}=0.7\\
F_{50}(100) =&\frac{45}{50}=0.9\\
F_{50}(1000) =&1.
\end{aligned}$$

The empirical cdf is

$$F_{50}(x)=\begin{cases}\frac{x}{4}, & 0\le x\le 2\\ \frac{9}{20}+\frac{x}{40}, & 2\le x\le 10\\ \frac{61}{90}+\frac{x}{450}, & 10\le x\le 100\\ \frac{8}{9}+\frac{x}{9000}, & 100\le x\le 1000\\ undefined, & x>1000 \blacksquare\end{cases}$$

Note that $F_n'(x)$ exists for all $x\neq c_j,\ j=0,1,\cdots,k$. Therefore, the density function can be obtained via the formula

$$f_n(x)=\frac{F_n(c_j)-F_n(c_{j-1})}{c_j-c_{j-1}}=\frac{n_j}{n(c_j-c_{j-1})},\ c_{j-1}\le x<c_j,\ j=1,2,\cdots,k.$$

The resulting graph is called a **histogram**.

Example 50.2
Find the density function in Example 50.1.

Solution.
The density function is

$$f_{50}(x)=\begin{cases}\frac{1}{4}, & 0\le x<2\\ \frac{1}{40}, & 2\le x<10\\ \frac{1}{450}, & 10\le x<100\\ \frac{1}{9000}, & 100\le x<1000\\ undefined, & x\ge 1000 \blacksquare\end{cases}$$

Example 50.3
Find $E(X\wedge 250)$ in Example 50.1

Solution.
We have

$$\begin{aligned}
E(X\wedge 250) =& \int_0^{250} xf_{50}(x)dx + 250[1-F_{50}(250)]\\
=& \int_0^2 \frac{x}{4}dx + \int_2^{10}\frac{x}{40}dx + \int_{10}^{100}\frac{x}{450}dx\\
+& \int_{100}^{250}\frac{x}{9000}dx + 250[1-0.9167]\\
=&0.5+1.2+11+2.9167+20.825\\
=&36.4417 \blacksquare
\end{aligned}$$

Example 50.4 ‡
You are given

Claim size	Number of claims
$(0, 25]$	25
$(25, 50]$	28
$(50, 100]$	15
$(100, 200]$	6

Assume a uniform distribution of claim sizes within each interval.
Estimate $E(X^2) - E[(X \wedge 150)^2]$.

Solution.
We have

$$\begin{aligned}
E(X^2) - E[(X\wedge 150)^2] =& \int_0^{200} x^2 f_{74}(x)dx + \int_0^{150} x^2 f_{74}(x)dx + 150^2\int_{150}^{200} f_{74}(x)dx\\
=& \int_{150}^{200}(x^2-150^2)f_{74}(x)dx\\
=& \int_{150}^{200}(x^2-150^2)\left(\frac{6}{7400}\right)dx\\
=& \left(\frac{x^3}{3}-150^2x\right)\Big|_{150}^{200} = 337.84 \blacksquare
\end{aligned}$$

Practice Problems

Problem 50.1 ‡
You are given

Claim size	Number of claims
$(0, 25]$	30
$(25, 50]$	32
$(50, 100]$	20
$(100, 200]$	8

Assume a uniform distribution of claim sizes within each interval.
(a) Estimate the mean of the claim size distribution.
(b) Estimate the second raw moment of the claim size distribution.

Problem 50.2 ‡
For 500 claims, you are given the following distribution:

Claim size	Number of claims
$(0, 500]$	200
$(500, 1000]$	110
$(1000, 2000]$	x
$(2000, 5000]$	y
$(5000, 10000]$	?
$(10000, 25000]$	?
$(25000, \infty)$	?

You are also given the following values taken from the ogive: $F_{500}(1500) = 0.689$ and $F_{500}(3500) = 0.839$. Determine y.

Problem 50.3 ‡
A random sample of payments from a portfolio of policies resulted in the following:

Claim size	Number of claims
$(0, 50]$	36
$(50, 150]$	x
$(150, 250]$	y
$(250, 500]$	84
$(500, 1000]$	80
$(1000, \infty)$	0

Suppose that $F_n(90) = 0.21$ and $F_n(210) = 0.51$. Find the value of x.

Problem 50.4 ‡
You are given the following information regarding claim sizes for 100 claims:

Claim size	Number of claims
$(0, 1000]$	16
$(1000, 3000]$	22
$(3000, 5000]$	25
$(5000, 10000]$	18
$(10000, 25000]$	10
$(25000, 50000]$	5
$(50000, 100000]$	3
$(100000, \infty)$	1

Use the ogive to estimate the probability that a randomly chosen claim is between 2000 and 6000.

Problem 50.5 ‡
You are given:
(i)

Interval	Number of observations
$(0, 50]$	30
$(50, 100]$	36
$(100, 200]$	18
$(200, 400]$	16

(ii) Claim sizes within each interval are uniformly distributed.
(iii) The second moment of the uniform distribution on $[a, b)$ is $\frac{b^3-a^3}{3(b-a)}$.
Estimate $E[(X \wedge 350)^2]$, the second moment of the claim size distribution subject to a limit of 350.

Estimation of Incomplete Data

Complete data means that each outcome in a random sample is observed or known. In contrast, in an **incomplete data** or **modified data** we may only know that some observations were above a certain value or below a certain value. Incomplete data are modified by either truncation or censoring, concepts to be introduced in Section 51. Our objective is to be able to estimate distributional quantities of modified data.

51 The Risk Set of Incomplete Data

An important element of the estimation is the risk set which is the subject of this section.

What do we mean by modified or incomplete data? An observation can be modified by truncation and/or censoring. An observation is

• **left truncated** at d if it is not recorded whenever its value is less than or equal to d but when it is above d , it is recorded at its observed value.

• **right truncated** at u, if when it is greater than or equal to u it is not recorded, but when it is below u it is recorded at its observed value.

• **left censored** at d if it is recorded as d if its value is less than or equal to d, and recorded at its observed value otherwise.

• **right censored** at u if it is recorded as u if its value is greater than or equal to u, and recorded at its observed value otherwise.

The most common modified data are the left truncated and right censored observations. Left truncation usually occurs when a policy has an ordinary deductible d (see Section 31). Right censoring occurs with a policy limit (see Section 34). In what follows we will just use the term truncated to refer to left truncated observation and we use the term censored to mean right censored.

As pointed out earlier, an important element of the estimation procedure is the concept of risk set. Some notations are first in order. For an individual data, let d_j denote the truncation point of the j-th observation with $d_j = 0$ in the absence of truncation. Let u_j denote the censored point of the j-th observation. Let x_j denote the uncensored value (loss amount).[12] Uncensored observations can be repeated (in other words, multiple loss amounts for a particular observation). Let $y_1 < y_2 < \cdots < y_k$ be the k unique (unrepeated) values of the x_is that appear in the sample. Let $s_j = \sum_i I(x_i = y_j)$ be the number of times the uncensored observation y_j appears in the sample.

[12] Each observation will have an assigned value of d and either (but not both) a value of x or u.

Now, for each unique observation y_j, the **risk set** is given by

$$
\begin{aligned}
r_j =& \sum_i I(x_i \geq y_j) + \sum_i I(u_i \geq y_j) - \sum_i I(d_i \geq y_j) \\
=& \sum_i I(d_i < y_j) - \sum_i I(x_i < y_j) - \sum_i I(u_i < y_j)
\end{aligned}
$$

where I is the indicator function.

The risk set can be interpreted as follows:

• For survival/mortality data[13], the risk set is the number of people observed alive at age y_j.

• For loss amount data, the risk set is the number of policies with observed loss amounts (either the actual amount or the maximum amount due to a policy limit) larger than or equal to y_j less those with deductibles greater than or equal to y_j.

From the above formulas of r_j, the following recursive formula holds

$$
r_j = r_{j-1} + \sum_i I(y_{j-1} \leq d_i < y_j) - \sum_i I(x_i = y_{j-1}) - \sum_i I(y_{j-1} \leq u_i < y_j)
$$

and $r_0 = 0$. That is,

$r_j = r_{j-1} -$ (# of $x's$ equal to y_{j-1})+ (# of observations left-truncated at y_{j-1}) $-$ (# of observations right-censored at y_{j-1}).

Example 51.1
You are given the following mortality table:

[13] In a typical mortality study, the following notation is used for an individual i : d_i will denote the time the individual joined the study; u_i will denote the time of withdrawal from the study; and x_i will denote the time of death of the individual.

Life	Time of Entry	Time of exit	Reason of exit
1	0	4	End of Study
2	0	0.5	Death
3	0	1	Surrunder
4	0	4	End of Study
5	1	4	End of Study
6	1.2	2	Death
7	1.5	2	Death
8	2	3	Surrunder
9	2.5	4	End of Study
10	3.1	3.2	Death

Complete the following table.

i	d_i	x_i	u_i
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Solution.

i	d_i	x_i	u_i
1	0	–	4
2	0	0.5	–
3	0	–	1
4	0	–	4
5	1	–	4
6	1.2	2	–
7	1.5	2	–
8	2	–	3
9	2.5	–	4
10	3.1	3.2	– ■

Example 51.2
Create a table summarizing y_j, s_j, and r_j of Example 51.1

Solution.
The table is given below.

j	y_j	s_j	r_j
1	0.5	1	$4+6-6=4$
2	2	2	$3+5-3=5$
3	3.2	1	$1+4-0=5$ ■

Practice Problems

Problem 51.1
For a ground up loss of amount X, an insurer pays the loss in excess of a deductible d, and with a policy limit of u. Which of the following statements is true regarding censoring and truncation.

(A) Ground up losses below the deductible amount are right-truncated.
(B) Ground up losses below the deductible amount are left-truncated.
(C) Ground up losses above amount u are right-censored.
(D) Ground up losses above amount u are left-censored.

Problem 51.2
Which of the following statements is true?

(A) Losses of an amount in excess of a policy limit are right-truncated losses.
(B) Losses of an amount in excess of a policy limit are left-truncated losses.
(C) Losses below a policy deductible are right-truncated losses.
(D) Losses below a policy deductible are left-truncated losses.
(E) All of A, B, C and D are false.

Problem 51.3
You are given the following mortality table:

Life	Time of Entry	Time of exit	Reason of exit
1	0	0.2	Lapse
2	0	0.3	Lapse
3	0	0.5	Lapse
4	0	0.5	Death
5	1	0.7	Lapse
6	1.2	1.0	Death
7	1.5	2.0	Lapse
8	2	2.5	Death
9	2.5	3.0	Lapse
10	3.1	3.5	Death
11	0	4.0	Expiry of Study
12	0	4.0	Expiry of Study
13	0	4.0	Expiry of Study
14	0	4.0	Expiry of Study
15	0	4.0	Expiry of Study

16	0	4.0	Expiry of Study
17	0	4.0	Expiry of Study
18	0	4.0	Expiry of Study
19	0.5	4.0	Expiry of Study
20	0.7	1.0	Death
21	1.0	3.0	Death
22	1.0	4.0	Expiry of Study
23	2.0	2.5	Death
24	2.0	2.5	Lapse
25	3.0	3.5	Death

Complete the table similar to Example 51.1.

Problem 51.4
You are given the following

i	d_i	x_i	u_i
1	1	–	6
2	0	4	–
3	2	4	–
4	6	8	–
5	0	–	5

Create a table summarizing y_j, s_j, and r_j.

Problem 51.5
You are given the following

j	d_j	x_j	u_j
1	0	0.9	–
2	0	–	1.2
3	0	1.5	–
4	0	–	1.5
5	0	–	1.6
6	0	1.7	–
7	0	–	1.7
8	1.3	2.1	–
9	1.5	2.1	–
10	1.6	–	2.3

Create a table summarizing y_j, s_j, and r_j.

52 The Kaplan-Meier and Nelson-Åalen Estimators

In this section, we consider the question of estimating the survival function of incomplete data. We look at two estimators, the Kaplan-Meier product-limit estimator and the Nelson-Åalen estimator.

In a typical mortality study, the following notation is used for an individual i : d_i will denote the time when an individual has joined the study; u_i will denote the time of withdrawal from the study; and x_i will denote the time of death of the individual. The values s_i and r_i are as defined in the previous section.

The **Kaplan-Meier (product-limit)** estimate for the survival function is given by

$$S_n(t) = \hat{S}(t) = \begin{cases} 1, & 0 \le t < y_1 \\ \displaystyle\prod_{i=1}^{j-1}\left(1 - \frac{s_i}{r_i}\right), & y_{j-1} \le t < y_j,\ j = 2, 3, \cdots, k \\ \displaystyle\prod_{i=1}^{k}\left(1 - \frac{s_i}{r_i}\right) \text{ or } 0, & t \ge y_k. \end{cases}$$

Example 52.1
Find the Kaplan-Meier estimate of the survival function for the data in Example 51.1

Solution.
We have

$$S_{10}(t) = \begin{cases} 1, & 0 \le t < 0.5 \\ 1 - \frac{s_1}{r_1} = 0.75, & 0.5 \le t < 2 \\ 0.75\left(1 - \frac{s_2}{r_2}\right) = 0.45, & 2 \le t < 3.2 \\ 0.45\left(1 - \frac{s_3}{r_3}\right) = 0.36, & t \ge 3.2 \ \blacksquare \end{cases}$$

The Kaplan-Meier applied to complete data is the same as the empirical survival function defined by

$$S_n(x) = 1 - \frac{\text{number of observation } \le \text{ x}}{n}.$$

Example 52.2
Below are the losses suffered by policyholders of an insurance company:

$$49, 49, 50, 50, 50, 60, 75, 80, 120, 130.$$

Let X be the random variable representing the losses incurred by the policyholders. The observations are not truncated or censored. Use a Kaplan-Meier product-limit estimator to approximate the survival function for this data.

Solution.
We have

$$S_9(t) = \begin{cases} 1, & 0 \leq t < 49 \\ 1 - \frac{2}{10} = 0.8, & 49 \leq t < 50 \\ 0.8\left(1 - \frac{3}{8}\right) = 0.5, & 50 \leq t < 60 \\ 0.5\left(1 - \frac{1}{5}\right) = 0.4, & 60 \leq t < 75 \\ 0.4\left(1 - \frac{1}{4}\right) = 0.3, & 75 \leq t < 80 \\ 0.3\left(1 - \frac{1}{3}\right) = 0.2, & 80 \leq t < 120 \\ 0.2\left(1 - \frac{1}{2}\right) = 0.1, & 120 \leq t < 130 \\ 0.1\left(1 - \frac{1}{1}\right) = 0, & t \geq 130. \end{cases}$$

Note that this is the same as the empirical survival function for this data set ■

The Kaplan-Meier estimate can be used to evaluate probabilities. If we let T denote the time until death or failure (or size of loss), then $\Pr(T > y_1) = 1 - \frac{s_1}{r_1}$ or equivalently $\Pr(T \leq y_1) = \frac{s_1}{r_1}$. The probability of those who were still alive at time y_1 and who are at risk of death at time y_2 and survive is $\Pr(T > y_2 | T > y_1) = 1 - \frac{s_2}{r_2}$. Thus,

$$S_n(y_2) = \Pr(T > y_1)\Pr(T > y_2|T > y_1) = \left(1 - \frac{s_1}{r_1}\right)\left(1 - \frac{s_2}{r_2}\right).$$

Likewise,

$$\begin{aligned} S_n(y_3) =& \Pr(T > y_1)\Pr(T > y_2|T > y_1)\Pr(T > y_3|T > y_2) \\ =& \left(1 - \frac{s_1}{r_1}\right)\left(1 - \frac{s_2}{r_3}\right)\left(1 - \frac{s_3}{r_3}\right) \end{aligned}$$

and so on.

Example 52.3 ‡
You are studying the length of time attorneys are involved in settling bodily injury lawsuits. T represents the number of months from the time an attorney is assigned such a case to the time the case is settled. Nine cases were observed during the study period, two of which were not settled at the conclusion of the study. For those two cases, the time spent up to the

conclusion of the study, 4 months and 6 months, was recorded instead. The observed values of T for the other seven cases are as follows:

$$1 \;\; 3 \;\; 3 \quad 5 \;\; 8 \;\; 8 \;\; 9.$$

Estimate $\Pr(3 \le T \le 5)$ using the Product-Limit estimator.

Solution.
We have the following charts

j	d_j	x_j	u_j
1	0	1	–
2	0	3	–
3	0	3	–
4	0	–	4
5	0	5	–
6	0	–	6
7	0	8	–
8	0	8	–
9	0	9	–

j	y_j	s_j	r_j
1	1	1	9
2	3	2	8
3	5	1	5
4	8	2	3
5	9	1	1

The Kaplan-Meier estimator of $S(t)$ is given by

$$\hat{S}(t) = \begin{cases} 1, & 0 \le t < 1 \\ 1 - \frac{1}{9} = \frac{8}{9}, & 1 \le t < 3 \\ \frac{8}{9}\left(1 - \frac{2}{8}\right) = \frac{2}{3}, & 3 \le t < 5 \\ \frac{2}{3}\left(1 - \frac{1}{5}\right) = \frac{8}{15}, & 5 \le t < 8 \\ \frac{8}{15}\left(1 - \frac{2}{3}\right) = \frac{8}{45}, & 8 \le t < 9 \\ \frac{8}{45}\left(1 - \frac{1}{1}\right) = 0, & t \ge 9. \end{cases}$$

Now, we have

$$\begin{aligned} \Pr(3 \le T \le 5) =& \Pr(T \ge 3) - \Pr(T > 5) \\ =& \hat{S}(3^-) - \hat{S}(5) \\ =& \frac{8}{9} - \frac{8}{15} = 0.356 \blacksquare \end{aligned}$$

Example 52.4 ‡
The claim payments on a sample of ten policies are:

$$2 \;\; 3 \;\; 3 \;\; 5 \;\; 5^+ \;\; 6 \;\; 7 \;\; 7^+ \;\; 9 \;\; 10^+$$

where the "+" indicates that the loss exceeded the policy limit.
Using the Product-Limit estimator, calculate the probability that the loss on a policy exceeds 8.

Solution.
We have the following charts

j	d_j	x_j	u_j
1	0	2	–
2	0	3	–
3	0	3	–
4	0	5	–
5	0	–	5
6	0	6	–
7	0	7	–
8	0	–	7
9	0	9	–
10	0	–	10

j	y_j	s_j	r_j	$\hat{S}(y_j)$
1	2	1	10	$1 - \frac{1}{10} = 0.9$
2	3	2	9	$0.9(1 - \frac{2}{9}) = 0.7$
3	5	1	7	$0.7(1 - \frac{1}{7}) = 0.6$
4	6	1	5	$0.6(1 - \frac{1}{5}) = 0.48$
5	7	1	4	$0.48(1 - \frac{1}{4}) = 0.36$
6	9	1	2	$0.36(1 - \frac{1}{2}) = 0.18$

Since, $y_5 \le 8 < y_6$, we have $\hat{S}(8) = \hat{S}(7) = 0.36$ ■

An alternative estimator for the survival function is to estimate the cumulative hazard rate function via the **Nelson-Åalen** given by

$$\hat{H}(t) = \begin{cases} 0, & t < y_1 \\ \sum_{i=1}^{j-1} \frac{s_i}{r_i}, & y_{j-1} \le t < y_j,\ j = 2, 3, \cdots, k \\ \sum_{i=1}^{k} \frac{s_i}{r_i}, & t \ge y_k \end{cases}$$

and then set $\hat{S}(t) = e^{-\hat{H}(t)}$.

Example 52.5
Determine the Nelson-Åalen estimate of the survival function for the data in Example 51.1.

Solution.
We first use the Nelson-Åalen estimate of the cumulative hazard rate function.

$$\hat{H}(t) = \begin{cases} 0, & 0 \le t < 0.5 \\ \frac{1}{4} = 0.25, & 0.5 \le t < 2 \\ 0.25 + \frac{2}{5} = 0.65, & 2 \le t < 3.2 \\ 0.65 + \frac{1}{5} = 0.85, & t \ge 3.2. \end{cases}$$

An estimate of the survival function is

$$\hat{S}(t) = \begin{cases} 1, & 0 \le t < 0.5 \\ e^{-0.25} = 0.7788, & 0.5 \le t < 2 \\ e^{-0.65} = 0.5220, & 2 \le t < 3.2 \\ e^{-0.85} = 0.4274, & t \ge 3.2 \ ■ \end{cases}$$

Example 52.6
For a mortality study with right-censored data, you are given:

Time t_j	Number of deaths s_j	Number at risk r_j
3	1	50
5	3	49
6	5	k
10	7	21

You are also told that the Nelson-Åalen estimate of the survival function at time 10 is 0.575. Determine k.

Solution.
From $0.575\hat{S}(10) = e^{-\hat{H}(10)}$ we find $\hat{H}(10) = -\ln 0.575 = 0.5534$. Thus,

$$0.5534 = \sum_{t_i \le 10} \frac{s_i}{r_i} = \frac{1}{50} + \frac{3}{49} + \frac{5}{k} + \frac{7}{21}.$$

Solving this equation, we find $k = 36$ ■

Example 52.7 ‡
You are given:
(i) A mortality study covers n lives.
(ii) None were censored and no two deaths occurred at the same time.
(iii) t_k = time of the k^{th} death.
(iv) A Nelson-Åalen estimate of the cumulative hazard rate function is $\hat{H}(t_2) = \frac{39}{380}$.
Determine the Kaplan-Meier product-limit estimate of the survival function at time t_9.

Solution.
We have $\hat{H}(t_1) = \frac{s_1}{r_1} = \frac{1}{n}$ and $\hat{H}(t_2) = \frac{s_1}{r_1} + \frac{s_2}{r_2} = \frac{1}{n} + \frac{1}{n-1}$. The assumption $\hat{H}(t_2) = \frac{39}{280}$ leads to the equation

$$39n^2 - 799n + 380 = 0.$$

Solving this quadratic equation we find the roots $n = 0.487$(discard) and $n = 20$. Thus, the Kaplan-Meier product-limit estimate of the survival function at time t_9 is

$$\begin{aligned} S_n(t_9) =& \prod_{i=1}^{9} \left(1 - \frac{s_i}{r_i}\right) \\ =& \frac{19}{20} \cdot \frac{18}{19} \cdots \frac{11}{12} = \frac{11}{20} = 0.55 \ ■ \end{aligned}$$

Example 52.8 ‡
You are given the following times of first claim for five randomly selected auto insurance policies observed from time $t = 0$:

$$1 \quad 2 \quad 3 \quad 4 \quad 5$$

You are later told that one of the five times given is actually the time of policy lapse (i.e., terminated), but you are not told which one.
The smallest Product-Limit estimate of S(4), the probability that the first claim occurs after time 4, would occur at the lapse time t_0. Find t_0.

Solution.
If the time of policy lapse is at $t = 1$, then the at risk group at death time 2 is 4, at death time 3 is 3 and at death time 4 is 2, so that

$$\hat{S}(4) = \prod_{i=1}^{4} = \left(1 - \frac{0}{5}\right)\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{2}\right) = 0.250.$$

If the time of policy lapse is at $t = 2$, then the at risk group at death time 1 is 5, at death time 3 is 3 and at death time 4 is 2, so that

$$\hat{S}(4) = \prod_{i=1}^{4} = \left(1 - \frac{1}{5}\right)\left(1 - \frac{0}{4}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{2}\right) = 0.267.$$

If the time of policy lapse is at $t = 3$, then the at risk group at death time 1 is 5, at death time 2 is 4 and at death time 4 is 2, so that

$$\hat{S}(4) = \prod_{i=1}^{4} = \left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{4}\right)\left(1 - \frac{0}{3}\right)\left(1 - \frac{1}{2}\right) = 0.300.$$

If the time of policy lapse is at $t = 4$, then the at risk group at death time 1 is 5, at death time 2 is 4 and at death time 3 is 3, so that

$$\hat{S}(4) = \prod_{i=1}^{4} = \left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{0}{2}\right) = 0.400.$$

If the time of policy lapse is at $t = 5$, then the at risk group at death time 1 is 5, at death time 2 is 4, at death time 3 is 3, and at death time 4 is 2 so that

$$\hat{S}(4) = \prod_{i=1}^{4} = \left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{2}\right) = 0.200.$$

$\hat{S}(4)$ is the smallest at time $t = 5$, which is the policy lapse time ■

Practice Problems

Problem 52.1 ‡
For a mortality study with right-censored data, you are given:

Time t_j	Number of deaths s_j	Number at risk r_j
5	2	15
7	1	12
10	1	10
12	2	6

Calculate the estimate of $\hat{S}(12)$ based on the Nelson-Åalen estimate of $\hat{H}(12)$.

Problem 52.2 ‡
You are given:
(i) The following data set:

2500 2500 2500 3617 3662 4517 5000 5000 6010 6932 7500 7500

(ii) $\hat{H}_1(7000)$ is the Nelson-Åalen estimate of the cumulative hazard rate function calculated under the assumption that all of the observations in (i) are uncensored.
(iii) $\hat{H}_2(7000)$ is the Nelson-Åalen estimate of the cumulative hazard rate function calculated under the assumption that all occurrences of the values 2500, 5000 and 7500 in (i) reflect right-censored observations and that the remaining observed values are uncensored.

Calculate $|\hat{H}_1(7000) - \hat{H}_2(7000)|$.

Problem 52.3 ‡
For a mortality study of insurance applicants in two countries, you are given:
(i)

	Country A		*Country B*	
y_j	s_j	r_j	s_j	r_j
1	20	200	15	100
2	54	180	20	85
3	14	126	20	65
4	22	112	10	45

(ii) r_j is the number at risk over the period (y_{j-1}, y_j). Deaths during the period (y_{j-1}, y_j) are assumed to occur at y_j.
(iii) $S^T(t)$ is the Product-Limit estimate of $S(t)$ based on the data for all study participants.
(iv) $S^B(t)$ is the Product-Limit estimate of $S(t)$ based on the data for country B.

Calculate $|S^T(4) - S^B(4)|$.

Problem 52.4 ‡
For observation j of a survival study:
- d_j is the left truncation point
- x_j is the observed value if not right-censored
- u_j is the observed value if right-censored.

You are given:

j	d_j	x_j	u_j
1	0	0.9	–
2	0	–	1.2
3	0	1.5	–
4	0	–	1.5
5	0	–	1.6
6	0	1.7	–
7	0	–	1.7
8	1.3	2.1	–
9	1.5	2.1	–
10	1.6	–	2.3

Determine the Kaplan-Meier Product-Limit estimate of $S_{10}(1.6)$.

Problem 52.5 ‡
You are given:
(i) All members of a mortality study are observed from birth. Some leave the study by means other than death.
(ii) $s_3 = 1, s_4 = 3$.
(iii) The following Kaplan-Meier product-limit estimates were obtained: $S_n(y_3) = 0.65,\ S_n(y_4) = 0.50,\ S_n(y_5) = 0.25$.
(iv) Between times y_4 and y_5, six observations were censored.
(v) Assume no observations were censored at the times of deaths.

Determine s_5.

Problem 52.6 ‡
In a study of claim payment times, you are given:
(i) The data were not truncated or censored.
(ii) At most one claim was paid at any one time.
(iii) The Nelson-Åalen estimate of the cumulative hazard function, $H(t)$, immediately following the second paid claim, was 23/132.

Determine the Nelson-Åalen estimate of the cumulative hazard function, $H(t)$, immediately following the fourth paid claim.

Problem 52.7 ‡
You are given:
(i) The following is a sample of 15 losses: 11, 22, 22, 22, 36, 51, 69, 69, 69, 92, 92, 120, 161, 161, 230.
(ii) $\hat{H}_1(x)$ is the Nelson-Åalen empirical estimate of the cumulative hazard rate function.
(iii) $\hat{H}_2(x)$ is the maximum likelihood estimate of the cumulative hazard rate function under the assumption that the sample is drawn from an exponential distribution.

Calculate $|\hat{H}_2(75) - \hat{H}_1(75)|$.

Problem 52.8 ‡
For the data set

$$200 \quad 300 \quad 100 \quad 400 \quad X$$

you are given:
(i) $k = 4$
(ii) $s_2 = 1$
(iii) $r_4 = 1$
(iv) The NelsonÅalen Estimate $\hat{H}(410) > 2.15$.

Determine X.

53 Mean and Variance of Empirical Estimators with Complete Data

In this section, we calculate the variance of empirical survival function and empirical probability estimate for complete data.

Individual Data
We first consider finding the variance of the empirical survival function of individual data. Suppose that the sample is of size n. Let $S_n(t)$ be the empirical estimate of the survival function $S(x)$ defined by

$$S_n(x) = \frac{\text{number of observation greater than } x}{n} = \frac{Y}{n}$$

where Y is the number of observations in the sample that are greater than x. If we regard an observation as a trial then we define a success to be the observation that is greater than x and which occurs with probability $S(x)$. Then Y is a binomial random variable with parameters n and $S(x)$ and with mean $E(Y) = nS(x)$ and variance $\text{Var}(Y) = nS(x)(1 - S(x))$. Thus,

$$E[S_n(x)] = \frac{E(Y)}{n} = \frac{nS(x)}{n} = S(x).$$

This shows that the empirical estimate $S_n(x)$ is unbiased (see Section 46).

In the same token, we have

$$\text{Var}[S_n(x)] = \frac{1}{n^2}\text{Var}(Y) = \frac{S(x)(1 - S(x))}{n} \to 0$$

as $n \to \infty$. This shows that $S_n(t)$ is consistent.

If $S(x)$ is unknown, then we can estimate the variance of $S_n(x)$ using $S_n(x)$ itself in the formula

$$\widehat{\text{Var}}[S_n(x)] = \frac{S_n(x)(1 - S_n(x))}{n}.$$

Example 53.1
Let X be a discrete random variable and $p = \Pr(a < X \le b)$. An estimate of p is $\hat{p} = S_n(a) - S_n(b)$. Show that $\hat{p}$ is unbiased and consistent.

Solution.
We first note that $\hat{p} = \frac{Y}{n}$ where Y is a binomial random variable with parameters n and $S(a) - S(b)$. Thus,

$$E(\hat{p}) = \frac{1}{n}[n(S(a) - S(b))] = S(a) - S(b) = p.$$

This shows that $\hat{p}$ is unbiased. Next, we show that $\hat{p}$ is consistent. Indeed, we have

$$\text{Var}(\hat{p}) = \frac{1}{n^2}[n(S(a) - S(b))(1 - S(a) + S(b))] = \frac{p(1-p)}{n} \to 0$$

as $n \to \infty$. This shows that $\hat{p}$ is consistent. When p is unknown, we have

$$\widehat{\text{Var}}(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n} \blacksquare$$

Next, we consider survival probability estimators. Recall (see [3]) that the probability of a life aged x to attain age $x + t$ is the conditional probability

$$_tp_x = \Pr(X > x + t | X > x)$$

where X is the age-at-death random variable.

Let n_x and n_y be the observed number of survivors past time x and y respectively. The the probability of a person aged x to reach age y is

$$_{y-x}p_x = \frac{_yp_0}{_xp_0}$$

where $S(x) = {}_xp_0$. Thus, an estimator of $_{y-x}p_x$ is

$$_{y-x}\hat{p}_x = \frac{S_n(y)}{S_n(x)} = \frac{n_y}{n_x}.$$

The variance is given by

$$\widehat{\text{Var}}(_{y-x}\hat{p}_x | n_x) = \widehat{\text{Var}}(_{y-x}\hat{q}_x | n_x) = \frac{(n_x - n_y)n_y}{n_x^3}.$$

Example 53.2
The following chart provides the time of death of 15 individuals under observation from time 0.

Time of Death	1	2	3	4	5
# of Deaths	1	3	2	4	5

(a) Estimate $S(3)$ and estimate the variance of the estimator.
(b) Estimate q_2, the conditional probability that death occurs not later than 3 given survival to time 2. Estimate the variance of this estimator.

Solution.
(a) We have $S_{15}(3) = \frac{9}{15} = 0.6$ and

$$\widehat{\text{Var}}(S_{15}(3) = \frac{0.6(0.4)}{15} = 0.016.$$

(b) We have

$$\hat{q}_2 = 1 - \frac{S_{15}(3)}{S_{15}(2)} = 1 - \frac{9}{11} = \frac{2}{11}$$

and

$$\widehat{\text{Var}}(\hat{q}_2|n_2) = \frac{(n_2 - n_3)n_3}{n_2^3} = 0.0135 \blacksquare$$

Example 53.3
The following random sample of 9 losses has been observed from the distribution of loss random variable X :

$$49, 50, 50, 50, 60, 75, 80, 120, 130.$$

(a) Find the estimated variance of the estimate of $\Pr(X > 60)$.
(b) Find the estimated variance of the estimate of $\Pr(75 < X \leq 120)$.
(c) Find the estimated variance of the estimate of $\Pr(X > 60|X > 50)$.

Solution.
(a) Let $p_1 = \Pr(X > 60)$. Then $\hat{p}_1 = \frac{4}{9}$ and $\widehat{\text{Var}}(\hat{p}_1) = \frac{\frac{4}{9}\left(1-\frac{4}{9}\right)}{9} = \frac{20}{729}$.
(b) Let $p_2 = \Pr(75 < X \leq 120)$. Then $\hat{p}_2 = S_9(75) - S_9(120) = \frac{3}{9} - \frac{1}{9} = \frac{2}{9}$ and $\widehat{\text{Var}}(\hat{p}_2) = \frac{\frac{2}{9}\left(1-\frac{2}{9}\right)}{9} = \frac{14}{729}$.
(c) Let $p_3 = \Pr(X > 60|X > 50) = {}_{10}p_{50}$. Then $\hat{p}_3 = \frac{S_9(60)}{S_9(50)} = \frac{4}{5}$ and $\widehat{\text{Var}}(\hat{p}_3) = \frac{4}{125} \blacksquare$

Grouped Data
We first use the ogive to find the variance of the estimator of the survival function and the density function. Let n observations be spread over the grouped data

$$(c_0, c_], (c_1, c_2], \cdots, (c_{k-1}, c_k]$$

Let n_j be the number of observations in $(c_{j-1}, c_j]$. We define the survival estimate at c_j by

$$S_n(c_j) = 1 - \frac{n_1 + n_2 + \cdots + n_j}{n}.$$

If $x \in (c_{j-1}, c_j]$ then $S_n(x)$ can be found via the interpolation equation

$$S_n(x) = \frac{c_j - x}{c_j - c_{j-1}} S_n(c_{j-1}) + \frac{x - c_{j-1}}{c_j - c_{j-1}} S_n(c_j).$$

Let Y be the number of observations up to c_{j-1}. That is, $Y = n_1 + n_2 + \cdots + n_{j-1}$. Let Z be the number of observations in $(c_{j-1}, c_j]$. That is, $Z = n_j$. Then the previous equation can be expressed as

$$S_n(x) = 1 - \frac{Y(c_j - c_{j-1}) + Z(x - c_{j-1})}{n(c_j - c_{j-1})}.$$

Note that Y can be regarded as a binomial random variable with parameters $(n, 1 - S(c_{j-1}))$ whereas Z can be regarded as a binomial random variable with parameters $(n, S(c_{j-1}) - S(c_j))$. Thus,

$$\begin{aligned}
E[S_n(x)] =& 1 - \frac{n[1 - S(c_{j-1})](c_j - c_{j-1}) + n[S(c_{j-1}) - S(c_j)](x - c_{j-1})}{n(c_j - c_{j-1})} \\
=& \frac{c_j - x}{c_j - c_{j-1}} S(c_{j-1}) + \frac{x - c_{j-1}}{c_j - c_{j-1}} S(c_j).
\end{aligned}$$

The variance of $S_n(x)$ is

$$\begin{aligned}
\mathrm{Var}[S_n(x)] =& \frac{(c_j - c_{j-1})^2 \mathrm{Var}(Y) + (x - c_{j-1})^2 \mathrm{Var}(Z)}{[n(c_j - c_{j-1})]^2} \\
&+ \frac{2(c_j - c_{j-1})(x - c_{j-1}) \mathrm{Cov}(Y, Z)}{[n(c_j - c_{j-1})]^2}
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{Var}(Y) =& nS(c_{j-1})[1 - S(c_{j-1})] \\
\mathrm{Var}(Z) =& n[S(c_{j-1}) - S(c_j)][1 - S(c_{j-1}) + S(c_j)] \\
\mathrm{Cov}(Y, Z) =& -n[1 - S(c_{j-1})][S(c_{j-1}) - S(c_j)].
\end{aligned}$$

Example 53.4
Using a histogram (see Section 50), the empirical density function can be expressed as

$$f_n(x) = \frac{Z}{n(c_j - c_{j-1})}$$

where $Z = n_j$. Find the mean of $f_n(x)$.

Solution.
The random variable Z is a binomial random variable with parameters $(n, S(c_{j-1}) - S(c_j))$. Thus,

$$E[f_n(x)] = \frac{E(Z)}{n(c_j - c_{j-1})} = \frac{S_n(c_{j-1}) - S_n(c_j)}{c_j - c_{j-1}} \blacksquare$$

Example 53.5
Find $\text{Var}[f_n(x)]$, where $f_n(x)$ is the empirical density function in the previous example.

Solution.
We have

$$\text{Var}[f_n(x)] = \frac{\text{Var}(Z)}{[n(c_j - c_{j-1})]^2} = \frac{[S_n(c_{j-1}) - S_n(c_j)][1 - S_n(c_{j-1}) + S_n(c_j)]}{n(c_j - c_{j-1})^2} \blacksquare$$

Example 53.6
Given the following grouped data.

Loss Range	Number of Losses
$(0, 2]$	25
$(2, 10]$	10
$(10, 100]$	10
$(100, 1000]$	5

Estimate the probability that a loss will be no more than 90, and find the estimated variance of that estimate.

Solution.
An estimate to the probability that a loss will be no more than 90 is

$$\begin{aligned}
1 - S_{50}(90) =& 1 - \frac{c_j - x}{c_j - c_{j-1}} S_n(c_{j-1}) - \frac{x - c_{j-1}}{c_j - c_{j-1}} S_n(c_j) \\
=& 1 - \frac{100 - 90}{100 - 10}\frac{15}{50} - \frac{90 - 10}{100 - 10}\frac{5}{50} \\
=& 1 - \frac{1}{9}\frac{3}{10} - \frac{8}{9}\frac{1}{10} \\
=& 0.8778. \\
\widehat{\text{Var}}(Y) =& nS_n(c_{j-1})[1 - S_n(c_{j-1})] \\
=& 50\frac{15}{50}[1 - \frac{15}{50}] = 10.5 \\
\widehat{\text{Var}}(Z) =& n[S_n(c_{j-1}) - S_n(c_j)][1 - S_n(c_{j-1}) + S_n(c_j)] \\
=& 50[\frac{10}{50}][1 - \frac{10}{50}] = 8 \\
\widehat{\text{Cov}}(Y, Z) =& - n[1 - S_n(c_{j-1})][S_n(c_{j-1}) - S_n(c_j)] \\
=& - 50[1 - \frac{15}{50}]\frac{10}{50} = -7 \\
\widehat{\text{Var}}[1 - S_{50}(90)] =& \widehat{\text{Var}}[S_{50}(90)] \\
=& \frac{(c_j - c_{j-1})^2\widehat{\text{Var}}(Y) + (x - c_{j-1})^2\widehat{\text{Var}}(Z)}{[n(c_j - c_{j-1})]^2} \\
&+ \frac{2(c_j - c_{j-1})(x - c_{j-1})\widehat{\text{Cov}}(Y, Z)}{[n(c_j - c_{j-1})]^2} \\
=& 0.00175 \blacksquare
\end{aligned}$$

Practice Problems

Problem 53.1
Let X be a discrete random variable. In a sample of n outcomes, let N_j denote the number of times the value x_j was observed in the sample with corresponding probability $p(x_j)$. N_j can be regarded as a binomial random variable with parameters $(n, p(x_j))$.

Show that the estimator
$$p_n(x_j) = \frac{N_j}{n}$$
is unbiased and consistent.

Problem 53.2
Given the following grouped data.

Loss Range	Number of Losses
$(0, 2]$	25
$(2, 10]$	10
$(10, 100]$	10
$(100, 1000]$	5

Estimate the density function of the loss random variable at $x = 90$, and find the estimated variance of the estimator.

Problem 53.3
Consider the following data

Number of accidents	Number of drivers
0	155
1	109
2	64
3	50
4 or more	8

Estimate the probability that a driver will have two accidents and find the estimate of the variance of that estimator.

Problem 53.4
Estimated variances can be used to create confidence intervals for the true

probability. The formula for the endpoints of a $1-\alpha$ confidence interval is as follows (see Section 47):

$$p_n \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p_n(1-p_n)}{n}}$$

where $\Phi(z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.

Construct approximate 95% confidence inervala for $p(2)$ of the previous exercise.

54 Greenwood Estimate for the Variance of the Kaplan-Meier Estimator

In this section we look at estimating the variance to the Kaplan-Meier limit-product estimator.

First, we recall the Kaplan-Meier limit-product (see Section 50) in mortality context: Let $0 < y_1 < y_2 < \cdots < y_k$ be the unique observed death times in the sample. The risk size r_j is the number of individuals in the sample that were alive (or at "risk") just before time y_j. Equivalently, r_j, is the number of individuals who are either alive and observed at time y_j or else who died at time y_j. The number of individuals who died at time y_j is denoted by s_j. The Kaplan-Meier estimate of the survival function is given by

$$S_n(t) = \hat{S}(t) = \begin{cases} 1, & 0 \le t < y_1 \\ \prod_{i=1}^{j-1}\left(1 - \frac{s_i}{r_i}\right), & y_{j-1} \le t < y_j,\ j = 2, 3, \cdots, k \\ \prod_{i=1}^{k}\left(1 - \frac{s_i}{r_i}\right) \text{ or } 0, & t \ge y_k. \end{cases}$$

In what follows, $y_0 < y_1$ will denote the smallest alive or observed age in the sample. Assume that the $r_j s$ and $y_j s$ are fixed. The number of individuals who died at time y_j is the only random quantity, which we denote by S_j. As a random variable, S_j has a binomial distribution based on a sample of r_j and a probability of success

$$\Pr(T \le y_j | T > y_{j-1}) = \frac{\Pr(y_{j-1} < T \le y_j)}{\Pr(T > y_{j-1})} = \frac{S(y_{j-1}) - S(y_j)}{S(y_{j-1})}$$

where T is the time until death. For $1 \le j \le k$, we have

$$E\left(1 - \frac{S_j}{r_j}\right) = 1 - \frac{[S(y_{j-1}) - S(y_j)]}{S(y_{j-1})} = \frac{S(y_j)}{S(y_{j-1})} \tag{54.1}$$

and

$$\text{Var}\left(1 - \frac{S_j}{r_j}\right) = \frac{1}{r_j^2}\text{Var}(S_j) = \frac{[S(y_{j-1}) - S(y_j)]S(y_j)}{r_j S(y_{j-1})^2}.$$

Example 54.1
Assume that the S_js are independent. Show that

$$E[S_n(y_j)] = \frac{S(y_j)}{S(y_0)} \tag{54.2}$$

Solution.
We have

$$\begin{aligned}
E[S_n(y_j)] =&E\left[\prod_{i=1}^{j}\left(1-\frac{S_i}{r_i}\right)\right] = \prod_{i=1}^{j}E\left(1-\frac{S_i}{r_i}\right)\\
=&\prod_{i=1}^{j}\frac{S(y_i)}{S(y_{i-1})} = \frac{S(y_j)}{S(y_0)}.
\end{aligned}$$

It is important to keep in mind that this estimate of $S_n(y_j)$ is conditioned on being alive at time or age y_0.

Using Problems 54.1 and 54.2, we have

$$\begin{aligned}
\text{Var}[S_n(y_j)] =&\text{Var}\left[\prod_{i=1}^{j}\left(1-\frac{S_j}{r_j}\right)\right]\\
=&\prod_{i=1}^{j}\left[\frac{S(y_i)^2}{S(y_{i-1})^2}+\frac{[S(y_{i-1})-S(y_i)]S(y_i)}{r_iS(y_{i-1})^2}\right]-\frac{S(y_j)^2}{S(y_0)^2}\\
=&\prod_{i=1}^{j}\left[\frac{S(y_i)^2}{\frac{r_iS(y_{i-1})^2}{+}[S(y_{i-1})-S(y_i)]S(y_i)}r_iS(y_{i-1})^2\right]-\frac{S(y_j)^2}{S(y_0)^2}\\
=&\prod_{i=1}^{j}\left[\frac{S(y_i)^2}{S(y_{i-1})^2}\frac{r_iS(y_i)+[S(y_{i-1})-S(y_i)]}{r_iS(y_i)}\right]-\frac{S(y_j)^2}{S(y_0)^2}\\
=&\prod_{i=1}^{j}\frac{S(y_i)^2}{S(y_{i-1})^2}\prod_{i=1}^{j}\left[\frac{r_iS(y_i)+[S(y_{i-1})-S(y_i)]}{r_iS(y_i)}\right]-\frac{S(y_j)^2}{S(y_0)^2}\\
=&\frac{S(y_j)^2}{S(y_0)^2}\left\{\prod_{i=1}^{j}\left[1+\frac{S(y_{i-1})-S(y_i)}{r_iS(y_i)}\right]-1\right\}.
\end{aligned}$$

If we assume $r_iS(y_i)$ to be large for $i = 1, 2, \cdots, j$ we can estimate the product as

$$\prod_{i=1}^{j}\left[1+\frac{S(y_{i-1})-S(y_i)}{r_iS(y_i)}\right] = 1+\sum_{i=1}^{j}\frac{S(y_{i-1})-S(y_i)}{r_iS(y_i)}$$

and thus we obtain

$$\text{Var}[S_n(y_j)] \approx \left[\frac{S(y_j)}{S(y_0)}\right]^2 \sum_{i=1}^{j} \frac{S(y_{i-1}) - S(y_i)}{r_i S(y_i)}. \tag{54.3}$$

Because the survival function is usually unknown, we use (54.1) and (54.2) to write

$$\left[\tfrac{S(y_j)}{S(y_0)}\right]^2 \approx [S_n(y_j)]^2 \text{ and } \tfrac{S(y_i)}{S(y_{i-1})} \approx 1 - \tfrac{s_i}{r_i}$$

Using these estimates in (54.3) we obtain

$$\text{Var}[S_n(y_j)] \approx [S_n(y_j)]^2 \sum_{i=1}^{j} \frac{s_i}{r_i(r_i - s_i)}. \tag{54.4}$$

Equation (54.4) is known as **Greenwood's approximation**.

Remark 54.1
For non-death ages, the convention is to take the sum up to the last death age that is less than or equal to to the age under consideration.

Example 54.2 ‡
For a survival study with censored and truncated data, you are given:

Time (t)	Number at risk at time t	Failures at time t
1	30	5
2	27	9
3	32	6
4	25	5
5	20	4

The probability of failing at or before Time 4, given survival past Time 1, is ${}_3q_1$. Calculate Greenwood's approximation of the variance of the estimator ${}_3\hat{q}_1$.

Solution.
Let ${}_3p_1$ be the probability that an individual which has survived past time 1 will also survive past time 4. Then ${}_3p_1 = 1 - {}_3q_1$ which implies

$$\widehat{\text{Var}}({}_3\hat{q}_1) = \widehat{\text{Var}}({}_3\hat{p}_1).$$

Now, by (54.3), we have

$$_3\hat{p}_1 \approx \frac{S(4)}{S(1)} \approx S_5(4).$$

Since we are concerned with the individuals who survived time 1, then $y_0 = 1$ so that the next death time is 2. Hence, the calculation in Greenwood's approximation must start from time 2 and not time 1. Having this in mind, we have

$$_3\hat{p}_1 = \left(1 - \frac{9}{27}\right)\left(1 - \frac{6}{32}\right)\left(1 - \frac{5}{25}\right) = \frac{13}{30}$$

and

$$\widehat{\text{Var}}(_3\hat{q}_1) = \left(\frac{13}{30}\right)^2 \left[\frac{9}{27(27-9)} + \frac{6}{32(32-6)} + \frac{5}{25(25-5)}\right] = 0.0067 \blacksquare$$

Example 54.3 ‡
The following is a sample of 10 payments:

$$4 \;\; 4 \;\; 5^+ \;\; 5^+ \;\; 5^+ \;\; 8 \;\; 10^+ \;\; 10^+ \;\; 12 \;\; 15$$

where + indicates that a loss exceeded the policy limit. Determine Greenwood's approximation to the variance of the product-limit estimate $S_{10}(11)$.

Solution.
We first create the following table summarizing y_j, r_j, and s_j.

y_j	r_j	s_j
4	10	2
8	5	1
12	2	1
15	1	1

Since 11 is an uncensored value (loss), the largest loss less than or equal to 11 is 8. Thus,

$$\widehat{\text{Var}}(S_{10}(11)) = \widehat{\text{Var}}(S_{10}(8)).$$

The Kaplan-Meier estimate is

$$S_{10}(8) = \left(1 - \frac{2}{10}\right)\left(1 - \frac{1}{5}\right) = \frac{16}{25}.$$

Hence

$$\widehat{\text{Var}}(S_{10}(11)) = \left(\frac{16}{25}\right)^2 \left[\frac{2}{10(10-2)} + \frac{1}{5(5-1)}\right] = 0.03072 \blacksquare$$

Practice Problems

Problem 54.1
Let $X_1, X_2, \cdots, X_n$ be indepedent random variable. Show that

$$\text{Var}(X_1X_2\cdots X_n) = \prod_{i=1}^{n}(\mu_i^2 + \sigma_i^2) - \prod_{i=1}^{n}\mu_i^2.$$

Problem 54.2
Show that if $0 < a_i << 1$ for $i = 1, 2, \cdots, n$ then

$$(1 + a_1)(1 + a_2)\cdots(1 + a_n) \approx 1 + a_1 + a_2 + \cdots + a_n.$$

Problem 54.3
For a mortality study with right-censored data, you are given:

Time t_j	Number of deaths s_j	Number at risk r_j
5	2	15
7	1	12
10	1	10
12	2	6

Find the approximate variance of the Product-Limit estimate of $\hat{S}(10)$.

Problem 54.4
For a mortality study with right-censored data, you are given:

Time t_j	Number of deaths s_j	Number at risk r_j
5	2	15
7	1	12
10	1	10
12	2	6

Find the approximate variance of the Product-Limit estimate of ${}_7\hat{p}_5$.

Problem 54.5 ‡
For 200 auto accident claims you are given:
(i) Claims are submitted t months after the accident occurs, where $t = 0, 1, 2, \cdots$.
(ii) There are no censored observations.

(iii) $\hat{S}(t)$ is calculated using the Kaplan-Meier product limit estimator.
(iv) $c_S^2(t) = \frac{\widehat{\mathrm{Var}}(\hat{S}(t))}{\hat{S}(t)^2}$, where $\widehat{\mathrm{Var}}(\hat{S}(t))$ is calculated using Greenwood's approximation.
(v) $\hat{S}(8) = 0.22,\ \hat{S}(9) = 0.16,\ c_S^2(9) = 0.02625,\ c_S^2(10) = 0.04045.$

Determine the number of claims that were submitted to the company 10 months after an accident occurred.

55 Variance Estimate of the Nelson-Åalen Estimator and Confidence Intervals

We start this section by deriving an estimate formula for the variance of the Nelson-Åalen estimator. We assume r_i to be fixed so that the random quantity is s_i. We also assume that S_i has a Poisson distribution with parameter s_i. In this case, we can write

$$\widehat{\mathrm{Var}}(\hat{H}(y_j)) = \widehat{\mathrm{Var}}\left(\sum_{i=1}^{j} \frac{S_i}{r_i}\right) = \sum_{i=1}^{j} \frac{\widehat{\mathrm{Var}}(S_i)}{r_i^2} = \sum_{i=1}^{j} \frac{s_i}{r_i^2} \qquad (55.1)$$

where we assume that the S_js are independent.

Example 55.1
For a survival study with censored and truncated data, you are given:

Time (t)	Number at risk at time t	Failures at time t
1	30	5
2	27	9
3	32	6
4	25	5
5	20	4

Estimate the variance of $\hat{H}(4)$.

Solution.
The estimated variance is given by

$$\widehat{\mathrm{Var}}(\hat{H}(y_4)) = \sum_{i=1}^{4} \frac{s_i}{r_i^2} = \frac{5}{30^2} + \frac{9}{27^2} + \frac{6}{32^2} + \frac{5}{25^2} = 0.0318 \blacksquare$$

We next look at computing the confidence intervals (see Section 47) of both the Kaplan-Meier estimator and the Nelson-Åalen estimator. But first we consider the following example.

Example 55.2
You are given:
- The Kaplan-Meier estimator: $\hat{S}(3) = 0.8667$.
- The Greenwwod approximation: $\widehat{\mathrm{Var}}(\hat{S}(3)) = 0.0077$.
- $\hat{S}(t)$ is approximated by a normal distribution.

Find the 95% confidence interval of $\hat{S}(3)$.

Solution.
Using normal approximation, the 95% confidence interval can be expressed as
$$0.8667 \pm 1.96\sqrt{0.0077}$$
or in interval notation as
$$(0.6947, 1.0387).$$
Confidence intervals of this type are referred to as **linear confidence intervals** ■

A disadvantage of the above estimate is that the computed confidence interval may fall outside the range $(0,1)$ which is not desirable for confidence intervals involving probabilities and survival functions. One way to remedy this drawback is to use the so-called **log-transformed confidence intervals** which we state without justification: A $100(1-\alpha)\%$ confidence interval of $S_n(t)$ is given by
$$(S_n(t)^{\frac{1}{U}}, S_n(t)^U)$$
where
$$U = \exp\left[z_{\frac{\alpha}{2}}\frac{\sqrt{\widehat{\mathrm{Var}}(\hat{S}_n(t))}}{S_n(t)\ln S_n(t)}\right].$$
Log-transformed confidence intervals never result in probabilities outside the range from 0 to 1, which is extremely desirable for confidence intervals involving probabilities and survival functions.

Example 55.3
Obtain the log-transformed confidence interval to $\hat{S}(3)$ as in Example 55.2.

Solution.
We have
$$U = \exp\left[1.96\frac{\sqrt{0.0077}}{0.8667\ln 0.8667}\right] \approx 0.2498.$$
Thus, the confidence interval is
$$(0.8667^{\frac{1}{0.2498}}, 0.8667^{0.2498}) = (0.564, 0.9649) ■$$

Similar results are available for the Nelson-Åalen estimators. We define the **linear** $(1-\alpha)$ **confidence interval** for the cumulative hazard rate function by
$$\hat{H}(t) \pm z_{\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{H}(y_j))},\ y_j \le t < y_{j+1}.$$

The corresponding log-transformed confidence interval is given by

$$\left(\frac{\hat{H}(t)}{U}, \hat{H}(t)U\right)$$

where

$$U = \exp\left[z_{\frac{\alpha}{2}}\frac{\sqrt{\widehat{\mathrm{Var}}(\hat{H}(y_j))}}{\hat{H}(t)}\right].$$

Example 55.4
You are given:

y_j	r_j	s_j
1	50	4
2	53	5
3	32	9
4	45	11
5	20	2

(i) Find $\hat{H}(5)$ and $\widehat{\mathrm{Var}}(\hat{H}(5))$.
(ii) Obtain the 95% linear confidence interval to $H(5)$.
(iii) Obtain the 95% log-transformed confidence interval to $H(5)$.

Solution.
(i) We have

$$\hat{H}(5) = \hat{H}(y_5) = \sum_{j=1}^{5}\frac{s_j}{r_j} = 0.8$$

and

$$\widehat{\mathrm{Var}}(\hat{H}(5)) = \widehat{\mathrm{Var}}(\hat{H}(y_5)) = \sum_{j=1}^{5}\frac{s_j}{r_j} = 0.0226.$$

(ii)The linear confidence interval is

$$(0.8 - 1.96\sqrt{0.0226}, 0.8 + 1.96\sqrt{0.0226}) = (0.5053, 1.09467).$$

(iii) The log-transformed confidence interval is

$$(0.8e^{-1.96\frac{\sqrt{0.0266}}{0.8}}, 0.8e^{1.96\frac{\sqrt{0.0266}}{0.8}}) = (0.5365, 1.1562)\ \blacksquare$$

Example 55.5 ‡
A survival study gave $(1.63, 2.55)$ as the 95% linear confidence interval for the cumulative hazard function $H(t_0)$.
Calculate the 95% log-transformed confidence interval for $H(t_0)$.

Solution.
The interval $(1.63, 2.55)$ has endpoints that can be written as 2.09 ± 0.46. Thus, $\hat{H}(t_0) = 2.09$ and $1.96\hat{\sigma} = 0.46$. Hence, $\hat{\sigma} = 0.2347$. For the log-transformed confidence interval, we first find

$$U = e^{1.96\left(\frac{0.2347}{2.09}\right)} = 1.2462.$$

Thus, the confidence interval is

$$\left(\frac{\hat{H}(t)}{U}, \hat{H}(t)U\right) = \left(\frac{2.09}{1.2462}, 2.09(1.2462)\right) = (1.68, 2.60) \blacksquare$$

Example 55.6 ‡
For a survival study, you are given:
(i) Deaths occurred at times $y_1 < y_2 < \cdots < y_9$.
(ii) The Nelson-Åalen estimates of the cumulative hazard function at y_3 and y_4 are

$$\hat{H}(y_3) = 0.4128 \text{ and } \hat{H}(y_4) = 0.5691$$

(iii) The estimated variances of the estimates in (ii) are:

$$\widehat{\mathrm{Var}}(\hat{H}(y_3)) = 0.009565 \text{ and } \widehat{\mathrm{Var}}(\hat{H}(y_4)) = 0.014448.$$

Determine the number of deaths at y_4.

Solution.
The number of deaths at time y_4 is s_4. From the Nelson-Åalen estimate fiormula, we have

$$\frac{r_4}{s_4} = \hat{H}(y_4) - \hat{H}(y_3) = 0.5691 - 0.4128 = 0.1563.$$

From Equation (55.1), we have

$$\frac{s_4}{r_4^2} = \widehat{\mathrm{Var}}(\hat{H}(y_4)) - \widehat{\mathrm{Var}}(\hat{H}(y_3)) = 0.014448 - 0.009565 = 0.004883.$$

Thus,

$$s_4 = \frac{\left(\frac{s_4}{r_4}\right)^2}{\frac{s_4}{r_4^2}} = \frac{0.1563^2}{0.004883} = 5 \blacksquare$$

Example 55.7 ‡
You are given:
(i) Eight people join an exercise program on the same day. They stay in the program until they reach their weight loss goal or switch to a diet program.
(ii) Experience for each of the eight members is shown below:

	Time at which...	
Member j	Reach Weight Loss Goal x_j	Switch to Diet Program u_j
1		4
2		8
3	8	
4	12	
5		12
6	12	
7	22	
8	36	

(iii) The variable of interest is time to reach weight loss goal.
Using the Nelson-Åalen estimator, calculate the upper limit of the symmetric 90% linear confidence interval for the cumulative hazard rate function $H(12)$.

Solution.
Reaching weight loss goal is equivalent to death in mortality theory and switching to a dieting program is considered a censored observation. We have the following chart

j	y_j	s_j	r_j
1	8	1	7
2	12	2	5
3	22	1	2
4	36	1	1

By the Nelson-Åalen estimation, we have

$$\hat{H}(12) = \frac{s_1}{r_1} + \frac{s_2}{r_2} = \frac{1}{7} + \frac{2}{5} = 0.5429.$$

The estimated variance of the Nelson-Åalen estimate is

$$\widehat{\text{Var}}[\hat{H}(12)] = \frac{s_1}{r_1^2} + \frac{s_2}{r_2^2} = \frac{1}{7^2} + \frac{2}{5^2} = 0.1004.$$

The upper limit of the symmetric 90% linear confidence interval of $\hat{H}(12)$ is

$$0.5429 + 1.645\sqrt{0.1004} = 1.06 \blacksquare$$

Practice Problems

Problem 55.1 ‡
You are given the following information

y_j	r_j	s_j
1	30	5
2	27	9
3	32	6
4	25	5
5	20	4

Find $\hat{H}(3)$ and $\widehat{\mathrm{Var}}(\hat{H}(3))$.

Problem 55.2 ‡
Obtain the 95% linear confidence interval and the 95% log-transformed confidence interval in Problem 55.1

Problem 55.3 ‡
The interval $(0.357, 0.700)$ is a 95% log-transformed confidence interval for the cumulative hazard rate function at time t, where the cumulative hazard rate function is estimated using the Nelson-Åalen estimator.

Determine the value of the Nelson-Åalen estimate of $S(t)$.

Problem 55.4 ‡
Twelve policyholders were monitored from the starting date of the policy to the time of the first claim. The observed data are as follows.

Time of first claim	1	2	3	4	5	6	7
Number of claims	2	1	2	2	1	2	2

Using the Nelson-Åalen estimator, calculate the 95% linear confidence interval for the cumulative hazard rate function $H(4.5)$.

Problem 55.5 ‡
For a survival study, you are given:
(i) The Product-Limit estimator $S_n(t_0)$ is used to construct confidence intervals for $S(t_0)$.
(ii) The 95% log-transformed confidence interval for $S(t_0)$ is (0.695,0.843).

Determine $S_n(t_0)$.

Problem 55.6 ‡
Obtain the 95% log-transformed confidence interval for $H(3)$ in Example 55.1, based on the Nelson-Åalen estimate.

Problem 55.7 ‡
Fifteen cancer patients were observed from the time of diagnosis until the earlier of death or 36 months from diagnosis. Deaths occurred during the study as follows:

Time in Months Since Diagnosis	Number of Deaths
15	2
20	3
24	2
30	d
34	2
36	1

The Nelson-Åalen estimate $\hat{H}(35)$ is 1.5641.

Calculate the Åalen estimate of the variance of $\hat{H}(35)$.

56 Kernel Density Estimation

We have had a first encounter with kernel density models in Section 17. The objective of this section is to be able to create a continuous distribution whose pdf will approximate a given (discrete) empirical distribution consisting of a complete individual data set such as $\{y_1, y_2, \cdots, y_k\}$. This process is referred to as **kernel smoothing**.

The process of kernel smoothing goes as follows: For each observation $y_j,\ j = 1, 2, \cdots, k$, we let $p(y_j)$ be the probability assigned by the empirical distribution. We also create a continuous random variable X_j with mean y_j and with a pdf denoted by $k_{y_j}(x)$ and which we call the **kernel function**. The corresponding cdf will be denoted by $K_{y_j}(x)$.

Next, we define the **kernel density estimator**, also known as the **kernel smoothed estimate**, of the distribution function by

$$\hat{F}(x) = \sum_{j=1}^{k} p(y_j) K_{y_j}(x)$$

and the kernel density estimator of the density function by

$$\hat{f}(x) = \sum_{j=1}^{k} p(y_j) k_{y_j}(x).$$

In [1], three types of kernels are only considered: uniform, triangular, and Gamma. The **uniform kernel with bandwith** b is given by

$$k_y(x) = \begin{cases} 0, & x < y - b, \\ \frac{1}{2b}, & y - b \le x \le y + b, \\ 0, & x > y + b. \end{cases}$$

The corresponding cdf is given by

$$K_y(x) = \begin{cases} 0, & x < y - b, \\ \frac{x-y+b}{2b}, & y - b \le x \le y + b, \\ 1, & x > y + b. \end{cases}$$

Note that $k_{y_j}(x)$ is the pdf of a uniform distribution in the interval $[y_j - b, y_j + b]$.

Example 56.1
You are given the following ages at time of death of 10 individuals:

$$25 \quad 30 \quad 35 \quad 35 \quad 37 \quad 39 \quad 45 \quad 47 \quad 49 \quad 55.$$

Use a uniform kernel of bandwidth 10 to estimate the probability density function at $x = 40$.

Solution.
With $b = 10$, we have that $k_y(40) = 0$ for $y = 25$ and $y = 55$. Thus, we have

$$k_{30}(40) = k_{35}(40) = k_{37}(40) = k_{39}(40) = k_{45}(40) = k_{47}(40) = k_{49}(40) = \frac{1}{20}.$$

Hence,

$$\begin{aligned}\hat{f}(40) =& p(30)k_{35}(40) + p(35)k_{35}(40) + p(37)k_{37}(40) + p(39)k_{39}(40)\\ &+p(45)k_{45}(40) + p(47)k_{47}(40) + p(49)k_{49}(40)\\ =&6\left(\frac{1}{10}\right)\left(\frac{1}{20}\right) + \left(\frac{2}{10}\right)\left(\frac{1}{20}\right)\\ =&\frac{1}{25} \blacksquare\end{aligned}$$

The **triangular kernel with bandwith** b is given by

$$k_y(x) = \begin{cases} 0, & x < y - b, \\ \frac{x-y+b}{b^2}, & y - b \le x \le y, \\ \frac{y+b-x}{b^2}, & y \le x \le y + b, \\ 0, & x > y + b. \end{cases}$$

The corresponding cdf is given by

$$K_y(x) = \begin{cases} 0, & x < y - b, \\ \frac{(x-y+b)^2}{2b^2}, & y - b \le x \le y, \\ 1 - \frac{(y+b-x)^2}{2b^2}, & y \le x \le y + b, \\ 1, & x > y + b. \end{cases}$$

Example 56.2
You are given the following ages at time of death of 10 individuals:

$$25 \quad 30 \quad 35 \quad 35 \quad 37 \quad 39 \quad 45 \quad 47 \quad 49 \quad 55.$$

Using a triangular kernel with bandwith 10, find the kernel smoothed density estimate $\hat{f}(40)$.

Solution.
The triangular kernel with bandwith 10 is

$$k_y(x) = \begin{cases} 0, & x < y-10, \\ \frac{x-y+10}{100}, & y-10 \le x \le y, \\ \frac{y+10-x}{100}, & y \le x \le y+10, \\ 0, & x > y+10. \end{cases}$$

We first create the following chart:

$y-b$	y	$y+b$
15	25	35
20	30	40
25	35	45
27	37	47
29	39	49
35	45	55
37	47	57
39	49	59
45	55	65

We have

$$\begin{aligned}\hat{f}(40) =& p(30)k_{30}(40) + p(35)k_{35}(40) + p(37)k_{37}(40) + p(39)k_{39}(40) \\ &+ p(45)k_{45}(40) + p(47)k_{47}(40) + p(49)k_{49}(40) \\ =& \frac{1}{10}\frac{0}{100} + \frac{2}{10}\frac{5}{100} + \frac{1}{10}\frac{7}{100} + \frac{1}{10}\frac{9}{100} \\ &+ \frac{1}{10}\frac{5}{100} + \frac{1}{10}\frac{3}{100} + \frac{1}{10}\frac{1}{100} \\ =& 0.035 \blacksquare\end{aligned}$$

The **Gamma kernel** has a Gamma distribution with parameters α and $\frac{y}{\alpha}$. That is,

$$k_y(x) = \frac{x^{\alpha-1}e^{-\frac{x\alpha}{y}}}{(y/\alpha)^\alpha \Gamma(\alpha)}.$$

Example 56.3
You are given the following ages at time of death of 10 individuals:

$$25 \quad 30 \quad 35 \quad 35 \quad 37 \quad 39 \quad 45 \quad 47 \quad 49 \quad 55.$$

Using a triangular kernel with $\alpha = 1$, find the kernel smoothed density estimate $\hat{f}(40)$.

Solution.
With $\alpha = 1$, the kernel is expressed as follows:

$$k_y(x) = \frac{e^{-\frac{x}{y}}}{y}.$$

Thus,

$$\begin{aligned}\hat{f}(40) =& p(25)k_{25}(40) + p(30)k_{30}(40) + p(35)k_{35}(40) + p(37)k_{37}(40)\\ &+p(39)k_{39}(40) + p(45)k_{45}(40) + p(47)k_{47}(40)\\ &+p(49)k_{49}(40) + p(55)k_{55}(40)\\ =&\frac{1}{10}\frac{e^{-\frac{40}{25}}}{25} + \frac{1}{10}\frac{e^{-\frac{40}{30}}}{30} + \frac{2}{10}\frac{e^{-\frac{40}{35}}}{35}\\ &+\frac{1}{10}\frac{e^{-\frac{40}{37}}}{37} + \frac{1}{10}\frac{e^{-\frac{40}{39}}}{29} + \frac{1}{10}\frac{e^{-\frac{40}{45}}}{45}\\ &+\frac{1}{10}\frac{e^{-\frac{40}{47}}}{47} + \frac{1}{10}\frac{e^{-\frac{40}{49}}}{49} + \frac{1}{10}\frac{e^{-\frac{40}{55}}}{55}\\ =&0.0491 \blacksquare\end{aligned}$$

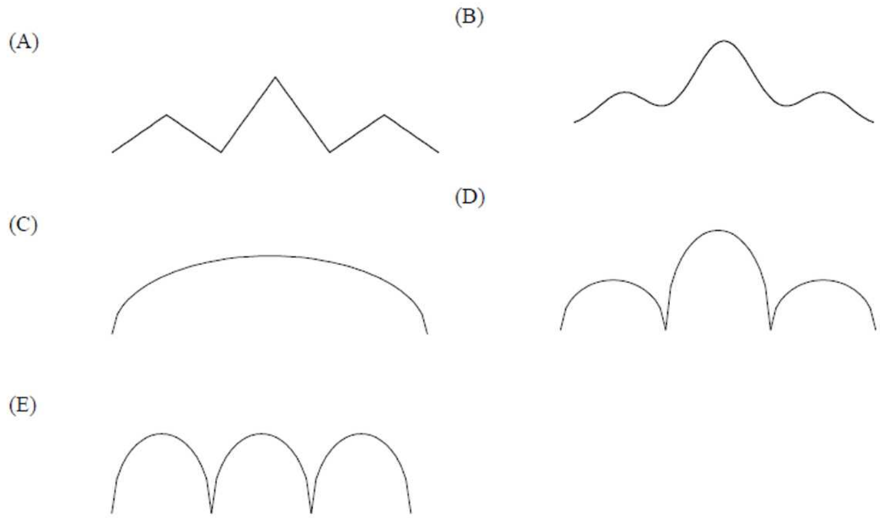
Example 56.4 ‡
You are given the kernel:

$$k_y(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-(x-y)^2}, & y-1 \le x \le y+1\\ 0, & \text{otherwise.}\end{cases}$$

You are also given the following random sample:

$$1 \quad 3 \quad 3 \quad 5$$

Determine which of the following graphs shows the shape of the kernel density estimator.



Solution.
We are given that $y_1 = 1,\ y_2 = 3$, and $y_3 = 5$. The empirical probabilities are $p(y_1) = p(y_3) = 0.25$ and $p(y_2) = 0.5$. The kernel densities are

$$k_1(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-(x-1)^2}, & 0 \le x \le 2 \\ 0, & \text{otherwise.} \end{cases}$$

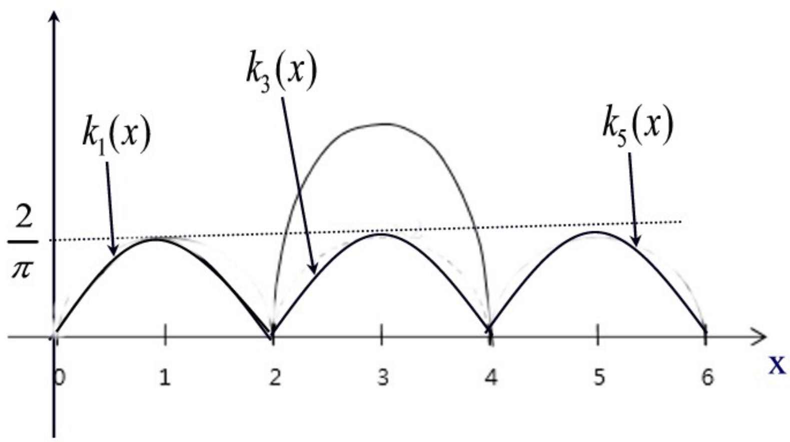
$$k_3(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-(x-3)^2}, & 2 \le x \le 4 \\ 0, & \text{otherwise.} \end{cases}$$

$$k_5(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-(x-5)^2}, & 4 \le x \le 6 \\ 0, & \text{otherwise.} \end{cases}$$

The graphs of $k_1(x),\ k_3(x)$, and $k_5(x)$ are shown below. The kernel density estimator is

$$\hat{f}(x) = \sum_{i=1}^{3} p(y_i)k_{y_i}(x) = 0.25k_1(x) + 0.5k_3(x) + 0.25k_5(x).$$

We see that the middle curve has double the coefficient as the curves on the left and right, so the middle curve is doubled.



Thus, the answer is (D) ■

Example 56.5 ‡
You are given:
(i) The sample:

$$1 \quad 2 \quad 3 \quad 3 \quad 3 \quad 3 \quad 3 \quad 3 \quad 3 \quad 3$$

(ii) $\hat{F}_1(x)$ is the kernel density estimator of the distribution function using a uniform kernel with bandwidth 1.
(iii) $\hat{F}_2(x)$ is the kernel density estimator of the distribution function using a triangular kernel with bandwidth 1.
Determine the interval where $\hat{F}_1(x) = \hat{F}_2(x)$.

Solution.
The empirical distribution of the data set is given by: $p(1) = \frac{1}{10} = 0.1$, $p(2) = \frac{1}{10} = 0.1$, and $p(3) = \frac{8}{10} = 0.8$.
The kernel density estimator of the distribution function using a uniform kernel with bandwidth 1 is

$$\hat{F}_1(x) = 0.1K_1^u(x) + 0.1K_2^u(x) + 0.8K_3^u(x)$$

where

$$K_1^u(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \le x \le 2 \\ 1, & x > 2 \end{cases}, K_2^u(x) = \begin{cases} 0, & x < 1 \\ \frac{x-1}{2}, & 1 \le x \le 3 \\ 1, & x > 3 \end{cases}, K_3^u(x) = \begin{cases} 0, & x < 2 \\ \frac{x-2}{2}, & 2 \le x \le 4 \\ 1, & x > 4. \end{cases}$$

Thus,

$$\hat{F}_1(x) = \begin{cases} 0, & x < 0 \\ 0.05x, & 0 \le x \le 1 \\ 0.1x - 0.05, & 1 \le x \le 2 \\ 0.45x - 0.75, & 2 \le x \le 3 \\ 0.4x - 0.6, & 3 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

The kernel density estimator of the distribution function using a triangular kernel with bandwidth 1 is

$$\hat{F}_2(x) = 0.1K_1^t(x) + 0.1K_2^t(x) + 0.8K_3^t(x)$$

where

$$K_1^t(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{2}, & 0 \le x \le 1 \\ 1 - \frac{(2-x)^2}{2}, & 1 \le x \le 2 \\ 1, & x > 2 \end{cases}, K_2^t(x) = \begin{cases} 0, & x < 1 \\ \frac{(x-1)^2}{2}, & 1 \le x \le 2 \\ 1 - \frac{(3-x)^2}{2}, & 2 \le x \le 3 \\ 1, & x > 3 \end{cases}$$

and

$$K_3^t(x) = \begin{cases} 0, & x < 2 \\ \frac{(x-2)^2}{2}, & 2 \le x \le 3 \\ 1 - \frac{(4-x)^2}{2}, & 3 \le x \le 4 \\ 1, & x > 4. \end{cases}$$

Thus,

$$\hat{F}_2(x) = \begin{cases} 0, & x < 0 \\ 0.05x^2, & 0 \le x \le 1 \\ 0.1x - 0.05, & 1 \le x \le 2 \\ 0.35x^2 - 1.3x + 1.35, & 2 \le x \le 3 \\ -0.4x^2 + 3.2x - 5.4, & 3 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

It follows that $\hat{F}_1(x) = \hat{F}_2(x)$ for all $1 \le x \le 2$ ■

Practice Problems

Problem 56.1 ‡
You are given the following ages at time of death of 10 individuals:

$$25 \quad 30 \quad 35 \quad 35 \quad 37 \quad 39 \quad 45 \quad 47 \quad 49 \quad 55.$$

Using a uniform kernel of bandwidth 10, determine the kernel density estimate of the probability of survival to age 40.

Problem 56.2 ‡
From a population having distribution function F, you are given the following sample:

$$2.0 \quad 3.3 \quad 3.3 \quad 4.0 \quad 4.0 \quad 4.7 \quad 4.7 \quad 4.7$$

Calculate the kernel density estimate of $F(4)$ using the uniform kernel with bandwidth 1.4.

Problem 56.3 ‡
You use a uniform kernel density estimator with $b = 50$ to smooth the following workers compensation loss payments:

$$82 \quad 126 \quad 161 \quad 294 \quad 384.$$

If $\hat{F}(x)$ denotes the estimated distribution function and $F_5(x)$ denotes the empirical distribution function, determine $|\hat{F}(150) - F_5(150)|$.

Problem 56.4 ‡
You study five lives to estimate the time from the onset of a disease to death. The times to death are:

$$2 \quad 3 \quad 3 \quad 3 \quad 7$$

Using a triangular kernel with bandwidth 2, estimate the density function at 2.5.

Problem 56.5 ‡
You are given:
(i) The sample: 1 2 3 3 3 3 3 3 3 3.
(ii) $\hat{F}_1(x)$ is the kernel density estimator of the distribution function using a uniform kernel with bandwidth 1.
(iii) $\hat{F}_2(x)$ is the kernel density estimator of the distribution function using a triangular kernel with bandwidth 1.

Determine the interval(s) where $\hat{F}_1(x) = \hat{F}_2(x)$.

Problem 56.6 ‡
You study five lives to estimate the time from the onset of a disease to death. The times to death are:

$$2 \quad 3 \quad 3 \quad 3 \quad 7$$

Using a triangular kernel with bandwidth 2, estimate the density function at 2.5.

57 The Kaplan-Meier Approximation for Large Data Sets

In this section, we consider modifying the Kaplan-Meier approximation of survival function for small data sets to sets with considerably a large number of data.

Following the notation of [1], suppose the sample data can be split into k intervals with boundary points $c_0 < c_1 < \cdots < c_k$. Let d_j denote the number of observations that are left-truncated at some value within the interval $[c_j, c_{j+1})$. In mortality terms, d_j is the number of lives that were first observed at an age in the given range. Let u_j be the number of observations that are right censored (individual leaving the study for reason other than death) at some value within the interval $(c_j, c_{j+1}]$. Note that the intervals for d_j and u_j differ by which endpoints are included and which are omitted. This is because left-truncation is not possible at the right end of a closed interval, while right-censoring is not possible at the left end of a closed interval. Let x_j be the number of uncensored observations (observed deaths) within the interval $(c_j, c_{j+1}]$. With these notation, the sample size can be axpressed as $n = \sum_{j=1}^{k-1} d_j = \sum_{j=1}^{k-1}(u_j + x_j)$.

In developping the Kaplan-Meier approximation, the following assumptions are made:

(1) All truncated values occur at the left-endpoint of the intervals and all censored values occur at the right-endpoints.

(2) None of the uncensored values fall at the endpoints of the intervals.

(3) $\hat{S}(c_0) = 1$.

With these assumptions, the number at risk[14] for the first interval is $r_0 = d_0 =$ all the new entrants for the first interval. For this interval, $\hat{S}(c_1) = 1 - \frac{x_0}{d_0}$.

For the second interval, the number at risk is

$$r_1 = d_0 + d_1 - x_0 - u_0.$$

[14] Recall that the risk set is the number of observations available at a given time that could produce an uncensored observation at that time.

This is the survivors from the first interval $(d_0 - x_0 - u_0)$ plus the new entrants d_1 in the second interval. For this interval,

$$\hat{S}(c_2) = \left(1 - \frac{x_0}{r_0}\right)\left(1 - \frac{x_1}{r_1}\right).$$

The general formulas are:

$$\begin{aligned} r_0 =& d_0 \\ r_j =& \sum_{i=0}^{j} d_i - \sum_{i=0}^{j-1}(x_i + u_i) \\ \hat{S}(c_0) =& 1 \\ \hat{S}(c_j) =& \prod_{i=0}^{j-1}\left(1 - \frac{x_i}{r_i}\right) \end{aligned}$$

where $j = 1, 2, \cdots, k$. Note that the survival function estimate is valued at the endpoints. For non-endpoints, a linear interpolation between endpoints is used.

We can use the above developed approximation to estimate the probability that, given someone is alive at age c_j, that person does not survive past age c_{j+1}. This estimation is given by

$$\begin{aligned} \hat{q}_j \approx& \Pr(T \le c_{j+1} | T > c_j) \approx \frac{\hat{S}(c_j) - \hat{S}(c_{j+1})}{\hat{S}(c_j)} \\ =& \frac{\prod_{i=0}^{j-1}\left(1 - \frac{x_i}{r_i}\right) - \prod_{i=0}^{j}\left(1 - \frac{x_i}{r_i}\right)}{\prod_{i=0}^{j-1}\left(1 - \frac{x_i}{r_i}\right)} \\ =& 1 - \left(1 - \frac{x_j}{r_j}\right) = \frac{x_j}{r_j}. \end{aligned}$$

That is,

$$\hat{q}_j = \frac{\text{number of deaths in time period}}{\text{number of lives considered during that time period}}.$$

Remark 57.1
The reader needs to be aware of the difference of notations for x_j, d_j, u_j, and r_j used in this section and the ones used in Section 52. See Problem 57.4.

Example 57.1
Below is the data for a 5-year mortality study.

j	Age c_j	Number of persons joining the study at this age (d_j)	Number of persons withdrawing at or before the next age u_j	Number of deaths x_j
0	45	800	85	10
1	46	50	65	8
2	47	65	55	6
3	48	45	35	4
4	49	30	25	2

Assume that the activities of withdrawing and joining occur at integral ages. Use the Kaplan-Meier approximation for large data sets to estimate the probability of survival to age 47.

Solution.
We first create the following table:

j	c_j	d_j	u_j	x_j	r_j
0	45	800	85	10	800
1	46	50	65	8	755
2	47	65	55	6	747
3	48	45	35	4	731
4	49	30	25	2	722

Thus,

$$\hat{S}(2) = {}_2\hat{p}_{45} = \left(1 - \frac{x_0}{r_0}\right)\left(1 - \frac{x_1}{r_1}\right)\left(1 - \frac{x_2}{r_2}\right) = 0.9692 \blacksquare$$

Example 57.2 ‡
Loss data for 925 policies with deductibles of 300 and 500 and policy limits of 5,000 and 10,000 were collected. The results are given below:

Loss range	300 deductible	Policy type (II)
(300,500]	50	–
(500,1000]	50	75
(1000,5000]	150	150
(5000,10000]	100	200
at 5000	40	80
at 10000	10	20
Total	400	525

The ground-up loss distribution for both types of policy is assumed to be the same. Using the Kaplan-Meier approximation for large data sets to estimate $F(5000)$.

Solution.
The boundaries of the intervals are: $c_0 = 300,\ c_1 = 500,\ c_2 = 1000,\ c_3 = 5000$ and $c_4 = 10000$. Recall that d_j is the number observations that are left-truncated at some value in $[c_j, c_{j+1}); u_j$ is the number of observations that are right censored at some value in $(c_j, c_{j+1}]$; and x_j is the number of uncensored observations in (c_j, c_{j+1}). We have the following chart

j	d_j	u_j	x_j	r_j
0	400	–	50	400
1	525	–	125	875
2	0	120	300	750
3	0	30	300	330
4	0	–	–	0

Using the Kaplan-Meier approximation we find

$$\hat{S}(5000) = \hat{S}(c_3) = \left(1 - \frac{x_0}{r_0}\right)\left(1 - \frac{x_1}{r_1}\right)\left(1 - \frac{x_2}{r_2}\right) = .45$$

Hence,

$$\hat{F}(5000) = 1 - 0.45 = 0.55 \ \blacksquare$$

Example 57.3 ‡
The following table was calculated based on loss amounts for a group of motorcycle insurance policies:

c_j	d_j	u_i	x_j	$P_j = \sum_{i=0}^{j-1}(d_i - u_i - x_i)$
250	6	0	1	0
500	6	0	2	5
1000	7	1	4	9
2750	0	1	7	11
5500	0	1	1	3
6000	0	0	1	1
10,000	0	0	0	0

Estimate the probability that a policy with a deductible of 500 will have a claim payment in excess of 5500.

Solution.
First note that the risk set is $r_j = P_j + d_j$ with $P_0 = 0$. The insurance will pay over 5500 if the insured claim is above 6000 (because of the deductible). Thus, we are asked to estimate $\Pr(X > 6000|X > 500)$ which by Bayes' theorem is

$$\Pr(X > 6000|X > 500) = \frac{\Pr(X > 6000)}{\Pr(X > 500)} = \frac{S(6000)}{S(500)}.$$

We have

$$\begin{aligned}
\hat{S}(500) =&\hat{S}(c_1) = 1 - \frac{x_0}{r_0}\\
=&1 - \frac{1}{6} = 0.83333\\
\hat{S}(6000) =&\hat{S}(c_5) = \prod_{j=0}^{4}\left(1 - \frac{x_j}{r_j}\right)\\
=&\left(\frac{5}{6}\right)\left(\frac{9}{11}\right)\left(\frac{12}{16}\right)\left(\frac{4}{11}\right)\left(\frac{2}{3}\right)\\
=&0.12397.
\end{aligned}$$

Hence,

$$\Pr(X > 6000|X > 500) \approx \frac{0.12397}{0.83333} = 0.14876 \blacksquare$$

In life table applications, $\hat{q}_j$ is a an example of a **single-decrement probabilities**. When there are multiple causes of decrement, we can express life table functions pertaining to all causes of decrement with a right superscript such as $q_j^{(\tau)}$. For a single cause of decrement we will use the notation $q_j^{'(i)}$.

Suppose that there are n causes for a decrement, then the following is true:

$$q_j^{(\tau)} = 1 - \prod_{i=1}^{n}(1 - p_j^{'(i)}).$$

Hence, Kaplan-Meier approximation of a survival function can be applied in the case of multiplt-decrement probabilities.

Example 57.4 ‡
For a double-decrement study, you are given:
(i) The following survival data for individuals affected by both decrements (1) and (2):

j	c_j	$q_j^{(\tau)}$
0	0	0.100
1	20	0.182
2	40	0.600
3	60	1.000

(ii) $q_j^{'(2)} = 0.05$ for all j.
(iii) Group A consists of 1000 individuals observed at age 0.
(iv) Group A is affected by only decrement (1).
Determine the Kaplan-Meier multiple-decrement estimate of the expected number of individuals in Group A that survive to be at least 40 years old.

Solution.
First, the notation $q_j^{(\tau)}$ stands for the probability that a person at age c_j will departs due to some decrement by time c_{j+1}. Also,

$$q_j^{(\tau)} = 1 - (1 - q_j^{'(1)})(1 - q_j^{'(2)}).$$

From these equations, we find

$$0.1 = q_0^{(\tau)} = 1-(1-q_0^{'(1)})(1-q_0^{'(2)}) = 1-(1-q_0^{'(1)})(1-0.05) \Longrightarrow 1-q_0^{'(1)} = 0.9474.$$

Likewise,

$$0.182 = q_1^{(\tau)} = 1-(1-q_1^{'(1)})(1-q_1^{'(2)}) = 1-(1-q_1^{'(1)})(1-0.05) \Longrightarrow 1-q_1^{'(1)} = 0.8611.$$

Since Group A is affected only by decrement 1, the survival probability to age 40= c_2 for Group A is

$$(1 - q_0^{'(1)})(1 - q_0^{'(1)}) = (0.9474)(0.8611) = 0.8158.$$

The expected number of survivors to age 40 from 1000 Group A individuals observed at 0 is $1000(0.8158) \approx 816$ ■

Life Table Approach
A second approach for estimating the survival function that is mentioned in [1] is the **life table approach**. In this approach, assumptions (1)-(2) are replaced by the following assumptions:

(1)' Truncation points and censoring points occur uniformly in each interval.

(2)' Uncensored observations all occur at the midpoint of the interval.

For such an approach, the risk set r_i for an interval $(c_i, c_{i+1}]$ is found by the formula

r_i = number of observations with value each greater than c_{i_1}+ number of uncensored observations in $(c_i, c_{i+1}]$ + half the total number of censored and truncated observations in $(c_i, c_{i+1}]$.

We illustrate this approach in the next example.

Example 57.5 ‡
You are given the following information about a group of 10 claims:

Claim Size Interval	Number of Claims in Interval	Number of Claims Censored in Interval
$(0 - 15,000]$	1	2
$(15,000 - 30,000]$	1	2
$(30,000 - 45,000]$	4	0

Assume that claim sizes and censorship points are uniformly distributed within each interval.
Estimate, using the life table methodology, the probability that a claim exceeds 30,000.

Solution.
We are asked to find $\hat{S}(30,000)$ which is given by

$$\hat{S}(30,000) = \left(1 - \frac{x_0}{r_0}\right)\left(1 - \frac{x_1}{r_1}\right)$$

where $x_0 = x_1 = 1$ (uncensored observations). By the life table methodology, we have

$$\begin{aligned} r_0 =& 7 + 1 + 1 = 9 \\ r_1 =& 4 + 1 + 1 = 6. \end{aligned}$$

Thus,

$$\hat{S}(3000) = \left(1 - \frac{1}{9}\right)\left(1 - \frac{1}{6}\right) = 0.741 \blacksquare$$

Practice Problems

Problem 57.1
Below is the data for a 5-year mortality study.

j	Age c_j	Number of persons joining the study at this age (d_j)	Number of persons withdrawing at or before the next age u_j	Number of deaths x_j
0	45	800	85	10
1	46	50	65	8
2	47	65	55	6
3	48	45	35	4
4	49	30	25	2

Assume that the activities of withdrawing and joining occur at integral ages.
(a) What is the sample size?
(b) What is the probability that a person does not survive past age 47, given that he was alive at age 45?

Problem 57.2
Show that $r_j = r_{j-1} + d_j - (x_{j-1} + u_{j-1})$.

Problem 57.3 ‡
Loss data for 925 policies with deductibles of 300 and 500 and policy limits of 5,000 and 10,000 were collected. The results are given below:

Loss range	300 deductible	Policy type (II)
(300,500]	50	–
(500,1000]	50	75
(1000,5000]	150	150
(5000,10000]	100	200
at 5000	40	80
at 10000	10	20
Total	400	525

The ground-up loss distribution for both types of policy is assumed to be the same.

Using the Kaplan-Meier approximation for large data sets to estimate $S(1000)$.

Problem 57.4 ‡
Loss data for 925 policies with deductibles of 300 and 500 and policy limits of 5,000 and 10,000 were collected. The results are given below:

Loss range	300 deductible	Policy type (II)
(300,500]	50	–
(500,1000]	50	75
(1000,5000]	150	150
(5000,10000]	100	200
at 5000	40	80
at 10000	10	20
Total	400	525

The ground-up loss distribution for both types of policy is assumed to be the same.

Estimate the probability that a loss will be greater than 3000 using a Kaplan-Meier type approximation for large data sets.

Methods of parameter Estimation

The purpose of this chapter is to discuss methods for the estimation of parameters in parametric models. Below we present an example of a parameter estimation.

Example 58.1 ‡
For a sample of 15 losses, you are given:
(i)

Interval	Observed number of Losses
$(0,2]$	5
$(2,5]$	5
$(5,\infty)$	5

(ii) Losses follow the uniform distribution on $(0,\theta)$.
Estimate θ by minimizing the function $\sum_{i=1}^{3}\frac{(E_i-O_i)^2}{O_i}$ where E_i is the expected number of losses in the i^{rmth} interval and O_j is the observed number of losses in the i^{th} interval.

Solution.
Since there are losses in $(5,\infty)$, we must have $\theta>5$. Since there are 15 losses and the probability of a loss to be in $(0,2]$ is $\frac{2-0}{\theta}$, the expected number of losses in that interval is $E_1=\frac{2}{\theta}(15)=\frac{30}{\theta}$. Likewise, $E_2=\frac{45}{\theta}$ and $E_3=15-\frac{75}{\theta}$. Hence, the formula given in the problem reduces to

$$f(\theta)=\frac{1}{5}\left[\left(\frac{30}{\theta}-5\right)^2+\left(\frac{45}{\theta}-5\right)^2+\left(10-\frac{75}{\theta}\right)^2\right].$$

Taking the derivative and setting it 0, we find

$$\frac{1}{5}\left[-2(30\theta^{-1}-5)(30\theta^{-2})-2(45\theta^{-1}-5)(45\theta^{-2})+2(10-75\theta^{-1})(75\theta^{-2})\right]=0$$

Multiply both sides by $-\frac{5}{2}\theta^3$ and simplify the resulting equation to obtain

$$8550-1125\theta=0\Longrightarrow\theta=7.60.$$

Thus, $f(\theta)$ has only one critical value. Moreover, $f''(\theta)=10260\theta^{-4}$ and $f''(7.60)>0$ so that $f(\theta)$ is minimized at $\theta=7.60$. That is, $\hat{\theta}=7.60$ ■

58 Method of Moments and Matching Percentile

In this section, we want to estimate parameters in a parametric model by using finite random samples taken from the underlying distribution. We will look at two methods of estimation: the method of moments and the method of percentile matching.

Let X be a random variable with parameters $\theta_1, \theta_2, \cdots, \theta_p$. From the distribution of X, select randomly n independent observations. We will denote the cdf of the distribution by

$$F(x|\theta) \quad \text{where} \quad \theta^T = (\theta_1, \theta_2, \cdots, \theta_p).$$

We will also denote the kth raw moment of X by $\mu'_k(\theta) = E(x^k|\theta)$. For a sample of n independent observations $\{x_1, x_2, \cdots, x_n\}$ from this distribution we let $\tilde{\mu}'_k = \frac{1}{n}\sum_{i=1}^n x_i^k$ be the empirical estimate of the kth moment.

A **method-of-moments estimate** of the vector θ is any solution to the following system of p equations

$$\mu'_k(\theta) = \tilde{\mu}'_k, \ k = 1, 2, \cdots, p.$$

Example 58.2
For a normal distribution, derive expressions for the method of moment estimators for the parameters μ and σ^2.

Solution.
We have to solve the following system of two equations

$$\mu = E(X) = \frac{1}{n}\sum_{i=1}^n x_i = \overline{X}$$

and

$$\frac{1}{n}\sum_{i=1}^n x_i^2 = E(X^2) = \sigma^2 + \mu^2.$$

Thus,

$$\tilde{\mu} = \overline{X} \text{ and } \tilde{\sigma^2} = \tfrac{1}{n}\sum_{i=1}^n x_i^2 - \overline{X}^2 \ \blacksquare$$

Example 58.3
For a Gamma distribution with a shape parameter α and a scale parameter θ, derive expressions for their method of moment estimators.

Solution.
We have to solve the following system of two equations

$$\alpha\theta = E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and

$$\alpha(\alpha+1)\theta = E(X^2) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

We have

$$\frac{\alpha(\alpha+1)\theta^2}{\alpha\theta} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{\overline{X}} \Longrightarrow \tilde{\theta} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{\overline{X}} - \overline{X}$$

and

$$\tilde{\alpha} = \overline{X}\left[\frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{\overline{X}} - \overline{X}\right]^{-1} \blacksquare$$

Example 58.4
For a Pareto distribution, derive expressions for the method of moment estimators for the parameters α and θ.

Solution.
We have to solve the following system of two equations

$$\frac{\theta}{\alpha-1} = \overline{X}$$

and

$$\frac{2\theta^2}{(\alpha-1)(\alpha-2)} = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

We have

$$\left[\frac{2\theta^2}{(\alpha-1)(\alpha-2)}\right]\left[\frac{\theta}{\alpha-1}\right]^{-2} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{\overline{X}^2}.$$

Solving this last equation, we find

$$\tilde{\alpha} = \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2\right]\left[\frac{1}{2n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2\right]^{-1}.$$

Also,

$$\tilde{\theta} = \overline{X}\left\{\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2\right]\left[\frac{1}{2n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2\right]^{-1} - 1\right\} \blacksquare$$

Let $100g$th percentile be denoted by $\pi_g(\theta)$ where $F(\pi_g(\theta)|\theta) = g$. Let $\hat{\pi}_g$ denote the **smoothed empirical estimate** of the $100g$th percentile.

A **percentile matching estimate** of θ is any solution of the p equations

$$\pi_{g_k}(\theta) = \hat{\pi}_{g_k},\ k = 1, 2, \cdots, p$$

or

$$F(\hat{\pi}_{g_k}|\theta) = g_k,\ k = 1, 2, \cdots, p$$

where $g_1, g_2, \cdots, g_p$ are arbitrarily chosen percentiles. In this book, the possible p values are either 1 or 2.

The smoothed empirical estimate $\hat{\pi}_g$ of the $100g$th percentile is found in the following way:

(i) order the sample values from smallest to largest: $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$;
(ii) find the integer p such that

$$\frac{p}{n+1} \le g \le \frac{p+1}{n+1};$$

(iii) $\hat{\pi}_g$ is found by linear interpolation

$$\hat{\pi}_g = [p + 1 - (n+1)g]x_{(p)} + [(n+1)g - p]x_{(p+1)}.$$

Example 58.5 ‡
A random sample of 20 observations has been ordered as follows:

12	16	20	23	26	28	30	32	33	35
36	38	39	40	41	43	45	47	50	57

Determine the 60^{th} sample percentile using the smoothed empirical estimate.

Solution.
We want an integer p such that

$$p \le 0.6(21) \le p + 1 \Longrightarrow p = 12.$$

Thus,

$$\hat{\pi}_{0.6} = [12 + 1 - (20+1)(0.6)](38) + [(20+1)(0.6) - 12](39) = 38.6 \blacksquare$$

Example 58.6 ‡
You are given:
(i) Losses follow a Burr distribution with parameters $\gamma,\ \theta$ and $\alpha = 2$.
(ii) A random sample of 15 losses is

195 255 270 280 350 360 365 380 415 450 490 550 575 590 615.

Use the smoothed empirical estimates of the 30^{th} and 65^{th} percentiles matching to estimate the parameters γ and θ.

Solution.
The cdf of the Burr distribution is

$$F(x|\theta) = 1 - \left[\frac{1}{1 + (x/\theta)^\gamma}\right]^2.$$

We want an integer $p_{0.3}$ such that

$$p_{0.3} \le 0.3(16) \le p_{0.3} + 1 \Longrightarrow p_{0.3} = 4.$$

Likewise, we want an integer $p_{0.65}$ such that

$$p_{0.65} \le (0.65)(16) \le p_{0.65} + 1 \Longrightarrow p_{0.65} = 10.$$

Hence,

$$\hat{\pi}_{0.3} = [5 - 16(0.3)](280) + [16(0.3) - 4](350) = 336$$

and

$$\hat{\pi}_{0.65} = [11 - 16(0.65)](450) + [16(0.65) - 10](490) = 466.$$

Hence, we have

$$1 - \left[\frac{1}{1 + (336/\theta)^\gamma}\right]^2 = 0.3 \Longrightarrow \left(\frac{336}{\theta}\right)^\gamma = 0.1952$$

and

$$1 - \left[\frac{1}{1 + (466/\theta)^\gamma}\right]^2 = 0.65 \Longrightarrow \left(\frac{466}{\theta}\right)^\gamma = 0.6903.$$

Hence,

$$\left(\frac{466}{336}\right)^\gamma = \left(\frac{466}{\theta}\right)^\gamma \left(\frac{336}{\theta}\right)^{-\gamma} = 3.5364.$$

Hence, $\tilde{\gamma} = \ln 3.5364[\ln 466 - \ln 336]^{-1} = 3.86$ and $\tilde{\theta} = 512.96$ ■

Example 58.7 ‡
For a sample of dental claims $X_1, X_2, \cdots, X_{10}$, you are given:
(i) $\sum_{i=1}^{10} X_i = 3860$ and $\sum_{i=1}^{10} X_i^2 = 4,574,802$.
(ii) Claims are assumed to follow a lognormal distribution with parameters μ and σ.
(iii) μ and σ are estimated using the method of moments.
Calculate $E(X \wedge 500)$ for the fitted distribution.

Solution.
We have

$$e^{\mu+0.5\sigma^2} = E(X) = \frac{\sum_{i=1}^{10} X_i}{10} = 386 \Longrightarrow \mu + 0.5\sigma^2 = 5.9558$$

and

$$e^{2\mu+2\sigma^2} = E(X^2) = \frac{\sum_{i=1}^{10} X_i^2}{10} = 457,480.2 \Longrightarrow 2\mu + 2\sigma^2 = 13.0335.$$

Solving this system of equations, we find $\hat{\mu} = 5.3949$ and $\hat{\sigma}^2 = 1.1218$.
Next, using A.5.1.1 in Table C, we have

$$\begin{aligned}
E(X \wedge 500) =& e^{5.3949+0.5(1.1218)}\Phi\left(\frac{\ln 500 - 5.3949 - 1.1218}{\sqrt{1.1218}}\right) \\
&+500\left[1 - \Phi\left(\frac{\ln 500 - 5.3949}{\sqrt{1.1218}}\right)\right] \\
=& e^{5.3949+0.5(1.1218)}\Phi(-0.2852) + 500[1 - \Phi(0.7739) \\
=& e^{5.3949+0.5(1.1218)}(0.3877) + 500(1 - 0.7805) \\
=& 259.4.
\end{aligned}$$

Note that the values of the standard normal distribution were obtained using Excel ■

Example 58.8 ‡
You are given:
(i) Losses follow an exponential distribution with mean θ.
(ii) A random sample of losses is distributed as follows:

Loss Range	Number of Losses
$(0-100]$	32
$(100-200]$	21
$(200-400]$	27
$(400-750]$	16
$(750-1000]$	2
$(1000-1500]$	2

Estimate θ by matching at the 80^{th} percentile.

Solution.
The empirical distribution of the grouped data (see Section 50) is expressed as follows

c	100	200	400	750	1000	1500
$F(c)$	0.32	0.53	0.8	0.96	0.98	1.00

Thus,

$$0.8 = F(400) = 1 - e^{-\frac{400}{\theta}} \Longrightarrow \hat{\theta} = 248.53 \blacksquare$$

Example 58.9 ‡
You are given the following sample of claim counts:

$$0 \quad 0 \quad 1 \quad 2 \quad 2$$

You fit a binomial (m, q) model with the following requirements:
(i) The mean of the fitted model equals the sample mean.
(ii) The 33^{rd} percentile of the fitted model equals the smoothed empirical 33^{rd} percentile of the sample.
Determine the smallest estimate of m that satisfies these requirements.

Solution.
Let N denote the binomial random variable. Then (i) yields

$$mq = E(N) = \overline{X} = \frac{0+0+1+2+2}{5} = 1.$$

For the 33^{rd} percentile of the sample, we seek p such that

$$p \leq 0.33(6) \leq p+1 \Longrightarrow p = 1.$$

Since the first and the second terms in the sample are 0, we conclude that the 33^{rd} percentile of the sample is 0. By (ii) and the definition of the percentile introduced in Section 6, we must have $p_0 = \Pr(N = 0) \geq 0.33$ where

$$p_0 = \left(1 - \frac{1}{m}\right)^m.$$

We now have

m	p_0
1	0
2	0.25
3	0.2963
4	0.3154
5	0.3277
6	0.3349

Hence, the smallest value of m is $m = 6$ ■

Example 58.10 ‡
You are given:
(i) Losses on a certain warranty product in Year i follow a lognormal distribution with parameters μ_i and σ_i.
(ii) $\sigma_i = \sigma$ for $i = 1, 2, 3$.
(iii) The parameters μ_i vary in such a way that there is an annual inflation rate of 10% for losses.
(iv) The following is a sample of seven losses:

Year 1:	20	40	50	
Year 2:	30	40	90	120.

Using trended losses, determine the method of moments estimate of μ_3.

Solution.
Counting inflation in both Year 1 and Year 2, the amount of the 7 losses in Year 3 are:

Year 1	Year 2	Year 3
20	20(1.1)	$20(1.1)^2 = 24.2$
40	40(1.1)	$40(1.1)^2 = 48.4$
50	50(1.1)	$50(1.1)^2 = 60.5$
–	30	$30(1.1) = 33$
–	40	$40(1.1) = 44$
–	90	$90(1.1) = 99$
–	120	$120(1.1) = 132$

We have

$$\begin{aligned}
\mu_1' =& \frac{24.2 + 48.4 + 60.5 + 33 + 44 + 99 + 132}{7} = 63.014 \\
\mu_2' =& \frac{24.2^2 + 48.4^2 + 60.5^2 + 33^2 + 44^2 + 99^2 + 132^2}{7} = 5252.64 \\
\mu_1 =& e^{\mu_3 + 0.5\sigma_3^2} \\
\mu_2 =& e^{2\mu_3 + 2\sigma_3^2}.
\end{aligned}$$

By the method of moments, we have the following system of equations

$$e^{\mu_3 + 0.5\sigma_3^2} = 63.014 \Longrightarrow \mu_3 + 0.5\sigma_3^2 = \ln 63.014$$

$$e^{2\mu_3 + 2\sigma_3^2} = 5252, 64 \Longrightarrow 2\mu_3 + 2\sigma_3^2 = \ln 5252.64.$$

Solving this system, we find

$$\mu_3 = \frac{4 \ln 63.014 - \ln 5252.64}{2} = 4.00 \blacksquare$$

Practice Problems

Problem 58.1 ‡
The 20^{th} and 80^{th} percentiles of a sample are 5 and 12. Using the percentile matching method, estimate $S(8)$ assuming the population has a Weibull distribution.

Problem 58.2 ‡
You are given the following information about a sample of data:
(i) Mean = 35,000
(ii) Standard deviation = 75,000
(iii) Median = 10,000
(iv) 90^{th} percentile = 100,000
(v) The sample is assumed to be from a Weibull distribution.

Determine the percentile matching estimate of the parameter τ.

Problem 58.3 ‡
You are given the following sample of five claims:

$$4 \quad 5 \quad 21 \quad 99 \quad 421$$

You fit a Pareto distribution using the method of moments. Determine the 95^{th} percentile of the fitted distribution.

Problem 58.4 ‡
In year 1 there are 100 claims with an average size of 10,000, and in year 2 there are 200 claims with an average size of 12,500. Inflation increases the size of all claims by 10% per year. A Pareto distribution with $\alpha = 3$ and θ unknown is used to model the claim size distribution.

Estimate θ for year 3 using the method of moments.

Problem 58.5 ‡
The following 20 wind losses (in millions of dollars) were recorded in one year:

1	1	1	1	1	2	2	3	3	4
6	6	8	10	13	14	15	18	22	25

Determine the 75^{th} sample percentile using the smoothed empirical estimate.

Problem 58.6 ‡
You are given:
(i) Losses follow a loglogistic distribution with cumulative distribution function:

$$F(x) = \frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma}.$$

(ii) The sample of losses is:

$$10 \quad 35 \quad 80 \quad 86 \quad 90 \quad 120 \quad 158 \quad 180 \quad 200 \quad 210 \quad 1500$$

Calculate the estimate of θ by percentile matching, using the 40^{th} and 80^{th} empirically smoothed percentile estimates.

Problem 58.7 ‡
You are given:
(i) A sample $x_1, x_2, \cdots, , x_{10}$ is drawn from a distribution with probability density function:

$$f(x) = \frac{1}{2}\left(\frac{1}{\theta}e^{-\frac{x}{\theta}} + \frac{1}{\sigma}e^{-\frac{x}{\sigma}}\right), \; x > 0.$$

(ii) $\theta > \sigma$.
(iii) $\sum_{i=1}^{10} x_i = 150$ and $\sum_{i=1}^{10} x_i^2 = 5000$.
Estimate θ by matching the first two sample moments to the corresponding population quantities.

Problem 58.8 ‡
You are given the following claim data for automobile policies:

$$200 \quad 255 \quad 295 \quad 320 \quad 360 \quad 420 \quad 440 \quad 490 \quad 500 \quad 520 \quad 1020$$

Calculate the smoothed empirical estimate of the 45^{th} percentile.

Problem 58.9 ‡
You are given:

x	0	1	2	3
$\Pr(X = x)$	0.5	0.3	0.1	0.1

The method of moments is used to estimate the population mean μ, and variance σ^2, by $\overline{X}$ nad $s_4^2 = \frac{1}{4}\sum_{i=1}^{4}(X_i - \overline{X})^2$, respectively.

Calculate the bias of s_4^2.

Problem 58.10 ‡
You are given:
(i) Claim amounts follow a shifted exponential distribution with probability density function:

$$f(x) = \frac{1}{\theta}e^{-\frac{x-\delta}{\theta}}, \ x > \delta.$$

(ii) A random sample of claim amounts $X_1, X_2, \cdots, X_{10}$

$$5 \ \ 5 \ \ 5 \ \ 6 \ \ 8 \ \ 9 \ \ 11 \ \ 12 \ \ 16 \ \ 23.$$

(iii) $\sum_{i=1}^{10} x_i = 100$ and $\sum_{i=1}^{10} x_i = 1306.$

Estimate δ using the method of moments.

Problem 58.11 ‡
You are given the following random sample of 13 claim amounts:

$$99 \ \ 133 \ \ 175 \ \ 216 \ \ 250 \ \ 277 \ \ 651 \ \ 698 \ \ 735 \ \ 745 \ \ 791 \ \ 906 \ \ 947$$

Determine the smoothed empirical estimate of the 35^{th} percentile.

Problem 58.12 ‡
The parameters of the inverse Pareto distribution

$$F(x) = \left(\frac{x}{x+\theta}\right)^{\tau}$$

are to be estimated using the method of moments based on the following data:

$$15 \ \ 45 \ \ 140 \ \ 250 \ \ 560 \ \ 1340$$

Estimate θ by matching k^{th} moments with $k = -1$ and $k = -2$.

Problem 58.13 ‡
You are given:
(i) Losses are uniformly distributed on $(0, \theta)$ with $\theta > 150$.

(ii) The policy limit is 150.
(iii) A sample of payments is:

14 33 72 94 120 135 150 150

Estimate θ by matching the average sample payment to the expected payment per loss.

Problem 58.14 ‡
You are given the following data:

0.49 0.51 0.66 1.82 3.71 5.20 7.62 12.66 35.24

You use the method of percentile matching at the 40^{th} and 80^{th} percentiles to fit an Inverse Weibull distribution to these data.

Determine the estimate of θ.

Problem 58.15 ‡
The following claim data were generated from a Pareto distribution:

130 20 350 218 1822

Using the method of moments to estimate the parameters of a Pareto distribution, calculate the limited expected value at 500.

Problem 58.16 ‡
A random sample of claims has been drawn from a Burr distribution with known parameter $\alpha = 1$ and unknown parameters θ and γ. You are given:
(i) 75% of the claim amounts in the sample exceed 100.
(ii) 25% of the claim amounts in the sample exceed 500.

Estimate θ by percentile matching.

Problem 58.17 ‡
A random sample of observations is taken from a shifted exponential distribution with probability density function:

$$f(x) = \frac{1}{\theta}e^{-\frac{(x-\delta)}{\theta}}, \ \delta < x < \infty.$$

The sample mean and median are 300 and 240, respectively.

Estimate δ by matching these two sample quantities to the corresponding population quantities.

Problem 58.18 ‡
For a portfolio of policies, you are given:
(i) Losses follow a Weibull distribution with parameters θ and τ.
(ii) A sample of 16 losses is :

54 70 75 81 84 88 97 105 109 114 122 125 128 139 146 153

(iii) The parameters are to be estimated by percentile matching using the 20^{th} and 70^{th} smoothed empirical percentiles.

Calculate the estimate of θ.

Problem 58.19 ‡
You are modeling a claim process as a mixture of two independent distributions A and B. You are given:
(i) Distribution A is exponential with mean 1.
(ii) Distribution B is exponential with mean 10.
(iii) Positive weight p is assigned to distribution A.
(iv) The standard deviation of the mixture is 2.

Determine p using the method of moments.

Problem 58.20 ‡
You are given the following information about a study of individual claims:
(i) 20^{th} percentile $= 18.25$
(ii) 80^{th} percentile $= 35.80$
Parameters μ and σ of a lognormal distribution are estimated using percentile matching.

Determine the probability that a claim is greater than 30 using the fitted lognormal distribution.

59 Maximum Likelihood Estimation for Complete Data

Maximum likelihood is a relatively simple method of constructing an estimator for an unknown parameter θ in a parametric model.

In order to define the maximum likelihood estimator, we start with the following notations: Let the data set be grouped into events $A_1, A_2, \cdots, A_n$ where A_j is whatever observed for the jth observation. The event A_j can be a single point (single individual data) or an interval (grouped or censored data). Furthermore, the event A_j results from observing the random variable X_j. We assume that the random variables $X_1, X_2, \cdots, X_n$ are independent and their distributions depend on the same parameter that need to be estimated. The method of maximum likelihood is to estimate the parameter that maximizes the probability or the likelihood of getting the data we observed. This is done by maximizing the **likelihood function**

$$L(\theta) = \prod_{i=1}^{n} \Pr(X_j \in A_j|\theta).$$

We use the symbol "|" to indicate that the distribution also depends on a parameter θ, where θ could be a real-valued unknown parameter or a vector of parameters.

In the case of a complete individual data, consider a random sample with observed values $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$. Then the likelihood function is

$$L(\theta) = f(x_1, x_2, \cdots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

where $f(x_1, x_2, \cdots, x_n|\theta)$ is a joint density if the the X_is are continuous or a joint mass function if the X_is are discrete.

A **maximum likelihood estimate**(MLE) is the value of θ that maximizes $L(\theta)$.

Example 59.1
Consider the following discrete random variable X whose pmf is given below.

X	0	1	2	3
$p(x\|\theta)$	$\frac{\theta}{3}$	$\frac{1-\theta}{3}$	$\frac{2\theta}{3}$	$\frac{2(1-\theta)}{3}$

where $0 \leq \theta \leq 1$. Consider the following random sample of 7 observations taken from this distribution $\{2, 3, 0, 1, 1, 2, 3\}$. Determine $L(\theta)$.

Solution.
We have

$$\begin{aligned} L(\theta) =& p(2|\theta)p(3|\theta)p(0|\theta)p(1|\theta)p(1|\theta)p(2|\theta)p(3|\theta) \\ =& \left(\frac{2\theta}{3}\right)^2 \left(\frac{2(1-\theta)}{3}\right)^2 \left(\frac{1-\theta}{3}\right)^2 \left(\frac{\theta}{3}\right) \blacksquare \end{aligned}$$

Clearly, the likelihood function $L(\theta)$ is not easy to maximize. But maximizing $L(\theta)$ is equivalent to maximizing $\ln[L(\theta)]$ since $\ln[L(\theta)]$ is an increasing function of θ. We define the **loglikelihood function** as

$$\ell(\theta) = \ln[L(\theta)] = \sum_{i=1}^{n} f(x_i|\theta).$$

Example 59.2
Find the maximum likelihood estimate of θ in the previous example.

Solution.
Let us look at the log likelihood function

$$\begin{aligned} \ell(\theta) =& \ln[L(\theta)] \\ =& 2\ln\left(\frac{2\theta}{3}\right) + 2\ln\left(\frac{2(1-\theta)}{3}\right) + 2\ln\left(\frac{1-\theta}{3}\right) + \ln\left(\frac{\theta}{3}\right). \end{aligned}$$

Using calculus, we have

$$\frac{d\ell}{d\theta} = 0 \Longrightarrow \theta = \frac{3}{7}.$$

Also,

$$\left.\frac{d^2\ell}{d\theta^2}\right|_{\theta=\frac{3}{7}} = -\frac{343}{49} < 0.$$

Hence, $\hat{\theta} = \frac{3}{7}$ $\blacksquare$

Example 59.3
A random sample of 5 claims obtained from an exponential distribution with parameter θ is given as follows:

$$15 \quad 10 \quad 7 \quad 8 \quad 20.$$

Find the maximum likelihood estimate of θ.

Solution.
The density function of the distribution is $f(x|\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$. The likelihood function is given by

$$L(\theta) = \frac{1}{\theta^5}e^{-\frac{1}{\theta}(x_1+x_2+\cdots+x_5)}.$$

The loglikelihood function is

$$\ell(\theta) = -\frac{1}{\theta}(x_1 + x_2 + \cdots + x_5) - 5\ln\theta.$$

Let the derivative with respect to θ be zero:

$$\ell'(\theta) = 0 \Longrightarrow \theta = \frac{x_1 + x_2 + \cdots + x_5}{5} = 12.$$

Moreover,

$$\left.\frac{d^2\ell}{d\theta^2}\right|_{\theta=12} = -5 < 0.$$

Thus, the MLE is

$$\hat{\theta} = \frac{x_1 + x_2 + \cdots + x_5}{5} = 12 \blacksquare$$

For a complete and grouped data, the process of finding the naximum likelihood estimate goes as follows: Arrange the unique observation values in increasing order

$$c_0 < c_1 < \cdots < c_k$$

where c_0 is the smallest possible observation (often zero) and c_k is the largest possible observation (often infinity). For $j = 1, 2, \cdots, k$, let n_j the number of observations in $A_j = (c_{j-1}, c_j]$. The likelihood contribution of each value in the jth observation is

$$\Pr(X \in A_j) = \Pr(c_{j-1} < x \le c_j) = F(c_j|\theta) - F(c_{j-1}|\theta).$$

Thus, the likelihood function is

$$L(\theta) = \prod_{j=1}^{k}[F(c_j|\theta) - F(c_{j-1}|\theta)]^{n_j}$$

and the loglikelihood function is

$$\ell(\theta) = \sum_{j=1}^{k} n_j \ln\left[F(c_j|\theta) - F(c_{j-1}|\theta)\right].$$

Example 59.4 ‡
Suppose that a group of 20 losses resulted in the following

Loss Range	# of observations
$(0, 10]$	9
$(10, 25]$	6
$(25, \infty)$	5

Losses follow the distribution $F(x) = 1 - \frac{\theta}{x}$, where $x > \theta$. Calculate the maximum likelihood estimate of θ.

Solution.
The likelihood function is

$$\begin{aligned}L(\theta) =&[F(10)]^9[F(25) - F(10)]^6[1 - F(25)]^5\\ =&(1 - 0.1\theta)^9(0.06\theta)^6(0.04\theta)^5.\end{aligned}$$

The loglikelihood function is

$$\ell(\theta) = 9\ln(1 - 0.1\theta) + 6\ln 0.06 + 6\ln\theta + 5\ln 0.04 + 5\ln\theta.$$

Taking its derivative and setting it to zero, we find

$$\ell'(\theta) = -\frac{0.9}{1 - 0.1\theta} + \frac{11}{\theta} = 0 \Longrightarrow \theta = 5.5.$$

Furthermore,

$$\left.\frac{d^2\ell}{d\theta^2}\right|_{\theta=5.5} = -0.8081 < 0.$$

Hence, $\hat{\theta} = 5.5$ ■

It is possible that a data set contains a combination of individual and grouped data as suggested in the following example.

Example 59.5 ‡
The random variable X has survival function:

$$S_X(x) = \frac{\theta^4}{(\theta^2 + x^2)^2}.$$

Two values of X are observed to be 2 and 4. One other value exceeds 4. Calculate the maximum likelihood estimate of θ.

Solution.
The likelihood function is

$$L(\theta) = f_X(2|\theta)f_X(4|\theta)S_X(4|\theta)$$

where

$$f_X(x) = \frac{4x\theta^4}{(\theta^2 + x^2)^3}.$$

The loglikelihood function is

$$\ell(\theta) = \ln f_X(2|\theta) + \ln f_X(4|\theta) + \ln S_X(2|\theta) = 7\ln 2 + 12\ln\theta - 3\ln(4+\theta^2)^2 - 5\ln(16+\theta^2)^2.$$

Taking the derivative to obtain

$$\ell'(\theta) = \frac{-4\theta^4 + 104\theta^2 + 768}{\theta(\theta^2+4)(\theta^2+16)}.$$

Thus,

$$\ell'(\theta) = 0 \Longrightarrow -4\theta^4 + 104\theta^2 + 768 = 0 \Longrightarrow \theta^2 = 32 \Longrightarrow \theta = 4\sqrt{2}.$$

We leave it to the reader to check that $\ell''(4\sqrt{2}) < 0$ so that $\hat{\theta} = 4\sqrt{2}$ ■

Example 59.6 ‡
You have observed the following three loss amounts:

186 91 66

Seven other amounts are known to be less than or equal to 60. Losses follow an inverse exponential with distribution function

$$F(x) = e^{-\frac{\theta}{x}},\ x > 0.$$

Calculate the maximum likelihood estimate of the population mode.

Solution.
The pdf of the inverse exponential distribution is given by $f(x) = \theta x^{-2}e^{-\frac{\theta}{x}}$.
The likelihood function is

$$\begin{aligned}
L(\theta) =& f(186)f(91)f(66)F(60)^7 \\
=& \theta 186^{-2}e^{-\frac{\theta}{186}}\theta 91^{-2}e^{-\frac{\theta}{91}}\theta 66^{-2}e^{-\frac{\theta}{66}}(e^{-\frac{\theta}{60}})^7 \\
=& (186\cdot 91\cdot 66)^{-2}\theta^3 e^{-\theta(186^{-1}+91^{-1}+66^{-1}+7(60^{-1})}.
\end{aligned}$$

The loglikelihood function is

$$\ln(\theta) = -2\ln(186 \cdot 91 \cdot 66) + 3\ln\theta - \theta(186^{-1} + 91^{-1} + 66^{-1} + 7(60^{-1}).$$

Taking the first derivative and setting it to 0, we find

$$\ell'(\theta) = \frac{3}{\theta} - (186^{-1} + 91^{-1} + 66^{-1} + 7(60^{-1}) = 0 \Longrightarrow \hat{\theta} = 20.25.$$

From Table C, the mode of the inverse exponential function is

$$\frac{\hat{\theta}}{2} = \frac{20.25}{2} = 10.125 \blacksquare$$

Example 59.7 ‡
You are given:
(i) The distribution of the number of claims per policy during a one-year period for 10,000 insurance policies is:

Number of Claims per Policy	Number of Policies
0	5000
1	5000
2 or more	0

(ii) You fit a binomial model with parameters m and q using the method of maximum likelihood.
Determine the maximum value of the loglikelihood function when $m = 2$.

Solution.
The pdf of the binomial distribution with $m = 2$ is

$$f(x) = \binom{2}{x} q^x(1-q)^{2-x}.$$

The likelihood function is

$$\begin{aligned} L(q) =& f(0)^{5000}f(1)^{5000} = \left[\binom{2}{0}(1-q)^2\right]^{5000}\left[\binom{2}{1}q(1-q)\right]^{5000} \\ =& 2^{5000}q^{5000}(1-q)^{15000}. \end{aligned}$$

The loglikelihood function is

$$\ell(q) = 5000\ln 2 + 5000\ln q + 15000\ln(1-q).$$

Taking the derivative and setting it to zero, we find

$$\ell'(q) = \frac{5000}{q} - \frac{15000}{1-q} = 0 \Longrightarrow \hat{q} = 0.25.$$

Finally,

$$\ell(0.25) = 5000 \ln 2 + 5000 \ln 0.25 + 15000 \ln 0.75 = -7780.97 \blacksquare$$

Practice Problems

Problem 59.1
Suppose $X_1, X_2, \cdots, X_n$ are i.i.d random variables with a Gamma distribution with $\alpha = 2$ and θ.

Find the maximum likelihood estimate of θ.

Problem 59.2
Suppose $X_1, X_2, \cdots, X_n$ are i.i.d random variables with a uniform distribution is $(0, \theta)$.

Find the maximum likelihood estimate of θ.

Problem 59.3 ‡
You are given the following three observations:

$$0.74 \quad 0.81 \quad 0.95$$

You fit a distribution with the following density function to the data:

$$f(x) = (p+1)x^p, \ 0 < x < 1, \ p > -1.$$

Determine the maximum likelihood estimate of p.

Problem 59.4 ‡
The proportion of allotted time a student takes to complete an exam, X, is described by the following distribution:

$$f(x) = (\theta+1)x^\theta, 0 \le x \le 1, \ \theta > -1.$$

A random sample of five students produced the following observations:

Student	Proportion of Allotted Time
1	0.92
2	0.79
3	0.90
4	0.65
5	0.86

Using the sample data, calculate the maximum likelihood estimate of θ.

Problem 59.5 ‡
Let $X_1, X_2, \cdots, X_n$ be a random sample from the following distribution with pdf

$$f(x) = \begin{cases} e^{-x+\theta}, & \theta < x, \ -\infty < \theta < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Find the maximum likelihood estimate of θ.

Problem 59.6 ‡
You are given:
(i) Losses follow a Single-parameter Pareto distribution with density function:

$$f(x) = \frac{\alpha}{x^{\alpha+1}}, \ x > 1, \ 0 < \alpha < \infty.$$

(ii) A random sample of size five produced three losses with values 3, 6 and 14, and two losses exceeding 25.

Determine the maximum likelihood estimate of α.

Problem 59.7 ‡
You are given:
(i) Low-hazard risks have an exponential claim size distribution with mean θ.
(ii) Medium-hazard risks have an exponential claim size distribution with mean 2θ.
(iii) High-hazard risks have an exponential claim size distribution with mean 3θ.
(iv) No claims from low-hazard risks are observed.
(v) Three claims from medium-hazard risks are observed, of sizes 1, 2 and 3.
(vi) One claim from a high-hazard risk is observed, of size 15.

Determine the maximum likelihood estimate of θ.

Problem 59.8 ‡
A random sample of three claims from a dental insurance plan is given below:

225 525 950

Claims are assumed to follow a Pareto distribution with parameters $\theta = 150$ and α.

Determine the maximum likelihood estimate of α.

Problem 59.9 ‡
You are given:
(i) Losses follow an exponential distribution with mean θ.
(ii) A random sample of 20 losses is distributed as follows:

Loss Range	# of observations
$(0, 1000]$	7
$(1000, 2000]$	6
$(2000, \infty)$	7

Calculate the maximum likelihood estimate of θ.

Problem 59.10 ‡
You fit an exponential distribution to the following data:

$$1000 \quad 1400 \quad 5300 \quad 7400 \quad 7600$$

Determine the coefficient of variation of the maximum likelihood estimate of the mean, θ.

Problem 59.11 ‡
Losses come from a mixture of an exponential distribution with mean 100 with probability p and an exponential distribution with mean 10,000 with probability $1-p$. Losses of 100 and 2000 are observed.

Determine the likelihood function of p.

Problem 59.12 ‡
Let $x_1, x_2, \cdots, x_n$ and $y_1, y_2, \cdots, y_m$ denote independent random samples of losses from Region 1 and Region 2, respectively. Single-parameter Pareto distributions with $\theta = 1$, but different values of α, are used to model losses in these regions.
Past experience indicates that the expected value of losses in Region 2 is 1.5 times the expected value of losses in Region 1. You intend to calculate the maximum likelihood estimate of α_1 for Region 1, using the data from both regions.

Find $\frac{d}{d\alpha_1}\ell(\alpha_1)$.

Problem 59.13 ‡
You have observed the following claim severities:

11.0 15.2 18.0 21.0 25.8

You fit the following probability density function to the data:

$$f(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{(x-\mu)^2}{2x}}, \ x > 0, \ \mu > 0.$$

Determine the maximum likelihood estimate of μ.

Problem 59.14 ‡
Phil and Sylvia are competitors in the light bulb business. Sylvia advertises that her light bulbs burn twice as long as Phil's. You were able to test 20 of Phil's bulbs and 10 of Sylvia's. You assumed that the distribution of the lifetime (in hours) of a light bulb is exponential, and separately estimated Phil's parameter as $\theta_P = 1000$ and Sylvia's parameter as q$\theta_S = 1500$ using maximum likelihood estimation.

Determine θ^* , the maximum likelihood estimate of θ_P restricted by Sylvia's claim that $\theta_S = 2\theta_P$.

60 Maximum Likelihood Estimation for Incomplete Data

If some of the data in the sample has been right-censored, then each observation in the sample censored at u contributes a factor of $S(u|\theta) = 1 - F(u|\theta)$ to the likelihood function. This is because, $\Pr(X_j \in A_j) = \Pr(X_j > u) = 1 - F(u|\theta)$. Recall that when a ground-up loss random variable has a policy limit of u, every loss less than or equal to u will be recorded with its observed value while all losses above u will be recorded as u.

Example 60.1
A ground-up loss random variable X has a policy limit of 30. The following is a random sample of 6 insurance payment amounts:

$$20 \quad 25 \quad 27 \quad 28 \quad 30 \quad 30.$$

If X is assumed to have an exponential distribution, apply maximum likelihood estimation to estimate the mean of X.

Solution.
The likelihood function is

$$\begin{aligned}
L(\theta) =& f_X(20|\theta)f_X(25|\theta)f_X(27|\theta)f_X(28|\theta)[S_X(30|\theta)]^2\\
=&\frac{1}{\theta^4}e^{-\frac{1}{\theta}(20+25+27+28)}e^{-\frac{30\times 2}{\theta}}\\
=&\frac{1}{\theta^4}e^{-\frac{1}{\theta}(20+25+27+28+2\times 30)}\\
=&\frac{1}{\theta^4}e^{-\frac{160}{\theta}}.
\end{aligned}$$

The loglikelihood function is

$$\ell(\theta) = -\frac{160}{\theta} - 4\ln\theta$$

and

$$\ell'(\theta) = \frac{160}{\theta^2} - \frac{4}{\theta} = 0 \Longrightarrow \theta = \frac{160}{4} = 40.$$

Furthermore,

$$\ell''(\theta)(40) = -\frac{320}{\theta^3} + \frac{4}{\theta^2}\Big|_{\theta=40} = -0.0025 < 0.$$

Finally, $\hat{\theta} = 40$ ■

If some of the data in the sample has been left-truncated, for example a policy with an ordinary deductible d, then each observation in the sample truncated at d contributes a factor of $f(y+d|\theta)[1-F(d|\theta)]$ where y is recorded after the deductible d is applied. For example, suppose the jth observation is y_j, the loss amount after a deductible of d is applied. Then

$$\Pr(X_j \in A_j) = \Pr(X_j|X_j > d) = \frac{f(x_j|\theta)}{1-F(d|\theta)} = \frac{f(y_j+d|\theta)}{1-F(d|\theta)},\ x_j = y_j + d.$$

Example 60.2
A ground up loss X has a deductible of 7 applied. A random sample of 6 insurance payments (after deductible is applied) is given

$$3\ \ 6\ \ 7\ \ 8\ \ 10\ 12.$$

If X is assumed to have an exponential distribution, apply maximum likelihood estimation to estimate the mean of X.

Solution.
The loss amounts before the deductible is applied are:

$$10\ \ 13\ \ 14\ \ 15\ \ 17\ 19.$$

The likelihood function is

$$\begin{aligned} L(\theta) =& \frac{f(10|\theta)f(13|\theta)f(14|\theta)f(15|\theta)f(17|\theta)f(19|\theta)}{[1-F(7|\theta)]^6} \\ =& \frac{1}{\theta^6}e^{-\frac{1}{\theta}(10+13+14+15+17+19-6\times 7)} \\ =& \frac{1}{\theta^6}e^{-\frac{46}{\theta}}. \end{aligned}$$

The loglikelihood function is

$$\ell(\theta) = -\frac{46}{\theta} - 6\ln\theta$$

and

$$\ell'(\theta) = \frac{46}{\theta^2} - \frac{6}{\theta} = 0 \Longrightarrow \theta = \frac{46}{6} = 7.6667.$$

Furthermore,

$$\ell''(\theta)(7.667) = -\frac{92}{\theta^3} + \frac{6}{\theta^2}\bigg|_{\theta=7.6667} = -0.1021 < 0.$$

Finally, $\hat{\theta} = 7.6667$ ■

Instead of using the ground up losses in the above estimation process, one can use instead the payments after the deductible is applied (i.e. cost per payment). In this case, the problem reduces to the case of a complete individual data. We illustrate this point in the next example.

Example 60.3
A ground up loss X has a deductible of 7 applied. A random sample of 6 insurance payments (after deductible is applied) is given

$$3 \quad 6 \quad 7 \quad 8 \quad 10 \; 12.$$

(a) If X is assumed to have a uniform distribution in $(0, \theta)$, apply maximum likelihood estimation to estimate the mean of X.
(b) Assuming the cost per payment Y has a uniform distribution in $(0, \theta)$, estimate the mean of Y.

Solution.
(a) Note that the condition $x_i < \theta$ is equivalent to $\max\{x_1, x_2, \cdots, x_n\} < \theta$. Hence, an estimate of θ is $\hat{\theta} = \max\{10, 13, 14, 15, 17, 19\} = 19$ so that an estimate of the mean of X is $\frac{19}{2} = 9.5$.
(b) Similar to (a), an estimate of θ is $\hat{\theta} = \max\{3, 6, 7, 8, 10, 12\} = 12$ so that an estimate of the mean is $\frac{12}{2} = 6$ ■

Now, if a policy has an ordinary deductible and maximum loss cover u then the policy limit is $u - d$. If the cost per payment y is less than $u - d$ then the ground up loss amount is $x = y + d$ and this contributes a factor $f(x|\theta)[1 - F(d|\theta)]^{-1}$ to the likelihood function. If the cost per payment is $y = u - d$, then the ground up loss amount is at least u and this contributes a factor $[1 - F(u|\theta)][1 - F(d|\theta)]^{-1}$ to the likelihood function.

Example 60.4
A policy has a deductible of $d = 3$ and a maximum covered loss of $u = 14$. A random sample of 6 insurance payments is given:

$$1 \quad 3 \quad 7 \quad 9 \quad 11 \quad 11.$$

Suppose that the ground up loss distribution is exponential with parameter θ. Find the maximum likelihood estimate of θ.

Solution.
The policy limit is $u - d = 14 - 3 = 11$. The likelihood function is

$$\begin{aligned}L(\theta) =&\frac{f(4|\theta)f(6|\theta)f(10|\theta)f(12|\theta)}{[1-F(3|\theta)]^4}\left[\frac{1-F(14|\theta)}{1-F(3|\theta)}\right]^2\\ =&\frac{1}{\theta^4}e^{-\frac{1}{\theta}(4+6+10+12+2\times14-6\times3)}\\ =&\frac{1}{\theta^4}e^{-\frac{42}{\theta}}.\end{aligned}$$

The loglikelihood function is

$$\ell(\theta) = -\frac{42}{\theta} - 4\ln\theta$$

and

$$\ell'(\theta) = \frac{42}{\theta^2} - \frac{4}{\theta} = 0 \Longrightarrow \theta = 10.5.$$

Moreover,

$$\ell''(10.5) = -\frac{84}{\theta^3} + \frac{4}{\theta^2}\Big|_{\theta=10.5} = -0.0363.$$

Hence, $\hat{\theta} = 10.5$ ■

Example 60.5 ‡
You are given:
(i) The number of claims follows a Poisson distribution with mean λ.
(ii) Observations other than 0 and 1 have been deleted from the data.
(iii) The data contain an equal number of observations of 0 and 1.
Determine the maximum likelihood estimate of λ.

Solution.
Let N be the number of claims. Notice that we are trying to estimate λ so the data in our sample represent number of claims. We are told that N is right truncated at 1. Thus, we have the following

$$\begin{aligned}\Pr[N=0|N\le 1] =&\frac{\Pr(N=0)}{F_N(1)} = \frac{\Pr(N=0)}{\Pr(N=0)+\Pr(N=1)} = \frac{e^{-\lambda}}{e^{-\lambda}+\lambda e^{-\lambda}} = \frac{1}{1+\lambda}\\ \Pr[N=1|N\le 1] =&\frac{\Pr(N=1)}{\Pr(N=0)+\Pr(N=1)} = \frac{\lambda e^{-\lambda}}{e^{-\lambda}+\lambda e^{-\lambda}} = \frac{\lambda}{1+\lambda}.\end{aligned}$$

By (ii), the data contain an equal number of observations of 0 and 1. Let's say that our sample has n data points of 0's and 1's. Then the likelihood function is

$$L(\lambda) = \left(\frac{1}{1+\lambda}\right)^{n/2}\left(\frac{\lambda}{1+\lambda}\right)^{n/2} = \left(\frac{\lambda}{(1+\lambda)^2}\right)^{n/2}.$$

The loglikelihood function is

$$\ell(\lambda) = \frac{n}{2}\ln\lambda - n\ln(1+\lambda).$$

Differentiating and setting to 0, we find

$$\ell'(\lambda) = \frac{n}{2\lambda} - \frac{n}{1+\lambda} = 0 \Longrightarrow \hat{\lambda} = 1 \blacksquare$$

Example 60.6 ‡
You are given:
(i) At time 4 hours, there are 5 working light bulbs.
(ii) The 5 bulbs are observed for p more hours.
(iii) Three light bulbs burn out at times 5, 9, and 13 hours, while the remaining light bulbs are still working at time $4+p$ hours.
(iv) The distribution of failure times is uniform on $(0,\omega)$.
(v) The maximum likelihood estimate of ω is 29.
Determine p.

Solution.
Let T be the time to failure random variable. T has a uniform distribution on $(0,\omega)$. Its pdf is $f(t) = \frac{1}{\omega}$ and its sdf is $S(t) = 1 - \frac{t}{\omega}$. The likelihood function is

$$\begin{aligned} L(\omega) =& \frac{f(5)}{S(4)}\frac{f(9)}{S(4)}\frac{f(13)}{S(4)}\left[\frac{S(4+p)}{S(4)}\right]^2 \\ =& \frac{\frac{1}{\omega}\frac{1}{\omega}\frac{1}{\omega}\left(1-\frac{4+p}{\omega}\right)^2}{\left(1-\frac{4}{\omega}\right)^5} \\ =& \frac{(\omega-4-p)^2}{(\omega-4)^5}. \end{aligned}$$

The loglikelihood function is

$$\ell(\omega) = 2\ln(\omega-4-p) - 5\ln(\omega-4).$$

Taking the derivative and setting it to zero, we find

$$\ell'(\omega) = \frac{2}{\omega-4-p} - \frac{5}{\omega-4} = 0 \Longrightarrow \frac{2}{\hat{\omega}-4-p} = \frac{5}{\hat{\omega}-4}.$$

But $\hat{\omega} = 29$. Hence,

$$\frac{2}{29-4-p} = \frac{5}{29-4} \Longrightarrow p = 15 \blacksquare$$

Practice Problems

Problem 60.1 ‡
You observe the following five ground-up claims from a data set that is truncated from below at 100:

$$125 \quad 150 \quad 165 \quad 175 \quad 250$$

You fit a ground-up exponential distribution using maximum likelihood estimation. Determine the mean of the fitted distribution.

Problem 60.2 ‡
You are given:
(i) A sample of losses is: 600 700 900
(ii) No information is available about losses of 500 or less.
(iii) Losses are assumed to follow an exponential distribution with mean θ.

Determine the maximum likelihood estimate of θ.

Problem 60.3 ‡
A policy has an ordinary deductible of 100 and a policy limit of 1000. You observe the following 10 payments:

$$15 \quad 50 \quad 170 \quad 216 \quad 400 \quad 620 \quad 750 \quad 900 \quad 900 \quad 900.$$

An exponential distribution is fitted to the ground up distribution function, using the maximum likelihood estimate.

Determine the estimated parameter $\hat{\theta}$.

Problem 60.4
A ground-up loss random variable X has a policy limit of 2000. The following is a random sample of 3 insurance payment amounts:

$$300 \quad 1000 \quad 2000.$$

If X is assumed to have a uniform distribution in $(0, \theta)$, apply maximum likelihood estimation to estimate θ.

Problem 60.5
A ground-up loss random variable X has a policy limit of 1000. A random sample of two payments is obtained as follows: 200 and 1000. The loss X has a Pareto distribution with $\theta = 2500$.

Calculate the maximum likelihood estimate for α.

Problem 60.6 ‡
For a dental policy, you are given:
(i) Ground-up losses follow an exponential distribution with mean θ.
(ii) Losses under 50 are not reported to the insurer.
(iii) For each loss over 50, there is a deductible of 50 and a policy limit of 350.
(iv) A random sample of five claim payments for this policy is:

$$50 \quad 150 \quad 200 \quad 350^+ \quad 350^+$$

where + indicates that the original loss exceeds 400.

Determine the likelihood function

Problem 60.7 ‡
You are given:
(i) An insurance company records the following ground-up loss amounts, which are generated by a policy with a deductible of 100:

$$120 \quad 180 \quad 200 \quad 270 \quad 300 \quad 1000 \quad 2500$$

(ii) Losses less than 100 are not reported to the company.
(iii) Losses are modeled using a Pareto distribution with parameters $\theta = 400$ and α.

Use the maximum likelihood estimate of α to estimate the expected loss with no deductible.

Problem 60.8 ‡
You are given the following information about a group of policies:

Claim Payment	Policy Limit
5	50
15	50
60	100
100	100
500	500
500	1000

Determine the likelihood function.

Problem 60.9 ‡
You are given the following 20 bodily injury losses (before the deductible is applied):

Loss	Number of Losses	Deductible	Policy Limit
750	3	200	∞
200	3	0	10,000
300	4	0	20,000
>10,000	6	0	10,000
400	4	300	∞

Past experience indicates that these losses follow a Pareto distribution with parameters α and $\theta = 10,000$.

Determine the maximum likelihood estimate of α.

Problem 60.10 ‡
Personal auto property damage claims in a certain region are known to follow the Weibull distribution:

$$F(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^{0.2}},\ x > 0.$$

A sample of four claims is:

130 240 300 540

The values of two additional claims are known to exceed 1000.

Determine the maximum likelihood estimate of θ.

Problem 60.11 ‡
For a portfolio of policies, you are given:
(i) There is no deductible and the policy limit varies by policy.
(ii) A sample of ten claims is:

$$350 \quad 350 \quad 500 \quad 500 \quad 500^+ \quad 1000 \quad 1000^+ \quad 1000^+ \quad 1200 \quad 1500$$

where the symbol + indicates that the loss exceeds the policy limit.
(iii) $\hat{S}_1(1250)$ is the product-limit estimate of $S(1250)$.
(iv) $\hat{S}_2(1250)$ is the maximum likelihood estimate of $S(1250)$ under the assumption that the losses follow an exponential distribution.

Determine the absolute difference between $\hat{S}_1(1250)$ and $\hat{S}_2(1250)$.

Problem 60.12 ‡
You are given a sample of losses from an exponential distribution. However, if a loss is 1000 or greater, it is reported as 1000. The summarized sample is:

Reported Loss	Number	Total Amount
Less than 1000	62	28,140
1000	38	38,000
Total	100	66,140

Determine the maximum likelihood estimate of θ, the mean of the exponential distribution.

Problem 60.13 ‡
You are given the following claims settlement activity for a book of automobile claims as of the end of 1999:

Number of Claims Settled			
Year	Year Settled		
Reported	1997	1998	1999
1997	Unknown	3	1
1998		5	2
1999			4

$L =$ (Year Settled Year Reported) is a random variable describing the time lag in settling a claim. The probability function of L is

$$f_L(\ell) = p(1-p)^\ell,\ \ell = 0, 1, 2, \cdots .$$

Determine the maximum likelihood estimate of the parameter p.

Problem 60.14 ‡
You are given the following information about a random sample:
(i) The sample size equals five.
(ii) The sample is from a Weibull distribution with $\tau = 2$ and unknown θ.
(iii) Two of the sample observations are known to exceed 50, and the remaining three observations are 20, 30 and 45.
Calculate the maximum likelihood estimate of θ.

61 Asymptotic Variance of MLE

Estimators can always be regarded as random variables themselves. In particular, the maximum likelihood estimator is a random variable. Calculating the variance of mle is not an easy task. Instead, methods of approximating the variance exist. The goal of this section is to discuss one such a method.

Our method uses the following concept found in the statistics literature: Let X be a random variable whose distribution depends on a parameter θ and denote its pdf by $f(x|\theta)$. We define the **Fisher information** $I(X|\theta)$ in X by

$$I(X|\theta) = E[\ell'(x|\theta)^2] = \int [\ell'(x|\theta)]^2 f(x|\theta)dx$$

where $\ell(x|\theta) = \ln f(x|\theta)$. The prime symbol stands for the differentiation with respect to θ. We assume that we can exchange the order of differentiation and integration, then

$$\int f'(x|\theta)dx = \frac{d}{d\theta}\int f(x|\theta)dx = \frac{d}{d\theta}(1) = 0.$$

Similarly,

$$\int f''(x|\theta)dx = \frac{d^2}{d\theta^2}\int f(x|\theta)dx = \frac{d^2}{d\theta^2}(1) = 0.$$

With these properties of f, we have

$$E[\ell'(x|\theta)] = \int \ell'(x|\theta)f(x|\theta)dx = \int \frac{f'(x|\theta)}{f(x|\theta)}f(x|\theta)dx = \int f'(x|\theta)dx = 0.$$

Hence,

$$I(X|\theta) = \text{Var}[\ell'(x|\theta)].$$

Also, notice that

$$\ell''(x|\theta) = \frac{f''(x|\theta)f(x|\theta) - f'(x|\theta)^2}{f(x|\theta)^2} = \frac{f''(x|\theta)}{f(x|\theta)} - [\ell'(\theta)]^2$$

so that

$$E[\ell''(x|\theta)] = \int f''(x|\theta)dx - E[\ell'(x|\theta)^2] = -I(x|\theta).$$

From this, we obtain another way for calculating $I(\theta)$:

$$I(X|\theta) = -\int \left[\frac{\partial^2}{\partial\theta^2}\ln f(x|\theta)\right] f(x|\theta)dx.$$

Example 61.1
Let X be a normal random variable with parameters μ and σ^2. Suppose that σ^2 is knwon but μ is the unknown parameter. Find the Fisher information $I(X|\mu)$ in X.

Solution.
The pdf of X is

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}.$$

Thus,

$$\ell(x|\mu) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}.$$

Hence

$$\ell'(x|\mu) = \tfrac{(x-\mu)}{\sigma^2} \text{ and } \ell''(x|\mu) = -\tfrac{1}{\sigma^2}.$$

Thus,

$$I(X|\mu) = -E[\ell''(x|\mu)] = \frac{1}{\sigma^2} \blacksquare$$

Now, suppose we have a random sample $X_1, X_2, \cdots, X_n$ coming from a distribution for which the pdf is $f(x|\theta)$ where the value of the parameter θ is unknown. Assuming independence, the joint pdf is given by

$$L(\theta) = f_n(x_1, x_2, \cdots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta).$$

The loglikelihood function is

$$\ell(\theta) = \sum_{i=1}^{n} \ell(x_i|\theta) = \sum_{i=1}^{n} lnf(x_i|\theta)$$

and

$$\ell'(\theta) = \frac{f_n'(x_1, x_2, \cdots, x_n|\theta)}{f_n(x_1, x_2, \cdots, x_n|\theta)}.$$

We define the Fisher information $I(\theta)$ in the random sample $X_1, X_2, \cdots, X_n$ as

$$I(\theta) = E[\ell'(\theta)^2] = \int \cdots \int \ell_n'(\theta)^2 f_n(x_1, x_2, \cdots, x_n|\theta) dx_1 dx_2 \cdots dx_n.$$

We assume that we can exchange the order of differentiation and integration, then

$$\int f'(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n = \frac{d}{d\theta}\int f(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n = \frac{d}{d\theta}(1) = 0.$$

Similarly,

$$\int f''(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n = \frac{d^2}{d\theta^2}\int f(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n = \frac{d^2}{d\theta^2}(1) = 0.$$

With these properties of f, we have

$$\begin{aligned}E[\ell'(\theta)] =& \int \ell_n'(\theta)f(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n\\ =& \int \frac{f'(x_1, x_2, \cdots, x_n|\theta)}{f(x_1, x_2, \cdots, x_n|\theta)}f(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n\\ =& \int f'(x_1, x_2, \cdots, x_n|\theta)dx_1dx_2\cdots dx_n = 0.\end{aligned}$$

Hence,

$$I(\theta) = \text{Var}[\ell'(\theta)].$$

Also, notice that

$$E[\ell''(\theta)] = -I(\theta).$$

and

$$I(\theta) = -E[\sum_{i=1}^n \ell''(x_i|\theta)] = -\sum_{i=1}^n E[\ell''(x_i|\theta)] = nI(X|\theta).$$

In other words, the Fisher information in a random sample of size n is simply n times the Fisher information in a single observation.

Example 61.2
Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma^2)$ where σ^2 is known but μ is unknown. Find the Fisher information of this random sample.

Solution.
We have

$$I(\mu) = nI(X|\mu) = \frac{n}{\sigma^2} \blacksquare$$

Now, let $\hat{\theta}$ denote an arbitrary estimator of θ. It is proven in statistics that

$$\text{Var}(\hat{\theta}) \geq \frac{m'(\theta)}{I(\theta)}$$

where $E(\hat{\theta}) = m(\theta)$. If $\hat{\theta}$ is unbiased then $m(\theta) = \theta$ and the above inequality becomes

$$\operatorname{Var}(\hat{\theta}) \geq \frac{1}{I(\theta)}.$$

The right-hand side of the inequality is known as the **Cramér-Rao lower bound**. It is shown in Statistics that under certain conditions, no other unbiased estimator of the parameter θ based on an i.i.d. sample of size n can have a variance smaller than the Cramér-Rao lower bound.

Example 61.3
Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \theta)$ where μ is known and θ is unknown. Calculate the Cramér-Rao lower bound of variance for any unbiased estimator,

Solution.
The pdf of $N(\mu, \theta)$ is given by

$$f(x|\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}.$$

Thus,

$$\ell(x|\theta) = -\frac{(x-\mu)^2}{2\theta} - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\theta$$

and

$$\ell'(x|\theta) = \tfrac{(x-\mu)^2}{2\theta^2} - \tfrac{1}{2\theta} \text{ and } \ell''(x|\theta) = -\tfrac{(x-\mu)^2}{\theta^3} + \tfrac{1}{2\theta^2}.$$

Hence,

$$I(X|\theta) = -E_\theta[\ell''(\theta)] = \frac{1}{\theta^3}E[(X-\mu)^2] - \frac{1}{2\theta^2} = \frac{1}{2\theta^2}$$

and

$$I(\theta) = nI(X|\theta) = \frac{n}{2\theta^2}.$$

The Cramér-Rao lower bound is $\frac{2\theta^2}{n}$ ■

The following result whose proof is omitted is a direct application of the Central Limit Theorem.

Theorem 61.1
Let $X_1, X_2, \cdots, X_n$ be a random sample fron a distribution with pdf $f(x|\theta)$ and unknown parameter θ. Let θ denote the true value of the parameter and

$\hat{\theta}$ the MLE estimator of θ. Then the probability distribution of $\sqrt{I(\theta)}(\hat{\theta}-\theta)$ approaches the standard normal distribution as $n \to \infty$. That is,

$$\hat{\theta} \sim N\left(\theta, \frac{1}{I(\theta)}\right).$$

Since this is merely a limiting result, which holds as the sample size tends to infinity, we say that the MLE is **asymptotically unbiased** and refer to the variance of the limiting normal distribution as the **asymptotic variance of the MLE**.

Example 61.4
Let $X_1, X_2, \cdots, X_n$ be a random sample from an exponential distribution with mean θ. Find the asymptotic distribution of $\hat{\theta}$.

Solution.
We have

$$\begin{aligned}
\ell(x|\theta) =& -\frac{x}{\theta} - \ln\theta \\
\ell'(x|\theta) =& \frac{x}{\theta^2} - \frac{1}{\theta} \\
\ell''(x|\theta) =& -\frac{2x}{\theta^3} + \frac{1}{\theta^2} \\
I(x|\theta) =& -E[\ell''(x|\theta)] = E[\frac{2X}{\theta^3} - \frac{1}{\theta^2} \\
=& \frac{2}{\theta^3}E(X) - \frac{1}{\theta^2} = \frac{1}{\theta^2} \\
I(\theta) =& \frac{n}{\theta^2}.
\end{aligned}$$

Hence,

$$\hat{\theta} \sim N\left(\theta, \frac{\theta^2}{n}\right) \blacksquare$$

It follows that

$$\text{Var}(\hat{\theta}) = \frac{\theta^2}{n}$$

and since we don't know the exact value θ, we have

$$\widehat{\text{Var}}(\hat{\theta}) = \frac{\hat{\theta}^2}{n}.$$

Example 61.5
For an exponential distribution, the parameter θ is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$7 \quad 12 \quad 15 \quad 19.$$

Find the estimated variance of the parameter estimate.

Solution.
From Example 59.3, we know that $\hat{\theta} = \overline{X}$. Hence,

$$\widehat{\mathrm{Var}}(\hat{\theta}) = \frac{\hat{\theta}^2}{n} = \frac{\overline{X}^2}{n} = \frac{[(7+12+15+19)/4]^2}{4} = 43.890625 \blacksquare$$

Since $\hat{\theta}$ is approximately normal for large sampling distribution, we can find confidence intervals for the true value θ. A $100(1-\alpha)\%$ confidence interval for θ is given by

$$\left[\hat{\theta} - z_{\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{\theta})}, \hat{\theta} + z_{\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{\theta})}\right].$$

Example 61.6
For an exponential distribution, the parameter θ is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$7 \quad 12 \quad 15 \quad 19.$$

Find the 95% confidence interval for θ.

Solution.
We have $\hat{\theta} = \overline{X} = \frac{7+12+15+19}{4} = 13.25$ and $\widehat{\mathrm{Var}}(\hat{\theta}) = 43.890625$. Hence, the 95% confidence interval is

$$[13.25 - 1.96\sqrt{43.890625}, 13.25 + 1.96\sqrt{43.890625}] = [0.265, 26.2350] \blacksquare$$

Practice Problems

Problem 61.1
Let X have a Pareto distribution with parameter α and $\theta = 20$. The parameter α is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$12 \quad 15 \quad 17 \quad 19.$$

Find $\ell'(\alpha)^2$, the square of the first partial derivative of the loglikelihood function.

Problem 61.2
Let X have a Pareto distribution with parameter α and $\theta = 20$. The parameter α is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$12 \quad 15 \quad 17 \quad 19.$$

Find $\ell''(\alpha)$, the second partial derivative of the loglikelihood function.

Problem 61.3
Let X have a Pareto distribution with parameter α and $\theta = 20$. The parameter α is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$12 \quad 15 \quad 17 \quad 19.$$

Find the Fisher information associated with the maximum likelihood estimator.

Problem 61.4
Let X have a Pareto distribution with parameter α and $\theta = 20$. The parameter α is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$12 \quad 15 \quad 17 \quad 19.$$

Estimate the asymptotic variance of the maximum likelihood estimator.

Problem 61.5
Let X have a Pareto distribution with parameter α and $\theta = 20$. The parameter α is estimated via the method of maximum likelihood by analyzing data from the following sample:

$$12 \quad 15 \quad 17 \quad 19.$$

Find the 95% confidence interval for α.

Problem 61.6 ‡
The information associated with the maximum likelihood estimator of a parameter θ is $4n$, where n is the number of observations. Calculate the asymptotic variance of the maximum likelihood estimator of 2θ.

Problem 61.7 ‡
You fit an exponential distribution to the following data:

$$1000 \quad 1400 \quad 5300 \quad 7400 \quad 7600$$

Determine the coefficient of variation of the maximum likelihood estimate of the mean, θ.

Problem 61.8 ‡
A random sample of size n is drawn from a distribution with probability density function:

$$f(x) = \frac{\theta}{(\theta+x)^2} < \ 0 < x < \infty, \ 0 < \theta < \infty.$$

Determine the asymptotic variance of the maximum likelihood estimator of θ.

Problem 61.9 ‡
You are given:
(i) The distribution of the number of claims per policy during a one-year period for a block of 3000 insurance policies:

# of claims per policy	# of policies
0	1000
1	1200
2	600
3	200
4^+	0

(ii) You fit Poisson model to the number of claims per policy using the method maximum likelihood.
(iii) You construct the large-sample 90% confidence interval for the mean of the underlying Poisson model that is symmetric around the mean.

Determine the lower end-point of the confidence interval.

62 Information Matrix and the Delta Method

In this section, we consider the Fisher information for a multiple parameter vector $\theta = (\theta_1, \theta_2, \cdots, \theta_n)^T$. Each θ_i has an asymptotic normal distribution. In this case, the Fisher information is a $n \times n$ square matrix $I(\theta)$ whose ijth entry is given by

$$I(\theta)_{ij} = E\left[\frac{\partial}{\partial\theta_i}\ell(\theta)\frac{\partial}{\partial\theta_j}\ell(\theta)\right] = -E\left[\frac{\partial^2}{\partial\theta_j\partial\theta_i}\ell(\theta)\right].$$

We refer to $I(\theta)$ as the **Fisher information matrix**.

Example 62.1
Find the Fisher information matrix of the MLE for the lognormal distribution.

Solution.
Recall the pdf of the lognormal distribution

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}.$$

Thus, the likelihood function is

$$L(\mu,\sigma) = \prod_{i=1}^n \frac{1}{x_i\sigma\sqrt{2\pi}}e^{-\frac{(\ln x_i-\mu)^2}{2\sigma^2}}$$

and its loglikelihood function is

$$\ell(\mu,\sigma) = \sum_{i=1}^n \left[-\ln x_i - \ln\sigma - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2\right].$$

Taking first derivatives, we find

$$\frac{\partial\ell}{\partial\mu} = \sum_{i=1}^n \frac{\ln x_i - \mu}{\sigma^2} \text{ and } \frac{\partial\ell}{\partial\sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(\ln x_i-\mu)^2}{\sigma^3}.$$

Taking second derivatives, we find

$$\begin{aligned}
\frac{\partial^2\ell}{\partial\mu^2} &= -\frac{n}{\sigma^2}\\
\frac{\partial^2\ell}{\partial\sigma\partial\mu} &= -2\sum_{i=1}^n \frac{\ln x_i-\mu}{\sigma^3}\\
\frac{\partial^2\ell}{\partial\sigma^2} &= \frac{n}{\sigma^2} - 3\sum_{i=1}^n \frac{(\ln x_i-\mu)^2}{\sigma^4}.
\end{aligned}$$

Taking expectations, and recalling that $\ln X_i \sim N(\mu, \sigma^2)$, we have

$$E\left(\frac{\partial^2 \ell}{\partial \mu^2}\right) = -\frac{n}{\sigma^2}$$
$$E\left(\frac{\partial^2 \ell}{\partial \sigma \partial \mu}\right) = 0$$
$$E\left(\frac{\partial^2 \ell}{\partial \sigma^2}\right) = -\frac{2n}{\sigma^2}.$$

Hence, the Fisher information matrix is

$$I(\theta) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix} \blacksquare$$

The inverse matrix $I(\theta)^{-1}$, referred to as the **covariance matrix**, has the variance of the individual random variables on the main diagonal and covariances in the off-diagonal positions.

Example 62.2
Find the covariance matrix of Example 62.1.

Solution.
The covariance matrix is

$$I(\theta)^{-1} = \frac{1}{det[I(\theta)]} \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{bmatrix} \blacksquare$$

In many cases, taking both the derivatives of the loglikelihood function and the corresponding expectations are not always easy. A way to avoid this problem is to simply not take the expected value. So instead of taking the values that result from the expectation, we can just use the observed data points. This method will result in the **observed information**. We illustrate this concept in the next example.

Example 62.3
You model a loss process using a lognormal distribution with parameters μ and σ. You are given:
(i) The maximum likelihood estimates of μ and σ are $\hat{\mu} = 4.4654$ and $\hat{\sigma}^2 = 0.3842$.
(ii) The following five observations: 27 82 115 126 155.
Estimate the covariance matrix using the observed information.

Solution.
Substituting the observations into the second derivatives, we find

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \mu^2} &= -\frac{n}{\sigma^2} = -\frac{5}{\sigma^2}\\
\frac{\partial^2 \ell}{\partial \sigma \partial \mu} &= -2\sum_{i=1}^{n}\frac{\ln x_i - \mu}{\sigma^3} = -2\frac{22.3272 - 5\mu}{\sigma^3}\\
\frac{\partial^2 \ell}{\partial \sigma^2} &= \frac{n}{\sigma^2} - 3\sum_{i=1}^{n}\frac{(\ln x_i - \mu)^2}{\sigma^4} = \frac{5}{\sigma^2} - 3\frac{101.6219 - 44.6544\mu + 5\mu^2}{\sigma^4}.
\end{aligned}$$

Using $\hat{\mu} = 4.4654$ and $\hat{\sigma}^2 = 0.3842$ in the previous expressions, we find

$$I(\theta) = \begin{bmatrix} 13.0141 & 0 \\ 0 & 26.0307 \end{bmatrix}$$

and

$$\widehat{\text{Var}}(\hat{\mu}, \hat{\sigma}) = [I(\theta)]^{-1} = \begin{bmatrix} 0.0768 & 0 \\ 0 & 0.0384 \end{bmatrix} \blacksquare$$

The Delta Method
The **delta method** is a method for estimating the variance of a function of estimators. The method is based on Taylor series expansions.

Let X be a distribution depending on a parameter θ. Suppose $f(x)$ is a differentiable function of x. The Taylor approximation of order one around θ is

$$f(\hat{\theta})) \approx f(\theta) + (\hat{\theta} - \theta)f'(\theta).$$

Hence,

$$\text{Var}[f(\hat{\theta})] \approx [f'(\theta)]^2 \text{Var}(\hat{\theta}).$$

Example 62.4
Consider a random sample of size n from an exponential distribution with mean θ.
(a) Find the variance of the estimated distribution variance.
(b) Construct the approximate 95% confidence interval for $\Pr(X > m)$.

Solution.
(a) We have $\widehat{\text{Var}}(X) \approx \hat{\theta}^2$. Hence,

$$\text{Var}(\hat{\theta}^2) = ([\theta^2]')^2 \text{Var}(\hat{\theta}) = 4\theta^2 \cdot \frac{1}{I(\theta)} = 4\theta^2 \cdot \frac{\theta^2}{n} = \frac{4\theta^4}{n}.$$

(b) The estimated probability is $e^{-\frac{m}{\hat{\theta}}}$. Thus,

$$\widehat{\text{Var}}\left[e^{-\frac{m}{\hat{\theta}}}\right] = \frac{m^2}{\hat{\theta}^4}e^{-\frac{2m}{\hat{\theta}}}\frac{\hat{\theta}^2}{n} = \frac{m^2}{n\hat{\theta}^2}e^{-\frac{2m}{\hat{\theta}}}.$$

An approximate 95% confidence interval is

$$e^{-\frac{m}{\hat{\theta}}} \pm 1.96\frac{m}{\sqrt{n}\hat{\theta}}e^{-\frac{m}{\hat{\theta}}} \blacksquare$$

This method can be extended to a pair of related maximum likelihood estimates. Let X be a distribution depending on two parameters θ_1 and θ_2. Let $f(s,t)$ be a differentiable function in the variables s and t. The delta method asserts that

$$\widehat{\text{Var}}[f(\hat{\theta}_1,\hat{\theta}_2)] = \begin{bmatrix} f_{\theta_1} & f_{\theta_2} \end{bmatrix} \begin{bmatrix} \text{Var}(\hat{\theta}_1) & \text{Cov}(\hat{\theta}_1,\hat{\theta}_2) \\ \text{Cov}(\hat{\theta}_1,\hat{\theta}_2) & \text{Var}(\hat{\theta}_2) \end{bmatrix} \begin{bmatrix} f_{\theta_1} \\ f_{\theta_2} \end{bmatrix}.$$

Example 62.5 ‡

You model a loss process using a lognormal distribution with parameters μ and σ. You are given:

(i) The maximum likelihood estimates of μ and σ are $\hat{\mu} = 4.215$ and $\hat{\sigma} = 1.093$.

(ii) The estimated covariance matrix of $\hat{\mu}$ and $\hat{\sigma}$ is

$$\begin{bmatrix} 0.1195 & 0 \\ 0 & 0.0597 \end{bmatrix}$$

(iii) The mean of the lognormal distribution is $e^{\mu+\frac{\sigma^2}{2}}$.

Estimate the variance of the maximum likelihood estimate of the mean of the lognormal distribution, using the delta method.

Solution.

Let $f(\mu,\sigma) = e^{\mu+\frac{\sigma^2}{2}}$. We have

$$\begin{aligned}
\widehat{\text{Var}}[f(\hat{\mu},\hat{\sigma})] &= \begin{bmatrix} e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} & \hat{\sigma}e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} \end{bmatrix} \begin{bmatrix} \frac{\hat{\sigma}^2}{n} & 0 \\ 0 & \frac{\hat{\sigma}^2}{2n} \end{bmatrix} \begin{bmatrix} e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} \\ \hat{\sigma}e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} \end{bmatrix} \\
&= \begin{bmatrix} 123.02 & 134.46 \end{bmatrix} \begin{bmatrix} 0.1195 & 0 \\ 0 & 0.0597 \end{bmatrix} \begin{bmatrix} 123.02 \\ 134.46 \end{bmatrix} \\
&= \begin{bmatrix} 14.70089 & 8.027262 \end{bmatrix} \begin{bmatrix} 123.02 \\ 134.46 \end{bmatrix} \\
&= 2887.85 \blacksquare
\end{aligned}$$

Example 62.6 ‡
You are given:
(i) Fifty claims have been observed from a lognormal distribution with unknown parameters μ and σ.
(ii) The maximum likelihood estimates are $\hat{\mu} = 6.84$ and $\hat{\sigma} = 1.49$.
(iii) The covariance matrix of $\hat{\mu}$ and $\hat{\sigma}$ is

$$\begin{bmatrix} 0.0444 & 0 \\ 0 & 0.0222 \end{bmatrix}$$

(iv) The partial derivatives of the lognormal cumulative distribution function are:

$$\frac{\partial F}{\partial \mu} = -\frac{\phi(z)}{\sigma} \text{ and } \frac{\partial F}{\partial \sigma} = -\frac{z\phi(z)}{\sigma}.$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$.
(v) An approximate 95% confidence interval for the probability that the next claim will be less than or equal to 5000 is: $[P_L, P_H]$.

Determine P_L.

Solution.
Let

$$F(\mu, \sigma) = \Pr(X \le 5000) = \Phi\left(\frac{ln5000 - \mu}{\sigma}\right).$$

The point estimate is

$$\hat{F}(6.84, 1.49) = \Phi\left(\frac{ln5000 - 6.84}{1.49}\right) = \Phi(1.125) = 0.87.$$

For the delta method, we need

$$\frac{\partial F}{\partial \mu} = -\frac{\phi(1.125)}{1.49} = -0.1422$$
$$\frac{\partial F}{\partial \sigma} = -\frac{1.125\phi(1.125)}{1.49} = -0.16.$$

Thus,

$$\widehat{\text{Var}}[F(\hat{\mu}, \hat{\sigma})] = \begin{bmatrix} -0.1422 & -0.16 \end{bmatrix} \begin{bmatrix} 0.0444 & 0 \\ 0 & 0.0222 \end{bmatrix} \begin{bmatrix} -0.1422 \\ -0.16 \end{bmatrix}$$
$$= 0.001466.$$

The lower limit of the 95% confidence interval is

$$P_L = 0.87 - 1.96\sqrt{0.001466} = 0.79496 \blacksquare$$

Practice Problems

Problem 62.1
For a Pareto distribution, the parameters α and θ are both estimated via the method of maximum likelihood by analyzing data from a random sample of n observations.

Find the loglikelihood function for this distribution.

Problem 62.2
For a Pareto distribution, the parameters α and θ are both estimated via the method of maximum likelihood by analyzing data from a random sample of n observations.

Find the second partial derivatives of the loglikelihood function.

Problem 62.3
For a Pareto distribution, the parameters α and θ are both estimated via the method of maximum likelihood by analyzing data from a random sample of n observations.

Find the information matrix.

Problem 62.4
For a Pareto distribution, the parameters α and θ are both estimated via the method of maximum likelihood by analyzing data from a random sample of n observations.

Find the covariance matrix.

Problem 62.5
You model a loss process using a Pareto distribution with parameters α and $\theta = 10$. You are given:
(i) The maximum likelihood estimate of $\alpha : \hat{\alpha} = 1.26$.
(ii) The following eight observations: 3 4 8 10 12 18 22 35.

Estimate $I(\alpha, \theta)$ using the observed information.

Problem 62.6
You model a loss process using a Pareto distribution with parameters α and $\theta = 10$. You are given:

(i) The maximum likelihood estimate of $\alpha : \hat{\alpha} = 1.26$.
(ii) The following eight observations: 3 4 8 10 12 18 22 35.

Estimate the covariance matrix using the observed information.

Problem 62.7 ‡
You are given:
(i) Loss payments for a group health policy follow an exponential distribution with unknown mean.
(ii) A sample of losses is: 100 200 400 800 1400 3100.

Use the delta method to approximate the variance of the maximum likelihood estimator of $S(1500)$.

Problem 62.8 ‡
The time to an accident follows an exponential distribution. A random sample of size two has a mean time of 6. Let Y denote the mean of a new sample of size two.

(a) Using moment generating functions, show that the sum of two independent exponential random variables is a Gamma random variables.
(b) Determine the maximum likelihood estimate of $\Pr(Y > 10)$.
(c) Use the delta method to approximate the variance of the maximum likelihood estimator of $F_Y(10)$.

Problem 62.9 ‡
A survival study gave $(0.283, 1.267)$ as the symmetric linear 95% confidence interval for $H(5)$.

Using the delta method, determine the symmetric linear 95% confidence interval for $S(5)$.

Problem 62.10 ‡
You have modeled eight loss ratios as $Y_t = \alpha + \beta t + \epsilon_t,\ t = 1, 2, \cdots, 8$, where Y_t is the loss ratio for year t and ϵ_t is an error term.
You have determined:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.02 \end{bmatrix}$$

$$\text{Var}\left(\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}\right) = \begin{bmatrix} 0.00055 & -0.00010 \\ -0.00010 & 0.00002 \end{bmatrix}$$

Estimate the standard deviation of the forecast for year 10, $\hat{Y}_{10} = \hat{\alpha} + 10\hat{\beta}$, using the delta method.

Problem 62.11 ‡
A sample of ten observations comes from a parametric family $f(x, y, \theta_1, \theta_2)$ with loglikelihood function

$$\ln\left[L(\theta_1, \theta_2)\right] = \sum_{i=1}^{10} \ln\left[f(x_i, y_i, \theta_1, \theta_2)\right] = -2.5\theta_1^2 - 3\theta_1\theta_2 - \theta_2^2 + 5\theta_1 + 2\theta_1 + k$$

where k is a constant.

Determine the estimated covariance matrix of the maximum likelihood estimator, $\begin{bmatrix} \hat{\theta_1} \\ \hat{\theta_2} \end{bmatrix}$.

63 Non-Normal Confidence Intervals for Parameter Estimation

Up to this point, confidence intervals for the maximum likelihood estimators have been based on the assumption that the estimators are normally distributed. Such an assumption is true for large sampling but may not be true for small or even moderate samples. In this section, we consider a method for finding confidence intervals that does not require the normality assumption.

We first recall the reader of the **chi-square distribution**. The chi-square distribution with k degrees of freedom is the distribution of the aggregate random variable

$$Q = Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

where the $Z_i's$ are standard normal distributions. The chi-square, denoted by χ^2, is a one parameter distribution with parameter k.

A generalization of a confidence interval is a confidence region. We define the $100(1-\alpha)\%$ **confidence region** to be the set of values of a parameter θ such that

$$\ell(\theta) \geq c = \ell(\hat{\theta}) - \frac{1}{2}\chi^2_{1-\alpha}$$

where $\chi^2_{1-\alpha}$ is the $1-\alpha$ percentile from the chi-square distribution with degrees of freedom equal to the number of estimated parameters.

Example 63.1
Write the inequality that will result in the 95% non-normal confidence region for the mean θ of the exponential distribution.

Solution.
The loglikelihood function is

$$\ell(\theta) = -n\ln\theta - \sum_{i=1}^{n}\frac{x_i}{\theta}.$$

The maximum likelihood estimator is

$$\hat{\theta} = \overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Thus,

$$\ell(\hat{\theta}) = -n - n\ln\overline{x}.$$

Now, from the table of chi-square distribution, $\chi^2_{0.95} = 3.841$. Hence, the confidence region is

$$-n\ln\theta - \frac{n\overline{x}}{\theta} \geq -n - n\ln\overline{x} - 1.9205$$

which must be evaluated numerically ■

Remark 63.1
Since numerical methods are needed for determining most of the non-normal confidence intervals and the exam candidates are not provided with the tools (such as Excel) for solving these problems, it is very unlikely to see this type of questions on the exam.

Example 63.2
Suppose that $n = 20$ and $\overline{x} = 1424.40$ in the previous example. Use Excel's Solver in finding the confidence interval by solving the confidence region

$$-\frac{28488}{\theta} - 20\ln\theta \geq -167.15.$$

Solution.
Using Excel's Solver, one solves the equation

$$-28488 - 20\theta\ln\theta + 167.15\theta = 0.$$

The two roots are 946.85 and 2285.05. Hence, the confidence interval is $[946.85, 2285.05]$ ■

Practice Problems

Problem 63.1
You are given:
• $\ell(\theta) = 5\ln\theta - 2.55413(\theta - 1)$.
• $\hat{\theta} = 1.95762$.

Find the value of c to be used in the construction of a 95% non-normal confidence region.

Problem 63.2
Use Excel's Solver in finding the confidence interval by solving the confidence region

$$5\ln\theta - -2.55413(\theta - 1) \geq -1.0077.$$

Problem 63.3
You are given:
• $\ell(\theta) = -\theta^2 + 11\theta - 24$.
• $\hat{\theta} = 3$.

Find a 95% non-normal confidence interval for θ.

64 Basics of Bayesian Inference

There are two approaches to parameter estimation. The **frequentest approach** assumes that there is an unknown but fixed parameter θ. It chooses the value of θ that maximizes the likelihood of observed data. In other words, making the available data as likely as possible. A common example is the maximum likelihood estimator. In the frequentest approach, probabilities are defined as the frequency of successful trials over the total number of trials in an experiment. To elaborate, suppose that the parameter to be estimated is the probability of getting a head when flipping a coin. we flip a fair coin 100 times and it comes out head 30 times and tail 70 times. In this case, we say that the probability of getting a head is 30% and that of a tail is 70%.

In contrast, the **Bayesian approach** allows probability to represent subjective uncertainty or subjective belief. In the coin example, our subjective belief tells us that if head and tail are equally likely to occur than the probability of each to occur must be 50%. This approach fixes the data and instead assumes possible values of θ. In other words, Bayesians treat the unknown model parameters as random variables and assign probabilities to the subsets of the parameter space.

We next introduce some Bayesian inference terminology: A **prior distribution** of a parameter θ is the probability distribution over the space of possible values of the parameter. It is denoted by $\pi(\theta)$ and represents your uncertainty about the parameter before the current data are examined.
A prior distribution $\pi(\theta)$ is said to be **improper** if $\pi(\theta) \geq 0$ and

$$\int \pi(\theta)d\theta = \infty.$$

Example 64.1
Show that the uniform prior distribution on the real line $\pi(\theta) = 1$ for $-\infty < \theta < \infty$ is improper.

Solution.
This follows from

$$\int_{-\infty}^{\infty} 1d\theta = \infty \blacksquare$$

Let the observable random variables $X_1, X_2, \cdots, X_n$ be i.i.d and denote $\mathbf{X} = (X_1, X_2, \cdots, X_n)^T$. The probability distribution of $\mathbf{X}$ given a particular value of a parameter is called the **model distribution** and is given by

$$f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta) = f_{\mathbf{X}|\Theta}(x_1|\theta)f_{\mathbf{X}|\Theta}(x_2|\theta)\cdots f_{\mathbf{X}|\Theta}(x_n|\theta)$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$.

The **joint distribution** of $\mathbf{X}$ and Θ is given by

$$f_{\mathbf{X},\Theta}(\mathbf{x}, \theta) = f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta).$$

The **marginal distribution** of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \int f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta)d\theta.$$

The **posterior distribution** is the conditional probability distribution of the parameters, given the observed data. It is denoted $\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})$ and is given by Bayes Theorem

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \frac{f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta)}{f_{\mathbf{X}}(\mathbf{x})}.$$

The **predictive distribution** is the conditional probability distribution of a new observation y, given the data $\mathbf{x}$. It is denoted $f_{Y|\mathbf{X}}(y|\mathbf{X})$ and is given by

$$f_{Y|\mathbf{X}}(y|\mathbf{X}) = \int f_{Y|\Theta}(y|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta.$$

Example 64.2 ‡
You are given:
(i) A portfolio consists of 100 identically and independently distributed risks.
(ii) The number of claims for each risk follows a Poisson distribution with mean λ.
(iii) The prior distribution of λ is:

$$\pi(\lambda) = \frac{(50\lambda)^4 e^{-50\lambda}}{6\lambda},\ \lambda > 0.$$

During Year 1, the following loss experience is observed:

# of claims	# of risks
0	90
1	7
2	2
3	1
Total	100

(a) Find the model distribution.
(b) Find the joint density of $\mathbf{X}$ and Λ.
(c) Find the marginal density at $\mathbf{x}$.
(d) Find the posterior distribution of Λ.
(e) Find the predictive distribution.

Solution.
(a) The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda) = \left(\frac{e^{-\lambda}\lambda^0}{0!}\right)^{90}\left(\frac{e^{-\lambda}\lambda^1}{1!}\right)^{7}\left(\frac{e^{-\lambda}\lambda^2}{2!}\right)^{2}\left(\frac{e^{-\lambda}\lambda^3}{3!}\right) = \frac{e^{-100\lambda}\lambda^{14}}{24}.$$

(b) The joint density is

$$f_{\mathbf{X},\lambda}(\mathbf{x},\lambda) = \frac{e^{-100\lambda}\lambda^{14}}{24}\frac{(50\lambda)^4 e^{-50\lambda}}{6\lambda} = \frac{390625}{9}e^{-150\lambda}\lambda^{17}.$$

(c) The marginal density of $\mathbf{X}$ is

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) =& \frac{390625}{9}\int_0^\infty e^{-150\lambda}\lambda^{17}d\lambda = \frac{390625}{9}\frac{\Gamma(18)}{150^{18}}\int_0^\infty \frac{e^{-\frac{\lambda}{\frac{1}{150}}}\left(\frac{\lambda}{\frac{1}{150}}\right)^{18}}{\lambda\Gamma(18)}d\lambda \\ =& \frac{390625}{9}\frac{\Gamma(18)}{150^{18}}. \end{aligned}$$

(d) The posterior distribution of Λ is

$$\pi_{\Lambda|\mathbf{X}}(\lambda|\mathbf{x}) = \frac{150^{18}}{17!}e^{-150\lambda}\lambda^{17}.$$

Note that the posterior distribution is a Gamma distribution with $\alpha = 18$ and $\theta = \frac{1}{150}$.
(e) The predictive distribution is

$$\begin{aligned} f_{Y|\mathbf{X}}(y|\mathbf{x}) =& \int_0^\infty \frac{e^{-\lambda}\lambda^y}{y!}\frac{150^{18}}{17!}e^{-150\lambda}\lambda^{17}d\lambda = \frac{150^{18}}{y!17!}\int_0^\infty \lambda^{17+y}e^{-151\lambda}d\lambda \\ =& \frac{150^{18}}{y!17!}\frac{(18+y)!}{151^{18+y}} \blacksquare \end{aligned}$$

Example 64.3 ‡
You are given:

(i) Losses on a company's insurance policies follow a Pareto distribution with probability density function:

$$f(x|\theta) = \frac{\theta}{(x+\theta)^2}, \; x > 0.$$

(ii) For half of the company's policies $\theta = 1$, while for the other half $\theta = 3$. For a randomly selected policy, losses in Year 1 were 5. Determine the posterior probability that losses for this policy in Year 2 will exceed 8.

Solution.
Let X_n denote the losses in Year n. We are asked to find $\Pr(X_2 > 8|X_1 = 5)$. Recall from Section 3, that probabilities can be found by conditioning. Thus, we can write

$$\begin{aligned}\Pr(X_2 > 8|X_1 = 5) =& \Pr(X_2 > 8|\theta = 1)\Pr(\theta = 1|X_1 = 5)\\ &+ \Pr(X_2 > 8|\theta = 3)\Pr(\theta = 3|X_1 = 5).\end{aligned}$$

Since $\Theta|X$ is discrete, we have

$$\begin{aligned}\Pr(\theta = 1|X_1 = 5) =& \pi(\theta = 1|x = 5) = \frac{f(\theta = 1|X_1 = 5)\pi(1)}{f(\theta = 1|X_1 = 5)\pi(1) + f(\theta = 3|X_1 = 5)\pi(3)}\\ =& \frac{\frac{1}{36}\frac{1}{2}}{\frac{1}{36}\frac{1}{2} + \frac{3}{64}\frac{1}{2}} = \frac{16}{43}.\end{aligned}$$

Thus,

$$\Pr(\theta = 3|X_1 = 5) = 1 - \Pr(\theta = 1|X_1 = 5) = 1 - \frac{16}{43} = \frac{27}{43}.$$

On the other hand,

$$\Pr(X_2 > 8|\theta = 1) = \int_8^\infty \frac{dx}{(x+1)^2} = \frac{1}{9}.$$

Likewise,

$$\Pr(X_2 > 8|\theta = 3) = \int_8^\infty \frac{3}{(x+3)^2} dx = \frac{3}{11}.$$

Finally,

$$\Pr(X_2 > 8|X_1 = 5) = \left(\frac{1}{9}\right)\left(\frac{16}{43}\right) + \left(\frac{3}{11}\right)\left(\frac{27}{43}\right) = 0.2126 \blacksquare$$

Example 64.4 ‡
You are given:
(i) Each risk has at most one claim each year.
(ii)

i	Type of Risk(θ_i)	$\pi(\theta_i)$	Annual claim probability
1	I	0.7	0.1
2	II	0.2	0.2
3	III	0.1	0.4

One randomly chosen risk has three claims during Years 1-6. Determine the posterior probability of a claim for this risk in Year 7.

Solution.
Let X_n denote the number of claims in year n. Then $X_n|\Theta$ is a Bernoulli random vraible with probability of a claim (success) p. Let $S = \sum_{i=1}^{6} X_i$. Then $S|\Theta$ is a binomial distribution with 6 trials (number of years) and probability of a claim p. Note that, $S|Theta = 3$ is the number of claims in 6 years. We are asked to find $\Pr(X_7 = 1||(S = 3)$ for a given θ which by Bayes theorem is

$$\Pr(X_7 = 1|S = 3) = \frac{\Pr[(X_7 = 1) \cap (S = 3)]}{\Pr(S = 3)}.$$

We have, for a given θ,

$$\Pr(S = 3|\theta_1) = \left(\begin{array}{c} 6 \\ 3 \end{array} \right)(0.1)^3(0.9)^3 = 0.01458$$
$$\Pr(S = 3|\theta_2) = \left(\begin{array}{c} 6 \\ 3 \end{array} \right)(0.2)^3(0.8)^3 = 0.08192$$
$$\Pr(S = 3|\theta_3) = \left(\begin{array}{c} 6 \\ 3 \end{array} \right)(0.4)^3(0.6)^3 = 0.27648$$
$$\begin{aligned}\Pr(S = 3) =&\Pr(S = 3|\theta_1)\pi(\theta_1) + \Pr(S = 3|\theta_2)\pi(\theta_2) + \Pr(S = 3|\theta_3)\pi(\theta_3)\\ =&0.01458(0.7) + 0.08192(0.2) + 0.27648(0.1) = 0.054238.\end{aligned}$$

Now, $(X_7 = 1) \cap (S = 3)$ is the intersection of two independent events that depend on θ. Thus, we have

$$\begin{aligned}
\Pr[(X_7 = 1) \cap (S = 3)] =& \Pr[(X_7 = 1) \cap (S = 3)|\theta_1]\pi(\theta_1) \\
&+ \Pr[(X_7 = 1) \cap (S = 3)|\theta_2]\pi(\theta_2) \\
&+ \Pr[(X_7 = 1) \cap (S = 3)|\theta_3]\pi(\theta_3)) \\
=& \Pr(X_7 = 1|\theta_1)\Pr(S = 3|\theta_1)\pi(\theta_1) \\
&+ \Pr(X_7 = 1|\theta_2)\Pr(S = 3|\theta_2)\pi(\theta_2) \\
&+ \Pr(X_7 = 1|\theta_3)\Pr(S = 3|\theta_3)\pi(\theta_3) \\
=& (0.1)(0.01458)(0.7) + (0.2)(0.08192)(0.2) + (0.4)(0.27648)(0.1) \\
=& 0.0153566.
\end{aligned}$$

Finally,

$$\Pr(X_7 = 1|S = 3) = \frac{0.0153566}{0.054238} = 0.283 \blacksquare$$

Example 64.5 ‡

You are given:

(i) For $Q = q$, the random variables $X_1, X_2, \cdots, X_m$ are independent, identically distributed Bernoulli random variables with parameter q.

(ii) $S_m = X_1 + X_2 + \cdots + X_m$.

(iii) The prior distribution of Q is beta with $a = 1, b = 99$, and $\theta = 1$.

Determine the smallest value of m such that the mean of the marginal distribution of S_m is greater than or equal to 50.

Solution.

Since $X_1, X_2, \cdots, X_m$ are independent and identically distributed with common distribution a Bernoulli distribution with parameter q, $S_m|Q$ is a binomial distribution with parameters (m, q). The mean of the marginal distribution is (using the double expectation property, Theorem 78.2) is

$$E(S_m) = E[E(S_M|Q)] = E(mQ) = mE(Q) = \frac{m}{1+b} = 0.01m.$$

The inequality $E(S_m) \geq 50$ implies $m \geq 5000 \blacksquare$

Example 64.6 ‡

You are given:

(i) The number of claims made by an individual in any given year has a binomial distribution with parameters $m = 4$ and q.

(ii) The prior distribution of q has probability density function

$$\pi(q) = 6q(1-q), \ 0 < q < 1.$$

(iii) Two claims are made in a given year. Determine the mode of the posterior distribution of q.

Solution.
We have

$$\begin{aligned} f(2,q) =& f(2|q)\pi(q) = 36q^3(1-q)^3 \\ f(2) =& \int_0^1 36q^3(1-q)^3 dq \\ \pi(q|2) =& \frac{1}{f(2)}q^3(1-q)^3. \end{aligned}$$

The mode is the value of q that maximizes the posterior distribution. Taking the derivative and setting it to zero, we find

$$3q^2(1-q)^3 - 3q^3(1-q)^2 = 0 \Longrightarrow q = 0.5.$$

Moreover,

$$\pi''(2|0.5) = -6(0.5)^4/f(2) < 0$$

so that the posterior distribution is maximizex when $q = 0.5$. That is, the mode of the posterior distribution is 0.5 ■

Example 64.7 ‡
The observation from a single experiment has distribution:

$$\Pr(D = d|G = g) = g^{(1-d)}(1-g)^d, \quad \text{for} d = 0, 1.$$

The prior distribution of G is:

$$\Pr\left(G = \tfrac{1}{5}\right) = \tfrac{3}{5} \text{ and } \Pr\left(G = \tfrac{1}{3}\right) = \tfrac{2}{5}.$$

Calculate $\Pr\left(G = \frac{1}{3}|D = 0\right)$.

Solution.
We have

$$\begin{aligned} \Pr\left(G = \frac{1}{3}|D = 0\right) =& \frac{\Pr\left(D = 0|G = \frac{1}{3}\right)\Pr\left(G = \frac{1}{3}\right)}{\Pr\left(D = 0|G = \frac{1}{3}\right)\Pr\left(G = \frac{1}{3}\right) + \Pr\left(D = 0|G = \frac{1}{5}\right)\Pr\left(G = \frac{1}{5}\right)} \\ =& \frac{(2/3)(1/5)}{(2/3)(1/5) + (1/5)(3/5)} = \frac{10}{19} \blacksquare \end{aligned}$$

Practice Problems

Problem 64.1
Show that the prior distribution $\pi(\theta) = \frac{1}{\theta},\ 0 < \theta < \infty$ is improper.

Problem 64.2 ‡
You are given:
(i) In a portfolio of risks, each policyholder can have at most two claims per year.
(ii) For each year, the distribution of the number of claims is:

Number of claims	Probability
0	0.10
1	$0.90 - q$
2	q

(iii) The prior density is:

$$\pi(q) = \frac{q^2}{0.039},\ 0.2 < q < 0.5.$$

A randomly selected policyholder had two claims in Year 1 and two claims in Year 2. For this insured, determine
(a) the model distribution
(b) the joint distribution
(c) the marginal distribution
(d) the posterior distribution.

Problem 64.3 ‡
You are given:
(i) The annual number of claims for a policyholder follows a Poisson distribution with mean λ.
(ii) The prior distribution of Λ is Gamma with probability density function:

$$f(\lambda) = \frac{(2\lambda)^5 e^{-2\lambda}}{24\lambda},\ \lambda > 0.$$

An insured is selected at random and observed to have $x_1 = 5$ claims during Year 1 and $x_2 = 3$ claims during Year 2.

Determine the posterior distribution.

Problem 64.4 ‡
You are given:
(i) Annual claim frequencies follow a Poisson distribution with mean λ.
(ii) The prior distribution of Λ has probability density function:

$$\pi(\lambda) = \frac{0.4}{6}e^{-\frac{\lambda}{6}} + \frac{0.6}{12}e^{-\frac{\lambda}{12}}, \ \lambda > 0.$$

Ten claims are observed for an insured in Year 1.

Determine the posterior distribution.

Problem 64.5 ‡
You are given:
(i) In a portfolio of risks, each policyholder can have at most one claim per year.
(ii) The probability of a claim for a policyholder during a year is q.
(iii) The prior density is $\pi(q) = \frac{q^3}{0.07}$ $0.6 < q < 0.8$.
A randomly selected policyholder has one claim in Year 1 and zero claims in Year 2.

For this policyholder, determine the posterior probability that $0.7 < q < 0.8$.

Problem 64.6 ‡
You are given:
(i) The probability that an insured will have exactly one claim is θ.
(ii) The prior distribution of Θ has probability density function:

$$\pi(\theta) = 1.5\sqrt{\theta}, \ 0 < \theta < 1.$$

A randomly chosen insured is observed to have exactly one claim.

Determine the posterior probability that Θ is greater than 0.60.

Problem 64.7 ‡
You are given:
(i) The prior distribution of the parameter Θ has probability density function:

$$\pi(\theta) = \frac{1}{\theta^2}, \ \theta > 1.$$

Given $\Theta = \theta$, claim sizes follow a Pareto distribution with parameters $\alpha = 2$ and θ.

A claim of 3 is observed.

Calculate the posterior probability that Θ exceeds 2.

Problem 64.8 ‡
You are given:
(i) The number of claims observed in a 1-year period has a Poisson distribution with mean θ.
(ii) The prior density is:

$$\pi(\theta) = \frac{e^{-\theta}}{1-e^{-k}},\ 0 < \theta < k.$$

(iii) The unconditional probability of observing zero claims in 1 year is 0.575.

Determine k.

Problem 64.9 ‡
You are given: (i) Conditionally, given β, an individual loss X follows the exponential distribution with probability density function:

$$f(x|\beta) = \frac{1}{\beta}e^{-\frac{x}{\beta}},\ 0 < x < \infty.$$

(ii) The prior distribution of β is inverse gamma with probability density function:

$$\pi(\beta) = \frac{c^2}{\beta^3}e^{-\frac{c}{\beta}},\ 0 < \beta < \infty.$$

(iii) $\int_0^\infty \frac{1}{y^n}e^{-\frac{a}{y}}dy = \frac{(n-2)!}{a^{n-1}},\ n = 2, 3, 4, \cdots$.

Given that the observed loss is x, calculate the mean of the posterior distribution of β.

Problem 64.10 ‡
You are given:
(i) Annual claim counts follow a Poisson distribution with mean λ.
(ii) The parameter λ has a prior distribution with probability density function:

$$f(\lambda) = \frac{1}{3}e^{-\frac{\lambda}{3}}.$$

Two claims were observed during the first year.

Determine the variance of the posterior distribution of λ.

Problem 64.11 ‡
You are given:
(i) For $Q = q$, the random variables $X_1, X_2, \cdots, X_m$ are independent, identically distributed Bernoulli random variables with parameter q.
(ii) $S_m = X_1 + X_2 + \cdots + X_m$.
(iii) The prior distribution of Q is beta with $a = 1, b = 99$, and $\theta = 1$.

Determine the variance of the marginal distribution of S_{101}.

Problem 64.12 ‡
You are given the following information about workers compensation coverage:
(i) The number of claims for an employee during the year follows a Poisson distribution with mean $(100 - p)/100$, where p is the salary (in thousands) for the employee.
(ii) The distribution of p is uniform on the interval $(0, 100]$.
An employee is selected at random. No claims were observed for this employee during the year.

Determine the posterior probability that the selected employee has salary greater than 50 thousand.

Problem 64.13 ‡
Prior to observing any claims, you believed that claim sizes followed a Pareto distribution with parameters $\theta = 10$ and $\alpha = 1$, 2 or 3 , with each value being equally likely.
You then observe one claim of 20 for a randomly selected risk.

Determine the posterior probability that the next claim for this risk will be greater than 30.

65 Bayesian Parameter Estimation

Among the objectives of Bayesian inference is the estimation of a model parameter. In this section, we consider the approach discussed in [1].

A **Bayes estimate** of a parameter θ is that value that minimizes the expected **loss function** given the posterior distribution of θ. The three most commonly used loss functions are:

• **Squared-error loss** : $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$.
• **Absolute loss**: $L(\hat{\theta}, \theta) = |\hat{\theta} - \theta|$.
• **Zero-one loss**:

$$L(\hat{\theta}, \theta) = \begin{cases} 0, & \hat{\theta} = \theta, \\ 1, & \text{otherwise} \end{cases}$$

The following theorem provides the Bayes estimate of each of the above loss functions.

Theorem 65.1
For squared-error loss, the Bayes estimate is the mean of the posterior distribution; for absolute loss, it is a median; for zero-one loss, it is a mode.

Example 65.1
The posterior distribution of θ is given by $\pi_{\Theta|\mathbf{X}}(\theta|x) = \frac{1}{x}e^{-\frac{\theta}{x}}$. Determine the Bayes estimate of θ using
(a) the squared-error loss function;
(b) the absolute loss function;
(c) the zero-one loss function.

Solution.
(a) $\hat{\theta} = x$.
(b) The median of the posterior distribution is the number M such that $\int_0^M \frac{1}{x}e^{-\frac{\theta}{x}}d\theta = 0.5$. Solving this equation, we find $M = x\ln 2$. Thus, $\hat{\theta} = x\ln 2$.
(c) The value that maximizes the posterior distribution is $\theta = 0$. That is, the mode of the posterior distribution is 0. Hence, $\hat{\theta} = 0$ ■

The expected value of the predictive distribution provides a point-estimate of the $(n+1)$st observation if the first n observations and the prior distribution are known. The expected value of the predictive distribution can be expressed as follows:

$$\begin{aligned}E(Y|\mathbf{x}) &= \int y f_{Y|\mathbf{X}}(y|\mathbf{x})dy = \int y \int f_{Y|\Theta}(y|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta dy \\ &= \int \pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) \int y f_{Y|\Theta}(y|\theta)dyd\theta = \int E(Y|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta.\end{aligned}$$

Example 65.2 ‡
You are given:
(i) The annual number of claims for a policyholder has a binomial distribution with probability function:

$$p(x|q) = \binom{2}{x} q^x(1-q)^{2-x},\ x = 0, 1, 2.$$

(ii) The prior distribution is:

$$\pi(q) = 4q^3,\ 0 < q < 1.$$

This policyholder had one claim in each of Years 1 and 2. Determine the Bayesian estimate of the number of claims in Year 3.

Solution.
Let X be the previously observed data and let Y be the number of claims in Year 3. Then $E(Y|q) = nq = 2q$. The joint distribution is

$$f_{\mathbf{X},Q}(\mathbf{x}, q) = 4q^2(1-q)^2(4q^3) = 16q^5(1-q)^2.$$

The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^1 16q^5(1-q)^2 dq = \frac{16}{168}.$$

The posterior distribution is

$$\pi_{Q|\mathbf{X}}(q|\mathbf{x}) = 168q^5(1-q)^2.$$

Thus,

$$E(Y|\mathbf{x}) = \int_0^1 E(Y|q)\pi_{Q|\mathbf{X}}(q|\mathbf{x})dq = \int_0^1 (2q)[168q^5(1-q)^2]dq = \frac{4}{3} \blacksquare$$

Example 65.3 ‡
For a group of insureds, you are given:
(i) The amount of a claim is uniformly distributed but will not exceed a certain unknown limit θ.
(ii) The prior distribution of Θ is

$$\pi(\theta) = \frac{500}{\theta^2},\ \theta > 500.$$

Two independent claims of 400 and 600 are observed.
Determine the probability that the next claim will exceed 550.

Solution.
We are asked to find $\Pr(X_3 > 550|X_1, X_2)$. We have

$$\begin{aligned}\Pr(X_3 > 550|X_1, X_2) =& \int_{550}^{\infty} f_{X_3|\mathbf{X}}(x_3|\mathbf{x})dx_3 = \int_{550}^{\infty}\int f_{X_3|\Theta}(x_3|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta dx_3\\ =& \int \pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})\int_{550} f_{X_3|\Theta}(x_3|\theta)dx_3 d\theta = \int \Pr(X_3 > 550|\Theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta.\end{aligned}$$

Since $X_3|\Theta$ has a uniform distribution on $(0, \theta)$, we have

$$\Pr(X_3 > 550|\Theta) = \frac{550 - \theta}{\theta}.$$

The model distribution is

$$f(\mathbf{x}|\theta) = \frac{1}{\theta^2},\ \theta > 600.$$

The joint distribution is

$$f(\mathbf{x}|\theta) = f(\mathbf{x}|\theta)\pi(\theta) = \frac{500}{\theta^4},\ \theta > 600.$$

The marginal distribution is

$$f(\mathbf{x}) = \int_{600}^{\infty} \frac{500}{\theta^4} d\theta = \frac{1}{3(600)^2}.$$

The posterior distribution is

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \frac{3(600)^3}{\theta^4}.$$

Hence,

$$\Pr(X_3 > 550|X_1, X_2) = \int_{600}^{\infty} \left(\frac{550 - \theta}{\theta}\right)\frac{3(600)^3}{\theta^4} d\theta = 0.3125 \blacksquare$$

A $100(1-\alpha)\%$ **credibility interval** for θ is an interval $[a,b]$ such that the posterior probability

$$\Pr(a \le \Theta \le b|\mathbf{x}) \ge 1-\alpha.$$

It is possible to have many $100(1-\alpha)\%$ credibility interval for a parameter θ. The following result provides the shortest such interval.

Theorem 65.2
If the posterior random variable $\Theta|\mathbf{X}$ is continuous and unimodal, then the $100(1-\alpha)\%$ credibility interval with smallest width $b-a$ is the unique solution to the following system of equations:

$$\begin{aligned}\int_a^b \pi_{\Theta|\mathbf{X}}(\theta,\mathbf{x})d\theta =& 1-\alpha\\ \pi_{\Theta|\mathbf{X}}(a,\mathbf{x}) =& \pi_{\Theta|\mathbf{X}}(b,\mathbf{x}).\end{aligned}$$

Example 65.4
You are given:
(i) The probability that an insured will have at least one loss during any year is p.
(ii) The prior distribution for p is uniform on $[0,0.5]$.
(iii) An insured is observed for 8 years and has at least one loss every year.
Develop a non-zero width 95% credibility interval for the posterior probability that the insured will have at least one loss during Year 9.

Solution.
In Problem 65.2, we find

$$\pi_{P|\mathbf{X}}(p,\mathbf{x}) = 4608p^8 \text{ and } p = 0.45.$$

Thus, according to Theorem 65.2, one of the equations of the system that we will need to solve is

$$\int_a^b \pi_{\Theta|\mathbf{X}}(\theta,\mathbf{x})d\theta = 1-\alpha \Longrightarrow \int_a^b 4608p^8dp = 0.95.$$

The other equation is

$$\pi_{\Theta|\mathbf{X}}(a,\mathbf{x}) = \pi_{\Theta|\mathbf{X}}(b,\mathbf{x}) \Longrightarrow 4608a^8 = 4608b^8.$$

The second equation can only be solved either such that $a = b$ or that $a = -b$. The solution $a = b$ would give a width of zero. The solution $a = -b$ implies the following:

$$\int_{-b}^{b} 4608p^8 dp = 1024b^9 = 0.95 \Longrightarrow b = 0.4603.$$

Thus, the credibility interval is $[-0.4603, 0.4603]$ ■

The interval in Theorem 65.2 is a special case of a highest posterior density (HPD) credibility set: For a given posterior distribution, the $100(1-\alpha)\%$ **HPD credibility set** for a parameter θ is a set defined by

$$\mathcal{C} = \{\theta : \pi_{\Theta|\mathbf{X}}(\theta, \mathbf{x}) \geq k\}$$

where k is the largest number satisfying

$$\int_{\mathcal{C}} \pi_{\Theta|\mathbf{X}}(\theta, \mathbf{x}) d\theta = 1 - \alpha.$$

Thus, $\mathcal{C}$ is the region with the highest posterior density. The HPD credible interval may not be unique and may in fact be a union of intervals. Also, it can be shown that the $100(1-\alpha)\%$ is the shortest among all $100(1-\alpha)\%$ credible intervals. HPD regions are difficult to determine analytically and we shall not pursue this topic any further.

Now, for large sampling, there is a version of the Central Limit Theorem which we call the **Bayesian Central Limit Theorem**:
If $\pi(\theta)$ (the prior distribution) and $f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)$ (the model distribution) are both twice differentiable in the elements of θ and other commonly satisfied assumptions hold, then the posterior distribution of Θ given $\mathbf{X} = \mathbf{x}$ is asymptotically normal.

Example 65.5
Redo Example 65.4 using the Bayesian Central Limit Theorem.

Solution.
Both the prior distribution function and the model distribution are twice differentiable so that we can assume that the posterior distribution is approximetely normal. We have

$$\begin{aligned} E(p) =& 0.45 \\ \text{Var}(p) =& E(p^2) - E(p)^2 \\ =& \int_0^{0.5} 4608p^{10} dp - 0.45^2 = 0.002. \end{aligned}$$

Thus, the credibility interval is

$$[0.45 - 1.96\sqrt{0.002}, 0.45 + 1.96\sqrt{0.002}] = [0.3623, 0.5377] \blacksquare$$

Practice Problems

Problem 65.1
The true value of a parameter is $\theta = 50$. Suppose that the Bayes estimate of θ using the squared-error loss function is $\hat{\theta} = 52$. Find the absolute value of the difference between the squared-error loss function and the absolute loss function.

Problem 65.2 ‡
You are given:
(i) The probability that an insured will have at least one loss during any year is p.
(ii) The prior distribution for p is uniform on $[0, 0.5]$.
(iii) An insured is observed for 8 years and has at least one loss every year.

Determine the posterior probability that the insured will have at least one loss during Year 9.

Problem 65.3
You are given:
(i) The probability that an insured will have at least one loss during any year is p.
(ii) The prior distribution for p is uniform on $[0, 0.5]$.
(iii) An insured is observed for 8 years and has at least one loss every year.

Using the Bayesian Central Limit Theorem, estimate the posterior probability that $p > 0.6$.

Problem 65.4 ‡
You are given:
(i) The amount of a claim, X, is uniformly distributed on the interval $[0, \theta]$.
(ii) The prior density of θ is $\pi(\theta) = \frac{500}{\theta^2},\ \theta > 500$.
Two claims, $x_1 = 400$ and $x_2 = 600$, are observed. You calculate the posterior distribution as:

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = 3\left(\frac{600^3}{\theta^4}\right),\ \theta > 600.$$

Calculate $E(X_3|\mathbf{X})$.

Problem 65.5 ‡
You are given:

(i) In a portfolio of risks, each policyholder can have at most two claims per year.
(ii) For each year, the distribution of the number of claims is:

Number of claims	Probability
0	0.10
1	$0.90 - q$
2	q

(iii) The prior density is:

$$\pi(q) = \frac{q^2}{0.039},\ 0.2 < q < 0.5.$$

A randomly selected policyholder had two claims in Year 1 and two claims in Year 2.

For this insured, determine the Bayesian estimate of the expected number of claims in Year 3.

Problem 65.6 ‡
You are given:
(i) The annual number of claims for each policyholder follows a Poisson distribution with mean θ.
(ii) The distribution of θ across all policyholders has probability density function:

$$f(\theta) = \theta e^{-\theta},\ \theta > 0.$$

(iii) $\int_0^\infty \theta e^{-n\theta} d\theta = \frac{1}{n^2}$.
A randomly selected policyholder is known to have had at least one claim last year.

Determine the posterior probability that this same policyholder will have at least one claim this year.

66 Conjugate Prior Distributions

In this section, we look at models where the prior and posterior distributions belong to the same family of distributions.

If a prior distribution combined with a model distribution result in a posterior distribution that belongs to the same family of functions as the prior distribution (perhaps with different parameters) then we call the prior distribution a **conjugate prior distribution**.

Example 66.1
You are given:
• The prior distribution Θ is Gamma with parameters α and β.
• The model distribution is exponential with parameter θ.
Show that Θ has a conjugate prior distribution.

Solution.
The model distribution is

$$f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)=\theta^n e^{-\theta\sum_{i=1}^n x_i}.$$

The joint distribution is

$$f_{\mathbf{X},\Theta}(\mathbf{x},\theta)=\theta^n e^{-\theta\sum_{i=1}^n x_i}\frac{\theta^{\alpha-1}e^{-\frac{\theta}{\beta}}}{\beta^\alpha\Gamma(\alpha)}.$$

The marginal distribution is

$$\begin{aligned}
f_{\mathbf{X}}(\mathbf{x})=&\int_0^\infty \frac{\theta^{n+\alpha-1}e^{-\theta\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)}}{\beta^\alpha\Gamma(\alpha)}d\theta\\
=&\frac{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-(n+\alpha)}}{\beta^\alpha}\frac{\Gamma(n+\alpha)}{\Gamma(\alpha)}\\
&\times\int_0^\infty \frac{\theta^{n+\alpha-1}e^{-\frac{\theta}{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-1}}}}{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-(n+\alpha)}\Gamma(\alpha+n)}d\theta\\
=&\frac{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-(n+\alpha)}}{\beta^\alpha}\frac{\Gamma(n+\alpha)}{\Gamma(\alpha)}.
\end{aligned}$$

The posterior distribution is

$$\pi_{\Theta|\mathbf{X}}(\theta,\mathbf{x})=\frac{\theta^{n+\alpha-1}e^{-\frac{\theta}{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-1}}}}{\left(\sum_{i=1}^n x_i+\frac{1}{\beta}\right)^{-(n+\alpha)}\Gamma(\alpha+n)}.$$

Hence, $\Theta|\mathbf{X}\sim\Gamma(\alpha+n,\left(\frac{1}{\beta}+\sum_{i=1}^n x_i\right)^{-1})$ ■

Example 66.2
You are given:
- The prior distribution Λ is Gamma with parameters α and θ.
- The model distribution $\mathbf{X}|\Lambda$ is Poisson with parameter λ.

Show that Λ has a conjugate pair distribution.

Solution.
The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda)=\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

The joint distribution is

$$f_{\mathbf{X},\Lambda}(\mathbf{x},\lambda)=\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\theta}}}{\theta^\alpha\Gamma(\alpha)}.$$

The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x})=\int_0^\infty\frac{\lambda^{\sum_{i=1}^n x_i+\alpha-1}e^{-\lambda\left(n+\frac{1}{\theta}\right)}}{\theta^\alpha\Gamma(\alpha)\prod_{i=1}^n x_i!}d\lambda=\left(\frac{\theta}{n\theta+1}\right)^{\sum_{i=1}^n x_i+\alpha}\frac{\Gamma(\sum_{i=1}^n x_i+\alpha)}{\prod_{i=1}^n x_i!\theta^\alpha\Gamma(\alpha)}.$$

The posterior distribution is

$$\pi_{\Lambda|\mathbf{X}}(\lambda,\mathbf{x})=\frac{\lambda^{\sum_{i=1}^n x_i+\alpha-1}e^{-\frac{\lambda(n\theta+1)}{\theta}}}{\left(\frac{\theta}{n\theta+1}\right)^{\sum_{i=1}^n x_i+\alpha}\Gamma(\sum_{i=1}^n x_i+\alpha)}.$$

Hence, $\Lambda|\mathbf{X}\sim Gamma\left(\sum_{i=1}^n x_i+\alpha,\frac{\theta}{n\theta+1}\right)$ ■

Example 66.3
You are given:
- The prior distribution Λ is normal with mean μ and variance a^2.
- The model distribution $\mathbf{X}|\Lambda$ is normal with mean λ and variance σ^2.

Show that Λ has a conjugate pair distribution.

Solution.
The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda) = \frac{1}{\sigma^n\sqrt{(2\pi)^n}} e^{-\frac{\sum_{i=1}^n (x_i-\lambda)^2}{2\sigma^2}}.$$

The joint distribution is

$$f_{\mathbf{X},\Lambda}(\mathbf{x},\lambda) = \frac{1}{\sigma^n\sqrt{(2\pi)^n}} e^{-\frac{\sum_{i=1}^n (x_i-\lambda)^2}{2\sigma^2}} \frac{1}{a\sqrt{2\pi}} e^{-\frac{(\lambda-\mu)^2}{2a^2}}.$$

The joint distribution can be simplified as follows:

$$\begin{aligned}
\frac{\sum_{i=1}^n (x_i-\lambda)^2}{\sigma^2} + \frac{(\lambda-\mu)^2}{a^2} =& \frac{\sum x_i^2 - 2\lambda\sum x_i + n\lambda^2}{\sigma^2} + \frac{\lambda^2 - 2\lambda\mu + \mu^2}{a^2} \\
=& \left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)\lambda^2 - 2\lambda\left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right) + \frac{\sum x_i^2}{\sigma^2} + \frac{\mu^2}{a^2} \\
=& \left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)\left[\lambda^2 - 2\lambda\left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}\right. \\
&\left. + \left(\frac{\sum x_i^2}{\sigma^2} + \frac{\mu^2}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}\right] \\
=& \frac{e^{\frac{1}{2}\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}\left(\frac{\sum x_i}{\sigma^2}+\frac{\mu}{a^2}\right)^2 - \frac{\sum x_i^2}{2\sigma^2} - \frac{\mu^2}{2a^2}}}{\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{1/2} a\sigma\sqrt{(2\pi)^n}} \frac{e^{-\frac{\left(\lambda - \left(\frac{\sum x_i}{\sigma^2}+\frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}\right)^2}{2\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}}}}{\sqrt{2\pi}\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1/2}}
\end{aligned}$$

The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X},\Lambda}(\mathbf{x},\lambda)d\lambda = \frac{e^{\frac{1}{2}\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}\left(\frac{\sum x_i}{\sigma^2}+\frac{\mu}{a^2}\right)^2 - \frac{\sum x_i^2}{2\sigma^2} - \frac{\mu^2}{2a^2}}}{\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{1/2} a\sigma\sqrt{(2\pi)^n}}.$$

The posterior distribution is

$$\pi_{\Lambda|\mathbf{X}}(\lambda,\mathbf{x}) = \frac{e^{-\frac{\left(\lambda - \left(\frac{\sum x_i}{\sigma^2}+\frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}\right)^2}{2\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1}}}}{\sqrt{2\pi}\left(\frac{n}{\sigma^2}+\frac{1}{a^2}\right)^{-1/2}}.$$

Thus, $\Lambda|\mathbf{X}$ has a normal distribution with mean $\left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}$ and variance $\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}$ ■

Practice Problems

Problem 66.1
You are given:
• The prior distribution Q is beta with parameters a, b and 1.
• The model distribution $\mathbf{X}|Q$ is binomial with parameters m and q.

Show that Q has a conjugate pair distribution.

Problem 66.2
You are given:
• The prior distribution Λ is Gamma with parameters α and θ.
• The model distribution $\mathbf{X}|\Lambda$ is inverse exponential with parameters λ.

Show that Λ has a conjugate pair distribution.

Problem 66.3
You are given:
• The prior distribution Λ is inverse Gamma with parameters α and θ.
• The model distribution $\mathbf{X}|\Lambda$ is exponential with mean λ.

Show that Λ has a conjugate pair distribution.

Problem 66.4
You are given:
• The prior distribution Λ is a single parameter Pareto with parameter α and with pdf $f(\lambda) = \frac{\alpha\theta^\alpha}{\lambda^{\alpha+1}}, \lambda > \theta$.
• The model distribution $\mathbf{X}|\Lambda$ is uniform in $[0, \lambda]$.

Show that Λ has a conjugate pair distribution.

67 Estimation of Class $(a, b, 0)$

In this section we use the methods of moments and the method of maximum likely estimate for the estimation of parmeters in $(a, b, 0)$ class.

Let N be a random variable in the class of $(a, b, 0)$ with probability function $p_k = \Pr(N = k),\ k = 0, 1, 2, \cdots$. Let n_k be the number of observations for which $N = k$ and n be the total number of observations or the sample size. Note that $n = \sum_{k=0}^{\infty} n_k$.

Example 67.1
Let N be a Poisson random variable with parameter λ. Then $E(N) = \lambda$ and $\text{Var}(N) = \lambda$.
(a) Estimate the Poisson parameter using the method of moments.
(b) Estimate the Poisson parameter using the method of maximum likelihood.
(c) Calculate $E(\hat{\lambda})$ and $\text{Var}(\hat{\lambda})$.
(d) Find the asymptotic variance of $\hat{\lambda}$.
(e) Construct a 95% confidence interval of the true value of λ.

Solution.
(a) The Poisson distribution parameter estimate by the method of moments is

$$\hat{\lambda} = \overline{x} = \frac{\sum_{k=1}^{\infty} kn_k}{\sum_{k=0}^{\infty} n_k} = \frac{\sum_{k=1}^{\infty} kn_k}{n}.$$

(b) The likelihood function is

$$L(\lambda) = \prod_{k=0}^{\infty} \left(\frac{e^{-\lambda}\lambda^k}{k!} \right)^{n_k}.$$

The loglikelihood function is

$$\ell(\lambda) = \sum_{k=0}^{\infty} n_k \ln \left(\frac{e^{-\lambda}\lambda^k}{k!} \right) = -\lambda \sum_{k=0}^{\infty} n_k + \sum_{k=1}^{\infty} kn_k \ln \lambda - \sum_{k=0}^{\infty} n_k \ln (k!).$$

The MLE of λ is found from

$$\frac{d}{d\lambda}\ell(\lambda) = -n + \frac{1}{\lambda}\sum_{k=1}^{\infty} kn_k = 0 \Longrightarrow \hat{\lambda} = \frac{\sum_{k=1}^{\infty} kn_k}{n}.$$

(c) We have

$$\begin{aligned} E(\hat{\lambda}) =& E(N) = \lambda \\ \operatorname{Var}(\hat{\lambda}) =& \frac{\operatorname{Var}(N)}{n} = \frac{\lambda}{n}, \end{aligned}$$

Thus, $\hat{\lambda}$ is unbiased and consistent.
(d) The asymptotic variance is found as follows:

$$\begin{aligned} I(\lambda) =& nI(N|\lambda) = -nE\left[\frac{\partial^2}{\partial\lambda^2}\ln\left(\frac{e^{-\lambda}\lambda^N}{N!}\right)\right] \\ =& nE\left(\frac{N}{\lambda^2}\right) = \frac{n}{\lambda} \\ \operatorname{Var}(\hat{\lambda}) =& \frac{\lambda}{n}. \end{aligned}$$

(e) The confidence interval is $(\hat{\lambda} - 1.96\sqrt{\operatorname{Var}(\hat{\lambda})}, \hat{\lambda} + 1.96\sqrt{\operatorname{Var}(\hat{\lambda})})$ ■

The following example describes the process of finding the likelihood function in the case of incomplete data.

Example 67.2
The distribution of accidents for 70 randomly selected policies is as follows:

Number of Accidents	Number of Policies
0	31
1	20
2	12
3^+	7
Total	70

Let N denote the number of accidents. Suppose $N \sim \text{Poisson}(\lambda)$. Find the likelihood function.

Solution.
The likelihood function is given by

$$L(\lambda) = p_0^{31} p_1^{20} p_2^{12} (1 - p_0 - p_1 - p_2)^7$$

where

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!} \blacksquare$$

Example 67.3
Estimate the negative binomial parameters by the method of moments.

Solution.
By the method of moments, we have the following system of two equations

$$r\beta = \frac{\sum_{k=1}^{\infty} kn_k}{n} = \overline{x}$$

and

$$r\beta(1+\beta) = \frac{\sum_{k=1}^{\infty} k^2 n_k}{n} - \left(\frac{\sum_{k=1}^{\infty} kn_k}{n}\right)^2 = s^2.$$

Solving this system for r and β we find

$$\hat{r} = \frac{\overline{x}}{\hat{\beta}} \text{ and } \hat{\beta} = \frac{s^2}{\overline{x}} - 1.$$

Note that, if $s^2 < \overline{x}$ then $\hat{\beta} < 0$ which is an indication that the negative binomial model is not a good representation for the data $\blacksquare$

Example 67.4
Estimate the negative binomial parameters by the method of maximum likelihood.

Solution.
Let

$$p_k = \left(\begin{array}{c} k+r-1 \\ k \end{array}\right) \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^k, \; k = 0, 1, 2, \cdots.$$

The likelihood function is

$$L(r, \beta) = \prod_k p_k$$

and the loglikelihood function is

$$\ell(r,\beta)=\sum_{k=0}^{\infty}n_k\left[\ln\begin{pmatrix} k+r-1\\ k\end{pmatrix}-r\ln(1+\beta)+k\ln\beta-k\ln(1+\beta)\right].$$

Taking first derivatives with respect to both r and β, we find

$$\begin{aligned}
\frac{\partial\ell}{\partial\beta}=&\sum_{k=0}^{\infty}n_k\left(\frac{k}{\beta}-\frac{r+k}{1+\beta}\right)\\
\frac{\partial\ell}{\partial r}=&-\sum_{k=0}^{\infty}n_k\ln(1+\beta)+\sum_{k=0}^{\infty}n_k\frac{\partial}{\partial r}\ln\frac{(r+k-1)\cdots r}{k!}\\
=&-n\ln(1+\beta)+\sum_{k=1}^{\infty}n_k\frac{\partial}{\partial r}\ln\prod_{m=0}^{k-1}(r+m)\\
=&-n\ln(1+\beta)+\sum_{k=1}^{\infty}n_k\frac{\partial}{\partial r}\sum_{m=0}^{k-1}\ln(r+m)\\
=&-n\ln(1+\beta)+\sum_{k=1}^{\infty}n_k\sum_{m=0}^{k-1}(r+m)^{-1}.
\end{aligned}$$

Setting $\ell_\beta(r,\beta)=0$, we find

$$\hat{r}\hat{\beta}=\frac{\sum_{k=0}^{\infty}kn_k}{n}=\overline{x}$$

and setting $\ell_r(r,\beta)=0$ we find

$$n\ln(1+\hat{\beta})=\sum_{k=1}^{\infty}n_k\sum_{m=0}^{k-1}(\hat{r}+m)^{-1}=n\ln\left(1+\frac{\overline{x}}{\hat{r}}\right).$$

The above equations are usually solved using numerical methods such as the Newton-Raphson method. Note that if r is given then $\hat{\beta}=\frac{\overline{x}}{r}$ ■

Example 67.5

Let N be a binomial random variable with parameters m and q. Then $E(N)=mq$ and $\mathrm{Var}(N)=mq(1-q)$.

(a) Estimate q using the method of moments, assuming that m is known.

(b) Estimate q using the method of moments, assuming that m is unknown.

Solution.
(a) We have

$$m\hat{q} = \frac{\sum_{k=1}^{\infty} kn_k}{n} \Longrightarrow \hat{q} = \frac{\sum_{k=1}^{\infty} kn_k}{nm}.$$

(b) We have to solve the system of two equations

$$\hat{m}\hat{q} = \frac{\sum_{k=1}^{\infty} kn_k}{n} \quad \text{and} \quad \hat{m}\hat{q}(1-\hat{q}) = \frac{\sum_{k=1}^{\infty} k^2 n_k}{n} - \left(\frac{\sum_{k=1}^{\infty} kn_k}{n}\right)^2.$$

We obtain

$$\hat{q} = 1 - \frac{\sum_{k=1}^{\infty} k^2 n_k}{\sum_{k=1}^{\infty} kn_k} + \frac{\sum_{k=1}^{\infty} kn_k}{n} \quad \text{and} \quad \hat{m} = \frac{\sum_{k=1}^{\infty} kn_k}{n\hat{q}}$$

Example 67.6
Let N be a binomial random variable with parameters m and q. Then $E(N) = mq$ and $\text{Var}(N) = mq(1-q)$.
Estimate q using the method of maximum likelihood, assuming that m is known.

Solution.
Let

$$p_k = \Pr(N = k) = \binom{m}{k} q^k (1-q)^{m-k}.$$

The likelihood function is

$$L(m, q) = \prod_{k=0}^{m} p_k^{n_k}.$$

The loglikelihood function is

$$\begin{aligned}\ell(m, q) &= \sum_{k=0}^{m} n_k \ln p_k \\ &= \sum_{k=0}^{m} n_k \left[\ln \binom{m}{k} + k \ln q + (m-k) \ln(1-q)\right].\end{aligned}$$

Since m known, we just need to maximize $\ell(m,q)$ with respect to q :

$$\frac{\partial \ell}{\partial q} = \frac{1}{q}\sum_{k=1}^{m} kn_k - \frac{1}{1-q}\sum_{k=0}^{m}(m-k)n_k.$$

Setting this expression to 0, we obtain

$$\hat{q} = \frac{\sum_{k=1}^{\infty} kn_k}{nm} \blacksquare$$

Remark 67.1
If m and q are both unknown, the estimation analysis is complex and hence is omitted.

Example 67.7 ‡
You are given:
(i) A hospital liability policy has experienced the following numbers of claims over a 10-year period:

$$10 \quad 2 \quad 4 \quad 0 \quad 6 \quad 2 \quad 4 \quad 5 \quad 4 \quad 2$$

(ii) Numbers of claims are independent from year to year.
(iii) You use the method of maximum likelihood to fit a Poisson model.
Determine the estimated coefficient of variation of the estimator of the Poisson parameter.

Solution.
Recall that

$$\hat{\lambda} = \overline{X} = \frac{10+2+4+0+6+2+4+5+4+2}{10} = 3.9.$$

We have

$$\begin{aligned} E(\hat{\lambda}) =& E(\overline{X}) = \lambda \\ \text{Var}(\hat{\lambda}) =& \frac{\lambda}{n} \\ \text{CV} =& \frac{\sqrt{\frac{\lambda}{n}}}{\lambda} = \frac{1}{\sqrt{\lambda n}} \\ =& \frac{1}{\sqrt{39}} = 0.16 \blacksquare \end{aligned}$$

Practice Problems

Problem 67.1
The distribution of accidents for 84 randomly selected policies is as follows:

Number of Accidents	Number of Policies
0	32
1	26
2	12
3	7
4	4
5	2
6	1
Total	84

(a) Estimate the Poisson parameter using the method of moments.
(b) Estimate the Poisson parameter using the method of maximum likelihood.
(c) Let N denote the number of accidents. Suppose $N \sim \text{Poisson}(\lambda)$. Calculate $E(\hat{\lambda})$ and $\text{Var}(\hat{\lambda})$.
(d) Find the asymptotic variance of $\hat{\lambda}$.
(e) Construct a 95% confidence interval of the true value of λ.

Problem 67.2 ‡
You are given the following observed claim frequency data collected over a period of 365 days:

# of claims/day	observed number of days
0	50
1	122
2	101
3	92
4^+	0

Fit a Poisson distribution to the above data, using the method of maximum likelihood.

Problem 67.3
You are given the following observed claim frequency data collected over a period of 365 days:

# of claims/day	observed number of days
0	50
1	122
2	101
3^+	92

Find the likelihood function of λ.

Problem 67.4
You are given:
(i)

k	0	1	2	3	4
n_k	30	35	20	10	5

(ii) The model is negative binomial.

Find the estimates of r and β using the method of moments.

Problem 67.5
You are given:
(i)

k	0	1	2	3	4
n_k	30	35	20	10	5

(ii) The model is negative binomial.
Find the equation $H(\hat{r}) = 0$, where $\hat{r} = \text{MLE}(r)$.

Problem 67.6 ‡
The number of claims follows a negative binomial distribution with parameters β and r, where β is unknown and r is known. You wish to estimate β based on n observations, where $\overline{x}$ is the mean of these observations.

Determine the maximum likelihood estimate of β.

Problem 67.7 ‡
The distribution of accidents for 100 randomly selected policies is as follows:

Number of Accidents	Number of Policies
0	6
1	17
2	26
3	21
4	18
5	7
6	5
Total	100

Estimate the binomial distribution parameters m and q using the method of moments.

Problem 67.8 ‡
You are given the following data for the number of claims during a one-year period:

Number of Claims	Number of Policies
0	157
1	66
2	19
3	4
4	2
5^+	0
Total	248

A geometric distribution is fitted to the data using maximum likelihood estimation.
Let P = probability of zero claims using the fitted geometric model.
A Poisson distribution is fitted to the data using the method of moments.
Let Q = probability of zero claims using the fitted Poisson model.

Calculate $|P - Q|$.

68 MLE with $(a, b, 1)$ Class

In this section[15], we estimate the parameters of the $(a, b, 1)$ class. The estimation procedure is similar to the one used for the $(a, b, 0)$ class.

Let N be a zero-modified random variable with probability function $p_k = \Pr(N = k),\ k = 0, 1, \cdots$. Let n_k be the number of observations for which $N = k$ and n be the total number of observations or the sample size. Recall that the probabilities of N are given by

$$p_k^M = \frac{1 - p_0^M}{1 - p_0} p_k = (1 - p_0^M) p_k^T,\ k = 1, 2, 3, \cdots$$

where

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1},\ k = 2, 3, \cdots.$$

and $p_0^M = \alpha,\ 0 \le \alpha < 1$. The parameters to be estimated are a, b, and p_0^M.

The likelihood function is

$$L(a, b, p_0^M) = (p_0^M)^{n_0} \prod_{k=1}^{\infty} (p_k^M)^{n_k} = (p_0^M)^{n_0} \prod_{k=1}^{\infty} [(1 - p_0^M) p_k^T]^{n_k}.$$

The loglikelihood function is

$$\begin{aligned}\ell(a, b, p_0^M) =& n_0 \ln p_0^M + \sum_{k=1}^{\infty} n_k [\ln(1 - p_0^M) + \ln p_k^T] \\ =& n_0 \ln p_0^M + \sum_{k=1}^{\infty} n_k \ln(1 - p_0^M) + \sum_{k=1}^{\infty} n_k [\ln p_k - \ln(1 - p_0)] = \quad \ell_0 + \ell_1\end{aligned}$$

where

$$\begin{aligned}\ell_0 =& n_0 \ln p_0^M + \sum_{k=1}^{\infty} n_k \ln(1 - p_0^M) \\ \ell_1 =& \sum_{k=1}^{\infty} n_k [\ln p_k - \ln(1 - p_0)].\end{aligned}$$

[15]This section has not appeared in any of the C exams.

The MLE of p_0^M is found by setting the first partial derivative of ℓ with respect to p_0^M to zero:

$$\frac{\partial \ell}{\partial p_0^M} = \frac{\partial \ell_0}{\partial p_0^M} = \frac{n_0}{p_0^M} - \sum_{k=1}^{\infty} \frac{n_k}{1 - p_0^M} = \frac{n_0}{p_0^M} - \frac{n - n_0}{1 - p_0^M} = 0$$

resulting in

$$\hat{p_0}^M = \frac{n_0}{n},$$

the proportion of observations that equal 0. This result is true for any zero-modified distribution.

Example 68.1
Let p_0^M be the zero-modified geometric probability function. Find the MLE for p_0^M and β.

Solution.
Recall that

$$p_k = \frac{\beta^k}{(1+\beta)^{k+1}}, \; p_0 = \frac{1}{1+\beta}.$$

The MLE of p_0^M is

$$\hat{p}_0^M = \frac{n_0}{n}.$$

For finding the MLE of β, we first find ℓ_1 :

$$\begin{aligned}
\ell_1 =& \sum_{k=0}^{\infty} n_k [\ln p_k - \ln(1 - p_0)] \\
=& \sum_{k=0}^{\infty} n_k [k \ln \beta - (k+1) \ln(1+\beta) - \ln \beta - \ln(1+\beta)] \\
=& \ln \beta \sum_{k=0}^{\infty} (kn_k - n_k) - \ln(1+\beta) \sum_{k=0}^{\infty} [(k+1)n_k - n_k] \\
\frac{\partial \ell_1}{\partial \beta} =& \frac{1}{\beta} \sum_{k=0}^{\infty} n_k (k-1) - \frac{1}{1+\beta} \sum_{k=0}^{\infty} kn_k.
\end{aligned}$$

Setting this last equation to zero, we find

$$\hat{\beta} = \frac{\sum_{k=0}^{\infty} kn_k}{n - n_0} - 1 = \frac{n\overline{x}}{n - n_0} - 1 \blacksquare$$

Example 68.2
Find ℓ_1 for the zero-modified Poisson distribution.

Solution.
We have

$$\begin{aligned}\ell_1 =&\sum_{k=1}^{\infty} n_k \left[\ln\left(\frac{e^{-\lambda}\lambda^k}{k!}\right) - \ln(1-e^{-\lambda})\right]\\ =& -(n-n_0)\lambda + \left(\sum_{k=1}^{\infty} kn_k\right)\ln\lambda - (n-n_0)\ln(1-e^{-\lambda}) + c\\ =& -(n-n_0)[\lambda + \ln(1-e^{-\lambda})] + n\overline{x}\ln\lambda + c\end{aligned}$$

where

$$c = -\sum_{k=1}^{\infty} n_k \ln k!$$

and

$$\overline{x} = \frac{\sum_{k=1}^{\infty} kn_k}{n} \blacksquare$$

To find the estimate of λ in Example 68.2, we set the first derivative of ℓ_1 with respect to λ to zero:

$$\frac{\partial \ell_1}{\partial\lambda} = -\frac{n-n_0}{1-e^{-\lambda}} + \frac{n\overline{x}}{\lambda} = 0$$

resulting in

$$\overline{x}(1-e^{-\lambda}) = \frac{n-n_0}{n}\lambda$$

which is solved numerically for λ. Note also that this last equation can be expressed as

$$\overline{x} = \frac{1-\hat{p}_0^M}{1-p_0}\lambda.$$

Example 68.3
Find ℓ_1 for the zero-modified binomial distribution.

Solution.
We have

$$\begin{aligned}
\ell_1 =& \sum_{k=1}^{m} n_k \left\{ \ln \left[\binom{m}{k} q^k (1-q)^{m-k} \right] - \ln\left[1-(1-q)^m\right] \right\} \\
=& \left(\sum_{k=1}^{m} k n_k \right) \ln q + \sum_{k=1}^{\infty} (m-k) n_k \ln(1-q) \\
&- \sum_{k=1}^{m} n_k \ln\left[1-(1-q)^m\right] + c \\
=& n\overline{x} \ln q + m(n-n_0) \ln(1-q) - n\overline{x} \ln(1-q) \\
&-(n-n_0) \ln\left[1-(1-q)^m\right] + c
\end{aligned}$$

where

$$c = \sum_{k=1}^{m} n_k \ln \binom{m}{k} \blacksquare$$

The equation

$$0 = \frac{\partial \ell}{\partial q} = \frac{n\overline{x}}{q} - \frac{m(n-n_0)}{1-q} + \frac{n\overline{x}}{1-q} - \frac{(n-n_0)m(1-q)^{m-1}}{1-1(1-q)^m}$$

results in

$$\overline{x} = \frac{1-\hat{p}_0^M}{1-p_0} mq$$

where $p_0 = (1-q)^m$. If m is known then the MLE of q is obtained by solving the above equation.

Practice Problems

Problem 68.1
You have the following observations of a discrete random variable.

Frequency (k)	n_k
0	9048
1	905
2	45
3	2
4^+	0

You are to fit these to a zero-modified geometric distribution using the maximum likelihood. Find the fitted values of p_0^M and β.

Problem 68.2
You have the following observations of a discrete random variable.

Frequency (k)	n_k
0	9048
1	905
2	45
3	2
4^+	0

You are to fit these to a zero-modified Poisson distribution using the maximum likelihood. Find the MLE for p_0^M and λ.

Problem 68.3
You have the following observations of a discrete random variable.

Frequency (k)	n_k
0	10
1	6
2	4

You are to fit these to a zero-truncated binomial distribution using the maximum likelihood. Find the MLE for p_0^M and q when $m = 3$.

Problem 68.4
You have the following observations of a discrete random variable.

Frequency (k)	n_k
0	10
1	6
2	4

You are to fit these to a zero-modified negative binomial distribution. Find ℓ_1 when $r = 2$ and $\beta = 3$.

Model Selection and Evaluation

The goal of this chapter is to evaluates models that fit the best to a given data set and compare competing models.

69 Assessing Fitted Models Graphically

Given a sample of size n, we let $F_n(x)$ and $f_n(x)$ denote the cumulative and the density functions of the empirical distribution. The distribution and the density functions of the estimated model to which we are trying to fit the data are denoted by $F^*(x)$ and $f^*(x)$. In this section, we assess how well an estimated model fit the original data graphically.

One method is to compare the plot of the empirical distribution to the distribution of the estimated model. A good fit is when the plot of $F^*(x)$ alternates above and below the histogram of $F_n(x)$.

One of the difficulties of the previous method is that the distinction is difficult to make when the heights between the two graphs is small. We consider two ways for magnifying small changes to better interpret the goodness of fit. The first method consists of graphing the difference function $D(x) = F_n(x) - F^*(x)$, known as the $D(x)-$plot. A fit is considered good, if the graph of $D(x)$ is close to the horizontal axis.

The second method is to create a **probability plot**, also known as **p-p plot**. The plot is created by ordering the observations in increasing order $x_1 \leq x_2 \leq \cdots \leq x_n$. For each $j = 1, 2, \cdots$, we plot the point $(F_n(x_j), F^*(x_j))$. According to [1], a model is a good fit if these points are close to the line $y = x$. This requires a new definition of $F_n(x_j)$. If F denotes the correct model and X_j is the $j-$th order statistic (i.e., the $j-$th smallest value) then it can be shown that $E[F(X_j)] = \frac{j}{n+1}$. Thus, it makes sense to define

$$F_n(x_j) = \frac{j}{n+1}.$$

Now, examining tails in the p-p plot: on the left-hand tail (near $y = 0$), if a point in the p-p plot is above the line $y = x$, then $F^*(x) > F_n(x)$ for all x to the left of that point. That is, the fitted distribution is thicker on the left than the empirical distribution. In probability terms, there is more probability on the fitted than the empirical. On the right-end tail (near $y = 1$), if a point in the p-p plot is above the line $y = x$, then $F^*(x) > F_n(x)$ for all x to the right of that point which implies $1 - F^*(x) < 1 - F_n(x)$ for all x to the right of that point. That is, to the right of that point there is less probability on the fitted than the empirical or in other words the fitted is thinner than the empirical. What about points not near end tails? Taking the $j-$th and the $(j + k)-$th points on the $p - p$ plot, the slope of the line crossing these points is

$$\frac{F(X_{j+k}) - F(X_j)}{k/n+1}$$

where F is the correct model, has an expected value of 1. So if this slope is greater than 1, the model puts more probability between those two order statistics. See Problem 69.2.

Example 69.1
You are given:
(i) The following observed data: $2, 3, 3, 3, 5, 8, 10, 13, 16$.
(ii) An exponential distribution is fit to the data using the maximum likelihood to estimate the mean of the exponential distribution.
(a) Plot $F_8(x)$ and $F^*(x)$ in the same window.
(b) Plot $D(x)$.
(c) Create a $p-p$ plot.

Solution.
The empirical mass function is

x	2	3	5	8	10	13	16
$p(x)$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

The empirical cumulative distribution is

$$F_9(x) = \begin{cases} 0, & x < 2 \\ \frac{1}{9}, & 2 \le x < 3 \\ \frac{4}{9}, & 3 \le x < 5 \\ \frac{5}{9}, & 5 \le x < 8 \\ \frac{2}{3}, & 8 \le x < 10 \\ \frac{7}{9}, & 10 \le x < 13 \\ \frac{8}{9}, & 13 \le x < 16 \\ 1, & x \ge 16. \end{cases}$$

By the maximum likelihood method, $\hat{\theta} = \overline{x} = 7$. The fitted distribution is given by

$$F^*(x) = 1 - e^{-\frac{x}{7}}.$$

(a) The plots of both $F_9(x)$ and $F^*(x)$ are shown in Figure 69.1.

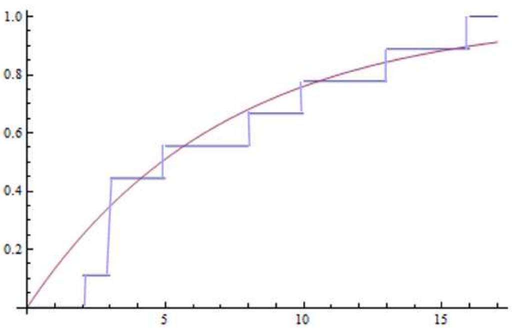


Figure 69.1

The plot indicates that the fitted model is a reasonable one.
(b) The difference function is

$$D(x) = F_9(x) - F^*(x) = \begin{cases} e^{-\frac{x}{7}} - 1, & x < 2 \\ e^{-\frac{x}{7}} - \frac{8}{9}, & 2 \le x < 3 \\ e^{-\frac{x}{7}} - \frac{5}{9}, & 3 \le x < 5 \\ e^{-\frac{x}{7}} - \frac{4}{9}, & 5 \le x < 8 \\ e^{-\frac{x}{7}} - \frac{1}{3}, & 8 \le x < 10 \\ e^{-\frac{x}{7}} - \frac{2}{9}, & 10 \le x < 13 \\ e^{-\frac{x}{7}} - \frac{1}{9}, & 13 \le x < 16 \\ e^{-\frac{x}{7}}, & x \ge 16. \end{cases}$$

The graph of $D(x)$ is shown in Figure 69.2.

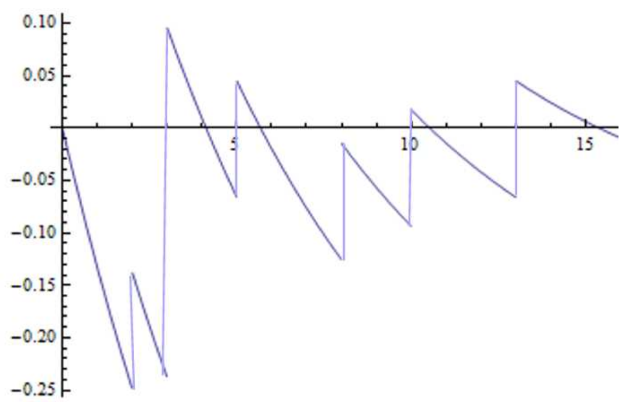


Figure 69.2

(c) We first create the points on the graph

j	x_j	$F_9(x_j) = \frac{j}{9}$	$F^*(x_j)$
1	2	0.1	0.249
2	3	0.2	0.349
3	3	0.3	0.349
4	3	0.4	0.349
5	5	0.5	0.510
6	8	0.6	0.681
7	10	0.7	0.760
8	13	0.8	0.844
9	15	0.9	0.883

The $p-p$ plot is shown in Figure 69.3 ■

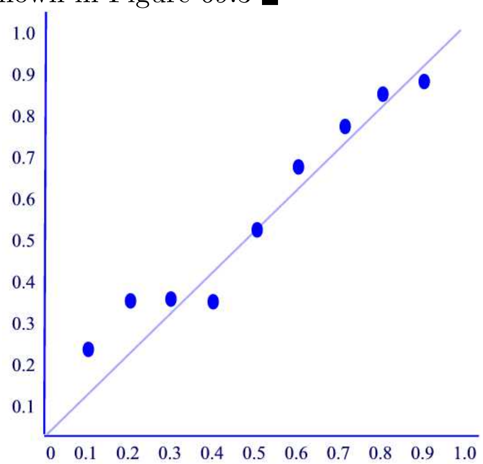


Figure 69.3

Example 69.2
You are given:
(i) The following are observed claim amounts:

$$400 \quad 1000 \quad 1600 \quad 3000 \quad 5000 \quad 5400 \quad 6200.$$

(ii) An exponential distribution with $\theta = 3300$ is hypothesized for the data.
(iii) The data are left-truncated at 500.
Write the formula for $F^*(x)$ and $f^*(x)$.

Solution.
The distribution function using a truncation of 500 is

$$F^*(x) = \begin{cases} 0, & x < 500 \\ \frac{F(x)-F(500)}{1-F(500)} = 1 - e^{-\frac{(x-500)}{3300}}, & x \geq 500 \end{cases}$$

and the pdf is

$$f^*(x) = \begin{cases} 0, & x < 500 \\ \frac{f(x)}{1-F(500)} = \frac{1}{3300}e^{-\frac{(x-500)}{3300}}, & x \geq 500 \ \blacksquare \end{cases}$$

Practice Problems

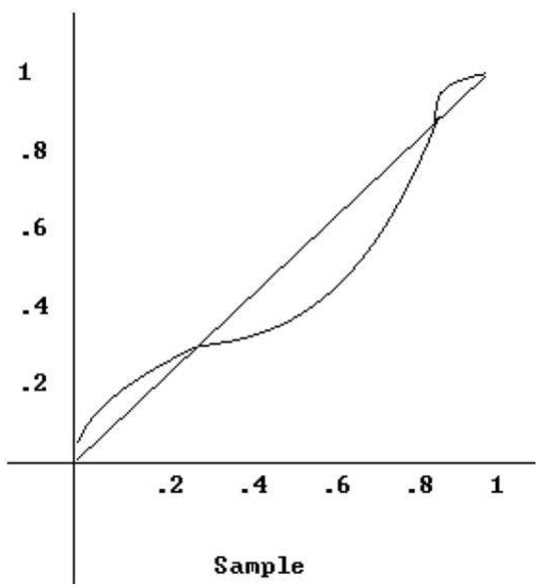
Problem 69.1 ‡
You are given:
(i) The following are observed claim amounts:

$$400 \quad 1000 \quad 1600 \quad 3000 \quad 5000 \quad 5400 \quad 6200.$$

(ii) An exponential distribution with $\theta = 3300$ is hypothesized for the data.
(iii) The goodness of fit is to be assessed by a $p-p$ plot and a $D(x)$ plot.

Let (s,t) be the coordinates of the $p-p$ plot for a claim amount of 3000. Determine $(s-t)-D(3000)$.

Problem 69.2 ‡
The graph below shows a $p-p$ plot of a fitted distribution compared to a sample.

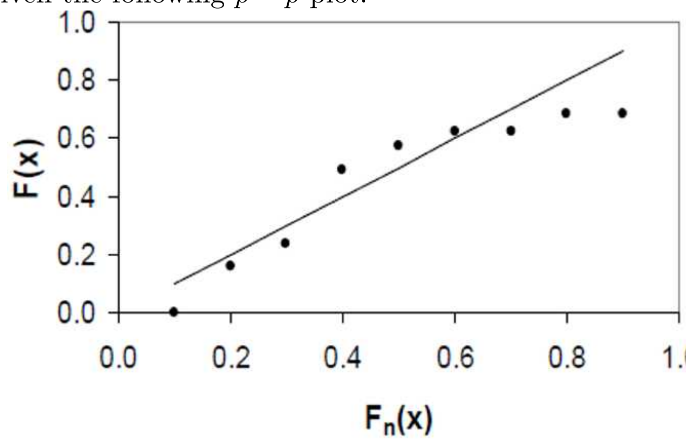


Which of the following is true?

(A) The tails of the fitted distribution are too thick on the left and on the right, and the fitted distribution has less probability around the median than the sample.
(B) The tails of the fitted distribution are too thick on the left and on the right, and the fitted distribution has more probability around the median

than the sample.
(C) The tails of the fitted distribution are too thin on the left and on the right, and the fitted distribution has less probability around the median than the sample.
(D) The tails of the fitted distribution are too thin on the left and on the right, and the fitted distribution has more probability around the median than the sample.
(E) The tail of the fitted distribution is too thick on the left, too thin on the right, and the fitted distribution has less probability around the median than the sample.

Problem 69.3 ‡
You are given the following $p-p$ plot:



The plot is based on the sample:

$$1 \quad 2 \quad 3 \quad 15 \quad 30 \quad 50 \quad 51 \quad 99 \quad 100$$

Which of the following is a possible fitted model underlying the $p-p$ plot?

(A) $F(x) = 1 - x^{-0.25}, x \geq 1$
(B) $F(x) = \frac{x}{1+x}, x > 0$
(C) Uniform on $[1, 100]$
(D) Exponential with mean 10 (E) Normal with mean 40 and standard deviation 40

70 Kolmogorov-Smirnov Hypothesis Test of Fitted Models

In this and the coming sections, we consider hypothesis tests of how well a fitted model fits the data. As such, consider the following hypothesis test problem:

H_0 : The data comes from the estimated model
H_a : The data does not come from the estimated model

The test statistic is usually a measure of how close the fitted distribution function is to the empirical distribution function. When the null hypothesis completely specifies the model (i.e. the parameters of the fitted distribution are given), critical values are well-known. In contrast, when the parameters of the distribution function in H_0 are to be estimated from the data, the test statistic tends to be smaller than it would be have been had the parameter values been prespecified. The estimation method tries to choose parameters that produce a distribution that is close to the data and this decreases the probability that the null hypothesis be rejected.

In this section, we look at the **Kolmogorov-Smirnov statistic**(KS) for testing whether an empirical distribution fits a hypothesized distribution well.

Consider a random sample of size n with order statistics $x_1 \leq x_2 \leq \cdots \leq x_n$ so that x_1 is the smallest observation and x_n is the largest observation. Let $F_n(x)$ denote the empirical distribution and $F^*(x)$ be the hypothesized distribution. The **Kolmogorov-Smirnov statistic** is defined by

$$D = \max_{x_1 \leq x \leq x_n} |F_n(x) - F^*(x)|.$$

But $F_n(x)$ is right-continuous and increasing step function and $F^*(x)$ is continuous and increasing so we only need to compare the differences at the observed data points. Moreover, at each order statistic x_i, one compares the differences $|F_n(x_{i-1}) - F^*(x_i)|$ and $|F_n(x_{i+1}) - F^*(x_i)|$. Hence, D can be written as

$$D = \max_{i=1,2,\cdots,n} \{|F_n(x_{i-1}) - F^*(x_i)|, |F_n(x_{i+1}) - F^*(x_i)|\}$$

where $F_n(x_0) = 0$.

Example 70.1
You are given:
(i) The following are observed claim amounts:

200 400 1000 1600 3000 5000 5400 6200

(ii) An exponential distribution with $\theta = 3300$ is hypothesized for the data. Find the value of D.

Solution.
We create the following chart.

x_j	$F_8(x_{j-1})$	$F_8(x_j)$	$F^*(x_j)$	Maximum of difference
200	0	0.125	0.05881	0.06619
400	0.125	0.25	0.1142	0.1358
1000	0.25	0.375	0.2614	0.1136
1600	0.375	0.5	0.3842	0.1158
3000	0.5	0.625	0.5971	0.0971
5000	0.625	0.75	0.7802	0.177
5400	0.75	0.875	0.8053	0.1552
6200	0.875	1.0	0.8472	0.1528

where $F^*(x) = 1 - e^{-\frac{x}{3300}}$. Hence, $D = 0.177$ ■

When $F^*(x)$ is completely specified with no unknown parameters, the critical values of D for some selected values of α are given as follows

Level of confidence α	0.10	0.05	0.01
Critical value	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

A value of D greater than the critical value will result in rejection of the null hypothesis.

Example 70.2
Determine whether the testing in Example 70.1 will result in rejection of the null hypothesis for a 10% level of confidence.

Solution.
For $\alpha = 0.10$, the critical value is $\frac{1.22}{\sqrt{8}} = 0.431$ which is smaller than D. Hence, the null hypothesis is rejected ■

Example 70.3
Suppose the data in Example 70.1 is right censored at 5100 and the estimated exponential mean is $\hat{\theta} = 3100$. Find the KS statistic.

Solution.
With censoring, we have $F^*(x) = 1 - e^{-\frac{x}{3000}}$ for $x \leq 5100$ and $F^*(x) = 1$ for $x > 5100$. Hence,

x_j	$F_8(x_{j-1})$	$F_8(x_j)$	$F^*(x_j)$	Maximum of difference
200	0	0.125	0.0623	0.0623
400	0.125	0.25	0.1211	0.0039
1000	0.25	0.375	0.2757	0.0257
1600	0.375	0.5	0.4032	0.0282
3000	0.5	0.625	0.6201	0.1201
5000	0.625	0.75	0.8007	0.1757
5100	0.75	0.75	0.8070	0.057

Hence, $D = 0.1757$ ■

Remark 70.1
According to [1], if the data is right-censored then the critical values should be smaller because there is less opportunity for the difference to become large.

Practice Problems

Problem 70.1 ‡
You are given:
(i) A sample of claim payments is:

$$29 \quad 64 \quad 90 \quad 135 \quad 182$$

(ii) Claim sizes are assumed to follow an exponential distribution.
(iii) The mean of the exponential distribution is estimated using the method of moments.

Calculate the value of the Kolmogorov-Smirnov test statistic.

Problem 70.2 ‡
You are given a random sample of observations:

$$0.1 \quad 0.2 \quad 0.5 \quad 0.7 \quad 1.3$$

You test the hypothesis that the probability density function is:

$$f(x) = \frac{4}{(1+x)^5}, \ x > 0.$$

Determine the KS test statistic.

Problem 70.3 ‡
You are given:
(i) A random sample of five observations:

$$0.2 \quad 0.7 \quad 0.9 \quad 1.1 \quad 1.3$$

(ii) You use the Kolmogorov-Smirnov test for testing the null hypothesis, H_0, that the probability density function for the population is:

$$f(x) = \frac{4}{(1+x)^5}, \ x > 0.$$

(iii) Critical values for the Kolmogorov-Smirnov test are:

Level of comfidence	0.10	0.05	0.025	0.01
Critical value	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.48}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Determine the result of the test.

Problem 70.4 ‡
The size of a claim for an individual insured follows an inverse exponential distribution with the following probability density function:

$$f(x|\theta) = \frac{\theta e^{-\frac{\theta}{x}}}{x^2},\ x > 0.$$

For a particular insured, the following five claims are observed:

$$1 \quad 2 \quad 3 \quad 5 \quad 13$$

Determine the value of the Kolmogorov-Smirnov statistic to test the goodness of fit of $f(x|\theta = 2)$.

Problem 70.5
You are given:
(i) The following are observed claim amounts:

$$200 \quad 400 \quad 1000 \quad 1600 \quad 3000 \quad 5000 \quad 5400 \quad 6200.$$

(ii) An exponential distribution with $\theta = 3300$ is hypothesized for the data.
(iii) The data are left-truncated at 500.
You conduct a Kolmogorov-Smirnov test at the 0.10 significance level. Do you reject or fail to reject the null hypothesis that the data came from a population having the left-truncated exponential distribution above? Perform the test and justify your answer.

Problem 70.6
You use the Kolmogorov-Smirnov goodness-of-fit test to assess the fit of the natural logarithms of $n = 200$ losses to a distribution with distribution function F^*.
You are given:
(i) The largest value of $|F^*(x) - F_n(x)|$ occurs for some x between 4.26 and 4.42.
(ii)

x	$F^*(x)$	$F_n(x-)$	$F_n(x)$
4.26	0.584	0.505	0.510
4.30	0.599	0.510	0.515
4.35	0.613	0.515	0.520
4.36	0.621	0.520	0.525
4.39	0.636	0.525	0.530
4.42	0.638	0.530	0.535

Commonly used large-sample critical values for this test are

Level of comfidence	0.10	0.05	0.025	0.01
Critical value	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.48}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Determine the result of the test.

71 Anderson-Darling Hypothesis Test of Fitted Models

The **Anderson-Darling test** is another test of the null hypothesis of the previous section. Like the Kolmogorov-Smirnov test, the test does not work for grouped data but rather on a complete individual data.

Consider a random sample $\{x_1, x_2, \cdots, x_n\}$. Some of the x_i's may be repeated. Suppose there are k distinct values which we arrange in increasing order

$$y_0 < y_1 < y_2 < \cdots < y_k < y_{k+1}$$

where $y_0 = 0$ if there is no truncation and $y_{k+1} = \infty$ if there is no censoring. If there is a left-truncation of d, we set $y_0 = d$. If there is right-censoring of u, we set $y_{k+1} = u$.

The Anderson-Darling test statistic, denoted by A^2, is defined by

$$A^2 = -nF^*(u) + n\left[\sum_{j=0}^{k}[1 - F_n(y_j)]^2 \ln\left(\frac{1 - F^*(y_j)}{1 - F^*(y_{j+1})}\right) + \sum_{j=1}^{k} F_n(y_j)^2 \ln\left(\frac{F^*(y_{j+1})}{F^*(y_j)}\right)\right].$$

When $F^*(x)$ is completely specified with no unknown parameters (see Remark below), the critical values of A^2 for some selected values of α are given as follows

Level of confidence α	0.10	0.05	0.01
Critical value	1.933	2.492	3.857

A value of A^2 greater than the critical value will result in rejection of the null hypothesis.

Example 71.1
You are given:
(i) The following are observed claim amounts:

200	400	1000	1600	3000	5000	5400	6200

(ii) An exponential distribution with $\theta = 3300$ is hypothesized for the data.
(a) Find the value of A^2.
(b) Determine the result of the test at the 10% level of confidence.

Solution.
We have $d = 0$ and $u = \infty$. Also, $F^*(x) = 1 - e^{-\frac{x}{3300}}$. Let

$$a = [1 - F_n(y_j)]^2 \ln\left(\frac{1-F^*(y_j)}{1-F^*(y_{j+1})}\right) \text{ and } b = F_n(y_j)^2 \ln\left(\frac{F^*(y_{j+1})}{F^*(y_j)}\right).$$

We create the following chart.

y_j	$F_8(y_j)$	$F^*(y_j)$	a	b
0	0	0	0.0606	0
200	0.125	0.0588	0.0464	0.0104
400	0.25	0.1142	0.1022	0.0517
1000	0.375	0.2614	0.0710	0.0541
1600	0.5	0.3842	0.1061	0.1102
3000	0.625	0.5971	0.0852	0.1045
5000	0.75	0.7802	0.0076	0.0178
5400	0.875	0.8053	0.0038	0.0388
6200	1.0	0.8472	0	0.1658
Total	–	–	0.4829	0.5533

Hence,

$$A^2 = -8 + 8(0.4829 + 0.5533) = 0.2896.$$

(b) At the 10% level of confidence, the critical value is $1.933 > A^2$ so the we fail to reject the null hypothesis ■

Example 71.2
You are given:
(i) Four policy claims:

$$300 \quad 300 \quad 800 \quad 1500$$

(ii) Policy limit $u = 2000$.
(iii) The claims are hypothezised by uniform distribution on $[0, 2500]$.
Calculate the Anderson-Darling test statistic.

Solution.
We have: $d = 0, u = 2000$, and $F^*(x) = \frac{x}{2500}$. We create the following chart.

y_j	$F_8(y_j)$	$F^*(y_j)$	a	b
0	0	0	0.1278	0
300	0.5	0.12	0.0644	0.2452
800	0.75	0.32	0.0332	0.3536
1500	1.0	0.6	0	0.2877
Total	–	–	0.2254	0.8865

Hence,

$$A^2 = -4(0.8) + 4(0.2254 + 0.8865) = 1.2476 \blacksquare$$

Remark 71.1

Keep in mind that the critical values for Kolomogorv-Smirnov test and the Anderson-Darling test are correct ONLY when the null hypothesis completely specifies the model[16].

[16]See [1], page 450.

Practice Problems

Problem 71.1
Batonic Inc. has workers' compensation claims during a month of:

$$100 \quad 350 \quad 550 \quad 1000$$

An actuary who works for the company believes that the claims are distributed exponentially with mean 500.

Determine the Anderson-Darling test statistic.

Problem 71.2
Using the previous problem, complete the following chart.

Level of confidence α	0.10	0.05	0.01
Critical value	1.933	2.492	3.857
Test Result			

Problem 71.3 ‡
Which of the following statements is true?

(A) For a null hypothesis that the population follows a particular distribution, using sample data to estimate the parameters of the distribution tends to decrease the probability of a Type II error.
(B) The Kolmogorov-Smirnov test can be used on individual or grouped data.
(C) The Anderson-Darling test tends to place more emphasis on a good fit in the middle rather than in the tails of the distribution.
(D) None of the above.

Problem 71.4 ‡
Which of the following is false?

(A) For the Kolmogorov-Smirnov test, when the parameters of the distribution in the null hypothesis are estimated from the data, the probability of rejecting the null hypothesis decreases.
(B) For the Kolmogorov-Smirnov test, the critical value for right censored data should be smaller than the critical value for uncensored data.
(C) The Anderson-Darling test does not work for grouped data.
(D) None of the above is true.

Problem 71.5
If the null hypothesis completely spacifies the fitted distribution, the critical values tend to zero as the sample size goes to infinity for which test?

(A) The Kolmogorov-Smirnov test.
(B) The Anderson-Darling test.
(C) None of the above.
(D) True for both tests.

72 The Chi-Square Goodness of Fit Test

The K-S test and the A-D are mostly applicable for complete individual data. In contrast, the **Chi-square goodness-of-fit test** is applicable to grouped data. To this end, data can be grouped into k intervals such as $(c_0, c_1], (c_1, c_2], \cdots .(c_{k-1}, c_k]$, where c_k may be infinity in which case the last interval is written (c_{k-1}, ∞). We let n_j denote the number of observations in the interval $(c_{j-1}, c_j]$. Thus, if the total number of observations is n then $n = n_1 + n_2 + \cdots + n_k$.
Also, we allow data to be grouped into categories as in the chart below.

Frequencey (k)	Number of Observations (n_k)
0	n_0
1	n_1
2	n_2
$\vdots$	$\vdots$

For the $j-$th category, let p_j denote the probability that an observation falls in the interval $(c_{j-1}, c_j]$. Then $\hat{p}_j = F^*(c_j) - F^*(c_{j-1})$. Let $p_{n_j} = F(c_j) - F(c_{j-1})$ be the same probability according to the empirical distribution. Thus, the expected number of observations based on $F^*(\cdot)$ is $E_j = n\hat{p}_j$ and that based on $F_n(\cdot)$ is $O_j = np_{n_j} = n_j$.

The Chi-square test statistic is given by

$$\chi^2 = \sum_{j=1}^{k} \frac{(E_j - O_j)^2}{E_j} = \sum_{j=1}^{k} \frac{n(\hat{p}_j - p_{n_j})^2}{\hat{p}_j} = \left(\sum_{j=1}^{k} \frac{n_j^2}{E_j} \right) - n.$$

If $F^*(x)$ is completely specified with no unknown parameters, the critical values for this test come from the Chi-square distribution with degrees of freedom equal to $k-r-1$ where r is the number of parameters that have been estimated in the model distribution, and k is the number of categories. If α is the level of significance or confidence, then the critical value $\chi^2_{k-r-1,1-\alpha}$ is the $100(1-\alpha)$th percentile of the Chi-square distribution with degrees of freedom $k - r - 1$. If $\chi^2 \geq \chi^2_{k-r-1,1-\alpha}$ then the null hypothesis is rejected. In such a case, the model does not fit the data well.

Example 72.1 ‡
You test the hypothesis that a given set of data comes from a known distribution with distribution function $F(x)$. The following data were collected:

Interval	$F(c_j)$	Number of Observations
$x < 2$	0.035	5
$2 \leq x < 5$	0.130	42
$5 \leq x < 7$	0.630	137
$7 \leq x < 8$	0.830	66
$8 \leq x$	1.000	50
Total		300

where c_j is the upper endpoint of each interval.
(a) Find the Chi-square statistic.
(b) Test the null hypothesis at the 5% level of significance.
(c) Test the null hypothesis at the 2.5% level of significance.

Solution.
We have

j	$\hat{p}_j$	E_j	O_j	$(E_j - O_j)^2$
1	0.035	10.5	5	30.25
2	0.095	28.5	42	182.25
3	0.5	150	137	169
4	0.2	60	66	36
5	0.17	51	50	1

(a) The Chi-square statistic is

$$\chi^2 = \frac{30.25}{10.5} + \frac{182.25}{28.5} + \frac{169}{150} + \frac{36}{60} + \frac{1}{51} = 11.02.$$

(b) Since no mention of any parameter estimation, we assume $r = 0$. Thus, $\chi^2_{4,0.95} = 9.488 < \chi^2$ so reject the null hypothesis.
(c) We have $\chi^2_{4,0.975} = 11.143 > \chi^2$ so fail to reject ■

Example 72.2 ‡
1000 workers insured under a workers compensation policy were observed for one year. The number of work days missed is given below:

Number of Days of Work Missed	Number of Workers
0	818
1	153
2	25
3^+	4
Total	1000
Total Number of Days Missed	230

The Chi-square goodness-of-fit test is used to test the hypothesis that the number of work days missed follows a Poisson distribution where:
(i) The Poisson parameter is estimated by the average number of work days missed.
(ii) Any interval in which the expected number is less than one is combined with the previous interval.
(a) Find the Chi-square statistic.
(b) Complete the following table.

Level of Significance	$\chi^2_{k-r-1,1-\alpha}$	Test result
10%		
5%		
2.5$		
1%		

Solution.
For the Poisson distribution, the mean is estimated as $\hat{\lambda} = \frac{230}{1000} = 0.23$. The probability that a worker missed 0 days is $p_0 = e^{-0.23} = 0.7945$. We can create the following table of information.

j	$\hat{p}_j$	E_j	O_j	$(E_j - O_j)^2$
0	0.7945	794.5	818	552.25
1	0.1827	182.7	153	882.09
2	0.210	21.0	25	16
3^+	0.0017	1.7	4	5.29

Note that $\hat{p}_4 = 1 - 0.210 - 0.1827 - 0.7945 = 0.0017$.
(a) The Chi-square statistic is

$$\chi^2 = \frac{552.25}{794.5} + \frac{882.09}{182.7} + \frac{16}{21} + \frac{5.29}{1.7} = 9.397.$$

(b) We have

Level of Significance	$\chi^2_{k-r-1,1-\alpha}$	Test result
10%	4.605	Reject
5%	5.991	Reject
2.5$	7.378	Reject
1%	9.210	Reject

where the degrees of freedom is $k - r - 1 = 4 - 1 - 1 = 2$ ■

Example 72.3 ‡
You are given:
(i) A computer program simulates $n = 1000$ pseudo$-U(0,1)$ variates.
(ii) The variates are grouped into $k = 20$ ranges of equal length.
(iii) $\sum_{j=1}^{20} O_j^2 = 51,850.$
(iv) The Chi-square goodness-of-fit test for $U(0,1)$ is performed.
Determine the result of the test.

Solution.
The hypothesized cdf is

$$F^*(x) = x.$$

The 20 ranges are of equal length each of length 0.05. Thus, $\hat{p}_{n_j} = F^*(c_j) - F^*(c_{j-1}) = c_j - c_{j-1} = 0.05$ and $E_j = n\hat{p}_{n_j} = 1000(0.05) = 50$. The Chi-square statistic is

$$\begin{aligned}\chi^2 =& \sum_{j=1}^{20} \frac{(E_j - O_j)^2}{E_j} \\ =&0.02\left[\sum_{j=1}^{20} O_j^2 - 100\sum_{j=1}^{20} O_j + 20(50^2)\right] \\ =&0.02[51,850 - 100(1000) + 50,000] = 37.\end{aligned}$$

There are $k-r-1 = 20-0-1 = 19$ degrees of freedom. Note that $r = 0$ since no model parameters are estimated. From the table of chi- square, we find $\chi^2_{19,0.99} = 36.191 < \chi^2$ so the null hypothesis is rejected at the 1% lelvel of confidence. At the 0.5% level of confidence, we have $\chi^2_{19,0.995} = 38.582 > \chi^2$, the null hypothesis is not rejected ■

Example 72.4 ‡
During a one-year period, the number of accidents per day was distributed as follows:

# of Accidents	Days
0	209
1	111
2	33
3	7
4	3
5	2

You use a chi-square test to measure the fit of a Poisson distribution with mean 0.60.
The minimum expected number of observations in any group should be 5.
The maximum possible number of groups should be used.
Determine the chi-square statistic.

Solution.
We have $E_j = 365e^{-0.6}\frac{0.6^j}{j!}$. We create the following table:

j	E_j	O_j	$(E_j - O_j)^2/E_j$
0	200.32	209	0.38
1	120.19	111	0.70
2	36.06	33	0.26
3	7.21	7	1.52*
4	1.08	3	
5	0.13	2	

$*$ We are told that the minimum expected number of observations for any group should be 5. Therefore, we combine groups 3, 4 and 5 to obtain

$$\frac{(E_j - O_j)^2}{E_j} = \frac{(12 - 8.42)^2}{8.42} = 1.52.$$

The chi-square test-statistic is

$$\chi^2 = 0.38 + 0.7 + 0.26 + 1.52 = 2.86 \blacksquare$$

Practice Problems

Problem 72.1 ‡
A particular line of business has three types of claims. The historical probability and the number of claims for each type in the current year are:

Type	Historical Probability	Number of Claims in Current Year
A	0.2744	112
B	0.3512	180
C	0.3744	138

You test the null hypothesis that the probability of each type of claim in the current year is the same as the historical probability.

Calculate the Chi-square goodness-of-fit test statistic.

Problem 72.2 ‡
You are given the following observed claim frequency data collected over a period of 365 days:

Number of Claims per Day	Observed Number of Days
0	50
1	122
2	101
3	92
4^+	0

(a) Fit a Poisson distribution to the above data, using the method of maximum likelihood.
(b) Regroup the data, by number of claims per day, into four groups:

$$0 \quad 1 \quad 2 \quad 3^+$$

Apply the Chi-square goodness-of-fit test to evaluate the null hypothesis that the claims follow a Poisson distribution. Determine the result of the chi-square test.

Problem 72.3 ‡
You are investigating insurance fraud that manifests itself through claimants who file claims with respect to auto accidents with which they were not involved. Your evidence consists of a distribution of the observed number of claimants per accident and a standard distribution for accidents on which fraud is known to be absent. The two distributions are summarized below:

Number of Claimants per Accident	Standard Probability	Observed Number of Accidents
1	0.25	235
2	0.35	335
3	0.24	250
4	0.11	111
5	0.04	47
6^+	0.01	22
Total	1.00	1000

Determine the result of a Chi-square test of the null hypothesis that there is no fraud in the observed accidents.

Problem 72.4 ‡
You are given the following random sample of 30 auto claims:

54	140	230	560	600	1,100	1,500	1,800	1,920	2,000
2,450	2,500	2,580	2,910	3,800	3,800	3,810	3,870	4,000	4,800
7,200	7,390	11,750	12,000	15,000	25,000	30,000	32,300	35,000	55,000

You test the hypothesis that auto claims follow a continuous distribution $F(x)$ with the following percentiles:

x	310	500	2,498	4,876	7,498	12,930
$F(x)$	0.16	0.27	0.55	0.81	0.90	0.95

You group the data using the largest number of groups such that the expected number of claims in each group is at least 5.

Calculate the Chi-square goodness-of-fit statistic.

Problem 72.5 ‡
Which of the following statements is true?

(A) For a null hypothesis that the population follows a particular distribution, using sample data to estimate the parameters of the distribution tends to decrease the probability of a Type II error.
(B) The Kolmogorov-Smirnov test can be used on individual or grouped data.
(C) The Anderson-Darling test tends to place more emphasis on a good fit in the middle rather than in the tails of the distribution.

(D) For a given number of cells, the critical value for the Chi-square goodness-of-fit test becomes larger with increased sample size.
(E) None of (A), (B), (C) or (D) is true.

Problem 72.6 ‡
Which of statements (A), (B), (C), and (D) is false?

(A) The Chi-square goodness-of-fit test works best when the expected number of observations varies widely from interval to interval.
(B) For the Kolmogorov-Smirnov test, when the parameters of the distribution in the null hypothesis are estimated from the data, the probability of rejecting the null hypothesis decreases.
(C) For the Kolmogorov-Smirnov test, the critical value for right censored data should be smaller than the critical value for uncensored data.
(D) The Anderson-Darling test does not work for grouped data.
(E) None of (A), (B), (C) or (D) is false.

73 The Likelihood Ratio Test

A **likelihood ratio test** is a statistical test used to compare the fit of two models, one of which (the null model) is a special case of the other (the alternative model). For example, the null model is the exponential model while the alternative model is the Gamma model. The hypothesis test problem can be expressed as follows:

H_0 : The data comes from a population with distribution A.
H_a : The data does not come from a population with distribution A.

The likelihood ratio test is conducted as follows: Let $L(\theta)$ denote the likelihood function. Let Θ_0 denote the set of all possible values of θ as specified in the null hypothesis. Suppose that the maximum of $L(\theta)$ occurs at some value $\theta_0 \in \Theta_0$ with maximum value $L_0 = L(\theta_0)$. Likewise, let Θ_a denote the set of all possible values of θ as specified in the alternative hypothesis. Suppose that the maximum of $L(\theta)$ occurs at some value $\theta_a \in \Theta_a$ with maximum value $L_a = L(\theta_a)$. Note that $\theta_a = MLE(\theta)$.

The likelihood ratio statistic, denoted by T, is defined as

$$T = 2\ln\left(\frac{L_1}{L_0}\right) = 2[\ln L_1 - \ln L_0].$$

Regarding T as a random variable, it is shown that, for large sampling, T is approximated by a Chi-square distribution with degrees of freedom r equals the number of free parameters under the alternative hypothesis minus the number of free parameters under the null hypothesis. Let c_α denote the critical value at a significance level α. That is, c_α is the $100(1-\alpha)$th percentile from the Chi-square distribution with degrees of freedom r. Then the null hypothesis is rejected if $T > c_\alpha$.

Example 73.1
A random sample of $n = 8$ values from an exponential random variable X is given:

$$3 \;\; 3 \;\; 4 \;\; 6 \;\; 7 \;\; 8 \;\; 10 \;\; 25.$$

You are performing the following hypothesis test:

$H_0 : \theta = 8.$
$H_1 : \theta \neq 8.$

(a) Find the likelihood ratio statistic for this hypothesis.
(b) Test the hypothesis at the 5% level of significance.

Solution.
(a) We have $\Theta_0 = \{8\}$ so that $\theta_0 = 8$. The likelihood function is

$$L(\theta) = \prod_{i=1}^{8} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \frac{1}{\theta^8} e^{-\frac{66}{\theta}}.$$

Hence, $L_0 = L(8) = \frac{1}{8^8} e^{-\frac{66}{8}}$.
On the other hand, $\Theta_a = \{\theta > 0 : \theta \neq 8\}$ and $\theta_1 = MLE(\theta) = \overline{x} = \frac{66}{8} = 8.25$ and therefore $L_1 = \frac{1}{8.25^8} e^{-\frac{66}{8.25}}$. The test statistic is

$$T = 2 \ln \left[\left(\frac{8}{8.25} \right)^8 e^{-\frac{66}{8.25} + \frac{66}{8}} \right] = 0.00765.$$

(b) The number of free parameters in the null hypothesis is 0 (θ is specified as 8) while it is 1 in the alternate hypothesis (θ is freely chosen to maximize $L(\theta)$). That is, $r = 1$. From the table of Chi-square distribution we find $c_{0.05} = 3.84$. Since $T < c_{0.05}$, do not reject the null hypothesis ■

Example 73.2 ‡
You are given:
(i) A random sample of losses from a Weibull distribution is:

595 700 789 799 1109

(ii) You use the likelihood ratio test to test the hypothesis:

$H_0 : \tau = 2.$
$H_a : \tau \neq 2.$

(iii) When $\tau = 2$, the maximum likelihood estimate of θ is 816.7.
(a) Find $\ln L_0$.
(b) Find the degrees of freedom.
(c) Complete the following table:

α	10%	5%	2.5%	1%
c_α				

Solution.
(a) The likelihood function is

$$L(\tau,\theta)=\prod_{j=1}^{5}\frac{\tau x_j^{\tau-1}e^{-\left(\frac{x_j}{\theta}\right)^{\tau}}}{\theta^{\tau}}.$$

The loglikelihood function is

$$\ell(\tau,\theta)=\sum_{j=1}^{5}\left[\ln\tau+(\tau-1)\ln x_j-\tau\ln\theta-\left(\frac{x_j}{\theta}\right)^{\tau}\right].$$

Thus,

$$\begin{aligned}\ln L_0 =&5\ln 2+\ln\left(595\cdot 700\cdot 789\cdot 799\cdot 1109\right)-10\ln 816.7\\ &-\left(\frac{595^2+700^2+789^2+799^2+1109^2}{816.7^2}\right)=-35.28.\end{aligned}$$

(b) The free parameter in the null hypothesis is θ. The free parameters in the alternate hypothesis is τ and θ. Hence, $r=2-1=1$.
(c) Using the Chi-square distribution with one degree of freedom, we find

α	10%	5%	2.5%	1%
c_α	2.706	3.841	5.024	6.635

Example 73.3 ‡
You are given:
(i) Twenty claim amounts are randomly selected from a Pareto distribution with $\alpha=2$ and an unknown θ.
(ii) The maximum likelihood estimate of θ is 7.0
(iii) $\sum ln(x_i+7.0)=49.01$
(iv) $\sum ln(x_i+3.1)=39.30$
You use the likelihood ratio test to test the hypothesis that $\theta=3.1$.
Determine the result of the test at the 1% level.

Solution.
The likelihood function is

$$L(\alpha,\theta)=\prod_{j=1}^{20}\frac{\alpha\theta^{\alpha}}{(x_i+\theta)^{\alpha+1}}$$

and its loglikelihood function is

$$\ell(\alpha, \theta) = 20 \ln \alpha + 20\alpha \ln \theta - (\alpha + 1) \sum_{j=1}^{20} ln(x_j + \theta).$$

The loglikelihood value under the null hypothesis ($\alpha = 2$ and $\theta = 3.1$) is

$$\ln L_0 = 20 \ln 2 + 20(2) \ln 3.1 - 3(39.30) = -58.781.$$

The loglikelihood value under the alternative hypothesis ($\alpha = 2$ and $\theta = 7.0$) is

$$\ln L_1 = 20 \ln 2 + 20(2) \ln 7.0 - 3(49.01) = -55.331.$$

There is one degree of freedom (the hypothesized distribution has no parameters estimated versus one estimated parameter in the alternative hypothesis). The test statistic for the Chi-square test is

$$2(\ln L_1 - \ln L_0) = 2(-55.331 + 58.781) = 6.9.$$

In the Chi-square table with 1 degree of freedom, we see that $\chi^2_{1,0.99} = 6.636 < 6.9$ so the null hypothesis is rejected at this level $\blacksquare$

Practice Problems

Problem 73.1
Consider the following hypothesis test problem:

H_0 : The data came from a Pareto distribution with $\alpha = 1.5$ and $\theta = 7.8$.
H_a : The data came from a Pareto distribution with $\alpha \neq 1.5$ and $\theta \neq 7.8$.

Find the degrees of freedom associated with the likelihood ratio test.

Problem 73.2
Consider the following hypothesis test problem:

H_0 : The data came from a Pareto distribution with $\alpha = 1.5$ and $\theta = 7.8$.
H_a : The data came from a Pareto distribution with $\alpha \neq 1.5$ and $\theta \neq 7.8$.

Using the likelihood ratio test, complete the following table:

α	5%	2.5%	1%	0.5%
c_α				

Problem 73.3 ‡
You fit a Pareto distribution to a sample of 200 claim amounts and use the likelihood ratio test to test the hypothesis that $\alpha = 1.5$ and $\theta = 7.8$.
You are given:
(i) The maximum likelihood estimates are $\hat{\alpha} = 1.4$ and $\hat{\theta} = 7.6$..
(ii) The natural logarithm of the likelihood function evaluated at the maximum likelihood estimates is -817.92.
(iii) $\sum ln(x_i + 7.8) = 607.64$.
Determine the result of the test.

Problem 73.4 ‡
You are given:
(i) A random sample of losses from a Weibull distribution is:

595 700 789 799 1109

(ii) At the maximum likelihood estimates of θ and τ, $\sum_{i=1}^{5} \ln f(x_i) = -33.05$.
(iii) When $\tau = 2$, the maximum likelihood estimate of θ is 816.7.
(iv) You use the likelihood ratio test to test the hypothesis:

$H_0 : \tau = 2.$
$H_a : \tau \neq 2.$

Determine the result of the test.

Problem 73.5 ‡
During a one-year period, the number of accidents per day was distributed as follows:

# of Accidents	Days
0	209
1	111
2	33
3	7
4	3
5	2

For these data, the maximum likelihood estimate for the Poisson distribution is $\hat{\lambda} = 0.60$, and for the negative binomial distribution, it is $\hat{r} = 2.9$ and $\hat{\beta} = 0.21$.
The Poisson has a nega tive loglikelihood value of 385.9, and the negative binomial has a negative loglikelihood value of 382.4.

Determine the likelihood ratio test statistic, treating the Poisson distribution as the null hypothesis.

74 Schwarz Bayesian Criterion

When selecting a model, one has to keep two things in mind:

• **Principle of Parsimony:** When selecting a model, a simpler model is always preferred unless there is considerable evidence to do otherwise. A reason of this preference is that a complex model may do a great job of matching the data but may fail in matching the population from which the data were sampled.

• Reduce, if possible, the universe of potential models.

In [1], two approaches for model selection are considered. The first one is judgement-based approach where the modeler's experience is critical. This is a feature of the second point mentioned above. The other approach is a score-based approach where a numerical value (a score) is assigned to a model and the model selected is the one with best score.

In this section, we consider a score-based selection method, known as the **Schwarz Bayesian Criterion**, denoted by SBC, and is defined by

$$SBC = \ln L(\hat{\theta}) - \frac{r}{2}\ln n$$

where $\ln L(\hat{\theta})$ is the maximum loglikelihood value, r is the number of parameters being estimated in the model, and n is the sample size. The quantity $\frac{r}{2}\ln n$ is referred to as the **Schwarz Bayesian adjustment**.

With multiple models, the SBC of each model is computed. The model preferred is the one with highest SBC.

Example 74.1
You are given that a particular model has a maximum loglikelihood value of -412. You also know that the model uses 2 parameters and is used to interpret a sample of 260 data points. Find the SBC of this model.

Solution.
The SBC is given by

$$SBC = -412 - \frac{2}{2}\ln 260 = -417.56 \blacksquare$$

Example 74.2
What is the limit of the Schwarz Bayesian adjustement as the sample size increases without bound?

Solution.
We have

$$\lim_{n\to\infty} \frac{r}{2} \ln n = \infty \blacksquare$$

Example 74.3
Four models are fitted to a sample of $n = 200$ observations with the following results:

Model	# of Parameters	Loglikelihood
I	3	-180.2
II	2	-181.4
III	2	-181.6
IV	1	-183

Determine the model favored by the Schwarz Bayesian criterion.

Solution.
We have the following

Model	# of Parameters	Loglikelihood	SBC
I	3	-180.2	-188.1
II	2	-181.4	-186.7
III	2	-181.6	-186.9
IV	1	-183	-185.6

Hence, the favored model is IV $\blacksquare$

Practice Problems

Problem 74.1
Which of the following statements is false?

(i) The principle of parsimony states that a more complex model is better because it will always match the data better.
(ii) In judgment-based approaches to determining a model, a modeler's experience is critical.
(iii) In score-based approaches, one assigns scores to the potential models.

Problem 74.2 ‡
If the proposed model is appropriate, which of the following tends to zero as the sample size goes to infinity?

(A) Kolmogorov-Smirnov test statistic
(B) Anderson-Darling test statistic
(C) Chi-square goodness-of-fit test statistic
(D) Schwarz Bayesian adjustment
(E) None of (A), (B), (C) or (D)

Problem 74.3 ‡
Five models are fitted to a sample of $n = 260$ observations with the following results:

Model	# of Parameters	Loglikelihood
I	1	-414
II	2	-412
III	3	-411
IV	4	-409
V	6	-409

Determine the model favored by the Schwarz Bayesian criterion.

Problem 74.4 ‡
You are given:
(i) Sample size = 100
(ii) The negative loglikelihoods associated with five models are:

Model	# of Parameters	Loglikelihood
Generalized Pareto	1	-414
Burr	2	-412
Pareto	3	-411
Lognormal	4	-409
Inverse exponential	6	-409

Determine the model favored by the Schwarz Bayesian criterion.

Credibility Theory

A framework of credibility problem can be described as follows: We consider a block of insurance policies, referred to as a **risk group**. This risk group is covered by an insurer over a period of time upon the payment of a premium. The value of the premium is decided based on a rate specified in the **manual rate** and on the specific risk characteristics of the group. An actuary studies the recent **claim experience** of the risk group to decide whether a revised premium for the next period is required. **Credibility theory** concerns the finding of this premium for the next period using the recent claim experience and the manual rate.
We will consider the following three approaches to credibility: The **limited fluctuation credibility approach**[17], the **Bayesian approach**, and the **Bühlmann's approach**.

[17]Also known as **classical credibility**.

75 Limited Fluctuation Credibility Approach: Full Credibility

Let $X_1, X_2, \cdots, X_n$ represent the total claims/losses[18] (or other entity) experienced by a policyholder in the past n years. That is, X_i is being experienced at time i. In the literature, an observed value of X_i is also referred to as an **exposure**[19] or **exposure unit** of X_i. The average claim severity is

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

The statistic $\overline{X}$ is the average of past experience for the given group[20].

We make the assumptions:

$$\begin{aligned} E(X_j) =& \xi \\ \text{Var}(X_j) =& \sigma^2 \end{aligned}$$

for all $j = 1, 2, \cdots, n$. That is ξ is the mean, which is the premium to be charged if known, across the members of the group−presumed to be stable over time. If the X_i' are independent, then we can write

$$\begin{aligned} E(\overline{X}) =& \xi \\ \text{Var}(\overline{X}) =& \frac{\sigma^2}{n}. \end{aligned}$$

An insurer's goal is to decide on the value of ξ. The choice can be done in one of three ways:

• Ignore past experience or data (no credibility) and charge the **manual premium**[21] M.

• Use only past data (full credibility) and charge $\overline{X}$ (the **observed pure premium**).

• Use a combination of M and $\overline{X}$ (partial credibility).

[18]Recall that the distribution of the number of claims is a frequency distribution and that of the size of a claim is a severity distribution

[19]In general, an exposure for a particular random variable is one observation of that random variable.

[20]In the literature, this is also referred to as **pure premium**

[21]A manual premium is a premium that comes from a manual (book) of premiums.

In general, insurers prefer to choose $\overline{X}$ because of its stability over time and is a good indicator of future results. However, in the presence of more variability, a manual premium is more suitable.

By assigning $\overline{X}$ to ξ (full credibility) we require that $\overline{X}$ be stable over time. In statistical terms, this means that the relative error $\left|\frac{\overline{X}-\xi}{\xi}\right|$ is small with high probability. That is, there are two numbers $r > 0$ (with r close to 0, a common choice is $r = 0.05$) and $0 < p < 1$(with p close to 1, a common choice is $p = 0.9$). Mathematically,

$$\Pr\left(\left|\frac{\overline{X}-\xi}{\xi}\right| \le r\right) \ge p.$$

As $n \to \infty$, by the Central Limit theorem, we have

$$\overline{X} \sim N\left(\xi, \frac{\sigma^2}{n}\right) \to \frac{\overline{X}-\xi}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1).$$

In this case, we have the equivalence

$$\Pr\left(\left|\frac{\overline{X}-\xi}{\xi}\right| \le r\right) \ge p \Longleftrightarrow \Pr\left(\left|\frac{\overline{X}-\xi}{\frac{\sigma}{\sqrt{n}}}\right| \le \frac{r\xi\sqrt{n}}{\sigma}\right) \ge p.$$

Now, let

$$y_p = \inf_y \left\{\Pr\left(\left|\frac{\overline{X}-\xi}{\xi}\right| \le y\right) \ge p\right\}.$$

If $\overline{X}$ is continuous then y_p satisfies

$$\Pr\left(\left|\frac{\overline{X}-\xi}{\xi}\right| \le y_p\right) = p.$$

Accordingly, the condition of full credibility is met when

$$\frac{r\xi\sqrt{n}}{\sigma} \ge y_p.$$

Letting $\lambda_0 = \left(\frac{y_p}{r}\right)^2$, we obtain

$$\frac{r\xi\sqrt{n}}{\sigma} \ge y_p \Longleftrightarrow \frac{\sigma}{\xi} \le \sqrt{\frac{n}{\lambda_0}}.$$

This says that full credibility is assigned if the coefficient of variation $\frac{\sigma}{\xi}$ of X_j is no larger than $\sqrt{\frac{n}{\lambda_0}}$.
Alternatively, full credibility occurs if

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n} \le \frac{\xi^2}{\lambda_0}.$$

The number of exposure units required for full credibility is

$$n \ge \lambda_0 \left(\frac{\sigma}{\xi}\right)^2. \tag{75.1}$$

We next describe how to find y_p when $\overline{X}$ is continuous. Recall that the cdf of the standard normal distribution Z is denoted by Φ. We have

$$\begin{aligned} p =&\Pr(|Z| \le y_p)\\ =&\Pr(-y_p \le Z \le y_p)\\ =&\Phi(y_p) - \Phi(-y_p)\\ =&\Phi(y_p) - 1 + \Phi(y_p)\\ =&2\Phi(y_p) - 1. \end{aligned}$$

Thus, $\Phi(y_p) = \frac{1+p}{2}$ so that y_p is the $100\left(\frac{1+p}{2}\right)-$th percentile of the standard normal distribution.

Example 75.1
Determine y_p and λ_0 if $p = 0.9$ and $r = 0.05$. Also, determine the full-credibility standard.

Solution.
The standard normal tables give $y_{0.9} = 1.645$. Thus, $\lambda_0 = \left(\frac{y_p}{r}\right)^2 = \left(\frac{1.645}{0.05}\right)^2 = 1082.41$. The full credibility standard is

$$n \ge 1082.41 \left(\frac{\sigma}{\xi}\right)^2 \quad \blacksquare$$

The right hand-side of Equation (75.1) is known as the **standard of full crediblitiy** when measuring it in terms of exposure units. If some other unit is desired, such as the expected number of claims or the expected amount of claims, it is usually sufficient to multiply both sides of (75.1) by an appropriate factor.

Example 75.2
Let $N_1, N_2, \cdots, N_n$ the number of claims in the past n years for a policy. That is, N_i is the total number of claims of year i. Let $X_1, X_2, \cdots, X_n$. be the losses in the past n years. That is, X_i is the total losses in year i. We assume that the X_i' and the $N_i's$ are independent with the $N_i's$ being iid with common distribution the Poisson distribution with mean λ.
Let Y_{ij} denote the j^{th} claim in year i. We assume that the Y_{ij} are iid with mean θ_Y and variance σ_Y^2. Then we have

$$X_i = \sum_{j=1}^{N_i} Y_{ij}.$$

Find the standard of full credibility based on
(a) the number of exposure units;
(b) the expected total number of claims;
(c) the expected total amount of claims.

Solution.
X_i is a compound Poisson distribution so that $E(X_i) = E(N_i)E(Y_{ij}) = \lambda\theta_Y$ and $\text{Var}(X_i) = \text{Var}(N_i)[E(Y_{ij})]^2 + E(N_i)\text{Var}(Y_{ij}) = \lambda(\theta_Y^2 + \sigma_Y^2)$.
(a) The standard of full credibility based on the number of exposure units is

$$\lambda_0 \left(\frac{\sigma_{X_i}}{E(X_i)}\right)^2 = \frac{\lambda_0}{\lambda}\left[1 + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right] \text{ exposures.}$$

(b) The expected total number of claims is $\sum_{i=1}^n E(N_i) = nE(N_i) = n\lambda$. The standard of full credibility based on the expected total number of claims is

$$E(N_i)\lambda_0 \left(\frac{\sigma_{X_i}}{E(X_i)}\right)^2 = \lambda_0 \left[1 + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right] \text{ claims.}$$

(c) The expected total amount of claims is $\sum_{i=1}^n E(X_i) = nE(X_i)$. The standard of full credibility based on the expected total amount of claims is

$$E(X_i)\lambda_0 \left(\frac{\sigma_{X_i}}{E(X_i)}\right)^2 = \lambda_0 \frac{\sigma_{X_i}^2}{E(X_i)} = \lambda_0 \left(\theta_Y + \frac{\sigma_Y^2}{\theta_Y}\right) \text{ dollars} \blacksquare$$

Example 75.3
Repeat the previous example by replacing the Poisson distribution with a binomial distribution with parameters $(50, p')$.

Solution.
X_i is a compound binomial distribution so that $E(X_i) = E(N_i)E(Y_{ij}) = 50p'\theta_Y$ and $\text{Var}(X_i) = \text{Var}(N_i)[E(Y_{ij})]^2 + E(N_i)\text{Var}(Y_{ij}) = \lambda(\theta_Y^2 + \sigma_Y^2) = 50p'[(1-p')\theta_Y^2 + \sigma_Y^2]$.
(a) The standard of full credibility based on the number of exposure units is

$$\lambda_0\left(\frac{\sigma_{X_i}}{E(X_i)}\right)^2 = \frac{\lambda_0}{50p'}\left[(1-p') + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right].$$

(b) The expected total number of claims is $\sum_{i=1}^n E(N_i) = nE(N_i) = 50np'$. The standard of full credibility based on the expected total number of claims is

$$\lambda_0\left[(1-p') + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right].$$

(c) The expected total amount of claims is $\sum_{i=1}^n E(X_i) = nE(X_i)$. The standard of full credibility based on the expected total amount of claims is

$$E(X_i)\lambda_0\left(\frac{\sigma_{X_i}}{E(X_i)}\right)^2 = \lambda_0\frac{\sigma_{X_i}^2}{E(X_i)} = \lambda_0\left[(1-p')\theta_Y + \frac{\sigma_Y^2}{\theta_Y}\right] \blacksquare$$

Example 75.4 ‡
You are given:
(i) The number of claims follows a negative binomial distribution with parameters r and $\beta = 3$.
(ii) Claim severity has the following distribution:

Claim Size	Probability
1	0.4
10	0.4
100	0.2

(iii) The number of claims is independent of the severity of claims.
Determine the expected number of claims needed for aggregate losses to be within 10% of expected aggregate losses with 95% probability.

Solution.
We are asked to find the full credibility standard based on the number of claims of aggregate losses. If N denote the number of claims then the full credibility standard is the right side of the inequality

$$n(3r) \geq (3r)\lambda_0\left(\frac{\sqrt{\text{Var}(S)}}{E(S)}\right)^2$$

where S is the aggregate loss and

$$\lambda_0 = \left(\frac{1.96}{0.1}\right)^2 = 384.16.$$

We have

$$\begin{aligned}
E(S) =& E(N)E(X) = 3r[1(0.4) + 10(0.4) + 100(0.2)] = 3r(24.4) = 73.2r \\
\text{Var}(S) =& E(N)\text{Var}(X) + \text{Var}(N)E(X)^2 \\
=& 3r(1^2(0.4) + 10^2(0.4) - 100^2(0.2) - 24.4^2) + 3r(1+3)24.4^2 \\
=& 11,479.44r.
\end{aligned}$$

Thus, the full credibility standard based on the number of claims of aggregate losses

$$3r(384.16)\left(\frac{11,479.44}{(73.2r)^2}\right) = 2469.06.$$

Hence, the expected number of claims needed is at least 2470 ■

Example 75.5 ‡
For an insurance portfolio, you are given:
(i) For each individual insured, the number of claims follows a Poisson distribution.
(ii) The mean claim count varies by insured, and the distribution of mean claim counts follows a gamma distribution.
(iii) For a random sample of 1000 insureds, the observed claim counts are as follows:

Number Of Claims, n	0	1	2	3	4	5
Number Of Insureds, f_n	512	307	123	41	11	6

$\sum nf_n = 750$ and $\sum n^2 f_n = 1494$.
(iv) Claim sizes follow a Pareto distribution with mean 1500 and variance 6,750,000.
(v) Claim sizes and claim counts are independent.
(vi) The full credibility standard is to be within 5% of the expected aggregate loss 95% of the time.
Determine the minimum number of insureds needed for the aggregate loss to be fully credible.

Solution.
We want the smallest positive integer n such that

$$n \geq \lambda_0 \frac{\text{Var}(S)}{E(S)^2}$$

where

$$\lambda_0 = \left(\frac{1.96}{0.05}\right)^2 = 1536.64.$$

We have

$$\begin{aligned}
\xi =& E(S) = E(N)E(X) = \frac{750}{1000}(1500) = 1125 \\
\mathrm{Var}(S) =& E(N)\mathrm{Var}(X) + \mathrm{Var}(N)E(X)^2 \\
=& 0.75(6,750,000) + \mathrm{Var}(N)(1500)^2 \\
\mathrm{Var}(N) =& \frac{\sum f_i(n_i - 0.75)^2}{1000 - 1} = 0.93243 \\
\mathrm{Var}(S) =& 0.75(6,750,000) + 0.93243(1500)^2 \\
=& 7,160,468.
\end{aligned}$$

Thus,

$$\lambda_0 \frac{\mathrm{Var}(S)}{E(S)^2} = 1536.64 \left(\frac{7,160,468}{1125^2}\right) = 8693.77$$

so that $n = 8694$ ■

Practice Problems

Problem 75.1 ‡
You are given:
(i) The number of claims has a Poisson distribution.
(ii) Claim sizes have a Pareto distribution with parameters $\theta = 0.5$ and $\alpha = 6$.
(iii) The number of claims and claim sizes are independent.
(iv) The observed pure premium should be within 2% of the expected pure premium 90% of the time.

Determine the expected number of claims needed for full credibility.

Problem 75.2 ‡
You are given the following information about a commercial auto liability book of business:
(i) Each insured's claim count has a Poisson distribution with mean λ, where λ has a gamma distribution with $\alpha = 1.5$ and $\theta = 0.2$.
(ii) Individual claim size amounts are independent and exponentially distributed with mean 5000.
(iii) The full credibility standard is for aggregate losses to be within 5% of the expected with probability 0.90.

Using classical credibility, determine the expected number of claims required for full credibility.

Problem 75.3 ‡
You are given:
(i) The number of claims has probability function:

$$p(x) = \binom{m}{x} q^x (1-q)^{m-x},\ x = 0, 1, \cdots, m.$$

(ii) The actual number of claims must be within 1% of the expected number of claims with probability 0.95.
(iii) The expected number of claims for full credibility is 34,574.

Determine q.

Problem 75.4 ‡
You are given:

(i) The number of claims follows a Poisson distribution.
(ii) Claim sizes follow a gamma distribution with parameters α (unknown) and $\theta = 10,000$.
(iii) The number of claims and claim sizes are independent.
(iv) The full credibility standard has been selected so that actual aggregate losses will be within 10% of expected aggregate losses 95% of the time.

Using limited fluctuation (classical) credibility, determine the expected number of claims required for full credibility.

Problem 75.5 ‡
A company has determined that the limited fluctuation full credibility standard is 2000 claims if:
(i) The total number of claims is to be within 3% of the true value with probability p.
(ii) The number of claims follows a Poisson distribution.
The standard is changed so that the total cost of claims is to be within 5% of the true value with probability p, where claim severity has probability density function:

$$f(x) = \frac{1}{10,000}, \ 0 \leq x \leq 10,000.$$

Using limited fluctuation credibility, determine the expected number of claims necessary to obtain full credibility under the new standard.

76 Limited Fluctuation Credibility Approach: Partial Credibility

When the premium charged under full credibility is not approapriate reflection of the actual premium, an alternative remedy is the partial credibility.

Under **partial credibility** the premium is a weighted average of the past experience $\overline{X}$ and the manual premium $M > 0$. Mathematically,

$$P_c = Z\overline{X} + (1 - Z)M, \ 0 \leq Z \leq 1$$

where Z is called the **credibility factor** and P_c is the **credibility premium**. Note that when $Z = 1$, the partial credibility coincides with the full credibility.

Example 76.1
(a) Find $E(P_c)$ and $\text{Var}(P_c)$.
(b) Find $\frac{P_c - E(P_c)}{\sqrt{\text{Var}(P_c)}}$.

Solution.
(a) We have

$$E(P_c) = E[Z\overline{X} + (1 - Z)M] == Z\xi + (1 - Z)M$$

and

$$\text{Var}(P_c) = \text{Var}[Z\overline{X} + (1 - Z)M] = \text{Var}(Z\overline{X}) = Z^2\text{Var}(\overline{X}) = Z^2\frac{\sigma^2}{n}.$$

(b) We have

$$\frac{P_c - E(P_c)}{\sqrt{\text{Var}(P_c)}} = \frac{Z\overline{X} - Z\xi}{Z\sigma/\sqrt{n}} = \frac{\overline{X} - \xi}{\sigma/\sqrt{n}} \blacksquare$$

How do we find Z? There are several ways for finding Z and all of them lead to the same result. We will follow an approach similar to Section 75: We want to choose Z such that $\left|\frac{Z\overline{X} - Z\xi}{\xi}\right|$ is small with high probability. Mathematically,

$$\Pr\left(\left|\frac{Z\overline{X} - Z\xi}{\xi}\right| \leq r\right) \geq p.$$

As $n \to \infty$, by the Central Limit theorem, we have

$$Z\overline{X} \sim N\left(Z\xi, Z^2\frac{\sigma^2}{n}\right) \to \frac{Z\overline{X} - Z\xi}{Z\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

In this case, we have the equivalence

$$\Pr\left(\left|\frac{Z\overline{X} - Z\xi}{\xi}\right| \le r\right) \ge p \Longleftrightarrow \Pr\left(\left|\frac{\overline{X} - \xi}{\frac{\sigma}{\sqrt{n}}}\right| \le \frac{r\xi\sqrt{n}}{Z\sigma}\right) \ge p.$$

Now, let

$$y_p = \inf_y \left\{\Pr\left(\left|\frac{\overline{X} - \xi}{\sigma/\sqrt{n}}\right| \le y\right) \ge p\right\}.$$

Accordingly, the condition of partial credibility is met when

$$\frac{r\xi\sqrt{n}}{Z\sigma} \ge y_p.$$

Letting $\lambda_0 = \left(\frac{y_p}{r}\right)^2$, we obtain

$$\frac{r\xi\sqrt{n}}{Z\sigma} \ge y_p \Longleftrightarrow Z \le \frac{\xi}{\sigma}\sqrt{\frac{n}{\lambda_0}}.$$

We choose Z such that

$$Z = \min\left\{\frac{\xi}{\sigma}\sqrt{\frac{n}{\lambda_0}}, 1\right\}.$$

Letting $n_F = \frac{\xi}{\sigma\sqrt{\lambda_0}}$ be the standard full credibility measured in terms of exposures then $Z = \sqrt{\frac{n}{n_F}}$ can be interpreted as the square root of the number of observations (exposures) available divided by the number of observations (exposures) required for full credibility.

Based on the total number of claims, Z is the square root of the number of available claims divided by the total number of claims required for full credibility. Based on total amout of claims, Z is the square root of the total amount of available claims divided by the total amount of claims required for the standard credibility.

Example 76.2
You are given:
(i) The number of claims per exposure follows a Poisson distribution with a mean of 10.
(ii) Claim size follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 1$.
(iii) The number of claims per exposure and claim sizes are independent.

(iv) The method of limited fluctuation credibility is used, and the full credibility standard has been selected so that total claim dollars per exposure will be within 10% of expected total claim dollars per exposure 95% of the time.
Find the credibility factor
(a) based on 45 exposures
(b) based on a total claim number of 120
(c) based on a total claim amount of 600.

Solution.
(a) The standard for full credibility based on the number of exposures is

$$\frac{\lambda_0}{\lambda}\left[1+\frac{\sigma_Y^2}{\theta_Y^2}\right]=\left(\frac{1.96}{0.1}\right)^2\frac{1}{10}\left[1+\frac{0.75^2}{0.5^2}\right]=124.852.$$

The credibility factor based on 45 exposures is

$$Z=\sqrt{\frac{45}{124.852}}=0.6004.$$

(b) The standard for full credibility based on the total number of claims is

$$\lambda_0\left[1+\frac{\sigma_Y^2}{\theta_Y^2}\right]=\left(\frac{1.96}{0.1}\right)^2\left[1+\frac{0.75^2}{0.5^2}\right]=1248.52.$$

The credibility factor based on 120 claims is

$$Z=\sqrt{\frac{120}{1248.52}}=0.31.$$

(c) The standard for full credibility based on the total amount of claims is

$$\lambda_0\left[\theta_Y+\frac{\sigma_Y^2}{\theta_Y}\right]=\left(\frac{1.96}{0.1}\right)^2\left[0.5+\frac{0.75^2}{0.5}\right]=624.26.$$

The credibility factor based on a total amount of 2500 is

$$Z=\sqrt{\frac{600}{624.26}}=0.9804 \blacksquare$$

Example 76.3
You are given:
(i) 350 claims with a total of 300,000.
(ii) The manual premium of $M=1000$.
(iii) The credibility factor $Z=0.809$.
Determine the credibility premium.

Solution.
The partial credibility pure premium is

$$\overline{X} = \frac{300000}{350} = 857.14.$$

The credibility premium is

$$P_c = Z\overline{X} + (1 - Z)M = 0.809(857.14) + (1 - 0.809)(1000) = 884.43 \blacksquare$$

Practice Problems

Problem 76.1 ‡
You are given the following information about a general liability book of business comprised of 2500 insureds:
(i) $X_i = \sum_{j=1}^{N_i} Y_{ij}$, is a random variable representing the annual loss of the i^{th} insured.
(ii) $N_1, N_2, \cdots, N_{2500}$, are independent and identically distributed random variables following a negative binomial distribution with parameters $r = 2$ and $\beta = 0.2$.
(iii) $Y_{i1}, Y_{i2}, \cdots, Y_{iN_i}$ are independent and identically distributed random variables following a Pareto distribution with $\alpha = 3.0$ and $\theta = 1000$.
(iv) The full credibility standard is to be within 5% of the expected aggregate losses 90% of the time.

Using classical credibility theory, determine the partial credibility of the annual loss experience for this book of business.

Problem 76.2 ‡
You are given:
(i) $X_{partial}$ = pure premium calculated from partially credible data.
(ii) $\mu = E[X_{partial}]$.
(iii) Fluctuations are limited to $\pm k\mu$ of the mean with probability P. (iv) Z = credibility factor.

Which of the following is equal to P?

(A) $\Pr(|X_{partial} - \mu| \leq k\mu)$
(B) $\Pr(|ZX_{partial} - Z\mu| \leq k)$
(C) $\Pr(|ZX_{partial} - Z\mu| \leq \mu)$
(D) $\Pr(|ZX_{partial} + (1-Z)\mu - 1| \leq k)$
(E) $\Pr(|ZX_{partial} + (1-Z)\mu - \mu| \leq k\mu)$.

Problem 76.3
You are given:
(i) 50 claim amounts.
(ii) Claim size is uniform in $[0, \theta]$.
(iii) The full credibility standard is to be within 5% of the expected claim amount 90% of the time.

Using partial credibility theory, calculate the credibility factor Z.

Problem 76.4 ‡
You are given:
(i) Claim counts follow a Poisson distribution.
(ii) Claim sizes follow a lognormal distribution with coefficient of variation 3.
(iii) Claim sizes and claim counts are independent.
(iv) The number of claims in the first year was 1000.
(v) The aggregate loss in the first year was 6.75 million.
(vi) The manual premium for the first year was 5.00 million.
(vii) The exposure in the second year is identical to the exposure in the first year.
(viii) The full credibility standard is to be within 5% of the expected aggregate loss 95% of the time.

Determine the limited fluctuation credibility net premium (in millions) for the second year.

Problem 76.5 ‡
You are given:
(i) Claim counts follow a Poisson distribution.
(ii) claim size follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 1$.
(iii) A full credibility standard is established so that the actual number of claims will be within 5% of the expected number of claims 95% of the time.

Determine the number of expected claims needed for 30% partial credibility for the distribution of number of claims.

Problem 76.6 ‡
You are given:
(i) The full credibility standard is 100 expected claims.
(ii) The square-root rule is used for partial credibility.
You approximate the partial credibility formula with a Bühlmann credibility formula by selecting a Bühlmann k value that matches the partial credibility formula when 25 claims are expected.

Determine the credibility factor for the Bühlmann credibility formula when 100 claims are expected.

77 Greatest Accuracy Credibility Approach

Greatest accuracy credibility theory is a credibility theory based on Bayesian inference. In this section, we consider this model-based approach to the solution of the following credibility problem:

Let $X_1, X_2, \cdots, X_n$ be the abserved losses or claims in the past n years of a particular policyholder. Past experience had shown that the manual rate μ is not an appropriate measure of the premium for the next year ($\mu \neq \overline{X}$). This leads to the question of whether the net premium of next year should be based on μ alone, $\overline{X}$ alone, or a combination of μ and $\overline{X}$.

Greatest accuracy credibility theory addresses the following two questions:

(1) Is the policyholder is different than what was assumed in figuring out the manual rate μ?

(2) Has it been random chance that has been responsible in the difference between μ and $\overline{X}$?

Now when an individual applies for an insurance policy, there is usually a strict underwriting process by the insurer. In this process, a policyholder is rated and placed into a rating or risk class. One thinks that there is homogeneity in the rating process among the members of the same risk class. Unfortunately, reality shows otherwise. That is, there is heterogeneity in risk ratings amongst policyholders of the same rating class.

This assumption that a policyholder is different from members of the same rating class leads to the question of how an underwriter chooses what is an appropriate rate for the policyholder. In order to answer this question, we make the following assumptions:

(1) Every policyholder is characterized by a risk level θ within the rating class.

(2) The value of θ varies amongst policyholders in the same risk class. This assumption allows us to quantify the difference between policyholders with respect to the risk characteristics.

(3) θ can be viewed as a random variable on the set of risk levels with

a probability function $\pi(\theta)$ and cumulative distribution function $\Pi(\theta)$. In words, $\Pi(\theta)$ represents the probability that a policyholder picked at random from the risk class has a risk parameter less than or equal to θ.

While the risk parameter θ associated with an individual policyholder is not known, we will assume that $\pi(\theta)$ is known.

Because risk level varies in the population, the experience of a policyholder picked at random from the population arises from a two-stage process:

(1) The risk parameter θ is selected from the distribution $\pi(\theta)$ (prior distribution).

(2) Claims or losses are selected from the conditional distribution $f_{X|\Theta}(x|\theta)$. (model distribution)

Example 77.1
The amount of a claim $X|\Theta$ has the exponential distribution with parameter $\frac{1}{\theta}$. The risk parameter Θ has a gamma distribution with parameters α and β. Provide a mathematical description of this model.

Solution.
For the risk parameter, we have

$$\pi(\theta) = \frac{\theta^{\alpha-1}e^{-\frac{\theta}{\beta}}}{\beta^\alpha\Gamma(\alpha)}.$$

For the claims, we have

$$f_{X|\Theta}(x|\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}} \blacksquare$$

Example 77.2
The amount of a claim $X|\Lambda$ has the inverse exponential distribution with parameter λ The risk parameter Λ has a gamma distribution with parameters α and θ. Provide a mathematical description of this model.

Solution.
For the risk parameter, we have

$$\pi(\lambda) = \frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\theta}}}{\theta^\alpha\Gamma(\alpha)}.$$

For the claims, we have

$$f_{X|\Lambda}(x|\lambda) = \frac{\lambda}{x^2}e^{-\frac{\lambda}{x}} \blacksquare$$

Practice Problems

Problem 77.1
The amount of a claim $X|\Lambda$ has the Poisson distribution with parameter λ. The risk parameter Λ has a gamma distribution with parameters α and β.

Provide a mathematical description of this model.

Problem 77.2
The amount of a claim $X|\Theta$ has the normal distribution with parameters θ and σ_1^2. The risk parameter Θ has a normal distribution with parameters μ and σ_2^2.

Provide a mathematical description of this model.

Problem 77.3
The amount of a claim $X|\mathbf{Q}$ has the Binomial distribution with parameters (m, q) where m is known. The risk parameter $\mathbf{Q}$ has a beta distribution with parameters $(a, b, 1)$.

Provide a mathematical description of this model.

Problem 77.4
The amount of a claim $X|\Lambda$ has the exponential distribution with parameter $\frac{1}{\lambda}$. The risk parameter Λ has an inverse Gamma distribution with parameters α and θ.

Provide a mathematical description of this model.

Problem 77.5
The amount of a claim $X|\Lambda$ has the uniform distribution in $[0, \lambda]$. The risk parameter Λ s single parameter Pareto distribution with α and θ.

Provide a mathematical description of this model.

78 Conditional Distributions and Expectation

The credibility model of the previous section together with the credibility models to follow in the coming sections require a good understanding of conditional distributions and conditional expectation. In this section, a discussion of these topics are presented.

Suppose X and Y are two continuous random variables with joint density $f_{XY}(x,y)$. Let $f_{X|Y}(x|y)$ denote the probability density function of X given that $Y=y$. The **conditional density function** of X given $Y=y$ is

$$f_{X|Y}(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)}$$

provided that $f_Y(y)>0$.
Compare this definition with the discrete case where

$$p_{X|Y}(x|y)=\frac{p_{XY}(x,y)}{p_Y(y)}.$$

The **marginal distribution** of X is obtained by integrating out the joint distribution,

$$f_X(x)=\int_{-\infty}^{\infty}f_{XY}(x,y)dy$$

or

$$f_X(x)=\int_{-\infty}^{\infty}f_{X|Y}(x|y)f_Y(y)dy.$$

Note that

$$\int_{-\infty}^{\infty}f_{X|Y}(x|y)dx=\int_{-\infty}^{\infty}\frac{f_{XY}(x,y)}{f_Y(y)}dx=\frac{f_Y(y)}{f_Y(y)}=1.$$

Example 78.1
Suppose X and Y have the following joint density

$$f_{XY}(x,y)=\begin{cases}\frac{1}{2} & |X|+|Y|<1\\ 0 & \text{otherwise}\end{cases}$$

(a) Find the marginal distribution of X. That is, $f_X(x)$.
(b) Find the conditional distribution of Y given $X=\frac{1}{2}$.

Solution.
(a) Clearly, X only takes values in $(-1, 1)$. So $f_X(x) = 0$ if $|x| \geq 1$. Let $-1 < x < 1$,

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2} dy = \int_{-1+|x|}^{1-|x|} \frac{1}{2} dy = 1 - |x|.$$

(b) The conditional density of Y given $X = \frac{1}{2}$ is then given by

$$f_{Y|X}(y|x) = \frac{f(\frac{1}{2}, y)}{f_X(\frac{1}{2})} = \begin{cases} 1 & -\frac{1}{2} < y < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

Thus, $f_{Y|X}$ follows a uniform distribution on the interval $\left(-\frac{1}{2}, \frac{1}{2}\right)$ ■

Theorem 78.1
Continuous random variables X and Y with $f_Y(y) > 0$ are independent if and only if

$$f_{X|Y}(x|y) = f_X(x).$$

Proof.
Suppose first that X and Y are independent. Then $f_{XY}(x, y) = f_X(x)f_Y(y)$. Thus,

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x).$$

Conversely, suppose that $f_{X|Y}(x|y) = f_X(x)$. Then $f_{XY}(x, y) = f_{X|Y}(x|y)f_Y(y) = f_X(x)f_Y(y)$. This shows that X and Y are independent ■

Example 78.2
Let X and Y be two continuous random variables with joint density function

$$f_{XY}(x, y) = \begin{cases} c & 0 \leq y < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $f_X(x), f_Y(y)$ and $f_{X|Y}(x|1)$.
(b) Are X and Y independent?

Solution.
(a) We have

$$f_X(x) = \int_0^x c dy = cx, \quad 0 \leq x \leq 2$$

$$f_Y(y) = \int_y^2 c dx = c(2 - y), \quad 0 \leq y \leq 2$$

and

$$f_{X|Y}(x|1) = \frac{f_{XY}(x,1)}{f_Y(1)} = \frac{c}{c} = 1, \quad 1 \le x \le 2.$$

(b) Since $f_{X|Y}(x|1) \neq f_X(x)$, X and Y are dependent ■

Example 78.3
Show that

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}.$$

Solution.
We have

$$f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)(= f_{XY}(x,y)).$$

Dividing by sides of this equation by $f_Y(y)$, the result follows ■

We now turn our attention to conditional expectation. Let X and Y be random variables. We define **conditional expectation** of X given that $Y = y$ by

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx$$

where

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}.$$

Example 78.4
Suppose that the joint density of X and Y is given by

$$f_{XY}(x,y) = \frac{e^{-\frac{x}{y}}e^{-y}}{y}, \quad x, y > 0.$$

Compute $E(X|Y = y)$.

Solution.
The conditional density is found as follows

$$\begin{aligned}
f_{X|Y}(x|y) =& \frac{f_{XY}(x,y)}{f_Y(y)} \\
=& \frac{f_{XY}(x,y)}{\int_{-\infty}^{\infty} f_{XY}(x,y)dx} \\
=& \frac{(1/y)e^{-\frac{x}{y}}e^{-y}}{\int_0^{\infty}(1/y)e^{-\frac{x}{y}}e^{-y}dx} \\
=& \frac{(1/y)e^{-\frac{x}{y}}}{\int_0^{\infty}(1/y)e^{-\frac{x}{y}}dx} \\
=& \frac{1}{y}e^{-\frac{x}{y}}
\end{aligned}$$

Hence,

$$\begin{aligned}
E(X|Y=y) =& \int_0^{\infty} \frac{x}{y}e^{-\frac{x}{y}}dx = -\left[xe^{-\frac{x}{y}}\Big|_0^{\infty} - \int_0^{\infty} e^{-\frac{x}{y}}dx \right] \\
=& -\left[xe^{-\frac{x}{y}} + ye^{-\frac{x}{y}} \right]_0^{\infty} = y \ \blacksquare
\end{aligned}$$

Notice that if X and Y are independent then $f_{X|Y}(x|y) = f_X(x)$ so that $E(X|Y=y) = E(X)$.

Theorem 78.2 *(Double Expectation Property)*

$$E(X) = E(E(X|Y))$$

Proof.
We give a proof in the case X and Y are continuous random variables.

$$\begin{aligned}
E(E(X|Y)) =& \int_{-\infty}^{\infty} E(X|Y=y)f_Y(y)dy \\
=& \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx \right) f_Y(y)dy \\
=& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y)dxdy \\
=& \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x,y)dydx \\
=& \int_{-\infty}^{\infty} xf_X(x)dx = E(X) \ \blacksquare
\end{aligned}$$

Example 78.5
Suppose that $X|\Theta$ has a Poisson distribution with parameter θ and Θ has a Gamma distribution with parameters α and β. Find $E(X)$.

Solution.
We have

$$E(X) = E[E(X|\Theta)] = E(\Theta) = \alpha\beta \blacksquare$$

Now, for any function $g(x, y)$, the conditional expected value of g given $Y = y$ is, in the continuous case,

$$E(g(X,Y)|Y = y) = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)dx$$

if the integral exists.

Example 78.6
Show that

$$E[E(g(X,Y)|Y)] = E[g(X,Y)].$$

Solution.
We have

$$\begin{aligned}
E[E(g(X,Y)|Y)] =& \int_{-\infty}^{\infty} E(g(X,Y)|Y = y) f_Y(y)dy \\
=& \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)dx \right) f_Y(y)dy \\
=& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y) f_Y(y)dxdy \\
=& \int_{-\infty}^{\infty} g(x,y) \int_{-\infty}^{\infty} f_{XY}(x,y)dydx \\
=& \int_{-\infty}^{\infty} g(x,y) f_X(x)dx = E[g(X,Y)] \blacksquare
\end{aligned}$$

The Conditional Variance
Next, we introduce the concept of conditional variance. Just as we have defined the conditional expectation of X given that $Y = y$, we can define the conditional variance of X given Y as follows

$$\text{Var}(X|Y = y) = E[(X - E(X|Y))^2|Y = y].$$

Note that the conditional variance is a random variable since it is a function of Y.

Proposition 78.1
Let X and Y be random variables. Then
(a) $\text{Var}(X|Y) = E(X^2|Y) - [E(X|Y)]^2$
(b) $E(\text{Var}(X|Y)) = E[E(X^2|Y) - (E(X|Y))^2] = E(X^2) - E[(E(X|Y))^2]$.
(c) $\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (E(X))^2$.
(d) Law of Total Variance: $\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E(X|Y))$.

Proof.
(a) We have

$$\begin{aligned}\text{Var}(X|Y) =&E[(X - E(X|Y))^2|Y]\\ =&E[(X^2 - 2XE(X|Y) + (E(X|Y))^2|Y]\\ =&E(X^2|Y) - 2E(X|Y)E(X|Y) + (E(X|Y))^2\\ =&E(X^2|Y) - [E(X|Y)]^2\end{aligned}$$

(b) Taking E of both sides of the result in (a) we find

$$E(\text{Var}(X|Y)) = E[E(X^2|Y) - (E(X|Y))^2] = E(X^2) - E[(E(X|Y))^2].$$

(c) Since $E(E(X|Y)) = E(X)$ we have

$$\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (E(X))^2.$$

(d) The result follows by adding the two equations in (b) and (c) ■

Example 78.7
Suppose that X and Y have joint distribution

$$f_{XY}(x,y) = \begin{cases} \frac{3y^2}{x^3} & 0 < y < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(X),\ E(X^2), Var(X), E(Y|X), Var(Y|X), E[Var(Y|X)], Var[E(Y|X)]$, and $Var(Y)$.

Solution.
First we find marginal density functions.

$$f_X(x) = \int_0^x \frac{3y^2}{x^3}dy = 1, \quad 0 < x < 1$$

$$f_Y(y) = \int_y^1 \frac{3y^2}{x^3}dx = \frac{3}{2}(1 - y^2), \quad 0 < y < 1$$

Now,

$$E(X) = \int_0^1 x dx = \frac{1}{2}$$

$$E(X^2) = \int_0^1 x^2 dx = \frac{1}{3}$$

Thus,

$$Var(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Next, we find conditional density of Y given $X = x$

$$f_{Y|X}(x|y) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{3y^2}{x^3}, \quad 0 < x < y < 1$$

Hence,

$$E(Y|X = x) = \int_0^x \frac{3y^3}{x^3} dx = \frac{3}{4}x$$

and

$$E(Y^2|X = x) = \int_0^x \frac{3y^4}{x^3} dx = \frac{3}{5}x^2$$

Thus,

$$Var(Y|X = x) = E(Y^2|X = x) - [E(Y|X = x)]^2 = \frac{3}{5}x^2 - \frac{9}{16}x^2 = \frac{3}{80}x^2$$

Also,

$$Var[E(Y|X)] = Var\left(\frac{3}{4}x\right) = \frac{9}{16}Var(X) = \frac{9}{16} \times \frac{1}{12} = \frac{3}{64}$$

and

$$E[Var(Y|X)] = E\left(\frac{3}{80}X^2\right) = \frac{3}{80}E(X^2) = \frac{3}{80} \times \frac{1}{3} = \frac{1}{80}.$$

Finally,

$$Var(Y) = Var[E(Y|X)] + E[Var(Y|X)] = \frac{19}{320} \blacksquare$$

Example 78.8 ‡

An actuary for an automobile insurance company determines that the distribution of the annual number of claims for an insured chosen at random is modeled by the negative binomial distribution with mean 0.2 and variance 0.4.

The number of claims for each individual insured has a Poisson distribution

and the means of these Poisson distributions are gamma distributed over the population of insureds.
Calculate the variance of this gamma distribution

Solution.
Let N be the annual number of claims. We are given that $E(N) = E(N|\Gamma) = E(\Gamma) = 0.2$. By the law of total variance, we have

$$0.4 = \text{Var}(N) = E[\text{Var}(N|\Gamma)] + \text{Var}(E(N|\Gamma)) = E(\Gamma) + \text{Var}(\Gamma).$$

Solving for $\text{Var}(\Gamma)$ we find $\text{Var}(\Gamma) = 0.4 - 0.2 = 0.2$ ■

Practice Problems

Problem 78.1
Suppose that X is uniformly distributed on the interval $[0,1]$ and that, given $X=x,\ Y$ is uniformly distributed on the interval $[1-x,1]$.
(a) Determine the joint density $f_{XY}(x,y)$.
(b) Find the probability $P(Y\geq\frac{1}{2})$.

Problem 78.2
The joint density of X and Y is given by

$$f_{XY}(x,y)=\begin{cases}\frac{15}{2}x(2-x-y) & 0\leq x,y\leq 1\\ 0 & \text{otherwise}\end{cases}$$

Compute the conditional density of X, given that $Y=y$ for $0\leq y\leq 1$.

Problem 78.3
The joint density function of X and Y is given by

$$f_{XY}(x,y)=\begin{cases}\frac{e^{-\frac{x}{y}}e^{-y}}{y} & x\geq 0,y\geq 0\\ 0 & \text{otherwise}\end{cases}$$

Compute $P(X>1|Y=y)$.

Problem 78.4
Let Y be a random variable with a density f_Y given by

$$f_Y(y)=\begin{cases}\frac{\alpha-1}{y^\alpha} & y>1\\ 0 & \text{otherwise}\end{cases}$$

where $\alpha>1$. Given $Y=y$, let X be a random variable which is Uniformly distributed on $(0,y)$.
(a) Find the marginal distribution of X.
(b) Calculate $E(Y|X=x)$ for every $x>0$.

Problem 78.5
Suppose that X and Y have joint distribution

$$f_{XY}(x,y)=\begin{cases}8xy & 0<x<y<1\\ 0 & \text{otherwise}\end{cases}$$

Find $E(X|Y)$ and $E(Y|X)$.

Problem 78.6
Suppose that X and Y have joint distribution

$$f_{XY}(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(Y)$ in two ways.

Problem 78.7
The stock prices of two companies at the end of any given year are modeled with random variables X and Y that follow a distribution with joint density function

$$f_{XY}(x,y) = \begin{cases} 2x & 0 < x < 1, x < y < x+1 \\ 0 & \text{otherwise} \end{cases}$$

What is the conditional variance of Y given that $X = x$?

Problem 78.8
Let X be a random variable with mean 3 and variance 2, and let Y be a random variable such that for every x, the conditional distribution of Y given $X = x$ has a mean of x and a variance of x^2.

What is the variance of the marginal distribution of Y?

Problem 78.9
The number of stops X in a day for a delivery truck driver is Poisson with mean λ. Conditional on their being $X = x$ stops, the expected distance driven by the driver Y is Normal with a mean of αx miles, and a standard deviation of βx miles.

Give the mean and variance of the numbers of miles she drives per day.

79 Bayesian Credibility with Discrete Prior

Let's recall the credibility problem in Section 77: For a particular policyholder, we have the observed past losses $X_1, X_2, \cdots, X_n$ and we are interested in setting the premium to cover the loss of the next exposure unit (next year) X_{n+1}. We assume that the risk parameter θ (which is unknown) associated with the policyholder comes from a **prior distribution** $\pi(\theta)$ and that the losses $X_1, X_2, \cdots, X_{n+1}$ are conditionally independent, that is the $X_i|\Theta$ are independent, but not necessarily identically distributed.

In order to predict the claim in the $(n+1)$ period we normally condition on θ, but θ for the $n+1$ period is unknown so we condition on X_{n+1}. The resulting distribution is the **predictive distribution** $f_{X_{n+1}|\mathbf{X}}(x_{n+1}(x_{n+1}|\mathbf{x})$ (introduced in Section 64) where $\mathbf{X} = (X_1, X_2, \cdots, X_n)^T$ and $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$.

Using the terms discussed in Section 64, the **model distribution** of $\mathbf{X}$ given Θ is

$$f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta) = \prod_{j=1}^{n} f_{X_j|\theta}(x_j|\theta).$$

The **joint distribution** of $\mathbf{X}$ given Θ is

$$f_{\mathbf{X},\Theta}(\mathbf{x}, \theta) = \left[\prod_{j=1}^{n} f_{X_j|\theta}(x_j|\theta)\right] \pi(\theta).$$

The **marginal distribution** of $\mathbf{X}$ is obtained by integrating the joint distribution with respect to θ to obtain

$$f_{\mathbf{X}}(\mathbf{x}) = \int \left[\prod_{j=1}^{n} f_{X_j|\theta}(x_j|\theta)\right] \pi(\theta) d\theta.$$

The **posterior distribution** is by Bayes' Theorem

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \frac{f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)}{f_{\mathbf{X}}(\mathbf{x})} = \frac{1}{f_{\mathbf{X}}(\mathbf{x})} \left[\prod_{j=1}^{n} f_{X_j|\theta}(x_j|\theta)\right] \pi(\theta).$$

Thus, the predictive distribution can be expressed as

$$\begin{aligned}
f_{X_{n+1}|\mathbf{X}}(x_{n+1}|\mathbf{x}) =& \frac{1}{f_{\mathbf{X}}(\mathbf{x})}\int\left[\prod_{j=1}^{n+1} f_{X_j|\theta}(x_j|\theta)\right]\pi(\theta) \\
=& \int f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\frac{\left[\prod_{j=1}^{n} f_{X_j|\theta}(x_j|\theta)\right]\pi(\theta)}{f_{\mathbf{X}}(\mathbf{x})}d\theta \\
=& \int f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta.
\end{aligned}$$

The mean of the predicitive distribution, also known as the **Bayesian premium** or **Bayesian estimate**, is what we would charge to cover the loss X_{n+1}. It is given by

$$\begin{aligned}
E[X_{n+1}|\mathbf{X}=\mathbf{x}] =& \int x_{n+1}f_{X_{n+1}|\mathbf{X}}(x_{n+1}|\mathbf{x})dx_{n+1} \\
=& \int x_{n+1}\left[\int f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta\right]dx_{n+1} \\
=& \int\left[\int x_{n+1}f_{X_{n+1}|\Theta}(x_{n+1}|\theta)dx_{n+1}\right]\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta \\
=& \int \mu_{n+1}(\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta
\end{aligned}$$

where

$$\mu_{n+1}(\theta) = E(X_{n+1}|\Theta=\theta) = \int x_{n+1}f_{X_{n+1}|\Theta}(x_{n+1}|\theta)dx_{n+1}$$

is the **hypothetical mean**, the premium we charge if we knew θ. Taking the expectation of the previous equation, we find

$$E(X_{n+1}) = E[E(X_{n+1}|\Theta)] = E[\mu_{n+1}(\Theta)].$$

This is the premium we charge if we knew nothing about the policyholder. It is called the **pure** or **collective premium.**. Note that this premium is independent of the risk parameter θ and does not depend on the data collected from the policyholder.

Remark 79.1
In the case Θ is discrete, the integrals above are replaced by sums.

Example 79.1 ‡
You are given the following for a dental insurer:
(I) Claim counts for individual insureds follow a Poisson distribution.
(ii) Half of the insureds are expected to have 2.0 claims per year.
(iii) The other half of the insureds are expected to have 4.0 claims per year.
A randomly selected insured has made 4 claims in each of the first two policy years. Determine the Bayesian estimate of this insured's claim count in the next (third) policy year.

Solution.
Let X be the claim count for an individual. We are given that conditional claim count $X|\Theta$ is Poisson with mean Θ. Let X_n be the number of claims in year n. We want to find $E(X_3|X_1, X_2)$ where $x_1 = x_2 = 4$.
The prior distribution is: $\pi(2) = 0.5$ and $\pi(4) = 0.5$.
The marginal distribution is

$$\begin{aligned}f_{\mathbf{X}}(4,4) =&f_{X_1|\Theta}(4|\theta=2)f_{X_2|\Theta}(4|\theta=2)\pi(2) + f_{X_1|\Theta}(4|\theta=4)f_{X_2|\Theta}(4|\theta=4)\pi(4)\\ =&\left(\frac{e^{-2}2^4}{4!}\right)^2(0.5) + \left(\frac{e^{-4}4^4}{4!}\right)^2(0.5)\\ =&0.02315.\end{aligned}$$

The posterior distribution is:

$$\pi_{\Theta|\mathbf{X}}(2|\mathbf{x}) = \frac{f_{\mathbf{X}|\Theta}(\mathbf{x}|2)\pi(2)}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\left(\frac{e^{-2}2^4}{4!}\right)^2(0.5)}{0.02315} = 0.1758$$

and

$$\pi_{\Theta|\mathbf{X}}(4|\mathbf{x}) = 1 - 0.1758 = 0.8242.$$

Finally, the Bayesian premium is

$$\begin{aligned}E(X_3|X_1,X_2) =&E(X_3|\Theta=2)\pi_{\Theta|\mathbf{X}}(2|\mathbf{x}) + E(X_3|\Theta=4)\pi_{\Theta|\mathbf{X}}(4|\mathbf{x})\\ =&2(0.1758) + 4(0.8242) = 3.6484 \blacksquare\end{aligned}$$

Example 79.2 ‡
You are given:

Class	Number of Insureds	Claim Count Probabilities 0	1	2	3	4
1	3000	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0
2	2000	0	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	0
3	1000	0	0	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$

A randomly selected insured has one claim in Year 1. Determine the expected number of claims in Year 2 for that insured.

Solution.
Let X_n denote the number of claims in Year n. We are asked to find $E(X_2|X_1=1)$. The parameter θ stands for the class. The prior distribution is $\pi(1)=\frac{3000}{6000}=\frac{1}{2}$, $\pi(2)=\frac{1}{3}$, and $\pi(3)=\frac{1}{6}$. The marginal distribution evaluated at $x_1=1$ is

$$\begin{aligned}f(1)=&f(1|1)\pi(1)+f(1|2)\pi(2)+f(1|3)\pi(3)\\=&\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{6}\cdot\frac{1}{3}=\frac{2}{9}.\end{aligned}$$

The posterior distribution is

$$\begin{aligned}\pi(1|1)=&\frac{f(1|1)\pi(1)}{f(1)}=\frac{\frac{1}{6}}{\frac{2}{9}}=\frac{3}{4}\\\pi(2|1)=&\frac{f(1|2)\pi(2)}{f(1)}=\frac{\frac{1}{18}}{\frac{2}{9}}=\frac{1}{4}\\\pi(3|1)=&\frac{f(3|1)\pi(1)}{f(1)}=\frac{0}{\frac{2}{9}}=0.\end{aligned}$$

Thus,

$$\begin{aligned}E(X_2|X_1=1)=&E(X_2|1)\pi(1|1)+E(X_2|2)\pi(2|1)+E(X_2|3)\pi(3|1)\\=&\left[1\left(\frac{1}{3}\right)+2\left(\frac{1}{3}\right)\right]\left(\frac{3}{4}\right)+\left[1\left(\frac{1}{6}\right)+2\left(\frac{2}{3}\right)+3\left(\frac{1}{6}\right)\right]\left(\frac{1}{4}\right)\\=&1.25\ \blacksquare\end{aligned}$$

Example 79.3 ‡
You are given the following information about six coins:

Coin (θ)	Probability of Heads
1 - 4	0.50
5	0.25
6	0.75

A coin is selected at random and then flipped repeatedly. Let X_i denote the outcome of the i^{th} flip, where "1" indicates heads and "0" indicates tails. The following sequence is obtained:

$$S=\{X_1,X_2,X_3,X_4\}=\{1,1,0,1\}.$$

Determine $E(X_5|S)$ using Bayesian analysis.

Solution.
The prior distribution is

$$\pi(\theta_1) = \tfrac{4}{6} \quad \pi(\theta_2) = \tfrac{1}{6} \quad \pi(\theta_3) = \tfrac{1}{6}.$$

The model distribution is

$$\begin{aligned}
f(S|\theta_1) =&(0.5)^4 = 0.0625\\
f(S|\theta_2) =&(0.25)^2(0.75)(0.25) = 0.011719\\
f(S|\theta_3) =&(0.75)^2(0.25)(0.75) = 0.105469.
\end{aligned}$$

The joint distribution is

$$\begin{aligned}
f(S,\theta_1) =&\frac{4}{6}(0.0625) = 0.04167\\
f(S,\theta_2) =&\frac{1}{6}(0.011719) = 0.00195\\
f(S,\theta_3) =&\frac{1}{6}(0.105469) = 0.01758.
\end{aligned}$$

The posterior distribution is

$$\begin{aligned}
f(\theta_1|S) =&\frac{0.04167}{0.04167 + 0.00195 + 0.01758} = 0.68088\\
f(\theta_2|S) =&\frac{0.00195}{0.04167 + 0.00195 + 0.01758} = 0.03186\\
f(\theta_3|S) =&\frac{0.01758}{0.04167 + 0.00195 + 0.01758} = 0.28726.
\end{aligned}$$

Thus,

$$\begin{aligned}
E(X_5|S) =&E(X_3|\theta_1)f(\theta_1|S) + E(X_3|\theta_2)f(\theta_2|S) + E(X_3|\theta_3)f(\theta_3|S)\\
=&0.5(0.68088) + 0.25(0.03186) + 0.75(0.28726) = 0.5639 \blacksquare
\end{aligned}$$

Example 79.4 ‡
For a particular policy, the conditional probability of the annual number of claims given $\Theta = \theta$, and the probability distribution of Θ are as follows:

Number of claims	0	1	2
Probability	2θ	θ	$1 - 3\theta$

θ	0.10	0.30
Probability	0.80	0.20

One claim is observed in Year 1. Calculate the Bayesian credibility estimate of the number of claims in Year 2.

Solution.
The marginal distribution is Let N denote the annual number of claims. We have

$$f_N(1) = f(1|0.10)\pi(0.10) + f(1|0.30)\pi(0.30) = (0.10)(0.80) + (0.30)(0.20) = 0.14.$$

The posterior distribution is

$$\begin{aligned}\pi_{\Theta|N}(0.10|1) =& \frac{0.80(0.10)}{0.14} = \frac{4}{7}\\ \pi_{\Theta|N}(0.30|1) =& \frac{0.20(0.30)}{0.14} = \frac{3}{7}.\end{aligned}$$

The Bayesian credibility estimate of the number of claims in Year 2 is

$$\begin{aligned}E(N_2|N_1) =& E(N_2|\Theta = 0.10)\pi_{\Theta|N}(0.10|1) + E(N_2|\Theta = 0.30)\pi_{\Theta|N}(0.30|1)\\ =& [0(2)(0.10) + 1(0.10) + 2(1 - 3 \times 0.10)](4/7)\\ &+ [0(2)(0.30) + 1(0.30) + 2(1 - 3 \times 0.30)](3/7) = 1.071 \blacksquare\end{aligned}$$

Example 79.5 ‡
You are given:
(i) The claim count and claim size distributions for risks of type A are:

Number of Claims	Probabilities
0	4/9
1	4/9
2	1/9

Claim Size	Probabilities
500	1/3
1235	2/3

(ii) The claim count and claim size distributions for risks of type B are:

Number of Claims	Probabilities
0	1/9
1	4/9
2	4/9

Claim Size	Probabilities
250	2/3
328	1/3

(iv) Claim counts and claim sizes are independent within each risk type.
A randomly selected risk is observed to have total annual losses of 500.
Determine the Bayesian premium for the next year for this same risk.

Solution.
The prior parameter represents the type of risk so that either $\Theta = A$ or $\Theta = B$. The prior distribution is $\Pr(A) = \Pr(B) = 0.5$. For each of the two classes, total annual loss L has a compound distribution. We want

$$E(L_2|L_1 = 500) = E(L_2|A)\Pr(A|L_1 = 500) + E(L_2|B)\Pr(B|L_1 = 500).$$

We have

$$\begin{aligned}
E(L|A) =& E(N|A)E(X|A) = [0(4/9) + 1(4/9) + 2(1/9)][500(1/3) + 1235(2/3)] = 660\\
E(L|B) =& E(N|B)E(X|B) = [0(1/9) + 1(4/9) + 2(4/9)][250(2/3) + 328(1/3)] = 368\\
\Pr(A|L = 500) =& \frac{\Pr(L = 500|A)\Pr(A)}{\Pr(L = 500|A)\Pr(A) + \Pr(L = 500|B)\Pr(B)}\\
\Pr(L = 500|A) =& \Pr(N = 1)\Pr(X) = 500)\\
=& (4/9)(1/3) = 4/27\\
\Pr(L = 500|B) =& \Pr(N = 2)[\Pr(X) = 250)]^2\\
=& (4/9)(2/3)^2 = 16/81\\
\Pr(A|L = 500) =& \frac{(4/27)(1/2)}{(4/27)(1/2) + (16/81)(1/2)} = 3/7\\
\Pr(B|L = 500) =& 1 - 3/7 = 4/7.
\end{aligned}$$

Thus,

$$\begin{aligned}
E(L_2|L_1 = 500) =& E(L_2|A)\Pr(A|L_1 = 500) + E(L_2|B)\Pr(B|L_1 = 500)\\
=& 660(3/7) + 368(4/7) = 493 \blacksquare
\end{aligned}$$

Example 79.6 ‡
Two eight-sided dice, A and B, are used to determine the number of claims for an insured. The faces of each die are marked with either 0 or 1, representing the number of claims for that insured for the year.

Die	Pr(Claims=0)	Pr(Claims=1)
A	1/4	3/4
B	3/4	1/4

Two spinners, X and Y, are used to determine claim cost. Spinner X has two areas marked 12 and c. Spinner Y has only one area marked 12.

Spinner	Pr(Cost=12)	Pr(Ccost=c)
X	1/2	1/2
Y	1	0

To determine the losses for the year, a die is randomly selected from A and B and rolled. If a claim occurs, a spinner is randomly selected from X and Y and spun. For subsequent years, the same die and spinner are used to determine losses.
Losses for the first year are 12. Based upon the results of the first year, you determine that the expected losses for the second year are 10.
Calculate c.

Solution.
The prior parameter Θ can be one of AX, BX, AY, and BY. The prior distribution is

$$\pi(AX) = \pi(AY) = \pi(BX) = \pi(BY) = \frac{1}{4}$$

We have

$$\begin{aligned}
10 =& E(L_2|L_1 = 12) = E(L_2|AX)\Pr(AX|L_1 = 12) + E(L_2|AY)\Pr(AY|L_1 = 12) \\
&+ E(L_2|BX)\Pr(BX|L_1 = 12) + E(L_2|BY)\Pr(BY|L_1 = 12) \\
E(L_2|AX) =& 12\left(\frac{1}{2}\right)\left(\frac{3}{4}\right) + c\left(\frac{1}{2}\right)\left(\frac{3}{4}\right) \\
=& \frac{3}{8}(12 + c) \\
E(L_2|AY) =& 12(1)\left(\frac{3}{4}\right) + c(0)\left(\frac{3}{4}\right) = 9 \\
E(L_2|BX) =& 12\left(\frac{1}{2}\right)\left(\frac{1}{4}\right) + c\left(\frac{1}{2}\right)\left(\frac{1}{4}\right) \\
=& \frac{1}{8}(12 + c) \\
E(L_2|BY) =& 12(1)\left(\frac{1}{4}\right) + c(0)\left(\frac{1}{4}\right) = 3 \\
\Pr(L_1 = 12) =& \Pr(L_1 = 12|AX)\Pr(AX) + \Pr(L_1 = 12|AY)\Pr(AY) \\
=& \Pr(L_1 = 12|BX)\Pr(BX) + \Pr(L_1 = 12|BY)\Pr(BY) \\
=& \left(\frac{3}{4}\right)\left(\frac{1}{2}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)(1)\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)\left(\frac{1}{2}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)(1)\left(\frac{1}{4}\right) \\
=& \frac{3}{8}
\end{aligned}$$

$$\Pr(AX|L_1=12)=\frac{\Pr(L_1=12|AX)\Pr(AX)}{\Pr(L_1=12)}=\frac{\left(\frac{3}{4}\right)\left(\frac{1}{2}\right)\left(\frac{1}{4}\right)}{\frac{3}{8}}=\frac{1}{4}$$
$$\Pr(AY|L_1=12)=\frac{\Pr(L_1=12|AY)\Pr(AY)}{\Pr(L_1=12)}=\frac{\left(\frac{3}{4}\right)(1)\left(\frac{1}{4}\right)}{\frac{3}{8}}=\frac{1}{2}$$
$$\Pr(BX|L_1=12)=\frac{\Pr(L_1=12|BX)\Pr(BX)}{\Pr(L_1=12)}=\frac{\left(\frac{1}{4}\right)\left(\frac{1}{2}\right)\left(\frac{1}{4}\right)}{\frac{3}{8}}=\frac{1}{12}$$
$$\Pr(BY|L_1=12)=\frac{\Pr(L_1=12|BY)\Pr(BY)}{\Pr(L_1=12)}=\frac{\left(\frac{1}{4}\right)(1)\left(\frac{1}{4}\right)}{\frac{3}{8}}=\frac{1}{6}.$$

Thus,

$$\frac{3}{8}(12+c)\left(\frac{1}{4}\right)+9\left(\frac{1}{2}\right)+\frac{1}{8}(12+c)\left(\frac{1}{12}\right)+3\left(\frac{1}{6}\right)=c.$$

Solving this equation, we find $c=36$ ■

Example 79.7 ‡
For a risk, you are given:
(i) The number of claims during a single year follows a Bernoulli distribution with mean p.
(ii) The prior distribution for p is uniform on the interval $[0,1]$.
(iii) The claims experience is observed for a number of years.
(iv) The Bayesian premium is calculated as $1/5$ based on the observed claims.
Which of the following observed claims data could have yielded this calculation?
(A) 0 claims during 3 years
(B) 0 claims during 4 years
(C) 0 claims during 5 years
(D) 1 claim during 4 years
(E) 1 claim during 5 years

Solution.
Let x_i be the number of claims in year i where $i=1,2,\cdots,n$ and $x_i=0,1$. Let $x=x_1+\cdots+x_n$ be the number of claims in n years. We have that $X_i|p$ is a Bernoulli distribution with probability function

$$f(X_i|p)=p^{x_i}(1-p)^{1-x_i},\ x_i=0,1.$$

The Bayesian premium is

$$E(X_{n+1}|x_1,x_2,\cdots,x_n)=\int_0^1 E(X_{n+1}|p)f(p|x_1,x_2,\cdots,x_n)dp=\int_0^1 pf(p|x_1,x_2,\cdots,x_n)dp.$$

We have

$$\begin{aligned}
f(x_1, x_2, \cdots, x_n|p) =& \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
=& p^x(1-p)^{n-x} \\
f(p|x_1, x_2, \cdots, x_n) =& \frac{f(x_1, x_2, \cdots, x_n|p)\pi(p)}{f(x_1, x_2, \cdots, x_n)} \\
=& \frac{p^x(1-p)^{n-x}}{f(x_1, x_2, \cdots, x_n)} \\
f(x_1, x_2, \cdots, x_n) =& \int_0^1 p^x(1-p)^{n-x}dp \\
=& \frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)} \int_0^1 \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} p^{x+1}(1-p)^{n-x+1-1}\frac{1}{p}dp \\
=& \frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)} \\
f(p|x_1, x_2, \cdots, x_n) =& \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} p^{x+1}(1-p)^{n-x+1-1}\frac{1}{p}.
\end{aligned}$$

Thus, the posterior distribution is a beta distribution with $a = x+1, b = n-x+1$, and $\theta = 1$. Moreover, $E(X_{n+1}|x_1, x_2, \cdots, x_n)$ is the mean of this distribution which implies

$$\frac{x+1}{n+2} = \frac{1}{5}.$$

This is satisfied for $x = 0$ and $n = 3$ so the answer is (A) ■

Practice Problems

Problem 79.1
Drivers are classified as good (G), average (A), or bad (B).
• Good drivers make up 70% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.65, 1 claim is 0.25, and 2 claims is 0.10.
• Average drivers make up 20% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.40, 1 claim is 0.40, and 2 claims is 0.20.
• Bad drivers make up 10% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.50, 1 claim is 0.30, and 2 claims is 0.20.
For a policyholder, the risk parameter is the classification of the individual as G, A, or B. For a particular policyholder, it has been observed that $x_1 = 1$ and $x_2 = 2$.
(a) Write the prior distribution of this model.
(b) Find the model distribution of $\mathbf{x} = (1, 2)^T$.

Problem 79.2
In Problem 79.1, answer the following questions:
(a) Find the marginal probability of $\mathbf{X}$.
(b) Find the joint distribution of X_1, X_2, X_3 given $\mathbf{x} = (1, 2)^T$.

Problem 79.3
In Problem 79.1, answer the following questions:
(a) Find the predictive distribution given $\mathbf{x} = (1, 2)^T$.
(b) Find the posterior probabilities.

Problem 79.4
In Problem 79.1, answer the following questions:
(a) Determine the hypothetical means.
(b) Determine the pure of the collective premium.

Problem 79.5
In Problem 79.1, answer the following questions:
(a) Determine the Bayesian premium without using the hypothetical means.
(b) Determine the Bayesian premium by using the hypothetical means.

Problem 79.6 ‡
In a certain town the number of common colds an individual will get in a

year follows a Poisson distribution that depends on the individual's age and smoking status. The distribution of the population and the mean number of colds are as follows:

	Proportion of population	Mean number of colds
Children (C)	0.30	3
Adult Non-Smokers (ANS)	0.60	1
Adult Smokers (AS)	0.10	4

Calculate the conditional probability that a person with exactly 3 common colds in a year is an adult smoker.

Problem 79.7 ‡
You are given:
(i) The annual number of claims on a given policy has the geometric distribution with parameter β.
(ii) One-third of the policies have $\beta = 2$, and the remaining two-thirds have $\beta = 5$.
A randomly selected policy had two claims in Year 1.
Calculate the Bayesian expected number of claims for the selected policy in Year 2.

Problem 79.8 ‡
An insurance company sells three types of policies with the following characteristics:

Type of Policy	Proportion of Total Policies	Annual Claim Frequency
I	5%	Poisson with $\lambda = 0.25$
II	20%	Poisson with $\lambda = 0.50$
III	75%	Poisson with $\lambda = 1.00$

A randomly selected policyholder is observed to have a total of one claim for Year 1 through Year 4.
For the same policyholder, determine the Bayesian estimate of the expected number of claims in Year 5.

Problem 79.9 ‡
You are given:
(i) Claim sizes follow an exponential distribution with mean θ.
(ii) For 80% of the policies, $\theta = 8$.

(iii) For 20% of the policies, $\theta = 2$.
A randomly selected policy had one claim in Year 1 of size 5.
Calculate the Bayesian expected claim size for this policy in Year 2.

Problem 79.10 ‡
You are given:
(i) Two classes of policyholders have the following severity distributions:

Amount of claim	Probability of claim amount for Class 1	Probability of claim amount for Class 2
250	0.5	0.7
2500	0.3	0.2
60000	0.2	0.1

(ii) Class 1 is twice as likely to be observed as Class 2.
A claim of 250 is observed.
Determine the Bayesian estimate of the expected value of a second claim from the same policyholder.

Problem 79.11 ‡
You are given:
(i) An individual automobile insured has annual claim frequencies that follow a Poisson distribution with mean λ.
(ii) An actuary's prior distribution for the parameter λ has probability density function:

$$f(\lambda = 0.5[5e^{-5\lambda} + \frac{1}{5}e^{-\frac{\lambda}{5}}].$$

(iii) In the first policy year, no claims were observed for the insured.
Determine the expected number of claims in the second policy year.

80 Bayesian Credibility with Continuous Prior

In this section, we consider the Bayesian credibility approach to models where the prior distribution is continuous.

Example 80.1
Claim amount is assumed to be exponential with mean $\frac{1}{\Theta}$. The prior distribution Θ is assumed to be Gamma with parameters $\alpha = 5$ and $\beta = 0.0005$. Suppose a person has claims in the amount of \$2000,\$1000, and \$3000.
(a) Provide a mathematical description of this model.
(b) Determine the predictive distribution of the fourth claim.
(c) Determine the posterior distribution of Θ.
(d) Determine the Bayesian premium without using the hypothetical means.
(e) Determine the Bayesian premium by using the hypothetical means.

Solution.
(a) The claims amount distribution (model distribution) is given by

$$f_{\mathbf{X}|\Theta}(x|\theta) = \theta e^{-\theta x}.$$

The risk parameter (prior) distribution is given by

$$\pi(\theta) = \frac{2000^5 \theta^4 e^{-2000\theta}}{24}.$$

(b) The marginal density at the observed values is

$$\begin{aligned}
f(2000, 1000, 3000) =& \int_0^\infty (\theta e^{-2000\theta})(\theta e^{-1000\theta})(\theta e^{-3000\theta}) \frac{\theta^1 2000^2 e^{-2000\theta}}{4!} d\theta \\
=& \frac{2000^5}{4!} \int_0^\infty \theta^7 e^{-8000\theta} d\theta \\
=& \frac{2000^5}{8000^8} \frac{7!}{4!} \underbrace{\int_0^\infty \frac{\theta^7 8000^8 e^{-8000\theta}}{\Gamma(8)} d\theta}_{1} \\
=& \frac{7!}{4!} \frac{2000^5}{8000^8}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
f(2000,1000,3000,x_4) =& \int_0^\infty (\theta e^{-2000\theta})(\theta e^{-1000\theta})(\theta e^{-3000\theta})(\theta e^{-x_4\theta})\frac{\theta^1 2000^2 e^{-2000\theta}}{4!}d\theta \\
=& \frac{2000^5}{4!}\int_0^\infty \theta^8 e^{-\theta(8000+x_4)}d\theta \\
=& \frac{2000^5}{(8000+x_4)^9}\frac{8!}{4!}\underbrace{\int_0^\infty \frac{(8000+x_4)^9\theta^8 e^{-\theta(8000+x_4)}}{\Gamma(9)}d\theta}_{1} \\
=& \frac{8!}{4!}\frac{2000^5}{(8000+x_4)^9}.
\end{aligned}$$

The predictive distribution is given by

$$f(x_4|2000,1000,3000) = \frac{\frac{8!}{4!}\frac{2000^5}{(8000+x_4)^9}}{\frac{7!}{4!}\frac{2000^5}{8000^8}} = \frac{8(8000^8)}{(8000+x_4)^9}$$

which is a type 2 Pareto distribution with parameters $\alpha 8$ and $\beta = 8000$.
(c) The posterior distribution of Θ is given by

$$\begin{aligned}
\pi(\theta|2000,1000,3000) =& \frac{f(2000,1000,3000,\theta)}{f(2000,1000,3000)} \\
=& \frac{(\theta e^{-2000\theta})(\theta e^{-1000\theta})(\theta e^{-3000\theta})\frac{\theta^1 2000^2 e^{-2000\theta}}{4!}}{\frac{7!}{4!}\frac{2000^5}{8000^8}} \\
=& \frac{\frac{2000^5}{4!}\theta^7 e^{-8000\theta}}{\frac{7!}{4!}\frac{2000^5}{8000^8}} \\
=& \frac{8000^8}{7!}\theta^7 e^{-8000\theta}.
\end{aligned}$$

(d) We have

$$\begin{aligned}
E[X_4|2000,1000,3000] =& \underbrace{\int_0^\infty x_4\frac{8(8000^8)}{(8000+x_4)^9}dx_4}_{\text{mean of Pareto}} \\
=& \frac{8000}{8-1} = \frac{8000}{7}.
\end{aligned}$$

(e) We have

$$\begin{aligned}
E[X_4|2000,1000,3000] =& \int_0^\infty \mu_5(\theta)\pi(\theta|2000,1000,3000)d\theta \\
=& \int_0^\infty \frac{1}{\theta}\frac{8000^8}{7!}\theta^7 e^{-8000\theta} \\
=& \frac{8000^8}{7!}\frac{6!}{8000^7}\underbrace{\int_0^\infty \frac{8000^7\theta^6 e^{-8000\theta}}{\Gamma(7)}d\theta}_{1} \\
=& \frac{8000}{7} \blacksquare
\end{aligned}$$

Example 80.2 ‡
You are given:
(i) The number of claims for each policyholder has a binomial distribution with parameters $m = 8$ and q.
(ii) The prior distribution of q is beta with parameters a (unknown), $b = 9$, and $\theta = 1$.
(iii) A randomly selected policyholder had the following claims experience:

Year	Number of Claims
1	2
2	k

(iv) The Bayesian credibility estimate for the expected number of claims in Year 2 based on the Year 1 experience is 2.54545.
(v) The Bayesian credibility estimate for the expected number of claims in Year 3 based on the Year 1 and Year 2 experience is 3.73333.

Determine k.

Solution.
By Problem 66.1, $Q|N$ has a beta distribution with parameters $a' = a + \sum_{i=1}^n x_i$, $b' = b + nm - \sum_{i=1}^n x_i$ and $\theta = 1$, where $x_1, x_2, \cdots, x_n$ are past data.
By (iv), we have $n = 1$ and $x_1 = 2$ so that $a' = a+2$ and $b' = 9+8-2 = 15$.

The Bayesian estimate is

$$\begin{aligned}
E(N_2|N_1) =& \int_0^\infty mq \frac{\Gamma(a+17)}{\Gamma(a+2)\Gamma(15)} q^{a+2}(1-q)^{14} dq \\
=& \frac{m(a+2}{a+17} \int_0^\infty mq \frac{\Gamma(a+18)}{\Gamma(a+3)\Gamma(15)} q^{a+3}(1-q)^{14} dq \\
=& \frac{m(a+2}{a+17}.
\end{aligned}$$

Thus,

$$\frac{8(a+2)}{a+17} = 2.54545 \Longrightarrow a \approx 5.$$

By Problem 66.1, $Q|N$ has a beta distribution with parameters $a' = a + \sum_{i=1}^n x_i$, $b' = b + nm - \sum_{i=1}^n x_i$ and $\theta = 1$, where $x_1, x_2, \cdots, x_n$ are past data.

By (v), we have $n = 2$ and $x_1 = 2$ and $x_2 = k$. Thus, $a' = a + 2 + k$ and $b' = 9 + 16 - 2 - k = 23 - k$. The Bayesian estimate is

$$\begin{aligned}
E(N_2|N_1) =& \int_0^\infty mq \frac{\Gamma(a+25)}{\Gamma(a+2+k)\Gamma(23-k)} q^{a+2+k}(1-q)^{22-k} dq \\
=& \frac{m(a+2+k}{a+25} \int_0^\infty mq \frac{\Gamma(a+26)}{\Gamma(a+3+k)\Gamma(23-k)} q^{a+3+k}(1-q)^{22-k} dq \\
=& \frac{m(a+2+k}{a+25}.
\end{aligned}$$

Thus,

$$\frac{8(5+2+k}{25+5} = 3.73333 \Longrightarrow k \approx 7 \blacksquare$$

Practice Problems

Problem 80.1
Suppose an individual's claim amounts are given by an exponential distribution with mean Λ where Λ is an inverse Gamma with parameters $\alpha = 2$ and $\theta = 15$. Last year claim was \$12.
(a) Provide a mathematical description of this model.
(b) Determine the predictive distribution of next year claim.
(c) Determine the posterior distribution of Λ.
(d) Determine the Bayesian premium.

Problem 80.2 ‡
You are given:
(i) The annual number of claims for a policyholder follows a Poisson distribution with mean Λ.
(ii) The Prior distribution of Λ is gamma with probability density function:

$$\pi(\lambda) = \frac{(2\lambda)^5 e^{-2\lambda}}{24\lambda}, \ \lambda > 0.$$

An insured is selected at random and observed to have $x_1 = 5$ claims during Year 1 and $x_2 = 3$ claims during Year 2. Determine $E(\Lambda | x_1 = 5, x_2 = 3)$.

Problem 80.3
You are given:
(i) $X|P$ is a binomial distribution with parameters $(10, p)$.
(ii) The prior distribution of P is

$$\pi(p) = 2p, 0 < p < 1.$$

Find the posterior distribution of P given $X_1 = 4$.

Problem 80.4
(i) Show that the negative binomial distribution with parameters r and β can be expressed in the form

$$p_k = \frac{\Gamma(r+k)}{\Gamma(r)\Gamma(k)} q^r (1-q)^k, \ k = 0, 1, 2, \cdots .$$

(ii) Suppose that $X|Q$ is negative binomial with parameters r and q. Suppose also that Q is beta with $(a, b, 1)$. Find the posterior distribution of Q.

Problem 80.5
The amount of a claim $X|\Lambda$ has the normal distribution with mean θ and known variance σ_1^2. The risk parameter Λ has a normal distribution with mean μ and and variance σ_2^2. Find the posterior distribution of Θ.

Problem 80.6 ‡
You are given:
(i) The parameter Λ has an inverse gamma distribution with probability density function:

$$g(\lambda) = 500\lambda^{-4}e^{-\frac{10}{\lambda}}.\ \lambda > 0.$$

(ii) The size of a claim has an exponential distribution with probability density function:

$$f(x|\lambda) = \lambda^{-1}e^{-\frac{x}{\lambda}},\ x > 0,\ \lambda > 0.$$

For a single insured, two claims were observed that totaled 50.
Determine the expected value of the next claim from the same insured.

81 Bühlman Credibility Premium

The purpose of this section is to find an estimate of the hypothetical mean $\mu_{n+1}(\theta) = E(X_{n+1}|\Theta)$. The estimation procedure disussed here is the one developed by Bühlmann.

The idea is to estimate $\mu_{n+1}(\theta)$ with a linear combination of the past data:

$$\alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n$$

where $\alpha_0, \alpha_1, \cdots, \alpha_n$ are to be determined.

The estimation is done with the linear least squares regression where the square of the distance between the $\mu_{n+1}(\theta)$ and the estimator is to be minimized. That is, we want to find the $\alpha_i's$ that minimize the function

$$Q = Q(\alpha_0, \alpha_1, \cdots, \alpha_n) = E[(\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i)^2].$$

To minimize Q, we take the first derivatives of Q with respect to the α_i and set them to zero. Thus,

$$\frac{\partial Q}{\partial \alpha_0} = E[2(\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i)(-1)] = 0$$

which leads to

$$E[\mu_{n+1}(\Theta)] = \alpha_0 + \sum_{i=1}^{n} \alpha_i E(X_i). \tag{81.1}$$

Equation (81.1) is known as the **unbiased equation** since

$$E[\mu_{n+1}(\Theta)] = E[E(X_{n+1}|\Theta)] = E(X_{n+1}).$$

Thus, we can write

$$E(X_{n+1}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i E(X_i). \tag{81.2}$$

Next, we have

$$\frac{\partial Q}{\partial \alpha_i} = E[2(\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i)(-X_i)] = 0$$

which leads to

$$E[\mu_{n+1}(\Theta)X_i] = \alpha_0 E(X_i) + \sum_{j=1}^{n} \alpha_j E(X_j X_i).$$

However, we have

$$\begin{aligned}E[\mu_{n+1}(\Theta)X_i] =& E\{E[X_i\mu_{n+1}(\Theta)|\Theta]\}\\ =& E[\mu_{n+1}(\Theta)E(X_i|\Theta)]\\ =& E[E(X_{n+1}|\Theta)E(X_i|\Theta)]\\ =& E[E(X_{n+1}X_i|\Theta)] \text{ (by independence)}\\ =& E(X_{n+1}X_i).\end{aligned}$$

Hence,

$$E(X_{n+1}X_i) = \alpha_0 E(X_i) + \sum_{j=1}^{n} \alpha_j E(X_j X_i). \tag{81.3}$$

Multiplying (81.2) by $E(X_i)$ and subtracting the resulting equation from (81.3) to obtain

$$\text{Cov}(X_i, X_{n+1}) = \sum_{j=1}^{n} \alpha_j \text{Cov}(X_i, X_j), \; i = 1, 2, \cdots, n. \tag{81.4}$$

Equations (81.2) and (81.4) are known as the **normal equations**. Solving these $n + 1$ equations to yield the credibility premium

$$\hat{\alpha}_0 + \sum_{i=1}^{n} \hat{\alpha}_i X_i. \tag{81.5}$$

Example 81.1
You are given:
(i) $E(X_i) = 2$ and $\text{Var}(X_i) = 3$ for $i = 1, 2, \cdots, 20$.
(ii) $\text{Cov}(X_i, X_j) = 1.5$ for all $i \neq j$.
Determine the credibility premium.

Solution.
The unbiasedness equation yields

$$\hat{\alpha}_0 + 2\sum_{j=1}^{20} \hat{\alpha}_j = 2.$$

This implies

$$\sum_{j=1}^{20} \hat{\alpha}_j = 1 - \frac{\hat{\alpha}_0}{2}.$$

On the other hand, for $i = 1, 2, \cdots, 20$, we have

$$\sum_{\substack{j = 1 \\ j \neq i}}^{20} \hat{\alpha}_j(1.5) + 3\hat{\alpha}_i = 1.5$$

or equivalently

$$\sum_{j=1}^{20} \hat{\alpha}_j(1.5) + 1.5\hat{\alpha}_i = 1.5.$$

Solving this last equation for $\hat{\alpha}_i$ we find

$$\hat{\alpha}_i = 1 - \left(1 - \frac{\hat{\alpha}_0}{2}\right) = \frac{\hat{\alpha}_0}{2}.$$

Summation over i from 1 to 20 yields

$$\sum_{1=1}^{20} \hat{\alpha}_i = 10\hat{\alpha}_0$$

or equivalently

$$1 - \frac{\hat{\alpha}_0}{2} = 10\hat{\alpha}_0.$$

Solving this equation for $\hat{\alpha}_0$ we find

$$\hat{\alpha}_0 = \frac{2}{21}.$$

Hence,

$$\hat{\alpha}_i = \frac{1}{21}$$

and the credibility premium is

$$\hat{\alpha}_0 + \sum_{i=1}^{20} \hat{\alpha}_i X_i = \frac{2}{21}(1 + 10\overline{X}) \blacksquare$$

Remark 81.1
It is easy to check that the values $\hat{\alpha}_0, \hat{\alpha}_1. \cdots, \hat{\alpha}_n$ also minimize

$$Q_1 = E\left\{\left[E(X_{n+1}|\mathbf{X}) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right]^2\right\}$$

and

$$Q_2 = E[(X_{n+1} - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i)^2].$$

That is, the credibility premium (81.5) is the best linear estimator of each the hypothetical mean $E(X_{n+1}|\Theta)$, the Bayesian premium $(X_{n+1}|\mathbf{X})$, and X_{n+1}.

Example 81.2 ‡
You are given the following information about a credibility model:

First observation (T)	$\Pr(X_1 = T)$	Bayesian estimate $E(X_2 \vert X_1 = T)$
1	1/3	1.50
2	1/3	1.50
3	1/3	3.00

Determine the Bühlmann credibility estimate of the second observation, given that the first observation is 1.

Solution.
Let X_1 be the outcome of the first observation. By Problem 81.5, the Bühlmann credibility estimate is of the form $ZX_1 + (1-Z)\mu$ where $X_1 = 1, 2,$ or 3.By Remark 8.1, the Bühlmann estimate is the least squares approximation to the Bayesian estimate. Thus, Z and μ are the minimizers of

$$f(Z, \mu) = \frac{1}{3}[1.50 - Z - (1-Z)\mu]^2 + \frac{1}{3}[1.50 - 2Z - (1-Z)\mu]^2 + \frac{1}{3}[3.00 - 3Z - (1-Z)\mu]^2.$$

Taking the derivative with respect to μ and setting it to zero, we find $\mu = 2$. Next, taking the derivative of f with respect to Z and setting it to zero, we find

$$2(-Z + 0.5)(-1) + 2(0.5)(0) + 2(Z - 1)(1) = 0 \Longrightarrow Z = 0.75.$$

Thus, the Bühlmann credibility estimate of the second observation, given that the first observation is 1, is

$$ZX_1 + (1 - Z)\mu = 0.75(1) + (1 - 0.75)(2) = 1.25 \blacksquare$$

Practice Problems

Problem 81.1
You are given:
(i) $E(X_j) = \mu$ and $\text{Var}(X_j) = \sigma^2$ for all $j = 1, 2, \cdots, n$.
(ii) $\text{Cov}(X_i, X_j) = \rho\sigma^2$ for all $i \neq j$, where ρ is the coefficient of correlation.
Use the unbiasedness equation to show that

$$\sum_{j=1}^{n} \hat{\alpha}_j = 1 - \frac{\hat{\alpha}_0}{\mu}.$$

Problem 81.2
With the assumptions of Problem 81.1, show that the n equations (81.4) lead to

$$\rho = \rho \sum_{j=1}^{n} \hat{\alpha}_j + \hat{\alpha}_i(1 - \rho), \ i = 1, 2, \cdots, n.$$

Problem 81.3
With the assumptions of Problem 81.1, show that

$$\hat{\alpha}_i = \frac{\rho \hat{\alpha}_0}{\mu(1 - \rho)}.$$

Problem 81.4
With the assumptions of Problem 81.1, show that

$$\hat{\alpha}_0 = \tfrac{(1-\rho)\mu}{1-\rho+n\rho} \ \text{ and } \ \hat{\alpha}_i = \tfrac{\rho}{1-\rho+n\rho}$$

Problem 81.5
With the assumptions of Problem 81.1, show that

$$\hat{\alpha}_0 + \sum_{j=1}^{n} \hat{\alpha}_j X_j = (1 - Z)\mu + Z\overline{X}$$

where Z to be determined.

82 The Bühlmann Model with Discrete Prior

The Bühlmann assumes that the past losses $X_1, X_2, \cdots, X_n$ have the same mean and variance and that $X_i|\Theta$ are iid. We define the following:
• The **hypothetical mean**: $\mu(\theta) = E[X_j|\Theta = \theta]$.
• The **process variance**: $v(\theta) = \text{Var}(X_j|\Theta = \theta)$.
• The **expected value of the hypothetical mean**[22]: $\mu = E[\mu(\Theta)]$.
• The **expected value of the process variance**: $v = E[v(\Theta)]$.
• The **variance of the hypothetical mean**: $a = \text{Var}[\mu(\Theta)] = E[\mu^2(\Theta)] - E[\mu(\Theta)]^2$.

Example 82.1
Find the (a) mean, (b) variance and (c) covariance of X_i.

Solution.
(a) The mean is given by

$$E(X_i) = E[E(X_i|\Theta)] = E[\mu(\Theta)] = \mu.$$

(b) The variance is given by

$$\begin{aligned}\text{Var}(X_i) =& E[\text{Var}(X_i|\Theta)] + \text{Var}[E(X_i|\Theta)]\\ =& E[v(\Theta)] + \text{Var}[\mu(\Theta)] = v + a.\end{aligned}$$

(c) For $i \neq j$, The covariance of X_i and X_j is given by

$$\begin{aligned}\text{Cov}(X_i, X_j) =& E(X_i, X_j) - E(X_i)E(X_j)\\ =& E[E(X_iX_j|\Theta)] - E[\mu(\Theta)]^2\\ =& E[E(X_i|\Theta)E(X_j|\Theta)] - E[\mu(\Theta)]^2 \quad \text{(by independence)}\\ =& E[\mu^2(\Theta)] - E[\mu(\Theta)]^2\\ =& \text{Var}[\mu(\Theta)] = a \blacksquare\end{aligned}$$

Using Problem 81.5 with $\sigma^2 = v + a$ and $\rho = \frac{a}{v+a}$, the **Bühlmann credibility factor** is given by

$$Z = \frac{n\rho}{1 - \rho + n\rho} = \frac{n}{n + k}$$

[22]Also known as the **collective premium**. It is the value that is used if there is no information about past claims data.

where

$$k = \frac{v}{a} = \frac{E[\text{Var}(X_j|\Theta)]}{\text{Var}[E(X_j|\Theta)]}.$$

The **Bühlmann credibility premium** is given by

$$BP = Z\overline{X} + (1 - Z)\mu.$$

Some of the features of the Bühlmann model:

• As more data is gathered, i.e., $n \to \infty$, the value of Z tends to 1 and the weighting average tends toward the sample mean rather than the collective premium, a feature that agrees with intuition.

• If the population is homogenepus with respect to Θ, then $\mu(\Theta)$ are very close in values so that they have very small variability. That is, a is very small which implies a very large k and consequently Z approaches 0.

• If the population is heterogeneous more weight will be given to the sample mean as it is better indicator of an individual's future claims.

Example 82.2

Drivers are classified as good (G), average (A), or bad (B).

• Good drivers make up 70% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.65, 1 claim is 0.25, and 2 claims is 0.10.

• Average drivers make up 20% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.40, 1 claim is 0.40, and 2 claims is 0.20.

• Bad drivers make up 10% of the population and for a driver in this class, the probability of having 0 claim in one year is 0.50, 1 claim is 0.30, and 2 claims is 0.20.

For a policyholder, the risk parameter is the classification of the individual as G, A, or B. For a particular policyholder, it has been observed that $x_1 = 1$ and $x_2 = 2$.

Determine the Bühlmann premium.

Solution.

By Problem 79.4(a), we have

$$\begin{array}{ccc} \mu(G) = E(X_i|G) = 0.45, & \mu(A) = E(X_i|A) = 0.80 & \mu(B) = E(X_i|B) = 0.70 \\ \pi(G) = 0.70, & \pi(A) = 0.20, & \pi(B) = 0.10. \end{array}$$

Hence,

$$\begin{aligned}
\mu =& \sum_{\theta} \mu(\theta)\pi(\theta) = 0.45(0.70) + 0.80(0.20) + 0.70(0.10) = 0.545 \\
E[\mu^2(\Theta)] =& \sum_{\theta} \mu(\theta)^2\pi(\theta) = 0.45^2(0.70) + 0.80^2(0.20) + 0.70^2(0.10) = 0.31875 \\
a =& E[\mu^2(\Theta)] - \mu^2 = 0.31875 - 0.545^2 = 0.021725.
\end{aligned}$$

For the process variance, we have

$$\begin{aligned}
v(G) =& \mathrm{Var}(X_i|G) = 0^2(0.65) + 1^2(0.25) + 2^2(0.10) - 0.45^2 = 0.4475 \\
v(A) =& \mathrm{Var}(X_i|A) = 0^2(0.40) + 1^2(0.40) + 2^2(0.20) - 0.80^2 = 0.56 \\
v(B) =& \mathrm{Var}(X_i|B) = 0^2(0.50) + 1^2(0.30) + 2^2(0.20) - 0.70^2 = 0.61 \\
v =& \sum_{\theta} v(\theta)\pi(\theta) = 0.4475(0.70) + 0.56(0.20) + 0.61(0.10) = 0.48625 \\
k =& \frac{v}{a} = \frac{0.48625}{0.021725} = 22.3820 \\
Z =& \frac{n}{n+k} = \frac{2}{2 + 22.3820} = 0.0820.
\end{aligned}$$

Hence, the Bühlmann premium is

$$BP = Z\overline{X} + (1 - Z)\mu = 0.0820(1.5) + (1 - 0.0820)(0.545) = 0.62331.$$

This is the best linear approximation to the Bayesian premium of 0.6297 found in Problem 79.5 ■

Example 82.3 ‡

You are given the following information on claim frequency of automobile accidents for individual drivers:

	Busienss use		Pleasure use	
	Expected claims	Claim variance	Expected claims	Claim variance
Rural	1.0	0.5	1.5	0.8
Urban	2.0	1.0	2.5	1.0
Total	1.8	1.06	2.3	1.12

You are also given:
(i) Each driver's claims experience is independent of every other driver's.
(ii) There are an equal number of business and pleasure use drivers.
Determine the Bühlmann credibility factor Z for a single driver.

Solution.
Let Θ be risk parameter with designations BR= Business Rural, BU= business urban, PR = Pleasure Rural, and PU = Pleasure Urban. By double expectation, we have

$$1.8 = E(B) = E[E(B|R,U)] = E(B|R)\Pr(R)+E(B|U)\Pr(U) = \Pr(R)+2\Pr(U).$$

Likewise,

$$2.3 = 1.5\Pr(R) + 2.5\Pr(U).$$

Solving these two equations, we find $\Pr(R) = 0.2$ and $\Pr(U) = 0.8$.
The prior distribution is

$$\begin{aligned}\pi(BR) =&\pi(PR) = (0.5)(0.2) = 0.10\\ \pi(BU) =&\pi(PU) = (0.5)(0.8) = 0.40.\end{aligned}$$

Let X represent the claim frequency of auto accidents of a randomly selected driver. Then, we have the following

$$\begin{aligned}\mu(BR) =&E(X|BR) = 1.0\\ \mu(BU) =&E(X|BU) = 2.0\\ \mu(PR) =&E(X|PR) = 1.5\\ \mu(PU) =&E(X|PU) = 2.5.\end{aligned}$$

Thus,

$$\mu = \sum_{\theta} \mu(\theta)\pi(\theta) = (0.10)(1.0) + (0.40)(2.0) + (0.10)(1.5) + (0.40(2.5) = 2.05$$

and

$$\begin{aligned}a =&\sum_{\theta} \mu(\theta)^2\pi(\theta) - \mu^2\\ =&(0.10)(1.0)^2 + (0.40)(2.0)^2 + (0.10)(1.5)^2 + (0.40(2.5)^2 - 2.05^2 = 0.2225.\end{aligned}$$

For the process variance, we have

$$\begin{aligned}v(BR) =&0.5\\ v(BU) =&1.0\\ v(PR) =&0.8\\ v(PU) =&1.0.\end{aligned}$$

Hence,

$$v = \sum_{\theta} v(\theta)\pi(\theta) = 0.5(0.10) + 1.0(0.40) + 0.8(0.10) + 1.0(0.40) = 0.93.$$

It follows that

$$k = \frac{v}{a} = \frac{0.93}{0.2225} = 4.18$$

and the credibility factor is

$$Z = \frac{n}{n+k} = \frac{1}{1+4.18} = 0.193 \blacksquare$$

Example 82.4 ‡
For a particular policy, the conditional probability of the annual number of claims given $\Theta = \theta$, and the probability distribution of Θ are as follows:

Number of claims	0	1	2
Probability	2θ	θ	$1-3\theta$

θ	0.05	0.30
Probability	0.80	0.20

Two claims are observed in Year 1. Calculate the Bühlmann credibility estimate of the number of claims in Year 2.

Solution.
We have

$$\begin{aligned}
E(\Theta) =&0.05(0.80) + 0.30(0.20) = 0.1\\
E(\Theta^2) =&0.05^2(0.80) + 0.30^2(0.20) = 0.02\\
\mu(\theta) =&E(N|\Theta) = 0(2\theta) + 1(\theta) + 2(1-3\theta) = 2-5\theta\\
\mu =&E(2-5\Theta) = 2-5E(\Theta) = 2-5(0.1) = 1.5\\
a =&\mathrm{Var}(2-5\Theta) = 25\mathrm{Var}(\Theta) = 25(0.02-0.1^2) = 0.25\\
v(\theta) =&\mathrm{Var}(N|\Theta) = 0^2(2\theta) + 1^2(\theta) + 2^2(1-3\theta) - (2-5\theta)^2 = 9\theta - 25\theta^2\\
v =&E(9\Theta - 25\Theta^2) = 9(0.1) - 25(0.02) = 0.4\\
k =&\frac{v}{a} = \frac{0.4}{0.25} = 1.6\\
Z =&\frac{1}{1+k} = \frac{5}{13}.
\end{aligned}$$

The required estimate is

$$2Z + (1-Z)\mu = 2\left(\frac{5}{13}\right) + \left(1 - \frac{5}{13}\right)(1.5) = 1.6923 \blacksquare$$

Example 82.5 ‡
For a group of policies, you are given:
(i) The annual loss on an individual policy follows a gamma distribution with parameters $\alpha = 4$ and θ.
(ii) The prior distribution of θ has mean 600.
(iii) A randomly selected policy had losses of 1400 in Year 1 and 1900 in Year 2.
(iv) Loss data for Year 3 was misfiled and unavailable.
(v) Based on the data in (iii), the Bühlmann credibility estimate of the loss on the selected policy in Year 4 is 1800.
(vi) After the estimate in (v) was calculated, the data for Year 3 was located. The loss on the selected policy in Year 3 was 2763.
Calculate the Bühlmann credibility estimate of the loss on the selected policy in Year 4 based on the data for Years 1, 2 and 3.

Solution.
Let the annual loss be denoted by X. Then $X|\Theta$ has a Gamma distribution with parameters $\alpha = 4$ and θ. We have

$$\begin{aligned}\mu(\theta) =& E(X|\Theta) = \alpha\theta = 4\theta\\ v(\theta) =& \mathrm{Var}(X|\Theta) = \alpha\theta^2 = 4\theta^2\\ \mu =& E[\mu(\Theta)] = 4E(\Theta) = 4(600) = 2400\\ v =& E[v(\Theta)] = 4E(\Theta^2)\\ a =& \mathrm{Var}[\mu(\Theta)] = 16\mathrm{Var}(\Theta).\end{aligned}$$

Based on Year 1 and Year 2, the average loss is

$$\overline{X} = \frac{1400 + 1900}{2} = 1650.$$

Based on (iv), we have

$$1650Z + (1 - Z)(2400) = 1800 \Longrightarrow Z = 0.8 = \frac{2}{2 + k} \Longrightarrow k = 0.5.$$

By (vi), we have

$$\overline{X} = \frac{1400 + 1900 + 2763}{3} = 2021$$

and

$$Z = \frac{3}{3 + 0.5} = \frac{6}{7}.$$

Thus, the Bühlmann credibility estimate of the loss on the selected policy in Year 4 based on the data for Years 1, 2 and 3

$$\left(\frac{6}{7}\right)(2021)+\left(1-\frac{6}{7}\right)(2400)=2075.14 \blacksquare$$

Example 82.6 ‡

You are given:

(i) The annual number of claims for an individual risk follows a Poisson distribution with mean λ.

(ii) For 75% of the risks, $\lambda = 1$.

(iii) For 25% of the risks, $\lambda = 3$.

A randomly selected risk had r claims in Year 1. The Bayesian estimate of this riskfs expected number of claims in Year 2 is 2.98.

Determine the Bühlmann credibility estimate of the expected number of claims for this risk in Year 2.

Solution.

Let X be the annual amount of claims and Λ be the prior parameter. Then $X|\Lambda$ has a Poisson distribution with mean λ. The prior distribution is: $\pi(1) = 0.75$ and $\pi(3) = 0.25$.

The posterior distribution is

$$\begin{aligned}
\pi(1|r) =&\frac{f(r|1)\pi(1)}{f(r|1)\pi(1)+f(r|3)\pi(3)} = \frac{\frac{e^{-1}}{r!}(0.75)}{\frac{e^{-1}}{r!}(0.75)+\frac{e^{-3}3^r}{r!}(0.25)}\\
=&\frac{0.2759}{0.2759+3^r(0.1245)}\\
\pi(3|r) =&1-\frac{0.2759}{0.2759+3^r(0.1245)} = \frac{3^r(0.1245)}{0.2759+3^r(0.1245)}.
\end{aligned}$$

Thus,

$$\begin{aligned}
2.98 =&E(X_2|X_1=r)=E(X_2|\Lambda=1)\pi(1|r)+E(X_2|\Lambda=3)\pi(3|r)\\
=&(1)\frac{0.2759}{0.2759+3^r(0.1245)}+(3)\frac{3^r(0.1245)}{0.2759+3^r(0.1245)}\\
=&\frac{0.2759+0.3735(3^r)}{0.2759+3^r(0.1245)}.
\end{aligned}$$

Cross multiply to obtain

$$0.82218+0.037103(3^r)=0.2759+0.03735(3^r).$$

Rearranging the terms and solving, we find

$$0.54628 = 0.00025(3^r) \Longrightarrow = r = 7.$$

Using Bühlmann theory, we find

$$\begin{aligned}
\mu(\Lambda) =& E(X|\Lambda) = \lambda \\
v(\Lambda) =& \text{Var}[X|\Lambda) = \lambda \\
\mu =& E(\Lambda) = 0.75(1) + 0.25(3) = 1.5 \\
v =& E(\Lambda) = 1.5 \\
a =& \text{Var}[\Lambda) = 0.75(1^2) + 0.25(3^2) - 1.5^2 = 0.75 \\
k =& \frac{v}{a} = \frac{1.5}{0.75} = 2 \\
Z =& \frac{1}{1+2} = \frac{1}{3}.
\end{aligned}$$

The Bühlmann credibility estimate of the expected number of claims for this risk in Year 2 is

$$\left(\frac{1}{3}\right)(7) + \left(\frac{2}{3}\right)(1.5) = 3.33 \blacksquare$$

Example 82.7 ‡

You are given:

(i) The claim count and claim size distributions for risks of type A are:

Number of Claims	Probabilities
0	4/9
1	4/9
2	1/9

Claim Size	Probabilities
500	1/3
1235	2/3

(ii) The claim count and claim size distributions for risks of type B are:

Number of Claims	Probabilities
0	1/9
1	4/9
2	4/9

Claim Size	Probabilities
250	2/3
328	1/3

(iv) Claim counts and claim sizes are independent within each risk type.

(v) The variance of the total losses is 296,962.

A randomly selected risk is observed to have total annual losses of 500.

Determine the Bühlmann premium for the next year for this same risk.

Solution.
Let L denote annual losses. The prior parameter represents the type of risk so that either $\Theta = A$ or $\Theta = B$. The prior distribution is $\Pr(A) = \Pr(B) = 0.5$. For each of the two classes, total annual loss L has a compound distribution.
We want

$$Z\overline{X} + (1 - Z)\mu.$$

Since there is a single observation ($n = 1$), the sample mean is $\overline{X} = 500$.
We have

$$\begin{aligned}
\mu(A) =& E(L|A) = E(N|A)E(X|A)\\
=&[0(4/9) + 1(4/9) + 2(1/9)][500(1/3) + 1235(2/3)] = 660\\
\mu(B) =& E(L|B) = E(N|B)E(X|B)\\
=&[0(1/9) + 1(4/9) + 2(4/9)][250(2/3) + 328(1/3)] = 368\\
\mu =& E[\mu(\Theta)] = 660(1/2) + 368(1/2) = 514\\
a =& \mathrm{Var}[\mu(\Theta)] = 660^2(1/2) + 368^2(1/2) - 514^2 = 21,316\\
\mathrm{Var}(L) =& \mathrm{Var}[E(L|\Theta)] + E[\mathrm{Var}(L|\Theta)]\\
296,962 =& \mathrm{Var}[\mu(\Theta)] + E[v(\Theta)\\
296,962 =& 21,316 + v\\
v =& 275,646\\
Z =& \frac{1}{1 + \frac{275{,}646}{21{,}316}} = 0.0718.
\end{aligned}$$

The Bühlmann premium is

$$0.0718(500) + (1 - 0.0718)(514) \approx 513 \blacksquare$$

Practice Problems

Problem 82.1 ‡
You are given:
(i) Two risks have the following severity distributions:

Amount of claim	Probability of claim amount for Risk 1	Probability of claim amount for Risk 2
250	0.5	0.7
2500	0.3	0.2
60000	0.2	0.1

(ii) Risk 1 is twice as likely to be observed as Risk 2.
A claim of 250 is observed.
Determine the Bühlmann credibility estimate of the second claim amount from the same risk.

Problem 82.2 ‡
You are given the following joint distribution:

	Θ	
X	0	1
0	0.4	0.1
1	0.1	0.2
2	0.1	0.1

For a given value of Θ and a sample of size 10 for $X : \sum_{i=1}^{10} x_i = 10$.
Determine the Bühlmann credibility premium.

Problem 82.3 ‡
An insurer writes a large book of home warranty policies. You are given the following information regarding claims filed by insureds against these policies:
(i) A maximum of one claim may be filed per year.
(ii) The probability of a claim varies by insured, and the claims experience for each insured is independent of every other insured.
(iii) The probability of a claim for each insured remains constant over time.
(iv) The overall probability of a claim being filed by a randomly selected insured in a year is 0.10.
(v) The variance of the individual insured claim probabilities is 0.01.
An insured selected at random is found to have filed 0 claims over the past 10 years.

Determine the Bühlmann credibility estimate for the expected number of claims the selected insured will file over the next 5 years.

Problem 82.4 ‡
You are given:
(i) Claim size, X, has mean μ and variance 500.
(ii) The random variable μ has a mean of 1000 and variance of 50.
(iii) The following three claims were observed: 750, 1075, 2000
Calculate the expected size of the next claim using Buhlmann credibility

Problem 82.5 ‡
You are given:
(i) A portfolio of independent risks is divided into two classes.
(ii) Each class contains the same number of risks.
(iii) For each risk in Class 1, the number of claims per year follows a Poisson distribution with mean 5.
(iv) For each risk in Class 2, the number of claims per year follows a binomial distribution with $m = 8$ and $q = 0.55$.
(v) A randomly selected risk has three claims in Year 1, r claims in Year 2 and four claims in Year 3.
The Bühlmann credibility estimate for the number of claims in Year 4 for this risk is 4.6019. Determine r.

Problem 82.6 ‡
For a portfolio of independent risks, the number of claims for each risk in a year follows a Poisson distribution with means given in the following table:

Class	Mean Number of Claims per risk	Number of Risks
1	1	900
2	10	90
3	20	10

You observe x claims in Year 1 for a randomly selected risk.
The Bühlmann credibility estimate of the number of claims for the same risk in Year 2 is 11.983.
Determine x.

Problem 82.7 ‡
An insurance company sells two types of policies with the following characteristics:

Type of Policy	Proportion of Total Policies	Poisson Annual Claim Frequency
I	θ	$\lambda = 0.50$
II	$1-\theta$	$\lambda = 1.50$

A randomly selected policyholder is observed to have one claim in Year 1. For the same policyholder, determine the Bhlmann credibility factor Z for Year 2.

Problem 82.8 ‡
You are given:
(i) Losses in a given year follow a gamma distribution with parameters α and θ, where θ does not vary by policyholder.
(ii) The prior distribution of α has mean 50.
(iii) The Buhlmann credibility factor based on two years of experience is 0.25.
Calculate $\text{Var}(\alpha)$.

Problem 82.9 ‡
For a portfolio of independent risks, you are given:
(i) The risks are divided into two classes, Class A and Class B.
(ii) Equal numbers of risks are in Class A and Class B.
(iii) For each risk, the probability of having exactly 1 claim during the year is 20% and the probability of having 0 claims is 80%. (iv) All claims for Class A are of size 2.
(v) All claims for Class B are of size c, an unknown but fixed quantity.
One risk is chosen at random, and the total loss for one year for that risk is observed. You wish to estimate the expected loss for that same risk in the following year.
Determine the limit of the Bühlmann credibility factor as c goes to infinity.

Problem 82.10 ‡
An insurance company writes a book of business that contains several classes of policyholders. You are given:
(i) The average claim frequency for a policyholder over the entire book is 0.425.
(ii) The variance of the hypothetical means is 0.370.
(iii) The expected value of the process variance is 1.793.
One class of policyholders is selected at random from the book. Nine policyholders are selected at random from this class and are observed to have produced a total of seven claims. Five additional policyholders are selected

at random from the same class.
Determine the Bühlmann credibility estimate for the total number of claims for these five policyholders.

83 The Bühlmann Model with Continuous Prior

In this section, we consider Bühlmann models with continuous prior.

Example 83.1
Let $X_1, X_2, \cdots, X_n$ be past claim amounts. Suppose that $X_i|\Theta$ are independent and identically exponentially distributed with mean Θ and Θ is Gamma distributed with parameters α and β. Determine the Bühlmann premium.

Solution.
The hypothetical mean is

$$\mu(\theta) = E(X_i|\Theta) = \theta.$$

Thus,

$$\mu = E[\mu(\Theta)] = E(\Theta) = \alpha\beta$$

and

$$a = \text{Var}[\mu(\Theta)] = \text{Var}(\Theta) = \alpha\beta^2.$$

Likewise,

$$v(\theta) = \text{Var}(X_i|\Theta) = \theta^2$$

and

$$v = E[v(\Theta)] = E[\Theta^2] = \text{Var}(\Theta) + E(\Theta)^2 = \alpha\beta^2 + \alpha^2\beta^2.$$

Hence,

$$k = \frac{v}{a} = \frac{\alpha\beta^2 + \alpha^2\beta^2}{\alpha\beta^2} = 1 + \alpha$$

and

$$Z = \frac{n}{n+k} = \frac{n}{n+\alpha+1}.$$

The Bühlmann premium is

$$BP = \frac{n}{n+\alpha+1}\overline{X} + \frac{\alpha+1}{n+\alpha+1}(\alpha\beta) \blacksquare$$

Example 83.2 ‡
You are given:
(i) Claim counts follow a Poisson distribution with mean θ.
(ii) Claim sizes follow an exponential distribution with mean 10θ.
(iii) Claim counts and claim sizes are independent, given θ.
(iv) The prior distribution has probability density function:

$$\pi(\theta) = \frac{5}{\theta^6},\ \theta > 1.$$

Calculate Bühlmann's k for aggregate losses.

Solution.
Let N be the Poisson claim count variable, let X be the claim size variable, and let S be the aggregate loss variable. Note that $S|\Theta$ is a compound Poisson distribution with primary distribution $N|\Theta$ and secondaru distribution $X|\Theta$.
The hypothetical mean is

$$\mu(\theta) = E(N|\Theta)E(X|\Theta) = \theta(10\theta) = 10\theta^2$$

with expected value

$$\mu = E[10\Theta^2] = \int_1^\infty 10\theta^2 \left(\frac{5}{\theta^6}\right) d\theta = \frac{50}{3}.$$

The variance of the hypothetical mean is

$$a = \int_1^\infty (10\theta^2)^2 \left(\frac{5}{\theta^6}\right) d\theta - \left(\frac{50}{3}\right)^2 = 222.2222.$$

The process variance is

$$\begin{aligned}v(\theta) =& \text{Var}(S|\Theta) = E(N|\Theta)\text{Var}(X|\Theta) + \text{Var}(N|\Theta)E(X|\Theta)^2\\ =&(\theta)((10\theta)^2 + \theta(10\theta)^2\\ =&200\theta^3.\end{aligned}$$

and its expected value is

$$v = E[v(\Theta)] = \int_1^\infty 200\theta^3 \left(\frac{5}{\theta^6}\right) d\theta = 500.$$

Hence,

$$k = \frac{v}{a} = \frac{500}{222.2222} = 2.25 \blacksquare$$

Example 83.3 ‡
You are given:
(i) The number of claims made by an individual insured in a year has a Poisson distribution with mean λ.
(ii) The prior distribution for Λ is Gamma with parameters $\alpha = 1$ and $\theta = 1.2$.
Three claims are observed in Year 1, and no claims are observed in Year 2. Using Bühlmann credibility, estimate the number of claims in Year 3.

Solution.
The hypothetical mean is

$$\mu(\lambda) = E(X|\Lambda) = \lambda$$

with expected value

$$\mu = E(\Lambda) = \alpha\theta = 1.2.$$

The variance of the hypothetical mean is

$$a = \mathrm{Var}[\mu(\Lambda)] = \mathrm{Var}(\Lambda) = \alpha\theta^2 = 1.44.$$

The process variance is

$$v(\lambda) = \mathrm{Var}(X|\Lambda) = \lambda$$

and its expected value is

$$v = E[v(\lambda)] = E(\Lambda) = 1.2.$$

Hence,

$$k = \frac{v}{a} = \frac{1.2}{1.44} = \frac{1}{1.2}$$

and the credibility factor is

$$Z = \frac{n}{n+k} = \frac{2}{2+\frac{1}{1.2}} = 0.706.$$

The expected the number of claims in Year 3 is

$$BP = Z\overline{X} + (1-Z)\mu = 0.706(1.5) + (1-0.706)(1.2) = 1.4118 \blacksquare$$

Example 83.4 ‡
You are given:
(i) The number of claims in a year for a selected risk follows a Poisson distribution with mean λ.
(ii) The severity of claims for the selected risk follows an exponential distribution with mean θ.
(iii) The number of claims is independent of the severity of claims.
(iv) The prior distribution of λ is exponential with mean 1.
(v) The prior distribution of θ is Poisson with mean 1.
(vi) A priori, λ and θ are independent.
Using Bühlmann's credibility for aggregate losses, determine k.

Solution.
We have

$$\begin{aligned}
E(S|\lambda,\theta) =& E(N)E(X) = \lambda\theta \\
\text{Var}(S|\lambda,\theta) =& E(N)E(X^2) = \lambda(2\theta^2) \\
a =& \text{Var}(\lambda\theta) = E(\lambda^2)E(\theta^2) - [E(\lambda)E(\theta)]^2 \\
=& 2(1^2+1) - [(1)(1)]^2 = 3 \\
v =& E(2\lambda\theta^2) = 2E(\lambda)E(\theta^2) = 2(1)(1^2+1) = 4 \\
k =& \frac{v}{a} = \frac{4}{3} \blacksquare
\end{aligned}$$

Example 83.5 ‡
You are given:
(i) The annual number of claims on a given policy has a geometric distribution with parameter β.
(ii) The prior distribution of β has the Pareto density function

$$\pi(\beta) = \frac{\alpha}{(\beta+1)^{\alpha+1}},\ 0 < \beta < \infty,$$

where α is a known constant greater than 2.
A randomly selected policy had x claims in Year 1. Determine the Bühlmann credibility estimate of the number of claims for the selected policy in Year 2.

Solution.
Note that the prior distribution is a Pareto distribution with parameters α and 1. We have

$$\begin{aligned}
\mu(\beta) =& E(X|\beta) = \beta \\
\mu =& E(\beta) = \frac{1}{\alpha-1} \\
a =& \text{Var}(\beta) = \frac{2}{(\alpha-1)(\alpha-2)} - \frac{1}{(\alpha-1)^2} = \frac{\alpha}{(\alpha-1)^2(\alpha-2)} \\
v(\beta) =& \text{Var}(X|\beta) = \beta(\beta+1) \\
v =& E[\beta(\beta+1)] = \frac{1}{\alpha-1} + \frac{2}{(\alpha-1)(\alpha-2)} = \frac{\alpha}{(\alpha-1)(\alpha-2)} \\
k =& \frac{v}{a} = \alpha - 1 \\
Z =& \frac{1}{1+k} = \frac{1}{\alpha}.
\end{aligned}$$

The required estimate is

$$xZ + (1-Z)\mu = \frac{x}{\alpha} + \frac{1}{\alpha} = \frac{x+1}{\alpha} \blacksquare$$

Example 83.6 ‡
For five types of risks, you are given:
(i) The expected number of claims in a year for these risks ranges from 1.0 to 4.0.
(ii) The number of claims follows a Poisson distribution for each risk.
During Year 1, n claims are observed for a randomly selected risk.
For the same risk, both Bayes and Bhlmann credibility estimates of the number of claims in Year 2 are calculated for $n = 0, 1, 2, \cdots, 9$.
Which graph represents these estimates?

Solution.
If X is the number of claims in Year 1, then the Bühlmann estimate is $ZX + (1-Z)\mu$ which is a linear function of X. This implies that (E) cannot be the answer. The Bayes estimate is given by

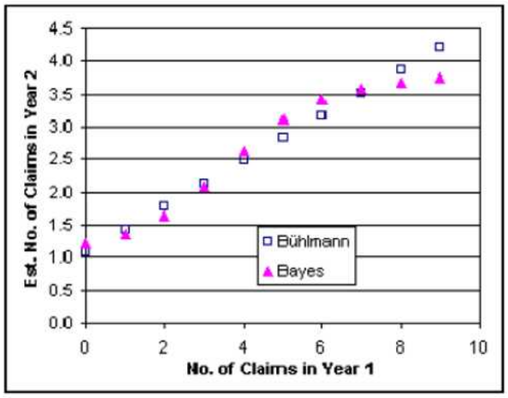
$$E(X_2|X_1 = n) = \int_1^4 \lambda\pi(\lambda|X_1 = n)d\lambda$$

where $1.0 \le \lambda \le 4$ by (i). Thus,

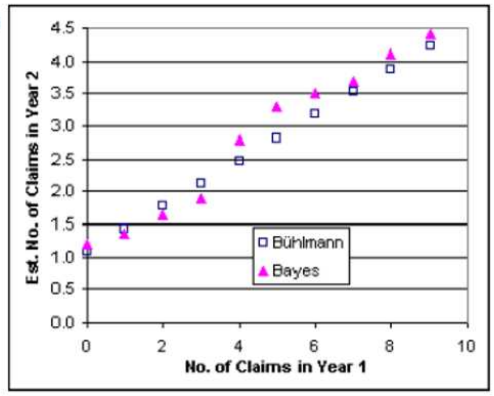
$$\pi(\lambda|X_1 = n) \le \lambda\pi(\lambda|X_1 = n) \le 4\lambda\pi(\lambda|X_1 = n) \Longrightarrow 1 \le E(X_2|X_1 = n) \le 4.$$

The graph (B) can not be the answer since $E(X_2|X_1 = 8)$ and $E(X_2|X_1 = 9)$ are greater than 4. Likewise, the graph (D) cannot be the answer since $E(X_2|X_1 = 0)$ and $E(X_2|X_1 = 1)$ are less than 1. Now, by Remark 81.1, the Bühlmann estimates are the linear least squares approximation to the Bayes estimates. We see from graph (C) that the Bayes estimates are consistently higher than the Bühlmann estimates and so it can not be the answer. Hence, (A) is the most appropriate answer for the problem $\blacksquare$

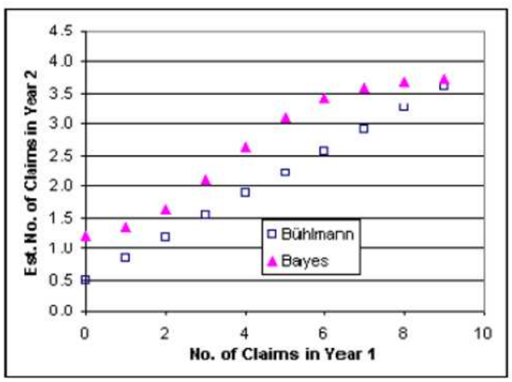
(A)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

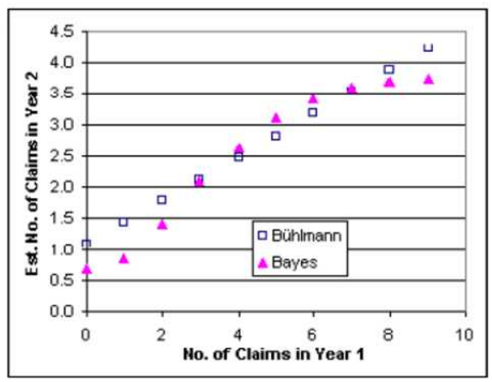
(B)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

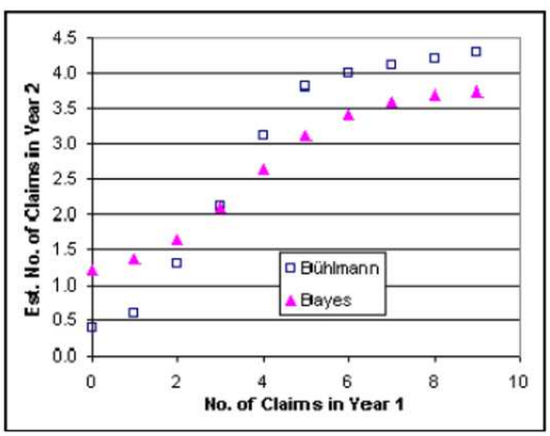
(C)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

(D)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

(E)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

Practice Problems

Problem 83.1
Let $X_1, X_2, \cdots, X_n$ be past claim amounts. Suppose that $X_i|\Theta$ are independent and identically Poisson distributed with mean Θ and Θ is Gamma distributed with parameters α and β. Determine the Bühlmann premium.

Problem 83.2 ‡
You are given:
(i) Annual claim frequency for an individual policyholder has mean λ and variance σ^2.
(ii) The prior distribution for λ is uniform on the interval $[0.5, 1.5]$.
(iii) The prior distribution for σ^2 is exponential with mean 1.25.
A policyholder is selected at random and observed to have no claims in Year 1.
Using Bühlmann credibility, estimate the number of claims in Year 2 for the selected policyholder.

Problem 83.3 ‡
You are given the following information about a book of business comprised of 100 insureds:
(i) $X_i = \sum_{j=1}^{N_i} Y_{ij}$ is a random variable representing the annual loss of the i^{rmth} insured.
(ii) $N_1, N_2, \cdots, N_{100}$ are independent random variables distributed according to a negative binomial distribution with parameters r (unknown) and $\beta = 0.2$.
(iii) Unknown parameter r has an exponential distribution with mean 2.
(iv) Y_{ij} are independent random variables distributed according to a Pareto distribution with $\alpha = 3.0$ and $\theta = 1000$.
Determine the Bühlmann credibility factor, Z, for the book of business.

Problem 83.4 ‡
You are given:
(i) The annual number of claims for an insured has probability function:

$$p(x) = \left(\begin{array}{c} 3 \\ x \end{array}\right) q^x (1-q)^{3-x},\ x = 0, 1, 2, 3.$$

(ii) The prior density is $\pi(q) = 2q,\ 0 < q < 1$.
A randomly chosen insured has zero claims in Year 1.
Using Bühlmann credibility, estimate the number of claims in Year 2 for the selected insured.

Problem 83.5 ‡
You are given:
(i) Claims are conditionally independent and identically Poisson distributed with mean Θ.
(ii) The prior distribution function of Θ is:

$$\pi(\theta) = 1 - \left(\frac{1}{1+\theta}\right)^{2.6}, \ \theta > 0.$$

Five claims are observed. Determine the Bühlmann credibility factor.

Problem 83.6 ‡
You are given:
(i) Claim counts follow a Poisson distribution with mean λ.
(ii) Claim sizes follow a lognormal distribution with parameters μ and σ.
(iii) Claim counts and claim sizes are independent.
(iv) The prior distribution has joint probability density function:

$$f(\lambda, \mu, \sigma) = 2\sigma, \ 0 < \lambda < 1, \ \ 0 < \mu < 1, \ 0 < \sigma < 1.$$

Calculate Bühlmann k for aggregate losses.

Problem 83.7
For a portfolio of policies, you are given:
(i) The annual claim amount on a policy has probability density function:

$$f(x|\theta) = \frac{2x}{\theta^2}, \ 0 < x < \theta.$$

(ii) The prior distribution of θ has density function:

$$\pi(\theta) = 4\theta^3, \ 0 < \theta < 1.$$

(iii) A randomly selected policy had claim amount 0.1 in Year 1.
Determine the Bühlmann credibility estimate of the claim amount for the selected policy in Year 2.

Problem 83.8 ‡
You are given the following information about workers compensation coverage:
(i) The number of claims for an employee during the year follows a Poisson distribution with mean $(100 - p)/100$, where p is the salary (in thousands) for the employee.

(ii) The distribution of p is uniform on the interval $(0, 100]$.
An employee is selected at random. During the last 4 years, the employee has had a total of 5 claims.
Determine the Bühlmann credibility estimate for the expected number of claims the employee will have next year.

84 The Bühlmann-Straub Credibility Model

The standard Bühlmann credibility model assumes that the conditional losses $X_i|\Theta$, where $i = 1, 2, \cdots, n$, of a policyholder in the past n years are independent and identically distributed with respect to each past year. A drawback of this assumption is that it does not allow for variations in exposure or size. For example, what happens if claims experience in year 1 for a policyholder reflected only a portion of a year? For a group insurance, what happens if the size of the group changed over time?

These issues can be resolved by modifying the Bühlmann model as follows: Let m_i be the number of exposure units in year i with exposure units $X_{i1}, X_{i2}, \cdots, X_{im_i}$ such that $X_{i1}|\Theta, X_{i2}|\Theta, , \cdots, X_{im_i}|\Theta$ are independent with the same mean $\mu(\theta)$ and variance $v(\theta)$. Let X_i be the average of the claims in year i. That is,

$$X_i = \frac{X_{i1} + X_{i2} + \cdots + X_{im_i}}{m_i}.$$

Note that $E(X_i|\Theta) = \mu(\theta)$ (the hypothetical mean) and $\text{Var}(X_i|\theta) = \frac{v(\theta)}{m_i}$ (the process variance). It is still assumed that there is independence from one period (or group) to another.

As in the Bühlmann model, we let

$$\mu = E[\mu(\Theta)], \quad v = E[v(\Theta)], \quad a = \text{Var}[\mu(\Theta)].$$

We then have the following

$$\begin{aligned}
E(X_i) =& E[E(X_i|\Theta)] = E[\mu(\Theta)] = \mu \\
\text{Cov}(X_i, X_j) =& E(X_i, X_j) - E(X_i)E(X_j) \\
=& E[E(X_iX_j|\Theta)] - E[\mu(\Theta)]^2 \\
=& E[E(X_i|\Theta)E(X_j|\Theta)] - E[\mu(\Theta)]^2 \quad \text{(by independence)} \\
=& E[\mu^2(\Theta)] - E[\mu(\Theta)]^2 \\
=& \text{Var}[\mu(\Theta)] = a, \ i \neq j \\
\text{Var}(X_i) =& E[\text{Var}(X_i|\Theta)] + \text{Var}[E(X_i|\Theta)] \\
=& E\left[\frac{v(\Theta)}{m_i}\right] + \text{Var}[\mu(\Theta)] \\
=& \frac{v}{m_i} + a
\end{aligned}$$

To find the credibility premium, we denote the total exposure by $m = \sum_{i=1}^n m_i$. Using the normal equations, we find

$$E(X_{n+1}) = \mu = \hat{\alpha}_0 + \sum_{i=1}^n \hat{\alpha}_i \mu \Longrightarrow \sum_{i=1}^n \hat{\alpha}_i = 1 - \frac{\hat{\alpha}_0}{\mu}. \tag{84.1}$$

For $i = 1, 2, \cdots, n$, (81.4) becomes

$$a = \sum_{\substack{j=1 \\ j \neq i}}^{n} \hat{\alpha}_j a + \hat{\alpha}_i \left(a + \frac{v}{m_i}\right) = \sum_{j=1}^n \hat{\alpha}_j a + \frac{v\hat{\alpha}_i}{m_i}.$$

Solving this equation for $\hat{\alpha}_i$, we find

$$\hat{\alpha}_i = \frac{a}{v} m_i \left(1 - \sum_{j=1}^n \hat{\alpha}_j\right) = \frac{a}{v}\frac{\hat{\alpha}_0}{\mu} m_i. \tag{84.2}$$

By (84.1) and (84.2), we find

$$1 - \frac{\hat{\alpha}_0}{\mu} = \sum_{j=1}^n \hat{\alpha}_j = \sum_{j=1}^n \frac{a}{v}\frac{\hat{\alpha}_0}{\mu} m_j = \frac{a}{v}\frac{\hat{\alpha}_0}{\mu} m$$

which leads to

$$\hat{\alpha}_0 = \frac{v/a}{m + v/a}\mu$$

and

$$\hat{\alpha}_i = \frac{m_i}{m + v/a}.$$

Letting $k = \frac{v}{a}$, the credibility premium can be expressed as

$$\hat{\alpha}_0 + \sum_{j=1}^n \hat{\alpha}_j X_j = Z\overline{X} + (1 - Z)\mu$$

where

$$Z = \frac{m}{m+k} \quad \text{and} \quad \overline{X} = \frac{m_1 X_1 + m_2 X_2 + \cdots + m_n X_n}{m}.$$

If X_i is interpreted to be the average loss/claims experienced by the m_i group members in year i, then $m_i X_i$ is the total loss/claims of the m_i group members in year i. Also, $\overline{X}$ is the overall loss/claims per group member over the n years. The credibility premium to be charged to the group in year $n+1$ is $m_{n+1}[Z\overline{X} + (1-Z)\mu]$ for the m_{n+1} members in the next year. Keep in mind that $Z\overline{X} + (1 - Z)\mu$ is the credibility premium per exposure unit (i.e., per occurrence of an individual X_{ij}).

Remark 84.1
If $m_i = 1$ for $i = 1, 2, \cdots, n$ then the Bühlmann-Straub model coincides with the original Bühlmann model.

Example 84.1
You are given:
(i) In year j, there are N_j claims for m_j policies.
(ii) An individual policy has a Poisson distribution with mean Λ.
(iii) Λ has a Gamma distribution with parameters α and β.
(a) Determine the Bühlmann-Straub estimate of the number of claims for one policyholder in year $n+1$.
(b) Determine the Bühlmann-Straub estimate of the number of claims in year $n+1$ if there will be m_{n+1} policies.

Solution.
We let $X_i = \frac{N_i}{m_i}$. Because N_i has a Poisson distribution with mean $m_i\lambda$ then $X_i|\Lambda$ has a Poisson distribution with mean λ. Thus,

$$\text{Var}(X_i|\Lambda) = \frac{1}{m_i^2}\text{Var}(N_i) = \frac{m_i\lambda}{m_i^2} = \frac{v(\lambda)}{m_i} \Longrightarrow v(\lambda) = \lambda.$$

We have

$$\begin{aligned}
\mu(\Lambda) =& E(X_i|\Lambda) = \lambda\\
\mu =& E[\mu(\Lambda)] = E(\Lambda) = \alpha\beta\\
a =& \text{Var}[\mu(\Lambda)] = \text{Var}(\Lambda) = \alpha\beta^2\\
v =& E[v(\Lambda)] = E(\Lambda) = \alpha\beta\\
k =& \frac{v}{a} = \frac{\alpha\beta}{\alpha\beta^2} = \frac{1}{\beta}\\
Z =& \frac{m}{m+1/\beta} = \frac{m\beta}{m\beta+1}.
\end{aligned}$$

(a) The Bühlmann-Straub estimate of the number of claims for one policyholder in year $n+1$ is

$$P_c = \left(\frac{m\beta}{m\beta+1}\right)\overline{X} + \left(\frac{1}{m\beta+1}\right)(\alpha\beta)$$

where $\overline{X} = \frac{1}{m}\sum_{i=1} nm_iX_i$.
(b) The Bühlmann-Straub estimate of the number of claims if there are m_{n+1} policies in year $n+1$ is $m_{n+1}P_c$ ■

Example 84.2 ‡
You are given:
(i) The number of claims incurred in a month by any insured has a Poisson distribution with mean λ.
(ii) The claim frequencies of different insureds are independent.
(iii) The prior distribution is Gamma with probability density function:

$$\pi(\lambda) = \frac{(100\lambda)^6 e^{-100\lambda}}{120\lambda}.$$

(iv)

Month	Number of Insureds	Number of Claims
1	100	6
2	150	8
3	200	11
4	300	?

Determine the Bühlmann-Straub credibility estimate of the number of claims in Month 4.

Solution.
Let $X_j = \frac{N_j}{m_j}$. Note that

$$\text{Var}(X_i|\Lambda) = \frac{1}{m_i^2}\text{Var}(N_i) = \frac{m_i\lambda}{m_i^2} = \frac{v(\lambda)}{m_i} \Longrightarrow v(\lambda) = \lambda.$$

We have

$$\begin{aligned}
\mu(\lambda) =& E(X_i|\Lambda) = \lambda\\
\mu =& E[\mu(\Lambda)] = E(\Lambda) = \alpha\beta = 0.06\\
a =& \text{Var}[\mu(\Lambda)] = \text{Var}(\Lambda) = \alpha\beta^2 = 0.0006\\
v =& E[v(\Lambda)] = E(\Lambda) = 0.06\\
k =& \frac{v}{a} = \frac{0.06}{0.0006} = 100\\
Z =& \frac{m}{m+k} = \frac{450}{450+100} = \frac{9}{11}\\
\overline{X} =& \frac{N_1+N_2+N_3}{m} = \frac{6+8+11}{450} = \frac{25}{450}.
\end{aligned}$$

The credibility estimate of the expected number of claims for one insured in month 4 is

$$P_c = Z\overline{X} + (1-Z)\mu = \frac{9}{11}\frac{25}{450} + \frac{2}{11}(0.06) = 0.056364.$$

For 300 insureds, the expected number of claims in Month 4 is $300(0.056364) = 16.9$ ■

Example 84.3 ‡
For a portfolio of insurance risks, aggregate losses per year per exposure follow a normal distribution with mean θ and standard deviation 1000, with θ varying by class as follows:

Class	θ	Percent of Risks in Class
X	2000	60%
Y	3000	30%
Z	4000	10%

A randomly selected risk has the following experience over three years:

Year	Number of Exposures	Aggregate Losses
1	24	24,000
2	30	36,000
3	26	28,000

Calculate the Bühlmann-Straub estimate of the mean aggregate losses per year per exposure in Year 4 for this risk.

Solution.
Let X_i be the aggregate losses per exposure in year i. That is, the average over all exposures of the total losses in year i. Thus,

$$\begin{aligned} X_1 =&\frac{24,000}{24} = 1,000 \\ X_2 =&\frac{36,000}{30} = 1,200 \\ X_1 =&\frac{28,000}{26} = \frac{14,000}{13}. \end{aligned}$$

We are given: $m_1 = 24,\ m_2 = 30,\ m_3 = 26$. The risk parameter Θ is the mean of the normal distribution. The prior distribution is

$$\begin{aligned} \pi(2000) =&0.6 \\ \pi(3000) =&0.3 \\ \pi(4000) =&0.1. \end{aligned}$$

We are given that X_i is the average of m_i (independent) random variables each normally distributed with mean θ and variance 1000^2, conditional on $\Theta = \theta$. Hence, we have the following

$$\begin{aligned}
\mu(\theta) =& E(X_i|\Theta) = \theta \\
v(\theta) =& 1,000,000 \\
\mu =& E(\Theta) = 2000(0.6) + 3000(0.3) + 4000(0.1) \\
=& 2500 \\
v =& 1,000,000 \\
a =& \mathrm{Var}(\Theta) = 2000^2(0.6) + 3000^2(0.3) + 4000^2(0.1) - 2500^2 \\
=& 450,000 \\
k =& \frac{v}{a} = \frac{1,000,000}{450,000} \\
Z =& \frac{80}{80 + \frac{1,000,000}{450,000}} = 0.97297 \\
\overline{X} =& \frac{1}{m}\sum_{i=1}^{3} m_i X_i \\
=& \frac{1}{80}\left[24(1000) + 30(1200) + 26\left(\frac{14,000}{13}\right)\right] \\
=& 1,100.
\end{aligned}$$

Hence, the The Bühlmann-Straub credibility estimate for the loss per exposure in Year 4 is

$$Z\overline{X} + (1 - Z)\mu = 0.97297(1,100) + (1 - 0.97297)(2500) = 1137.84 \blacksquare$$

Practice Problems

Problem 84.1
Let $X_1, X_2, \cdots, X_n$ be losses in the past n years. Suppose that they all have the same risk parameter θ. We assume that $X_i|\Theta$ are idependent with mean $\mu(\theta) = E(X_i|\Theta$ and variance $\text{Var}(X_i|\Theta) = w(\theta) + \frac{v(\theta)}{m_i}$. Such a credibility model is knwon as **Hewitt's model**.
(a) Show that

$$\text{Var}\left(\frac{m_iX_i + m_jX_j}{m_i+m_j}\middle|\Theta\right) = \frac{m_i^2+m_j^2}{(m_i+m_j)^2}w(\theta) + \frac{v(\theta)}{m_i+m_j}.$$

(b) Show that $\mu = E(X_i),\ \text{Cov}(X_i, X_j) = a$ for $i \neq j$ and $\text{Var}(X_i) = w + \frac{v}{m_j}$ where $v = E[v(\Theta)$ and $w = E[w(\Theta)]$.

Problem 84.2
Consider the Hewitt's model introduced above. Let $\hat{\alpha}_0 + \sum_{j=1}^n \hat{\alpha}_j X_j$ be the credibility premium. Using normal equations, show that

$$\hat{\alpha}_i = \tfrac{a\hat{\alpha}_0/\mu}{w+v/m_i} \quad \text{and} \quad \hat{\alpha}_0 = \tfrac{\mu}{1+am^*}$$

where

$$m^* = \sum_{j=1}^n \frac{m_j}{v+wm_j}.$$

Problem 84.3
Consider the Hewitt's model introduced above. Show that

$$\hat{\alpha}_0 + \sum_{j=1}^n \hat{\alpha}_j X_j = Z\overline{X} + (1-Z)\mu$$

where

$$Z = \tfrac{am^*}{1+am^*} \quad \text{and} \quad \overline{X} = \tfrac{1}{m^*}\sum_{j=1}^n \frac{m_j}{v+wm_j}X_j.$$

Problem 84.4 ‡
You are given four classes of insureds, each of whom may have zero or one claim, with the following probabilities:

Class	Number of claims	
	0	1
I	0.9	0.1
II	0.8	0.2
III	0.5	0.5
IV	0.1	0.9

A class is selected at random (with probability 0.25), and four insureds are selected at random from the class. The total number of claims is two.
If five insureds are selected at random from the same class, estimate the total number of claims using Bühlmann-Straub credibility.

Problem 84.5 ‡
You are given the following data on large business policyholders:
(i) Losses for each employee of a given policyholder are independent and have a common mean and variance.
(ii) The overall average loss per employee for all policyholders is 20.
(iii) The variance of the hypothetical means is 40.
(iv) The expected value of the process variance is 8000.
(v) The following experience is observed for a randomly selected policyholder:

Year	Average loss per Employee	Number of Employees
1	15	800
2	10	600
3	5	400

Determine the Bühlmann-Straub credibility premium per employee for this policyholder.

Problem 84.6 ‡
Members of three classes of insureds can have 0, 1 or 2 claims, with the following probabilities:

	Number of Claims		
Class	0	1	2
I	0.9	0.0	0.1
II	0.8	0.1	0.1
III	0.7	0.2	0.1

A class is chosen at random, and varying numbers of insureds from that class are observed over 2 years, as shown below:

Year	Number of Insureds	Number of Claims
1	20	7
2	30	10

Determine the Bühlmann-Straub credibility estimate of the number of claims in Year 3 for 35 insureds from the same class.

Problem 84.7 ‡
You are given:
(i) The number of claims incurred in a month by any insured follows a Poisson distribution with mean λ.
(ii) The claim frequencies of different insureds are independent.
(iii) The prior distribution of Λ is Weibull with $\theta = 0.1$ and $\tau = 2$.
(iv) Some values of the gamma function are:

$$\Gamma(0.5) = 1.77245,\ \Gamma(1) = 1,\ \Gamma(1.5) = 0.88623,\ \Gamma(2) = 1.$$

(v)

Month	Number of Insureds	Number of claims
1	100	10
2	150	11
3	250	14

Determine the Bühlmann-Straub credibility estimate of the number of claims in the next 12 months for 300 insureds.

Problem 84.8 ‡
For each policyholder, losses $X_1, \cdots, X_n$, conditional on Θ, are independently and identically distributed with mean,

$$\mu(\theta) = E(X_j|\Theta = \theta),\ j = 1, 2, \cdots, n$$

and variance,

$$v(\theta) = \mathrm{Var}(X_j|\Theta = \theta),\ j = 1, 2, \cdots, n.$$

You are given:
(i) The Bühlmann credibility assigned for estimating X_5 based on $X_1, \cdots, X_4$ is $Z = 0.4$.
(ii) The expected value of the process variance is known to be 8.
Calculate $\mathrm{Cov}(X_i, X_j),\ i \neq j$.

Problem 84.9 ‡

You are given n years of claim data originating from a large number of policies. You are asked to use the Bühlmann-Straub credibility model to estimate the expected number of claims in year $n+1$.

Which of conditions (A), (B), or (C) are required by the model?

(A) All policies must have an equal number of exposure units.

(B) Each policy must have a Poisson claim distribution.

(C) There must be at least 1082 exposure units.

(D) Each of (A), (B), and (C) is required.

(E) None of (A), (B), or (C) is required.

Problem 84.10 ‡

You are given the following information about a single risk:

(i) The risk has m exposures in each year.

(ii) The risk is observed for n years.

(iii) The variance of the hypothetical means is a.

(iv) The expected value of the annual process variance is $w + \frac{v}{m}$.

Determine the limit of the Bhlmann-Straub credibility factor as m approaches infinity.

85 Exact Credibility

The term **exact credibility** refers to the case when the credibility premium is equal to the Bayesian premium. Exact credibility arises in Bühlmann and Bühlmann-Struab models specifically in situations involving the linear exponential family members and their conjugate priors (see Sections 24 and 66) which we demonstrate next.

Let $X_i|\Theta$ be a member of the linear exponential family. Then its pdf can be expressed as

$$f_{X_i|\Theta}(x|\theta) = \frac{p(x)e^{r(\theta)x}}{q(\theta)}$$

and where the support of $X_i|\Theta$ does not depend on θ.

For a member in the linear exponential family, the mean is given by (see Example 24.2)

$$\mu(\theta) = E(X_i|\Theta) = \frac{q'(\theta)}{r'(\theta)q(\theta)}.$$

The variance (see Example 24.3) is given by

$$\mathrm{Var}(X_i|\Theta) = \frac{\mu'(\theta)}{r'(\theta)}.$$

Let the pdf of Θ (the prior distribution) be given by

$$\pi(\theta) = \frac{[q(\theta)]^{-k}e^{\mu k r(\theta)}r'(\theta)}{c(\mu, k)},\ \theta_0 < \theta < \theta_1.$$

We will show that the posterior distribution is of the same type as $\pi(\theta)$ so that the prior distribution is a conjugate prior distribution. Indeed, we have

$$\begin{aligned}
\int f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta)d\theta &= \int \frac{\left[\prod_{j=1}^n p(x_j)\right] e^{r(\theta)\sum_{j=1}^n x_j}}{[q(\theta)]^n} \times \frac{[q(\theta)]^{-k} e^{\mu k r(\theta)} r'(\theta)}{c(\mu,k)} d\theta \\
&= \left[\prod_{j=1}^n p(x_j)\right] \int \frac{[q(\theta)]^{-(k+n)} e^{r(\theta)\left(\frac{\mu k+\sum_{j=1}^n x_j}{k+n}\right)(k+n)} r'(\theta)}{c(\mu,k)} d\theta \\
&= \left[\prod_{j=1}^n p(x_j)\right] \int \frac{[q(\theta)]^{-k^*} e^{r(\theta)\mu^* k^*} r'(\theta)}{c(\mu,k)} \\
&= \left[\prod_{j=1}^n p(x_j)\right]
\end{aligned}$$

where

$$\mu^* = \tfrac{\mu k+\sum_{j=1}^n x_j}{k+n} \quad \text{and} \quad k^* = k+n.$$

The posterior distribution is

$$\begin{aligned}
\pi(\theta|\mathbf{x}) &= \frac{\frac{\left[\prod_{j=1}^n p(x_j)\right] e^{r(\theta)\sum_{j=1}^n x_j}}{[q(\theta)]^n} \times \frac{[q(\theta)]^{-k} e^{\mu k r(\theta)} r'(\theta)}{c(\mu,k)}}{\left[\prod_{j=1}^n p(x_j)\right]} \\
&= \frac{[q(\theta)]^{-(k+n)} e^{r(\theta)\left(\frac{\mu k+\sum_{j=1}^n x_j}{k+n}\right)(k+n)} r'(\theta)}{c(\mu,k)} \\
&= \frac{[q(\theta)]^{-k^*} e^{r(\theta)\mu^* k^*} r'(\theta)}{c(\mu,k)}.
\end{aligned}$$

From the expression of $\pi(\theta|\mathbf{x})$, we can write

$$\ln\left[\frac{\pi(\theta|\mathbf{x})}{r'(\theta)}\right] = -k^* \ln[q(\theta)] + \mu^* k^* r(\theta) - \ln[c(\mu,k)].$$

Differentiating with respect to θ yields

$$\frac{[\pi(\theta|\mathbf{x})/r'(\theta)]'}{\pi(\theta|\mathbf{x})/r'(\theta)} = -k^* \frac{q'(\theta)}{q(\theta)} + \mu^* k^* r'(\theta)$$

which can be rearranged as

$$\frac{d}{d\theta}\left[\frac{\pi(\theta|\mathbf{x})}{r'(\theta)}\right] = -k^*[\mu(\theta) - \mu^*]\pi(\theta|\mathbf{x}).$$

Integrating over the interval (θ_0, θ_1) yields

$$\frac{\pi(\theta_1|\mathbf{x})}{r'(\theta_1)} - \frac{\pi(\theta_0|\mathbf{x})}{r'(\theta_0)} = -k[E(X_{n+1}|\mathbf{X}) - \mu^*]$$

which yields the Bayesian premium

$$E(X_{n+1}|\mathbf{X} = \mu^* + \frac{\pi(\theta_0|\mathbf{x})}{kr'(\theta_0)} - \frac{\pi(\theta_1|\mathbf{x})}{kr'(\theta_1)}.$$

Assuming

$$\frac{\pi(\theta_0|\mathbf{x})}{r'(\theta_0)} = \frac{\pi(\theta_1|\mathbf{x})}{r'(\theta_1)}$$

we see that $E(X_{n+1}|\mathbf{X} = \mu^* = a_0 + \sum_{j=1}^n a_j X_j$.
Now recall that the Bühlmann premium is the minimizer of (see Remark 81.1)

$$Q_1 = E\left\{\left[E(X_{n+1}|\mathbf{X}) - \alpha_0 - \sum_{i=1}^n \alpha_i X_i\right]^2\right\}.$$

By letting, $\alpha_j = a_j$ for $j = 0, 1, \cdots, n$ we see that the credibility premium coincides with the Bayesian premium and credibility is exact.

The following table list some models where the Bayesian premium equals the credibility premium:

$f_{X\|\Theta}(\mathbf{x}\|\theta)$	$\pi(\theta)$	$\pi(\theta\|\mathbf{x})$
Poisson	Gamma	Gamma
Normal	Normal	Normal
Bernoulli	Beta	Beta
Exponential	Inverse Gamma	Inverse Gamma

Example 85.1
You are given:
(i) The model distribution $X|\Lambda$ is Poisson with parameter Λ.
(ii) The prior distribution of Λ is Gamma with parameters α and θ.
Show that this model satisfies exact credibility.

Solution.
Example 66.2 shows that the posterior distribution is a Gamma distribution with parameters $\alpha' = \alpha + n\overline{X}$ and $\theta' = \frac{\beta}{n\beta+1}$. The hypothetical mean is

$$\mu(\lambda) = E(X_i|\Lambda) = \lambda$$

and the Bayesian premium is

$$\begin{aligned}
E(X_{n+1}|\mathbf{X}) = &\int_0^\infty \lambda \frac{\lambda^{\sum_{i=1}^n x_i+\alpha-1}e^{-\frac{\lambda(n\theta+1)}{\theta}}}{\left(\frac{\theta}{n\theta+1}\right)^{\sum_{i=1}^n x_i+\alpha}\Gamma(\sum_{i=1}^n x_i+\alpha)}d\lambda\\
=&(n\overline{X}+\alpha)\frac{\theta}{n\theta+1}\int_0^\infty \frac{\lambda^{\sum_{i=1}^n x_i+\alpha+1-1}e^{-\frac{\lambda(n\theta+1)}{\theta}}}{\left(\frac{\theta}{n\theta+1}\right)^{\sum_{i=1}^n x_i+\alpha+1}\Gamma(\sum_{i=1}^n x_i+\alpha+1)}d\lambda\\
=&(n\overline{X}+\alpha)\frac{\theta}{n\theta+1}
\end{aligned}$$

Note that the Bayesian premium is a linear function of $X_1, X_2, \cdots, X_n$. Next, we find the Bühlmann credibility. We have

$$\begin{aligned}
\mu(\lambda) =&\lambda\\
\mu =&E(\Lambda) = \alpha\theta\\
v(\lambda) =&\mathrm{Var}(X_i|\Lambda) = \lambda\\
v =&E(\Lambda) = \alpha\theta\\
a =&\mathrm{Var}(\Lambda) = \alpha\theta^2\\
k =&\frac{v}{a} = \frac{1}{\theta}\\
Z =&\frac{n}{n+k} = \frac{n\theta}{n\theta+1}\\
P_c =&Z\overline{X} + (1-Z)\mu\\
=&\frac{n\theta}{n\theta+1}\overline{X} + \frac{1}{n\theta+1}(\alpha\theta)\\
=&(n\overline{X}+\alpha)\frac{\theta}{n\theta+1}.
\end{aligned}$$

Thus, the credibility premium equals the Bayesian premium ■

Example 85.2 ‡
You are given:
(i) The number of claims per auto insured follows a Poisson distribution with mean λ.
(ii) The prior distribution for Λ has the following probability density function:

$$f(\lambda) = \frac{(500\lambda)^{50}e^{-500\lambda}}{\lambda\Gamma(50)}.$$

(iii) A company observes the following claims experience:

	Year 1	Year 2
Number of claims	75	210
Number of autos insured	600	900

The company expects to insure 1100 autos in Year 3. Determine the expected number of claims in Year 3.

Solution.
The model distribution is Poisson with mean λ and the prior distribution is Gamma with $\alpha = 500$ and $\theta = \frac{1}{500}$. Thus, this model satisfies exact credibility.

The Bayesian Approach Solution
Let X_n denote the number of claims in Year n. We are asked to find $E(X_3|X_1, X_2)$. We have $E(X_3|\Lambda) = \lambda$. The model distribution is Poisson and the prior distribution is Gamma with parameters $\alpha = 50$ and $\theta = \frac{1}{500}$. By Example 66.2, the posterior distribution is Gamma with parameters $\alpha' = 75 + 210 + 50 = 335$ and $\theta' = \frac{\theta}{n\theta+1} = \frac{(1/500)}{1500(1/500)+1} = 0.0005$. Thus, $E(X_3|X_1, X_2) = E(\Lambda|\mathbf{X}) = \alpha'\theta' = (335)(0.0005) = 0.1675$. This is the expected number of claims per policy. The expected number of claims in the next year is $E(X_3|X_1, X_2) = 1100(0.1675) = 184.25$.

The Bühlmann Approach Solution
We have

$$\begin{aligned}
\mu(\lambda) =& E(\mathbf{X}|\Lambda) = \lambda\\
\mu =& E(\Lambda) = \alpha\theta = \frac{50}{500} = 0.1\\
a =& \mathrm{Var}(\Lambda) = \alpha\theta^2 = \frac{50}{500^2} = 0.0002\\
v(\lambda) =& \Lambda\\
v =& E(\Lambda) = 0.1\\
k =& \frac{v}{a} = \frac{0.1}{0.0002} = 500\\
Z =& \frac{n}{n+k} = \frac{1500}{1500+500} = 0.75\\
\overline{X} =& \frac{75+210}{600+900} = 0.19.
\end{aligned}$$

Thus,

$$E(X_3|X_1, X_2) = 1100[Z\overline{X} + (1-Z)\mu] = 1100[0.19(0.75) + 0.25(0.1)] = 184.25 \blacksquare$$

Practice Problems

Problem 85.1
You are given:
(i) The model distribution $X|\Lambda$ is exponential with mean Λ.
(ii) The prior distribution of Λ is inverse Gamma with parameters α and β.
Show that this model satisfies exact credibility.

Problem 85.2
You are given:
• The model distribution $\mathbf{X}|Q$ is binomial with parameters m and q.
• The prior distribution Q is beta with parameters a, b and 1.
Show that this model satisfies exact credibility.

Problem 85.3
You are given:
• The model distribution $\mathbf{X}|\Lambda$ is normal with mean Λ and variance σ^2.
• The prior distribution Λ is normal with mean μ and variance a^2.
Show that this model satisfies exact credibility.

Problem 85.4 ‡
You are given:
(i) The conditional distribution of the number of claims per policyholder is Poisson with mean λ.
(ii) The variable λ has a gamma distribution with parameters α and θ.
(iii) For policyholders with 1 claim in Year 1, the credibility estimate for the number of claims in Year 2 is 0.15.
(iv) For policyholders with an average of 2 claims per year in Year 1 and Year 2, the credibility estimate for the number of claims in Year 3 is 0.20.
Determine θ.

86 Non-parametric Empirical Bayes Estimation for the Bühlmann Model

The credibility models introduced earlier assume that the conditional distribution and the prior distribution are known to be of a certain form. More precisely, they are of parametric form. However, practical models necessitate the choice of parameters in such a way to ensure a close agreement between the model and reality.

What we will do next is to look at situations where the conditional distribution and the prior distribution are unknown. We refer to this situation as the **nonparametric case**. The unknown quantities of interest that are related to the prior distribution such as μ, v, and a (the ones to be estimated) are called **structural parameters**. The estimation process will be based on data at hand. We refer to this process as the **empirical Bayes estimation**.

The format of the data for our analysis has the following structure: Suppose we have $r \geq 1$ policyholders or groups of policyholders. For policyholder i there are n_i years of claim experience/observed exposure units. Let X_{ij} be the average number of losses/claims for policyholder i in year j. Let the claim vector for average number of losses/claims for policyholder i over all years be:

$$\mathbf{X}_i = (X_{i1}, \cdots, X_{in_i})^T, \ i = 1, \cdots, r.$$

We assume that $\mathbf{X}_1, \cdots, \mathbf{X}_r$ are independent, this is reasonable to think that different groups claims will be independent of each other.

Let θ_i denote the risk parameter for the i^{th} policyholder. We assume that the Θ_i are independent and identically distributed. We also assume that $X_{ij}|\Theta_i$ are independent for $j = 1, \cdots, n_i$.

Let m_{ij} denote the number of exposure units for policyholder i in year j. Then the total number of exposure units over all the years is

$$m_i = \sum_{j=1}^{n_i} m_{ij}, \ i = 1, \cdots, r.$$

The total exposure units for all policyholders over all the years is

$$m = \sum_{i=1}^{r} m_i.$$

The past average loss experience of policyholder i over all the years is

$$\overline{X}_i = \frac{1}{m_i} \sum_{j=1}^{n_i} m_{ij} X_{ij}, \ i = 1, \cdots, r.$$

The overall average losses are

$$\overline{X} = \frac{1}{m} \sum_{i=1}^{r} m_i \overline{X}_i.$$

Similary to the previous credibility premiums we can define the structural parameters:

$$\mu = E[\mu(\Theta_i)], \quad v = E[v(\Theta_i)], \quad a = \mathrm{Var}[\mu(\Theta_i)]$$

where $\mu(\theta_i) = E(X_{ij}|\Theta_i)$ and $v(\theta_i) = \mathrm{Var}(X_{ij}\Theta_i)$.
These parameters will be estimated from the data at hand. Let $\hat{\mu}, \hat{v}$, and $\hat{a}$ be the estimators of μ, v, and a respectively. Then the credibility premium estimate is

$$\hat{Z}_i \overline{X}_i + (1 - \hat{Z}_i)\hat{\mu}_i$$

where

$$\hat{Z}_i = \frac{m_i}{m_i + \hat{k}}, \quad \hat{k} = \frac{\hat{v}}{\hat{a}}.$$

The credibility premium to cover all m_{i,n_i+1} exposure units for policyholder i in the next year would be

$$m_{i,n_i+1}[\hat{Z}_i \overline{X}_i + (1 - \hat{Z}_i)\hat{\mu}_i], \ i = 1, \cdots, r.$$

Our goal is to estimate the parameters μ, v and a in the Bühlmann and Bühlmann-Straub models given the data. In this section, we cover the Bühlmann model.
For Bühlmann model, we assume that $m_{ij} = 1$ and $n_i = n$ for $i = 1, \cdots, r$ and $j = 1, \cdots, n$. Thus, $m_i = n$ and $m = nr$. Moreover,

$$\overline{X}_i = \frac{1}{n} \sum_{j=1}^{n} X_{ij}, \ i = 1, \cdots, r$$

and

$$\overline{X} = \frac{1}{r} \sum_{i=1}^{r} \overline{X}_i = \frac{1}{nr} \sum_{i=1}^{r} \sum_{j=1}^{n} X_{ij}.$$

For an estimator of μ, we let $\hat{\mu} = \overline{X}$. The expected value of $\hat{\mu}$ is

$$\begin{aligned}
E(\hat{u}) =& \frac{1}{rn}\sum_{i=1}^{r}\sum_{j=1}^{n}E(X_{ij}) = \frac{1}{rn}\sum_{i=1}^{r}\sum_{j=1}^{n}E[E(X_{ij}|\Theta_i)\\
=& \frac{1}{rn}\sum_{i=1}^{r}\sum_{j=1}^{n}E[\mu(\Theta_i)] = \frac{1}{rn}\sum_{i=1}^{r}\sum_{j=1}^{n}\mu = \mu.
\end{aligned}$$

Hence, $\hat{\mu}$ is an unbiased estimator of μ.
In order to estimate v, we first need the following result.

Theorem 86.1
Let $Y_1, Y_2, \cdots, Y_k$ be independent random variables with the same mean $\mu = E(Y_i)$ and the same variance $\sigma^2 = \text{Var}(Y_i)$. Let $\overline{Y} = \frac{1}{k}\sum_{i=1}^{k} Y_i$. Then
(a) $E(\overline{Y}) = \mu$
(b) $E\left[\frac{1}{k-1}\sum_{i=1}^{k}(Y_i - \overline{Y})^2\right] = \sigma^2$. That is, $\frac{1}{k-1}\sum_{i=1}^{k}(Y_i - \overline{Y})^2$ is an unbiased estimator of the variance of Y_i.

Proof.
(a) We have

$$E(\overline{Y}) = \frac{1}{k}\sum_{i=1}^{k}E(Y_i) = \frac{k\mu}{k} = \mu.$$

(b) We have

$$\begin{aligned}
\sum_{i=1}^{k}(Y_i - \overline{Y})^2 =& \sum_{i=1}^{k}[(Y_i\mu) + (\mu - \overline{Y})]^2\\
=& \sum_{i=1}^{k}[(Y_i - \mu)^2 + 2(Y_i - \mu)(\mu - \overline{Y}) + (\mu - \overline{Y})^2]\\
=& \sum_{i=1}^{k}(Y_i - \mu)^2 + 2(\mu - \overline{Y}))\sum_{i=1}^{k}(Y_i - \mu) + \sum_{i=1}^{k}(\overline{Y}) - \mu)^2\\
=& \sum_{i=1}^{k}(Y_i - \mu)^2 + 2(\mu - \overline{Y}))(k\overline{Y} - k\mu) + k(\overline{Y}) - \mu)^2\\
=& \sum_{i=1}^{k}(Y_i - \mu)^2 - 2k(\overline{Y} - \mu)^2 + k(\overline{Y}) - \mu)^2\\
=& \sum_{i=1}^{k}(Y_i - \mu)^2 - k(\overline{Y} - \mu)^2.
\end{aligned}$$

Taking the expectation of both sides yields

$$\begin{aligned} E\left[\sum_{i=1}^{k}(Y_i-\overline{Y})^2\right] =& \sum_{i=1}^{k} E[(Y_i-\mu)^2] - kE[(\overline{Y}-\mu)^2] \\ =& \sum_{i=1}^{k} \operatorname{Var}(Y_i) - k\operatorname{Var}(\overline{Y}) \\ =& k\sigma^2 - k\left(\frac{\sigma^2}{k}\right) = (k-1)\sigma^2. \end{aligned}$$

Hence,

$$E\left[\frac{1}{k-1}\sum_{i=1}^{k}(Y_i-\overline{Y})^2\right] = \sigma^2 \ \blacksquare$$

By the above theorem, an unbiased estimator to $\operatorname{Var}(X_{ij}|\Theta_i) = v(\theta_i)$ is

$$\hat{v}_i = \frac{1}{n-1}\sum_{j=1}^{n}(X_{ij}-\overline{X}_i)^2.$$

Since

$$E(\hat{v}_i) = E[E(\hat{v}_i|\Theta_i)] = E[v(\Theta_i)] = v$$

and

$$E\left(\frac{1}{r}\sum_{i=1}^{r}\hat{v}_i\right) = v$$

an unbiased estimator of v is

$$\hat{v} = \frac{1}{r}\sum_{i=1}^{r}\hat{v}_i = \frac{1}{r(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{n}(X_{ij}-\overline{X}_i)^2.$$

It remains to find an estimator of a. We begin with

$$E(\overline{X}_i|\Theta_i) = \frac{1}{n}\sum_{j=1}^{n} E(X_{ij}|\Theta_i) = \frac{1}{n}\sum_{j=1}^{n}\mu(\theta_i) = \mu(\theta_i).$$

unconditionally, we have

$$E(\overline{X}_i) = E[E(\overline{X}_i|\Theta_i)] = E[\mu(\Theta_i)] = \mu.$$

Moreover,

$$\begin{aligned}\mathrm{Var}(\overline{X}_i) =&\mathrm{Var}[E(\overline{X}_i|\Theta_i)] + E[\mathrm{var}(\overline{X}_i|\Theta_i)]\\ =&\mathrm{Var}[\mu(\Theta_i)] + E\left[\frac{v(\Theta_i)}{n}\right] = a + \frac{v}{n}.\end{aligned}$$

Let $\hat{a}$ be an estimator of a, $\hat{v}$ be an estimator of v. An unbiased estimator of $\mathrm{Var}(\overline{X}_i)$ is $\frac{1}{r-1}\sum_{i=1}^{r}(\overline{X}_i - \overline{X})^2$. Thus, we have

$$\frac{1}{r-1}\sum_{i=1}^{r}\overline{X}_i = \hat{a} + \frac{\hat{v}}{n}$$

which implies

$$\begin{aligned}\hat{a} =&\frac{1}{r-1}\sum_{i=1}^{r}(\overline{X}_i - \overline{X})^2 - \frac{\hat{v}}{n}\\ =&\frac{1}{r-1}\sum_{i=1}^{r}(\overline{X}_i - \overline{X})^2 - \frac{1}{rn(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{n}(X_{ij} - \overline{X}_i)^2.\end{aligned}$$

Remark 86.1

Due to the subtraction in the formula for $\hat{a}$, it is possible that $\hat{a} \le 0$. When this happens, it is customary to set $\hat{a} = \hat{Z} = 0$.

Example 86.1

You are given the losses of two policyholders over a period of three years:

Policyholder	Year 1	Year 2	Year 3
1	3	5	7
2	6	12	9

Determine the Bayes estimate of the Bühlmann premium for each policyholder for Year 4.

Solution.
We have

$$\begin{aligned}
\overline{X}_1 =&\frac{3+5+7}{3} = 5\\
\overline{X}_2 =&\frac{6+12+9}{3} = 9\\
\hat{\mu} =&\overline{X} = \frac{5+9}{2} = 7\\
\hat{v}_1 =&\frac{1}{3-1}\sum_{j=1}^{3}(X_{1j}-\overline{X}_1)^2\\
=&\frac{1}{3-1}[(3-5)^2+(5-5)^2+(7-5)^2] = 4\\
\hat{v}_2 =&\frac{1}{3-1}\sum_{j=1}^{3}(X_{2j}-\overline{X}_2)^2\\
=&\frac{1}{3-1}[(6-9)^2+(12-9)^2+(9-9)^2] = 9\\
\hat{v} =&\frac{\hat{v}_1+\hat{v}_2}{2} = \frac{9+4}{2} = 6.5\\
\hat{a} =&\frac{1}{r-1}\sum_{i=1}^{r}(\overline{X}_i-\overline{X})^2 - \frac{\hat{v}}{n}\\
=&\frac{1}{2-1}[(5-7)^2+(9-7)^2] - \frac{6.5}{3} = \frac{35}{6}\\
\hat{k} =&\frac{\hat{v}}{\hat{a}} = \frac{\frac{13}{2}}{\frac{35}{6}} = \frac{39}{35}\\
\hat{Z} =&\frac{3}{3+\frac{39}{35}} = \frac{35}{48}.
\end{aligned}$$

The estimated Bühlmann premium for policyholder 1 in year 4 is:

$$\hat{Z}\overline{X}_1 + (1-\hat{Z})\hat{\mu} = \frac{35}{48}(5) + \left(1-\frac{35}{48}\right)(7) = \frac{133}{24}.$$

The estimated Bühlmann premium for policyholder 2 in year 4 is:

$$\hat{Z}\overline{X}_2 + (1-\hat{Z})\hat{\mu} = \frac{35}{48}(9) + \left(1-\frac{35}{48}\right)(7) = \frac{203}{24} \blacksquare$$

Remark 86.2
Due to the subtraction in the formula for $\hat{a}$, it is possible that $\hat{a} \leq 0$. When this happens, it is customary to set $\hat{a} = \hat{Z} = 0$.

Example 86.2 ‡
Survival times are available for four insureds, two from Class A and two from Class B. The two from Class A died at times $t=1$ and $t=9$. The two from Class B died at times $t=2$ and $t=4$.
Nonparametric Empirical Bayes estimation is used to estimate the mean survival time for each class. Unbiased estimators of the expected value of the process variance and the variance of the hypothetical means are used.
Estimate Z, the Bühlmann credibility factor.

Solution.
We have: $r=n=2, X_{11}=1, X_{12}=9, X_{21}=2, X_{22}=4, \overline{X}_1=5$ and $\overline{X}_2=3$. Thus,

$$\begin{aligned}\hat{v} =& \frac{1}{r(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{n}(X_{ij}-\overline{X}_i)^2\\ =& \frac{1}{2(2-1)}[(1-5)^2+(9-5)^2+(2-3)^2+(4-3)^2]=17\end{aligned}$$

and

$$\begin{aligned}\hat{a} =& \frac{1}{r-1}\sum_{i=1}^{r}(\overline{X}_i-\overline{X})^2-\frac{1}{rn(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{n}(X_{ij}-\overline{X}_i)^2\\ =& \frac{1}{2-1}[(5-4)^2+(3-4)^2]-\frac{17}{2}=-6.5.\end{aligned}$$

Since $\hat{a}<0$, we obtain $\hat{Z}=0$ ■

Practice Problems

Problem 86.1 ‡
An insurer has data on losses for four policyholders for 7 years. The loss from the i^{th} policyholder for year j is X_{ij}.
You are given:

$$\sum_{i=1}^{4}\sum_{j=1}^{7}(X_{ij}-\overline{X}_i)^2 = 33.60$$

and

$$\sum_{i=1}^{4}(\overline{X}_i-\overline{X})^2 = 3.30.$$

Using nonparametric empirical Bayes estimation, calculate the Bühlmann credibility factor for an individual policyholder.

Problem 86.2 ‡
You are given total claims for two policyholders:

	Year			
Policyholder	1	2	3	4
X	730	800	650	700
Y	655	650	625	750

Using the nonparametric empirical Bayes method, determine the Bühlmann credibility premium for Policyholder Y.

Problem 86.3 ‡
Three individual policyholders have the following claim amounts over four years:

Policyholder	Year 1	Year 2	Year 3	Year 4
X	2	3	3	4
Y	5	5	4	6
Z	5	5	3	3

Using the nonparametric empirical Bayes procedure, calculate the estimated variance of the hypothetical means.

Problem 86.4 ‡
Three policyholders have the following claims experience over three months:

Policyholder	Month 1	Month 2	Month 3	Mean	Variance
I	4	6	5	5	1
II	8	11	8	9	3
III	5	7	6	6	1

Nonparametric empirical Bayes estimation is used to estimate the credibility premium in Month 4.
Calculate the credibility factor Z.

87 Non-parametric Empirical Bayes Estimation for the Bühlmann-Straub Model

In this section, we turn our attention to estimating the credibility factors μ, v, and a for the Bühlmann-Straub Model.
Recall that for this model, the hypothetical mean is

$$E(X_{ij}|\Theta_i) = \mu(\theta_i)$$

and the process variance is

$$m_{ij}\text{Var}(X_{ij}|\Theta_i) = v(\theta_i).$$

From

$$\overline{X} = \frac{1}{m}\sum_{i=1}^{r} m_i\overline{X}_i$$

we find

$$\begin{aligned}
E(\overline{X}) =&\frac{1}{m}\sum_{i=1}^{r} m_iE(\overline{X}_i) = \frac{1}{m}\sum_{i=1}^{r} m_iE[E(\overline{X}_i|\Theta_i)])\\
=&\frac{1}{m}\sum_{i=1}^{r} m_iE\left[\sum_{j=1}^{n_i}\frac{m_{ij}}{m_i}E(X_{ij}|\Theta_i)\right]\\
=&\frac{1}{m}\sum_{i=1}^{r} m_iE\left[\sum_{j=1}^{n_i}\frac{m_{ij}}{m_i}\mu(\theta_i)\right] = \frac{1}{m}\sum_{i=1}^{r} m_iE[\mu(\theta_i)]\\
=&\frac{1}{m}\sum_{i=1}^{r} m_i\mu = \mu.
\end{aligned}$$

Hence, $\hat{\mu} = \overline{X}$ is an unbiased estimator of μ.
Next, we turn our attention to estimating v. We begin with

$$E(\overline{X}_i|\Theta_i) = \frac{1}{m_i}\sum_{j=1}^{n_i} m_{ij}E(X_{ij}|\Theta_i) = \mu(\theta_i).$$

The conditional variance of $\overline{X}_i$ is

$$\begin{aligned}\mathrm{Var}(\overline{X}_i|\Theta_i) =& \frac{1}{m_i^2}\sum_{j=1}^{n_i} m_{ij}^2 \mathrm{Var}(X_{ij}|\Theta_i)\\ =& \frac{1}{m_i^2}\sum_{j=1}^{n_i} m_{ij}^2 \frac{v(\theta_i)}{m_{ij}}\\ =& \frac{v(\theta_i)}{m_i}.\end{aligned}$$

Similar to the argument in Section 86, one can show that

$$\sum_{j=1}^{n_i} m_{ij}(X_{ij}-\overline{X}_i)^2 = \sum_{j=1}^{n_i} m_{ij}(X_{ij}-\mu(\theta_i))^2 - m_i(\mu(\theta_i)-\overline{X}_i)^2.$$

Taking the expectations of both sides, we find

$$\begin{aligned}E\left[\sum_{j=1}^{n_i} m_{ij}(X_{ij}-\overline{X}_i)^2\right] =& \sum_{j=1}^{n_i} m_{ij}E[(X_{ij}-\mu(\theta_i)^2] - m_iE[(\mu(\theta_i)-\overline{X}_i)^2]\\ =& \sum_{j=1}^{n_i} m_{ij}\mathrm{Var}(X_{ij}|\Theta_i) - m_i\mathrm{Var}(\overline{X}_i|\Theta_i)\\ =& \sum_{j=1}^{n_i} m_{ij}\frac{v(\theta_i)}{m_{ij}} - m_i\frac{v(\theta_i)}{m_i}\\ =& n_iv(\theta_i) - v(\theta_i) = (n_i-1)v(\theta_i).\end{aligned}$$

Thus,

$$E\left[\frac{1}{n_i-1}\sum_{j=1}^{n_i} m_{ij}(X_{ij}-\overline{X}_i)^2\right] = v(\theta_i).$$

Thus, an unbiased estimator of $v(\theta_i)$ is

$$\hat{v}_i = \frac{1}{n_i-1}\sum_{j=1}^{n_i} m_{ij}(X_{ij}-\overline{X}_i)^2,\ i=1,2,\cdots,r.$$

By the double expectation formula, we have

$$E(\hat{v}_i) = E[E(\hat{v}_i|\Theta_i)] = E[v(\theta_i)] = v$$

so that $\hat{v}_i$ is an unbiased estimator of v. Another unbiased estimator of v is $\hat{v} = \sum_{i=1}^{r} w_i \hat{v}_i$ where $\sum_{i=1}^{r} w_i - 1$, One choice of the w_i is

$$w_i = \frac{n_i - 1}{\sum_{i=1}^{r}(n_i - 1)}.$$

Hence, an unbiased estimator to v is

$$\hat{v} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \overline{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}.$$

We now turn to estimation of a. we first note that

$$\begin{aligned}\operatorname{Var}(\overline{X}_i) =& \operatorname{Var}[E(\overline{X}_i|\Theta_i)] + E[\operatorname{Var}(\overline{X}_i|\Theta_i)] \\ =& \operatorname{Var}[\mu(\Theta_i)] + E\left[\frac{v(\Theta_i)}{m_i}\right] = a + \frac{v}{m_i}.\end{aligned}$$

Again, one can show

$$\sum_{i=1}^{r} m_i(\overline{X}_i - \overline{X})^2 = \sum_{i=1}^{r} m_i(\overline{X}_i - \mu)^2 - m(\overline{X} - \mu)^2.$$

Taking the expectations of both sides, we find

$$\begin{aligned}E\left[\sum_{i=1}^{r} m_i(\overline{X}_i - \overline{X})^2\right] =& \sum_{i=1}^{r} m_i E[(\overline{X}_i - \mu)^2] - mE[(\overline{X} - \mu)^2] \\ =& \sum_{i=1}^{r} m_i \operatorname{Var}(\overline{X}_i) - m\operatorname{Var}(\overline{X}) \\ =& \sum_{i=1}^{r} m_i\left(a + \frac{v}{m_i}\right) - m \cdot \frac{1}{m^2}\sum_{i=1}^{r} m_i^2\left(a + \frac{v}{m_i}\right) \\ =& ma + rv - \frac{1}{m}\left(a\sum_{i=1}^{r} m_i^2 + mv\right) \\ =& a\left(m - \frac{1}{m}\sum_{i=1}^{r} m_i^2\right) + v(r-1).\end{aligned}$$

Thus,

$$E\left\{\left(m - \frac{1}{m}\sum_{i=1}^{r} m_i^2\right)^{-1}\left[\sum_{i=1}^{r} m_i(\overline{X}_i - \overline{X})^2 - v(r-1)\right]\right\} = a$$

so that an unbiased estimator of a is

$$\hat{a} = \left(m - \frac{1}{m}\sum_{i=1}^{r} m_i^2\right)^{-1}\left[\sum_{i=1}^{r} m_i(\overline{X}_i - \overline{X})^2 - \hat{v}(r-1)\right].$$

With these estimators, for the policyholder i, we have

$$\hat{k} = \frac{\hat{v}}{\hat{a}} \text{ and } \hat{Z}_i = \frac{m_i}{m_i + \hat{k}}$$

Thus, the premium for each member of the policy is:

$$\hat{Z}_i\overline{X}_i + (1 - \hat{Z}_i)\hat{\mu}$$

and the credibility premium to cover all m_{i,n_i+1} exposure units for policyholder i in the next year would be

$$m_{i,n_i+1}[\hat{Z}_i\overline{X}_i + (1 - \hat{Z}_i)\hat{\mu}_i], \; i = 1, \cdots, r.$$

Remark 87.1
Note that the above equations provide unbiased estimators of μ, v, and a respectively. They are nonparametric, requiring no distributional assumptions. Also, due to the subtraction in the formula for $\hat{a}$, it is possible that $\hat{a} \leq 0$. When this happens, it is customary to set $\hat{a} = \hat{Z} = 0$.

Example 87.1
You are given:

	Policy Group	Year 1	Year 2	Year 3	Year 4
Total Losses	I	?	750	600	?
Number in Group		?	3	2	4
Total Losses	II	975	1200	900	?
Number in Group		5	6	4	5

(a) Calculate the unbiased estimates for μ, v, and a in the Bühlmann-Straub model.
(b) Determine the Bühlmann-Straub premium for each policyholder in Year 4.

Solution.
(a) We have

$$\begin{aligned}
\overline{X}_1 =& \frac{750+600}{3+2} = 270\\
\overline{X}_2 =& \frac{975+1200+900}{5+6+4} = 205\\
\overline{X} =& \frac{5}{200}\overline{X}_1 + \frac{15}{20}\overline{X}_2\\
=& \frac{5}{200}(270) + \frac{15}{20}(205) = 221.25\\
\hat{\mu} =& \overline{X} = 221.25\\
\hat{v}_1 =& \frac{3\left(\frac{750}{3}-270\right)^2 + 2\left(\frac{600}{2}-270\right)^2}{2-1} = 3000\\
\hat{v}_2 =& \frac{5\left(\frac{975}{5}-205\right)^2 + 6\left(\frac{1200}{6}-205\right)^2 + 4\left(\frac{900}{4}-205\right)^2}{3-1} = 1125\\
\hat{v} =& \frac{(2-1)\hat{v}_1 + (3-1)\hat{v}_2}{(2-1)+(3-1)} = \frac{3000+2(1125)}{3} = 1750\\
\hat{a} =& \frac{1}{20 - \frac{5^2+15^2}{20}}[5(270-221.25)^2 + 15(205-221.25)^2 - (2-1)(1750)] = 1879.1667.
\end{aligned}$$

(b) We have

$$\begin{aligned}
\hat{k} =& \frac{\hat{v}}{\hat{a}} = \frac{1750}{1879.1667} = 0.9313\\
\hat{Z}_1 =& \frac{m_1}{m_1+\hat{k}} = \frac{5}{5+0.9313} = 0.843\\
\hat{Z}_2 =& \frac{15}{15+0.9313} = 0.9415.
\end{aligned}$$

The premium in year 4 for a policyholder in Group I is

$$4[\hat{Z}_1\overline{X}_1 + (1-\hat{Z}_1)\hat{\mu}] = 4[0.8413(270) + (1-0.8413)(221.25)] = 1049.38.$$

The premium in year 4 for a policyholder in Group II is

$$5[\hat{Z}_2\overline{X}_2 + (1-\hat{Z}_2)\hat{\mu}] = 5[0.9415(205) + (1-0.9415)(221.25)] = 1029.75 \blacksquare$$

Now, the total losses (TL) on all policyholders in observed units is

$$TL = \sum_{i=1}^{r} m_i\overline{X}_i.$$

The total premium (TP) for all policyholders in year $n+1$ is:

$$TP = \sum_{i=1}^{r} m_i[\hat{Z}_i\overline{X}_i + (1-\hat{Z}_i)\hat{\mu}].$$

We can express (TP) in terms of (TL) as follows

$$TP = \sum_{i=1}^{r} m_i(1-\hat{Z}_i)(\hat{\mu}-\overline{X}_i) + \underbrace{\sum_{i=1}^{r} m_i\overline{X}_i}_{TL}.$$

It is often desirable for TP to equal TL. This leads to the following calculation

$$\begin{aligned}
0 =& \sum_{i=1}^{r} m_i(1-\hat{Z}_i)(\hat{\mu}-\overline{X}_i)\\
0 =& \sum_{i=1}^{r} \hat{k}\hat{Z}_i(\hat{\mu}-\overline{X}_i)\\
\hat{\mu}\sum_{i=1}^{r}\hat{Z}_i =& \sum_{i=1}^{r}\hat{Z}_i\overline{X}_i\\
\hat{\mu} =& \frac{\sum_{i=1}^{r}\hat{Z}_i\overline{X}_i}{\sum_{i=1}^{r}\hat{Z}_i}.
\end{aligned}$$

Hence, we have another estimator of $\hat{\mu}$. We refer to this process of estimating $\hat{\mu}$ as the **credibility weighted average method** or the method of preserving total losses/claims (See Problem 87.4).

Example 87.2
Redo Example 87.1(b) if μ is estimated by credibility weighted average.

Solution.
We have

$$\hat{\mu} = \frac{\hat{Z}_1\overline{X}_1 + \hat{Z}_2\overline{X}_2}{\hat{Z}_1+\hat{Z}_2} = \frac{0.843(270) + 0.9415(205)}{0.843 + 0.9415} = 235.7061.$$

The premium in year 4 for a policyholder in Group I is

$$4[\hat{Z}_1\overline{X}_1 + (1-\hat{Z}_1)\hat{\mu}] = 4[0.8413(270) + (1-0.8413)(235.7061)] = 1058.44.$$

The premium in year 4 for a policyholder in Group II is

$$5[\hat{Z}_2\overline{X}_2 + (1-\hat{Z}_2)\hat{\mu}] = 5[0.9415(205) + (1-0.9415)(235.7061)] = 1033.98 \blacksquare$$

Practice Problems

Problem 87.1 ‡
You are given the following commercial automobile policy experience:

	Company	Year 1	Year 2	Year 3
Total Losses	I	50,000	50,000	?
Number of Automobiles		100	200	?
Total Losses	II	?	150,000	150,000
Number of Automobiles		?	500	300
Total Losses	II	150,000	?	150,000
Number of Automobiles		50	?	150

Determine the nonparametric empirical Bayes credibility factor, Z, for Company III.

Problem 87.2 ‡
You are given:

	Group	Year 1	Year 2	Year 3	Total
Total Claims	1		10,000	15,000	25,000
Number in Group			50	60	110
Average			200	250	227.27
Total Claims	2	16,000	18,000		34,000
Number in Group		100	90		190
Average		160	200		178.95
Total Claims					59,000
Number in Group					300
Average					196.67

You are also given $\hat{a} = 651.03$..
Use the nonparametric empirical Bayes method to estimate the credibility factor for Group 1.

Problem 87.3 ‡
You are given the following data:

	Year 1	Year 2
Total losses	12,000	14,000
Number of Policyholders	25	30

The estimate of the variance of the hypothetical means is 254.
Determine the credibility factor for Year 3 using the nonparametric empirical Bayes method.

Problem 87.4 ‡
You are making credibility estimates for regional rating factors. You observe that the Bühlmann-Straub nonparametric empirical Bayes method can be applied, with rating factor playing the role of pure premium.
X_{ij} denotes the rating factor for region i and year j, where $i = 1, 2, 3$ and $j = 1, 2, 3, 4$.
Corresponding to each rating factor is the number of reported claims, m_{ij}, measuring exposure.
You are given:

i	$m_i = \sum_{i=1}^{4} m_i$	$\overline{X}_i = \frac{1}{m_i}\sum_{j=1}^{4} m_{ij}X_{ij}$	$\hat{v}_i = \frac{1}{3}\sum_{j=1}^{4} m_{ij}(X_{ij} - \overline{X}_i)^2$	$m_i(\overline{X}_i - \overline{X})^2$
1	50	1.406	0.536	0.887
2	300	1.298	0.125	0.191
3	150	1.178	0.172	1.348

Determine the credibility estimate of the rating factor for region 1 using the method that preserves $\sum_{i=1}^{3} m_i\overline{X}_i$.

Problem 87.5 ‡
You are given the following experience for two insured groups:

Group		Year			
		1	2	3	Total
1	Number of members	8	12	5	25
	Average loss per member	96	91	113	97
2	Number of members	25	30	20	75
	Average loss per member	113	111	116	113
Total	Number of members				100
	Average loss per member				109

$$\sum_{i=1}^{2}\sum_{j=1}^{3} m_{ij}(X_{ij} - \overline{X}_i)^2 = 2020 \quad \text{and} \quad \sum_{i=1}^{2} m_i(\overline{X}_i - \overline{X})^2 = 4800.$$

Determine the nonparametric Empirical Bayes credibility premium for group 1, using the method that preserves total losses.

Problem 87.6 ‡
You are given:
(i) A region is comprised of three territories. Claims experience for Year 1 is as follows:

A	10	4
B	20	5
C	30	3

(ii) The number of claims for each insured each year has a Poisson distribution.
(iii) Each insured in a territory has the same expected claim frequency.
(iv) The number of insureds is constant over time for each territory.
Determine the Bühlmann-Straub empirical Bayes estimate of the credibility factor Z for Territory A.

Problem 87.7 ‡
You are given the following information on towing losses for two classes of insureds, adults and youths:
Exposures

Exposure			
Year	Adult	Youth	Total
1996	2000	450	2450
1997	1000	250	1250
1998	1000	175	1175
1999	1000	125	1125
Total	5000	1000	6000

Pure Premium			
Year	Adult	Youth	Total
1996	0	15	2.755
1997	5	2	4.400
1998	6	15	7.340
1999	4	1	3.667
Weighted Average	3	10	4.167

You are also given that the estimated variance of the hypothetical means is 17.125.
Determine the nonparametric empirical Bayes credibility premium for the youth class.

88 Semiparametric Empirical Bayes Credibility Estimation

In this section, we look at **semiparamteric** methods where parametric assumptions concerning $X|\Theta$ are made, while the prior distribution of the risk parameter Θ remains unspecified and nonparametric (and thus the name semiparametric).

Let X_{ij} denote the average number of claims for policyholder i in year j where $i = 1, 2, \cdots, r$. Suppose that the number of claims, given Θ_i, $m_{ij}X_{ij}|\Theta_i$ has a Poisson distribution with parameter $m_{ij}\theta_i$. Under these assumptions, we have

$$\begin{aligned}\mu(\theta_i) =& E[X_{ij}|\Theta_i) = \frac{1}{m_{ij}}E(m_{ij}X_{ij}|\Theta_i) = \frac{m_{ij}\theta_i}{m_{ij}} = \theta_i \\ v(\theta_i) =& m_{ij}\mathrm{Var}(X_{ij}|\Theta_i) = \frac{1}{m_{ij}}\mathrm{Var}(m_{ij}X_{ij}|\Theta_i) = \frac{m_{ij}\theta_i}{m_{ij}} = \theta_i.\end{aligned}$$

It follows that

$$\mu = v = E(\Theta_i)$$

and an unbiased estimator for both μ and v is $\overline{X}$. That is, $\hat{\mu} = \hat{v} = \overline{X}$.

Next, we turn our attention to estimating a. We first note that

$$\begin{aligned}\mathrm{Var}(X_{ij}) =& \mathrm{Var}[E(X_{ij}|\Theta_i)] = E[\mathrm{Var}(X_{ij}|\Theta_i)] \\ =& \mathrm{Var}[\mu(\Theta_i)] + E[v(\Theta_i)] = a + v.\end{aligned}$$

Now, if we assume $n_i = m_{ij} = 1$, then (see Section 87)

$$E[\sum_{i=1}^{r}(X_{i1} - \overline{X})^2] = a(r-1) + v(r-1)$$
$$E\left[\frac{1}{r-1}\sum_{i=1}^{r}(X_{i1} - \overline{X})^2\right] = a + v.$$

Hence, $\frac{1}{r-1}\sum_{i=1}^{r}(X_{i1} - \overline{X})^2$ is an unbiased estimator of $a + v$ and therefore

$$\hat{a} = \frac{1}{r-1}\sum_{i=1}^{r}(X_{i1} - \overline{X})^2 - \hat{v}.$$

Example 88.1 *([1])*
In the past year, the distribution function of automobile insurance policyholders the number of claims is given by:

# of claims	# of insureds
0	1563
1	271
2	32
3	7
4	2
Total	1875

Estimate the credibility premium for the number of claims next year for a policyholder with 2 claims. Assume a conditional Poisson distribution function of the number of claims for each policyholders.

Solution.
We have: $r = 1875, n_i = 1, m_{i1} = 1$ and $X_{i1}|\Theta_i$ is Poisson with parameter θ_i. The estimators of $\mu, v,$ and a are found as follows:

$$\begin{aligned}
\overline{X} =& \frac{1}{1875}\sum_{i=1}^{1875} X_{i1} = \frac{0(1563) + 1(271) + 2(32) + 3(7) + 4(2)}{1875} = 0.194 \\
\hat{\mu} =& \hat{v} = 0.194 \\
\sum_{i=1}^{1875}(X_{i1} - \overline{X})^2 =& 1563(0 - 0.194)^2 + 271(1 - 0.194)^2 + 32(2 - 0.194)^2 \\
& +7(3 - 0.194)^2 + 2(4 - 0.194)^2 = 423.3355 \\
\hat{a} =& \frac{423.3355}{1874} - 0.194 = 0.032 \\
\hat{k} =& \frac{0.194}{0.032} = 6.06 \\
\hat{Z} =& \frac{1}{1 + 6.06} = 0.14.
\end{aligned}$$

The estimated credibility premium for the number of claims for each policyholder is

$$0.14X_{i1} + (0.86)(0.194).$$

For a policyholder with two claims, $X_{i1} = 2$ so that the premium is

$$0.14(2) + (0.86)(0.194) = 0.44684 \blacksquare$$

Practice Problems

Problem 88.1
Suppose that $m_{ij}X_{ij}|\Theta_i$ has a binomial distribution with parameters (m_{ij}, θ_i). Express a in terms of μ and v.

Problem 88.2 ‡
You are given:
(i) During a single 5-year period, 100 policie s had the following total claims experience:

Number of claims in Year 1 through Year 5	Number of Policies
0	46
1	34
2	13
3	5
4	2

(ii) The number of claims per year follows a Poisson distribution.
(iii) Each policyholder was insured for the entire period.
A randomly selected policyholder had 3 claims over the period.
Using semiparametric empirical Bayes estimation, determine the Bhlmann estimate for the number of claims in Year 6 for the same policyholder.

Problem 88.3 ‡
You are given:
(i) During a 2-year period, 100 policies had the following claims experience:

Number of claims in Year 1 through Year 2	Number of Policies
0	50
1	30
2	15
3	4
4	1

(ii) The number of claims per year follows a Poisson distribution.
(iii) Each policyholder was insured for the entire 2-year period.
A randomly selected policyholder had one claim over the 2-year period.

Using semiparametric empirical Bayes estimation, determine the Bühlmann estimate for the number of claims in Year 3 for the same policyholder.

Problem 88.4 ‡
You are given:
(i) Over a three-year period, the following claim experience was observed for two insureds who own delivery vans:

		Year		
Insured		1	2	3
A	Number of Vehicles	2	2	1
	Number of claims	1	1	0
B	Number of Vehicles	N/A	3	2
	Number of claims	N/A	2	3

(ii) The number of claims for each insured each year follows a Poisson distribution.
Determine the semiparametric empirical Bayes estimate of the claim frequency per vehicle for Insured A in Year 4.

Problem 88.5 ‡
For a group of auto policyholders, you are given:
(i) The number of claims for each policyholder has a conditional Poisson distribution.
(ii) During Year 1, the following data are observed for 8000 policyholders:

Number of claims	Number of Policies
0	5000
1	2100
2	750
3	100
4	50
5^+	0

A randomly selected policyholder had one claim in Year 1.
Determine the semiparametric empirical Bayes estimate of the number of claims in Year 2 for the same policyholder.

Problem 88.6 ‡
For a portfolio of motorcycle insurance policyholders, you are given:
(i) The number of claims for each policyholder has a conditional Poisson distribution.
(ii) For Year 1, the following data are observed:

# of claims	# of insureds
0	2000
1	600
2	300
3	80
4	20
Total	3000

Determine the credibility factor, Z, for Year 2.

Problem 88.7 ‡
The following information comes from a study of robberies of convenience stores over the course of a year:
(i) X_i is the number of robberies of the i^{th} store, with $i = 1, 2, \cdots, 500$.
(ii) $\sum_{i=1}^{500} X_i = 50$.
(iii) $\sum_{i=1}^{500} X_i^2 = 220$.
(iv) The number of robberies of a given store during the year is assumed to be Poisson distributed with an unknown mean that varies by store.
Determine the semiparametric empirical Bayes estimate of the expected number of robberies next year of a store that reported no robberies during the studied year.

Basics of Stochastic Simulation

In this chapter we look at some stochastic simulation procedures to imitate or simulate financial and insurance problems. The term "stochastic" is used so that to emphasize the step of the simulation procedure where values are randomnly chosen from probability distributions.

89 The Inversion Method for Simulating Random Variables

A procedure of simulation can consist of the following four steps:

Step 1. Select the appropriate model X.
Step 2. For $j = 1, 2, \cdots, n$ generate simulated(also known as pseudorandom) values $x_1, x_2, \cdots, x_n$.
Step 3. An approximation of the cdf of X is the cdf of the empirical distribution $F_n(s)$ generated by $x_1, x_2, \cdots, x_n$.
Step 4. Statistical quantities such as mean, variance, percentiles, etc. are found using the empirical cdf $F_n(s)$.

For such procedure, two important questions are in place: The first question deals with finding the pseudorandom values. The second question is the size of n.

One method for finding the pseudorandom values is the **inversion method**[23] which we discuss next.

Let X be a continuous random variable with cdf $F_X(x)$. Since F is strictly increasing, it has an inverse F^{-1} defined on the interval $(0, 1)$. Thus, for $0 < u < 1$, the equation $F_X(x) = u$ has a unique solution x. Let U be the uniform distribution on $(0, 1)$. Then

$$F_{F_X^{-1}(U)}(x) = \Pr(F_X^{-1}(U) \le x) = \Pr(U \le F_X(x)) = F_U(F_X(x)) = F_X(x).$$

Hence, $F_X^{-1}(U)$ has the distribution function F_X. Thus, we can simulate a realization of X by simulating a realization of $F_X^{-1}(U)$.

Example 89.1
A random number generated from a uniform distribution on $(0, 1)$ is 0.3102. Using the inversion method, calculate the simulated value of X assuming X to have the Pareto distribution with parameters $\alpha = 3$ and $\theta = 1000$.

Solution.
We are seeking x such that $F_X(x) = 0.3102$. That is, x satisfies the equation

$$1 - \left(\frac{1000}{1000 + x}\right)^3 = 0.3012.$$

[23] This procedure is sometimes referred to with the phrase small uniform random numbers correspond to small simulated values.

Solving this equation, we find $x = 131.772$ ■

Example 89.2 ‡
You wish to simulate a value, Y, from a two point mixture. With probability 0.3, Y is exponentially distributed with mean 0.5. With probability 0.7, Y is uniformly distributed on $[-3, 3]$. You simulate the mixing variable where low values correspond to the exponential distribution. Then you simulate the value of Y, where low random numbers correspond to low values of Y. Your uniform random numbers from $[0, 1]$ are 0.25 and 0.69 in that order. Calculate the simulated value of Y.

Solution.
The statement "low random numbers correspond to low values of Y" implies that the inversion method is to be used. The value 0.25 is used to simulate the mixture and the number 0.69 is used to simulate the value of Y. We are told that for low simulating mixing variable, the exponential distribution must be used. Since $0.25 < 0.3$, this satisfies the criterion. That is, the exponential distribution is to be used. By the inversion method, the simulated value of Y is the number y such that $\Pr(Y \leq y) = 0.69$. That is, y is the solution to the equation $1 - e^{-\frac{y}{0.5}} = 0.69$. Solving this equation, we find $y = 0.5855$ ■

The inversion method can also be used to generate discrete or mixed-type variables. In the case of a mixed-type distribution, suppose $F_X(x)$ has a jump at $x = c$. Let $F_X(c^-) = a$ and $F_X(c) = b > a$. For any uniform number $a \leq u < b$, the equation $F_X(x) = u$ has no solution. In this case, we choose c as the simulated value.

Example 89.3
The cdf of X is given by

$$F_X(x) = \begin{cases} 0.3x, & 0 \leq x < 1 \\ 0.3 + 0.35x, & 1 \leq x \leq 2. \end{cases}$$

Determine the simulated values of x resulting from the uniform numbers: $0 \leq u_1 < 0.3$, $0.3 \leq u_2 < 0.65$, and $0.65 \leq u_3 \leq 1$.

Solution.
To find the simulated value of $0 \leq u_1 < 0.3$, we solve the equation $0.3x_1 = u_1$ obtaining $x_1 = \frac{u_1}{0.3}$. For any uniform number $0.3 \leq u_2 < 0.65$, the simulated value is $x_2 = 1$. Note that $\Pr(0.3 \leq U < 0.65) = 0.65 - 0.3 = 0.35 =$

$\Pr(X = 1)$. Finally, to find the simulated value of $0.65 \le u_3 \le 1$, we solve the equation $0.3 + 0.35x_3 = u_3$ obtaining $x_3 = \frac{u_3 - 0.3}{0.35}$ ■

It is possible that $F_X(x)$ is constant for $a \le x \le b$, say $F_X(x) = c$ in that interval. In this case, every value in $a \le x \le b$ is a simulated value of c. One chooses b as the simulated value corresponding to $u = c$. See Remark 89.1.

Example 89.4
Suppose that

$$F_X(x) = \begin{cases} 0.5x, & 0 \le x < 1.2 \\ 0.6, & 1.2 \le x < 2.4 \\ 0.5x - 0.6, & 2.4 \le x \le 3.2. \end{cases}$$

Determine the simulated values of x resulting from the uniform numbers $0.5, 0.6$, and 0.8.

Solution.
For $0 \le x < 1.2$, we have $0 \le F_X(x) < 0.6$. Since 0.5 is in that range, the simulad value resulting from 0.5 is found by solving the equation $0.5x_1 = 0.5$ which implies $x_1 = 1$. Next, we see that $F_X(x) = 0.6$ for $1.2 \le x \le 2.4$, For thia case, $x_2 = 2.4$. Finally, $0.6 \le 0.8 \le 1$, so the corresponding simulated value is found by solving the equation $0.5x_3 - 0.6 = 0.8$ resulting in $x_3 = 2.8$ ■

Recall that a discrete distribution has jumps at the possible values of the random variable and is constant in between, two features covered in the previous two examples.

Example 89.5
Simulate values from a binomial distribution with $m = 2$ and $q = 0.3$ using uniform numbers.

Solution.
The cdf of X is given by

$$F_X(x) = \sum_{i=0}^{[x]} \binom{m}{i} q^i (1-q)^{m-i}$$

where $[x]$ is the greatest integer less than or equal to x. Thus, for $m = 2$ and $q = 0.3$, we have

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 0.49, & 0 \le x < 1 \\ 0.91, & 1 \le x < 2. \\ 1, & x \ge 2. \end{cases}$$

For $0 \le u < 0.49$, the simulated value is $x = 0$. For $0.49 \le u < 0.91$, the simulated value is $x = 1$. For $0.91 \le u < 1$, the simulated value is $x = 2$ ■

Remark 89.1
Note that $F_X(x) = 0.49$ for all $0 \le x < 1$. But $\Pr(0.49 \le U < 0.91) = 0.91 - 0.49 = \Pr(X = 1)$. This is the motivation for choosing the largest value in an interval where the cdf is constant.

The second question that we look at is the number of simulated values needed to achieve a certain objective such as estimating the mean of X. We illustrate this question in the next example.

Example 89.6 *([1])*
Suppose X has the Pareto distribution with parameters $\alpha = 3$ and $\theta = 1000$. Use simulation to estimate the mean of X. Stop the simulation when you are 95% confident that the simulated mean is within 1% of the population mean. Assume that the central limit theorem is applicable.

Solution.
The empirical estimate of $\mu = E(X)$ is $\overline{x}$. The central limit theorem tells us that $\overline{X}_n$ is approximately normal so that we can write

$$\begin{aligned} 0.95 =& \Pr(|\overline{X}_n - \mu| \le 0.01\mu) \\ =& \Pr(0.99\mu \le \overline{X}_n \le 1.01\mu) \\ =& \Pr\left(-\frac{0.01\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le \frac{0.01\mu}{\sigma/\sqrt{n}}\right) \\ =& \Pr\left(-\frac{0.01\mu}{\sigma/\sqrt{n}} \le Z \le \frac{0.01\mu}{\sigma/\sqrt{n}}\right). \end{aligned}$$

Our goal is achieved when

$$\frac{0.01\mu}{\sigma/\sqrt{n}} \ge 1.96 \Longrightarrow n \ge \frac{38,416\sigma^2}{\mu^2}.$$

Since we do not know σ^2 and μ^2, we estimate them with the sample variance and mean. Thus, we cease simulation when

$$n \geq \frac{38416s^2}{\overline{x}^2} \blacksquare$$

We can apply a similar sort of idea to estimating a probability.

Example 89.7 *([1])*
Suppose X has the Pareto distribution with parameters $\alpha = 3$ and $\theta = 1000$. Use simulation to estimate $F_X(1000)$. Stop the simulation when you are 95% confident that the simulated mean is within 1% of the population mean. Assume that the central limit theorem is applicable.

Solution.
Let $\frac{P_n}{n}$ be the empirical estimator of $F_X(1000)$ where P_n is the number of values at or below 1000 after n simulations (see Sections 49). The central limit theorem tells us that $\frac{P_n}{n}$ is approximately normal with mean $F_X(1000)$ and variance $F_X(1000)[1 - F_X(1000)]/n$. (See Section 53). Using $\frac{P_n}{n}$ as an estimator of $F_X(1000)$ and arguing as in the previous example, we arrive at

$$n \geq 38416\left(\frac{n - P_n}{P_n}\right) \blacksquare$$

Practice Problems

Problem 89.1 ‡
To estimate $E(X)$, you have simulated X_1, X_2, X_3, X_4, and X_5 with the following results:

$$1 \quad 2 \quad 3 \quad 4 \quad 5.$$

You want the standard deviation of the estimator of $E(X)$ to be less than 0.05. Estimate the total number of simulations needed.

Problem 89.2 ‡
A company insures 100 people age 65. The annual probability of death for each person is 0.03. The deaths are independent.
Use the inversion method to simulate the number of deaths in a year. Do this three times using:

$$\begin{aligned} u_1 =& 0.20 \\ u_2 =& 0.03 \\ u_3 =& 0.09. \end{aligned}$$

Calculate the average of the simulated values.

Problem 89.3 ‡
You simulate observations from a specific distribution $F(x)$, such that the number of simulations N is sufficiently large to be at least 95 percent confident of estimating $F(1500)$ correctly within 1 percent.
Let P represent the number of simulated values less than 1500. Determine which of the following could be values of N and P.

(A)	$N = 2000$	$P = 1890$
(B)	$N = 3000$	$P = 2500$
(C)	$N = 3500$	$P = 3100$
(D)	$N = 4000$	$P = 3630$
(E)	$N = 4500$	$P = 4020$

Problem 89.4 ‡
You are planning a simulation to estimate the mean of a non-negative random variable. It is known that the population standard deviation is 20% larger than the population mean.
Use the central limit theorem to estimate the smallest number of trials needed so that you will be at least 95% confident that the simulated mean is within 5% of the population mean.

Problem 89.5 ‡
Simulation is used to estimate the value of the cumulative distribution function at 300 of the exponential distribution with mean 100.
Determine the minimum number of simulations so that there is at least a 99% probability that the estimate is within ±1% of the correct value.

Problem 89.6 ‡
You are simulating a compound claims distribution:
(i) The number of claims, N, is binomial with $m = 3$ and mean 1.8.
(ii) Claim amounts are uniformly distributed on $\{1, 2, 3, 4, 5\}$.
(iii) Claim amounts are independent, and are independent of the number of claims.
(iv) You simulate the number of claims, N, then the amounts of each of those claims, $X_1, X_2, \cdots, X_N$. Then you repeat another N, its claim amounts, and so on until you have performed the desired number of simulations.
(v) When the simulated number of claims is 0, you do not simulate any claim amounts.
(vi) All simulations use the inverse transform method, with low random numbers corresponding to few claims or small claim amounts.
(vii) Your random numbers from (0, 1) are

$$0.7 \quad 0.1 \quad 0.3 \quad 0.1 \quad 0.9 \quad 0.5 \quad 0.5 \quad 0.7 \quad 0.3 \quad 0.1$$

Calculate the aggregate claim amount associated with your third simulated value of N.

90 Applications of Simulation in Actuarial Modeling

In this section, we look at how simulation is used in solving actuaruial models.

Example 90.1 ‡
Unlimited claim severities for a warranty product follow the lognormal distribution with parameters $\mu = 5.6$ and $\sigma = 0.75$.
You use simulation to generate severities. The following are six uniform $(0, 1)$ random numbers:

$$0.6179 \quad 0.4602 \quad 0.9452 \quad 0.0808 \quad 0.7881 \quad 0.4207.$$

Using these numbers and the inversion method, calculate the average payment per claim for a contract with a policy limit of 400.

Solution.
Let X be the lognormal random variable with $\mu = 5.6$ and $\sigma = 0.75$. Its cdf is given by

$$F(x) = \Phi\left(\frac{\ln x - 5.6}{0.75}\right)$$

where Φ is the cdf of the standard normal distribution. Using the table of the standard normal distribution, we find

$$\Phi\left(\frac{\ln x_1 - 5.6}{0.75}\right) = 1-0.6179 = 0.3821 \Longrightarrow \frac{\ln x_1 - 5.6}{0.75} = 0.3 \Longrightarrow x_1 = 338.66.$$

In a similar manner, we find

$$\Phi\left(\frac{\ln x_2 - 5.6}{0.75}\right) = 1-0.4602 = 0.5398 \Longrightarrow \frac{\ln x_2 - 5.6}{0.75} = -0.1 \Longrightarrow x_3 = 250.89.$$

$$\Phi\left(\frac{\ln x_3 - 5.6}{0.75}\right) = 1-0.9452 = 0.0548 \Longrightarrow \frac{\ln x_3 - 5.6}{0.75} = 1.6 \Longrightarrow x_3 = 897.85.$$

$$\Phi\left(\frac{\ln x_4 - 5.6}{0.75}\right) = 1-0.0808 = 0.9192 \Longrightarrow \frac{\ln x_4 - 5.6}{0.75} = -1.4 \Longrightarrow x_4 = 94.63.$$

$$\Phi\left(\frac{\ln x_5 - 5.6}{0.75}\right) = 1-0.7881 = 0.2119 \Longrightarrow \frac{\ln x_5 - 5.6}{0.75} = 0.8 \Longrightarrow x_5 = 492.75.$$

$$\Phi\left(\frac{\ln x_6 - 5.6}{0.75}\right) = 1-0.4207 = 0.5793 \Longrightarrow \frac{\ln x_1 - 5.6}{0.75} = -0.2 \Longrightarrow x_6 = 232.76.$$

Because of the policy limits, the payments are

$$338.66 \quad 250.89 \quad 400 \quad 94.63 \quad 400 \quad 232.76.$$

The average payment per claim is

$$\frac{338.66 + 250.89 + 400 + 94.63 + 400 + 232.76}{6} = 286.16 \blacksquare$$

Example 90.2 ‡
Total losses for a group of insured motorcyclists are simulated using the aggregate loss model and the inversion method.
The number of claims has a Poisson distribution with $\lambda = 4$. The amount of each claim has an exponential distribution with mean 1000.
The number of claims is simulated using $u = 0.13$. The claim amounts are simulated using $u_1 = 0.05, u_2 = 0.95$, and $u_3 = 0.10$ in that order, as needed. Determine the total losses.

Solution.
Finding the first three probabilities of the Poisson distribution, we obtain

$$\begin{aligned} p_0 =& \frac{e^{-4}(4^0)}{0!} = 0.0183 \\ p_1 =& \frac{e^{-4}(4^1)}{1!} = 0.0733 \\ p_2 =& \frac{e^{-4}(4^2)}{2!} = 0.1463. \end{aligned}$$

Thus, $F_X(x) = 0.0183$ for $0 \le x < 1$, $F_X(x) = 0.0916$ for $1 \le x < 2$ and $F_X(x) = 0.2381$ for $2 \le x < 3$. Since $u = 0.13$ falls in the interval $(0.0916, 0.2381)$, the simulated number of claims is 2.
For the simulated amount of claim corresponding to $u_1 = 0.05$, we have

$$1 - e^{-\frac{x_1}{1000}} = 0.05 \Longrightarrow x_1 = 51.29.$$

Likewise,

$$1 - e^{-\frac{x_2}{1000}} = 0.95 \Longrightarrow x_2 = 2995.73.$$

Since the simulated number of claims is 2, there is no need to consider u_3. In conclusion, the total losses are $51.29 + 2995.73 = 3047.02$ ■

Example 90.3 ‡
Losses for a warranty product follow the lognormal distribution with underlying normal mean and standard deviation of 5.6 and 0.75 respectively.

You use simulation to estimate claim payments for a number of contracts with different deductibles.
The following are four uniform $(0,1)$ random numbers:

$$0.6217 \quad 0.9941 \quad 0.8686 \quad 0.0485.$$

Using these numbers and the inversion method, calculate the average payment per loss for a contract with a deductible of 100.

Solution.
Let X be the lognormal random variable with $\mu = 5.6$ and $\sigma = 0.75$. Its cdf is given by

$$F(x) = \Phi\left(\frac{\ln x - 5.6}{0.75}\right)$$

where Φ is the cdf of the standard normal distribution. Using the table of the standard normal distribution, we find

$$\Phi\left(\frac{\ln x_1 - 5.6}{0.75}\right) = 1 - 0.6217 = 0.3783 \Longrightarrow \frac{\ln x_1 - 5.6}{0.75} = 0.31 \Longrightarrow x_1 = 341.21.$$

In a similar manner, we find

$$\Phi\left(\frac{\ln x_2 - 5.6}{0.75}\right) = 1 - 0.9941 = 0.0059 \Longrightarrow \frac{\ln x_2 - 5.6}{0.75} = 2.52 \Longrightarrow x_2 = 1790.05.$$

$$\Phi\left(\frac{\ln x_3 - 5.6}{0.75}\right) = 1 - 0.8686 = 0.1314 \Longrightarrow \frac{\ln x_3 - 5.6}{0.75} = 1.12 \Longrightarrow x_3 = 626.41.$$

$$\Phi\left(\frac{\ln x_4 - 5.6}{0.75}\right) = 1 - 0.0485 = 0.9515 \Longrightarrow \frac{\ln x_4 - 5.6}{0.75} = -1.66 \Longrightarrow x_4 = 77.87.$$

The amounts after the deductible are

$$241.21 \quad 1690.05 \quad 526.41 \quad 0.$$

The average payment per loss is

$$\frac{241.21 + 1690.05 + 526.41 + 0}{4} = 614.42 \blacksquare$$

Example 90.4 ‡
You are the consulting actuary to a group of venture capitalists financing a search for pirate gold.
It's a risky undertaking: with probability 0.80, no treasure will be found,

and thus the outcome is 0.
The rewards are high: with probability 0.20 treasure will be found. The outcome, if treasure is found, is uniformly distributed on $[1000, 5000]$.
You use the inverse transformation method to simulate the outcome, where large random numbers from the uniform distribution on $[0, 1]$ correspond to large outcomes.
Your random numbers for the first two trials are 0.75 and 0.85. Calculate the average of the outcomes of these first two trials.

Solution.
Let X denote the outcome of the hunt. We first find the cdf of X. We are told that $F(0) = 0.80$. Also, for $1000 \le x \le 5000$, we have

$$F(x) = 0.8 + \frac{0.2}{5000 - 1000}(x - 1000).$$

Note the presence of 0.2 in the second term of $F(x)$. Without it, $F(5000) = 1.8 > 1$ which contradicts the definition of $F(x)$ (i.e. $0 \le F(x) \le 1$). So, the cdf of X can be expressed as

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.8, & 0 \le x < 1000 \\ 0.75 + 0.00005x, & 1000 \le x \le 5000 \\ 1, & x > 5000 \end{cases}$$

By the inversion method, for $u = 0.75$ we find $F(0.75) = 0$. For $u = 0.85$, we find $F(0.85) = 2000$. Thus, the average of the two outcomes is $\frac{0+2000}{2} = 1000$ ■

Example 90.5 ‡
You are simulating the gain/loss from insurance where:
(i) Claim occurrences follow a Poisson process with $\lambda = \frac{2}{3}$ per year.
(ii) Times between successive claims follow an exponential distribution with mean 1.5.
(iii) Each claim amount is 1, 2 or 3 with $p(1) = 0.25, p(2) = 0.25$, and $p(3) = 0.50$.
(iv) Claim occurrences and amounts are independent. Successive time claims are independent.
(v) The annual premium equals expected annual claims plus 1.8 times the standard deviation of annual claims.
(vi) $i = 0$.
You use 0.25, 0.40, 0.60, and 0.80 from the unit interval and the inversion

method to simulate time between claims.
You use 0.30, 0.60, 0.20, and 0.70 from the unit interval and the inversion method to simulate claim size.
Calculate the gain or loss from the insurer's viewpoint during the first 2 years from this simulation.

Solution.
Let N be the number of claims, X the size of a claim, and S the total annual gain/loss. We have

$$\begin{aligned}
E(X) =&1(0.25) + 2(0.25) + 3(0.50) = 2.25\\
E(X^2) =&1^2(0.25) + 2^2(0.25) + 3^2(0.50) = 5.75\\
\text{Var}(X) =&5.75 - 2.25^2 = 0.6875\\
E(S) =&E(N)E(X) = \frac{2}{3}(2.25) = 1.5\\
\text{Var}(S) =&\frac{2}{3}(0.6875) + \frac{2}{3}(2.25)^2 = 3.8333.
\end{aligned}$$

The annual premium is

$$\frac{3}{2} + 1.8\sqrt{3.8333} = 5.0242$$

and the 2-year premium is $2(5.0242) = 10.05$.
The time between successive claims has an exponential distribution with mean 1.5. The simulated inter-claim times are t_i where $u_i = 1 - e^{-1.5t_i}$. The table below lists the times between claims.

u	0.25	0.40	0.60	0.80
t	0.43	0.77	1.37	2.41

We see that 2 claims occur before time 2. First claim at time 0.43 and second claim at time 1.2.
The cdf of X is defined by

$$F_X(x) = \begin{cases} 0, & x \le 0 \\ 0.25, & 0 < x \le 1 \\ 0.50, & 1 < x \le 2 \\ 1, & x \ge 1 \end{cases}$$

Let $u_1 = 0.30$. Since $0.25 < 0.30 \le 0.5$, we have $x_1 = 2$. Let $u_2 = 0.60$. Since $0.5 < 0.6 \le 1$, we find $x_2 = 3$. Finally, the gain to the insurer is $10.05 - (2 + 3) = 4.95$ ■

Practice Problems

Problem 90.1 ‡
A dental benefit is designed so that a deductible of 100 is applied to annual dental charges. The reimbursement to the insured is 80% of the remaining dental charges subject to an annual maximum reimbursement of 1000.
You are given:
(i) The annual dental charges for each insured are exponentially distributed with mean 1000.
(ii) Use the following uniform $(0,1)$ random numbers and the inversion method to generate four values of annual dental charges:

$$0.30 \quad 0.92 \quad 0.70 \quad 0.08.$$

Calculate the average annual reimbursement for this simulation.

Problem 90.2 ‡
For a warranty product you are given:
(i) Paid losses follow the lognormal distribution with $\mu = 13.294$ and $\sigma = 0.494$.
(ii) The ratio of estimated unpaid losses to paid losses, y, is modeled by

$$y = 0.801x^{0.851}e^{-0.747x}$$

where

$$x = 2006 - \text{contract purchase year}.$$

The inversion method is used to simulate four paid losses with the following four uniform (0,1) random numbers:

$$0.2877 \quad 0.1210 \quad 0.8238 \quad 0.6179.$$

Using the simulated values, calculate the empirical estimate of the average unpaid losses for purchase year 2005.

Problem 90.3 ‡
You are given:
(i) The cumulative distribution for the annual number of losses for a policyholder is:

n	$F_N(n)$
0	0.125
1	0.312
2	0.500
3	0.656
4	0.773
5	0.855
$\vdots$	$\vdots$

(ii) The loss amounts follow the Weibull distribution with $\theta = 200$ and $\tau = 2$.
(iii) There is a deductible of 150 for each claim subject to an annual maximum out-of-pocket of 500 per policy.
The inversion method is used to simulate the number of losses and loss amounts for a policyholder:
(a) For the number of losses use the random number 0.7654.
(b) For loss amounts use the random numbers: 0.2738 0.5152 0.7537 0.6481 0.3153.
Use the random numbers in order and only as needed.
Based on the simulation, calculate the insurer's aggregate payments for this policyholder.

Problem 90.4 ‡
The price of a non dividend-paying stock is to be estimated using simulation. It is known that:
(i) The price S_t follows the lognormal distribution: $\ln\left(\frac{S_t}{S_0}\right) \sim N\left[\left(\alpha - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right]$.
(ii) $S_0 = 50, \alpha = 0.15$, and $\sigma = 0.30$.
Using the following uniform $(0, 1)$ random numbers and the inversion method, three prices for two years from the current date are simulated.

$$0.9830 \quad 0.0384 \quad 0.7794$$

Calculate the mean of the three simulated prices.

Problem 90.5 ‡
You are given:
(i) For a company, the workers compensation lost time claim amounts follow the Pareto distribution with $\alpha = 2.8$ and $\theta = 36$.
(ii) The cumulative distribution of the frequency of these claims is:

n	$F_N(n)$
0	0.5556
1	0.8025
2	0.9122
3	0.9610
4	0.9827
5	0.9923
$\vdots$	$\vdots$

(iii) Each claim is subject to a deductible of 5 and a maximum payment of 30.
Use the uniform (0, 1) random number 0.981 and the inversion method to generate the simulated number of claims.
Use as many of the following uniform (0, 1) random numbers as necessary, beginning with the first, and the inversion method to generate the claim amounts.

0.571 0.932 0.303 0.471 0.878

Calculate the total of the company's simulated claim payments.

Problem 90.6 ‡
N is the random variable for the number of accidents in a single year. N follows the distribution:

$$\Pr(N=n)=0.9(0.1)^{n-1},\ n=1,2,\cdots.$$

X_i is the random variable for the claim amount of the ith accident. X_i follows the distribution:

$$g(x_i)=0.01e^{-0.01x_i},\ x_i>0,\ i=1,2,\cdots.$$

Let U and $V_1,V_2,\cdots$ be independent random variables following the uniform distribution on $(0,1)$. You use the inverse transformation method with U to simulate N and V_i to simulate X_i with small values of random numbers corresponding to small values of N and X_i.
You are given the following random numbers for the first simulation:

u	v_1	v_2	v_3	v_4
0.05	0.30	0.22	0.52	0.46

Calculate the total amount of claims during the year for the first simulation.

Problem 90.7 ‡
Annual dental claims are modeled as a compound Poisson process where the number of claims has mean 2 and the loss amounts have a two-parameter Pareto distribution with $\theta = 500$ and $\alpha = 2$.
An insurance pays 80% of the first 750 of annual losses and 100% of annual losses in excess of 750.
You simulate the number of claims and loss amounts using the inverse transform method with small random numbers corresponding to small numbers of claims or small loss amounts.
The random number to simulate the number of claims is 0.8. The random numbers to simulate loss amounts are 0.60, 0.25, 0.70, 0.10 and 0.80.
Calculate the total simulated insurance claims for one year.

91 Estimating Risk Measures Using Simulation

In this section, we use simulation to estimate the risk measures VaR and TVaR. Let $y_1 \le y_2 \le \cdots \le y_n$ be a simulated sample of size n of a random variable. For a percentile p, let $k = [pn] + 1$, where $[x]$ is the largest integer less than or equal to x. The estimators of VaR and TVaR are

$$\widehat{\text{VaR}}_p(X) = y_k \quad \text{and} \quad \widehat{\text{TVaR}}_p(X) = \tfrac{1}{n-k+1}\sum_{i=k}^{n} y_i.$$

Note that $\widehat{\text{TVaR}}_p(X)$ is the mean of the sample $\{y_k, y_{k+1}, \cdots, y_n\}$. The variance of this sample is

$$s_p^2 = \frac{1}{n-k}\sum_{i=k}^{n}(y_i - \widehat{\text{TVaR}}_p(X))^2.$$

It has been shown that an asymptotically unbiased estimator of the variance of $\widehat{\text{TVaR}}_p(X)$ is given by

$$\widehat{\text{Var}}(\widehat{\text{TVaR}}_p(X)) = \frac{s_p^2 + p[\widehat{\text{TVaR}}_p(X) - \widehat{\text{VaR}}_p(X)]^2}{n-k+1}.$$

Example 91.1
Consider the following sample of simulated values

$$45 \quad 107 \quad 210 \quad 81 \quad 153 \quad 189.$$

Find $\widehat{\text{VaR}}_p(X)$, $\widehat{\text{TVaR}}_p(X)$, and $\widehat{\text{Var}}(\widehat{\text{TVaR}}_p(X))$ for $p = 0.6$.

Solution.
Rearrnaging the values in increasing order to obtain

$$45 \quad 81 \quad 107 \quad 153 \quad 189 \quad 210.$$

We have the following sequence of calculation

$$\begin{aligned}
k =&[pn] + 1 = [3.6] + 1 = 4 \\
\widehat{\text{VaR}}_p(X) =&153 \\
\widehat{\text{TVaR}}_p(X) =&\frac{1}{n-k+1}\sum_{i=k}^{n} y_i \\
=&\frac{1}{6-4+1}(153 + 189 + 210) = 184 \\
s_p^2 =&\frac{1}{n-k}\sum_{i=k}^{n}(y_i - \widehat{\text{TVaR}}_p(X))^2 \\
=&\frac{1}{6-4}[(153-184)^2 + (189-184)^2 + (210-184)^2] = 831 \\
\widehat{\text{Var}}(\widehat{\text{TVaR}}_p(X)) =&\frac{s_p^2 + p[\widehat{\text{TVaR}}_p(X) - \widehat{\text{VaR}}_p(X)]^2}{n-k+1} \\
=&\frac{831 + 0.6(184-153)^2}{6-4+1} = 469.20 \blacksquare
\end{aligned}$$

Practice Problems

Problem 91.1
Consider the following sample of simulated values

$$93 \quad 109 \quad 120 \quad 123 \quad 150 \quad 153 \quad 189 \quad 190 \quad 195 \quad 200.$$

Find $\widehat{\text{VaR}}_p(X)$, $\widehat{\text{TVaR}}_p(X)$, and $\widehat{\text{Var}}(\widehat{\text{TVaR}}_p(X))$ for $p = 0.3$.

92 The Bootstrap Method for Estimating Mean Square Error

Let θ be a quantity related to a distribution X. This quantity can be the mean of the distribution, the variance, quantile, etc. Let $\hat{\theta}$ denote an estimator of θ. The **mean square error** of $\hat{\theta}$ is defined by

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2].$$

The **bootstrap method** is a way of estimating $MSE(\hat{\theta})$. This is how the method work in terms of simulation: We select a random sample of size n from the distribution X and we create from this sample the empirical distribution. Using this distribution, we calculate the estimator $\hat{\theta}$. The next step is to "resample" from the original sample. Since the size of the original sample is n, there are n^n possible bootstrap samples(with repetition allowed). For each bootstrap sample we calculate $\hat{\theta}_i,\ i = 1, 2, \cdots, n^n$. The next step is to calculate the square deviation $(\hat{\theta}_i - \hat{\theta})^2$. The bootstrap is the average of these deviations, that is an estimate to $MSE(\hat{\theta})$ is

$$\widehat{MSE}(\hat{\theta}) = \frac{1}{n^n}\sum_{i=1}^{n^n}(\hat{\theta}_i - \hat{\theta})^2.$$

Example 92.1
A sample of size 2 contains the values $x_1 = 2$ and $x_2 = 4$. Calculate the MSE of the unbiased estimator of the population mean using the bootstrap method.

Solution.
The original sample mean is $\overline{x} = \frac{2+4}{2} = 3$. Since the original sample is of size 2, there are $2^2 =$ bootstrap samples. The table below provides the various samples along with their mean and square deviation.

Sample	$\overline{X}_i$	$(\overline{X}_i - \overline{X})^2$
2,2	2	$(2-3)^2 = 1$
2,4	3	$(3-3)^2 = 0$
4,2	3	$(3-3)^2 = 0$
4,4	4	$(4-3)^2 = 1$
Total		2

Hence, the bootstrap estimate is given by

$$\widehat{MSE}(\hat{\mu}) = \frac{1+0+0+1}{4} = 0.5 \blacksquare$$

Example 92.2 ‡
You are given a random sample of two values from a distribution function $F: x_1 = 1$ and $x_2 = 3$.
You estimate $\theta(F) = \text{Var}(X)$ using the estimator $g(X_1, X_2) = \frac{1}{2}\sum_{i=1}^{2}(X_i - \overline{X})^2$ where $\overline{X} = \frac{X_1+X_2}{2}$.
Determine the bootstrap approximation to the mean square error.

Solution.
The estimator for the original sample is $g_0 = g(1,3) = 1$. We have the following table.

Sample	X_1	X_2	$\overline{X}_i$	$g_i = g(X_1, X_2)$	$(g_i - g_0)^2$
1	1	3	2	1	0
2	1	1	1	0	1
3	3	1	2	1	0
4	3	3	3	0	1
Total					2

Hence, the bootstrap estimate is given by

$$\widehat{MSE}(g) = \frac{2}{4} = 0.5 \blacksquare$$

Example 92.3 ‡
A sample of claim amounts is $\{300, 600, 1500\}$. By applying the deductible to this sample, the loss elimination ratio for a deductible of 100 per claim is estimated to be 0.125. You are given the following simulations from the sample:

Simulation	Claim		
1	600	600	1500
2	1500	300	1500
3	1500	300	600
4	600	600	300
5	600	300	1500
6	600	600	1500
7	1500	1500	1500
8	1500	300	1500
9	300	600	300
10	600	600	600

Determine the bootstrap approximation to the mean square error of the estimate.

Solution.
We have

Simulation	X_1	X_2	X_3	LER	$(LER - 0.125)^2$
1	600	600	1500	0.111111	0.000193
2	1500	300	1500	0.090909	0.001162
3	1500	300	600	0.125000	0.000000
4	600	600	300	0.200000	0.005625
5	600	300	1500	0.125000	0.000000
6	600	600	1500	0.111111	0.000193
7	1500	1500	1500	0.066667	0.003403
8	1500	300	1500	0.090909	0.001162
9	300	600	300	0.250000	0.015625
10	600	600	600	0.166667	0.001736
Total					0.029099

The bootstrap estimate to the mean sqaure error is $\frac{0.029099}{10} = 0.0029099$ ■

Practice Problems

Problem 92.1 ‡
You are given a random sample of two values from a distribution function $F : x_1 = 1$ and $x_2 = 3$.
You estimate $\theta(F) = \mathrm{Var}(X)$ using the estimator $g(X_1, X_2) = \sum_{i=1}^{2}(X_i - \overline{X})^2$ where $\overline{X} = \frac{X_1+X_2}{2}$.
Determine the bootstrap approximation to the mean square error.

Problem 92.2 ‡
With the bootstrapping technique, the underlying distribution function is estimated by which of the following?
(A) The empirical distribution function
(B) A normal distribution function
(C) A parametric distribution function selected by the modeler
(D) Any of (A), (B) or (C)
(E) None of (A), (B) or (C).

Problem 92.3 ‡
Three observed values of the random variable X are:

$$1 \quad 1 \quad 4$$

You estimate the third central moment of X using the estimator:

$$g(X_1, X_2, X_3) = \frac{1}{3}\sum_{i=1}^{3}(X_i - \overline{X})^3.$$

Determine the bootstrap estimate of the mean-squared error of g.

Problem 92.4 ‡
For a policy that covers both fire and wind losses, you are given:
(i) A sample of fire losses was 3 and 4.
(ii) Wind losses for the same period were 0 and 3.
(iii) Fire and wind losses are independent, but do not have identical distributions.
Based on the sample, you estimate that adding a policy deductible of 2 per wind claim will eliminate 20% of the insured loss.
Determine the bootstrap approximation to the mean square error of the estimate.

Problem 92.5 ‡
The random variable X has the exponential distribution with mean θ. Calculate the mean-squared error of X^2 as an estimator of θ^2.

Answer Key

Section 1

1.1 Deterministic

1.2 Stochastic

1.3 Stochastic

1.4 Stochastic

1.5 Mostly stochastic

Section 2

2.1 (a) $A^c = B,\ B^c = A$ and $C^c = \{1, 3, 4, 5, 6\}$
(b)

$$\begin{aligned} A \cup B &= \{1, 2, 3, 4, 5, 6\} \\ A \cup C &= \{2, 4, 6\} \\ B \cup C &= \{1, 2, 3, 5\} \end{aligned}$$

(c)

$$\begin{aligned} A \cap B &= \emptyset \\ A \cap C &= \{2\} \\ B \cap C &= \emptyset \end{aligned}$$

(d) A and B are mutually exclusive as well as B and C

2.2 Note that $\Pr(E) > 0$ for any event E. Moreover, if S is the sample space then

$$\Pr(S) = \sum_{i=1}^{\infty} \Pr(O_i) = \frac{1}{2}\sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i = \frac{1}{2} \cdot \frac{1}{1-\frac{1}{2}} = 1$$

Now, if $E_1, E_2, \cdots$ is a sequence of mutually exclusive events then

$$\Pr(\cup_{n=1}^{\infty} E_i) = \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \Pr(O_{nj}) = \sum_{n=1}^{\infty} \Pr(E_n)$$

where $E_n = \{O_{n1}, O_{n2}, \cdots\}$. Thus, Pr defines a probability function.

2.3 0.5

2.4 0.56

2.5 0.66

2.6 0.52

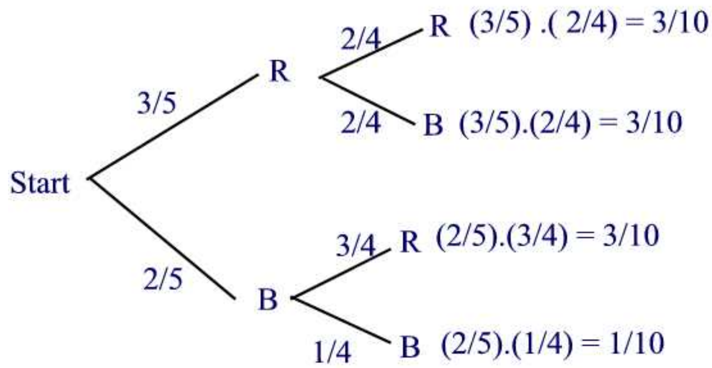
2.7 0.05

2.8 0.6

2.9 0.48

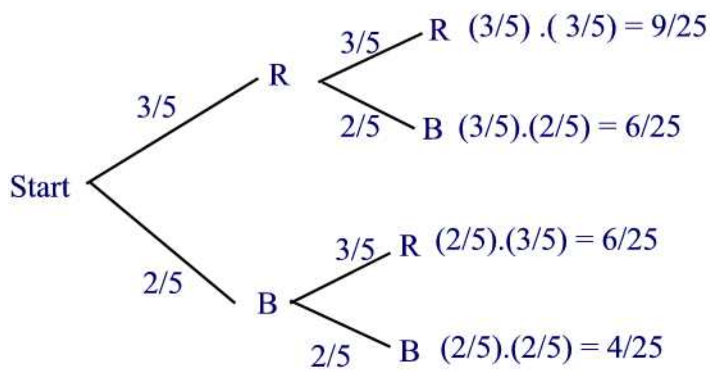
2.10 0.04

2.11 The probability is given is the figure below.



The probability that the first ball is red and the second ball is blue is $\mathrm{PR}(RB) = 0.3$.

2.12



The probability that the first ball is red and the second ball is blue is $\mathrm{PR}(RB) = 6/25$.

2.13 0.173

2.14 0.467

2.15 0.1584

2.16 0.0141

2.17 0.29

2.18 0.42

2.19 0.22

2.20 0.657

Section 3

3.1 (a) Discrete (b) Discrete (c) Continuous (d) Mixed.

3.2

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

3.3

$$p(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \\ 0, & x \neq 0, 1. \end{cases}$$

3.4 1/9

3.5 0.469

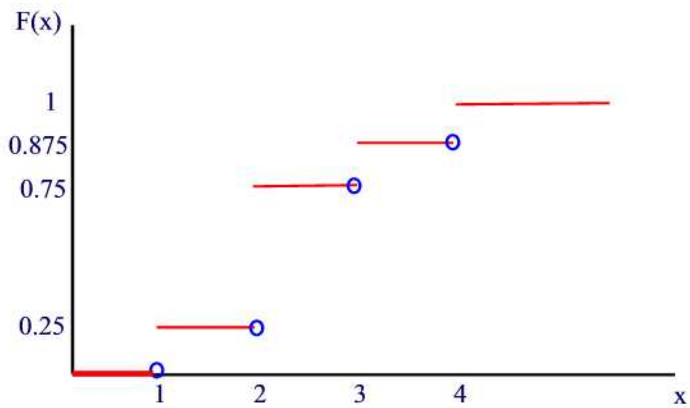
3.6 0.132

3.7 0.3

3.8 $\alpha = 1$.

3.9

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.25 & 1 \le x < 2 \\ 0.75 & 2 \le x < 3 \\ 0.875 & 3 \le x < 4 \\ 1 & 4 \le x \end{cases}$$



3.10 (a)

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{(1+x)^{a-1}}, & x \geq 0. \end{cases}$$

(b)

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-kx^{\alpha}}, & x \geq 0. \end{cases}$$

3.11

$$\begin{aligned} F(n) =& P(X \leq n) = \sum_{k=0}^{n} P(X = k) \\ =& \sum_{k=0}^{n} \frac{1}{3} \left(\frac{2}{3}\right)^{k} \\ =& \frac{1}{3} \frac{1 - \left(\frac{2}{3}\right)^{n+1}}{1 - \frac{2}{3}} \\ =& 1 - \left(\frac{2}{3}\right)^{n+1} \end{aligned}$$

3.12 (a) 0.135 (b) 0.233 (c)

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{5}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

3.13 $f(x) = F'(x) = \frac{e^x}{(1+e^x)^2}$

3.14 (a) We have that $S(0) = 1, S'(x) = -\frac{1}{20}(100 - x)^{-\frac{1}{2}} \leq 0$, $s(x)$ is right continuous, and $S(100) = 0$. Thus, S satisfies the properties of a survival function.
(b) $F(x) = 1 - S(x) = 1 - \frac{1}{10}(100 - x)^{\frac{1}{2}}$.
(c) 0.092

3.15 0.149

3.16 $F(x) = 1 - S(x) = \frac{x^2}{100},\ x \geq 0$

3.17 (a) 0.3 (b) 0.3

3.18 $h(x) = -\frac{S'(x)}{S(x)} = \frac{1}{2}(100 - x)^{-1}$

3.19 $S(x) = e^{-\mu x}, F(x) = 1 - e^{-\mu x}$, and $f(x) = F'(x) = \mu e^{-\mu x}$.

3.20 1/480

Section 4

4.1 (b) $np(1 - p)$

4.2 (b) λ

4.3 (c) $(1 - p)p^{-2}$

4.5 (c) $\frac{1}{\lambda^2}$

4.6 (a)

$$\begin{aligned}E(X) =&\frac{1}{\theta\Gamma(\alpha)}\int_0^\infty xe^{-\frac{x}{\theta}}x^{\alpha-1}dx\\ =&\frac{\theta}{\Gamma(\alpha)}\int_0^\infty \frac{1}{\theta}e^{-\frac{x}{\theta}}x^{\alpha}dx\\ =&\frac{\theta\Gamma(\alpha+1)}{\Gamma(\alpha)}\\ =&\alpha\theta\end{aligned}$$

(b)

$$\begin{aligned}E(X^2) =&\frac{1}{\Gamma(\alpha)}\int_0^\infty x^2e^{-\frac{x}{\theta}}\frac{\theta^\alpha\, {}^{\alpha-1}}{x}dx\\ =&\frac{1}{\Gamma(\alpha)}\int_0^\infty x^{\alpha+1}\frac{1}{\theta^\alpha}e^{-\frac{x}{\theta}}dx\\ =&\frac{\theta^2\Gamma(\alpha+2)}{\Gamma(\alpha)}\int_0^\infty \frac{x^{\alpha+1}}{\theta^{\alpha+2}\Gamma(\alpha+2)}e^{-\frac{x}{\theta}}dx\\ =&\frac{\theta^2\Gamma(\alpha+2)}{\Gamma(\alpha)}\end{aligned}$$

where the last integral is the integral of the pdf of a Gamma random variable with parameters $(\alpha+2,\theta)$. Thus,

$$E(X^2) = \frac{\theta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{\theta^2(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha)} = \theta^2\alpha(\alpha+1).$$

Finally,

$$Var(X) = E(X^2) - (E(X))^2 = \theta^2\alpha^2(\alpha+1) - \alpha^2\theta^2 = \alpha\theta^2$$

4.7 4

4.8 1,417,708,752

4.9 730,182,499.20

4.10 0

4.11 9

4.12 0.3284

4.13 We have

$$\begin{aligned}\mu_n' =& \int_0^\infty x^n f(x)dx\\ =&A\int_0^\infty x^{B+n}e^{-Cx}dx\\ =&A\left[-\frac{x^{B+n}e^{-Cx}}{C}\Big|_0^\infty + \frac{B+n}{C}\int_0^\infty x^{B+n-1}e^{-Cx}dx\right]\\ =&\frac{B+n}{C}\int_0^\infty Ax^{B+n-1}e^{-Cx}dx\\ =&\frac{B+n}{C}E(X^{n-1}).\end{aligned}$$

4.14 $\mu = \frac{B+1}{C}$ and $\mu_2' = \frac{(B+1)(B+2)}{C^2}$

4.15 $\frac{2}{\sqrt{B+1}}$

4.16 $\frac{3(B+3)}{B+1}$

4.17 0.5

4.18 (a) $F(x) = \int_{-\infty}^x 0.005tdt = \int_0^x 0.005tdt = \begin{cases} 0, & x < 0\\ 0.0025x^2, & 0 \le x \le 20\\ 1, & x > 20\end{cases}$
(b) The mean is $\frac{40}{3}$ and the variance is $\frac{200}{9}$
(c) 0.354

4.19 16

4.20 (a) $E(X^k) = \int_1^\infty \frac{ax^k}{x^{a+1}}dx = \frac{a}{k-1}x^{k-a}\Big|_1^\infty = \frac{a}{a-k},\ 0 < k < a.$
(b) $\frac{1}{\sqrt{a(a-2)}}$

4.21 -1.596

4.22 The mean is $\frac{\theta}{\alpha-1}$ and the variance is $\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$

4.23 $\sqrt{\frac{\alpha}{\alpha-2}}$

4.24 2

4.25 1.7

Section 5

5.1

Amount of loss	750	500	1200
Insurance payment	250	0	700

5.2

$$\begin{aligned}
\Pr(X \le 45) =& \frac{4}{12} = \frac{1}{3} \\
\Pr(X \le 67) =& \frac{5}{12} \\
\Pr(X \le 84) =& \frac{7}{12} \\
\Pr(X \le 93) =& \frac{8}{12} \\
\Pr(X \le 100) =& \frac{11}{12} \\
\Pr(X \le 102) =& 1.
\end{aligned}$$

5.3 $\mu_1' = 75.8333$ and $\mu_2' = 6312.8333$

5.4 0.509175

5.5 $\frac{1}{2}(100 - d)$ for $0 < d < 100$ and 0 otherwise.

5.6 108

5.7 $\frac{1}{\lambda}$

5.8 308,8728

5.9 $\frac{e^{-\lambda d}}{\lambda}$

5.10 $\frac{1}{\lambda^2}e^{-\lambda d}(2-e^{-\lambda d})$

5.11 $\frac{1}{160}$

5.12 94.84

5.13 88.4

5.14 $\theta - \frac{\theta}{\phi+1} - x - \frac{x^{\phi+1}}{\theta^\phi(\phi+1)}$

5.15

$$\begin{aligned}S(x) =&\Pr(X>x) = \int_x^\infty (1+2t^2)e^{-2t}dt\\ =&-(1+t+t^2)e^{-2t}\Big|_x^\infty = (1+x+x^2)e^{-2x}\end{aligned}$$

for $x \geq 0$ and 0 otherwise.
(b) $\frac{1+x+\frac{1}{2}x^2)}{1+x+x^2}$

5.16 We have

$$\begin{aligned}S_{Y^P}(y) =&1-F_{Y^P}(y) = 1 - \frac{1-S_X(y+d)-[1-S_X(d)]}{S_x(d)}\\ =&1-\frac{S_X(y+d)+S_X(d)}{S_X(d)} = \frac{S_X(y+d)}{S_X(d)}.\end{aligned}$$

5.17 (a)

$$\begin{aligned}f_{Y^P}(y) =&\frac{f(y+100)}{1-F(100)} = \frac{(0.001+0.00002(y+100))e^{-0.005(y+100)}}{1.4e^{-0.5}}\\ =&\frac{(0.0003+0.00002y)e^{-0.005y}}{1.4}\\ =&\left(\frac{3}{1400}+\frac{1}{70000}y\right)e^{-0.005y},\ y>0\end{aligned}$$

and 0 otherwise.
(b) $E(Y^P) = \frac{2200}{7}$ and $\mathrm{Var}(Y^P) = \frac{3560000}{49}$

5.18 $E[(X-10)_+] =$ and $\mathrm{Var}[(X-10)_+] = \frac{425}{36}$

5.19 $d = 6$.

5.20 175

5.21 1875

5.22 3.43

5.23 6.259

Section 6

6.1 we can either say 1120 is the twentieth percentile or 1120 is the one-fifth quantile

6.2 3

6.3 2

6.4 the median is 1 and the 70^{th} percentile is 2

6.5 The median is $M = 0.3466$. This means that half the people get in line less than 0.3466 minutes (about 21 seconds) after the previous person, while half arrive more than 0.3466 minutes later

6.6 0.693

6.7 998.72

6.8 $3\ln 2$

6.9 The median is 0.8409

6.10 3659

6.11 $a + 2\sqrt{\ln 2}$

6.12 $-\ln(1-p)$

6.13 $-\ln[2(1-p)]$

6.14 2

6.15 72.97

6.16 0.4472

6.17 6299.61

6.18 2.3811

6.19 50

6.20 2.71

Section 7

7.1 0.2119

7.2 0.9876

7.3 0.0094

7.4 0.692

7.5 0.1367

7.6 0.0088

7.7 0

7.8 23

7.9 0.0162

7.10 6,342,637.5

7.11 0.8185

7.12 16

7.13 0.1587

7.14 0.9887

7.15 0.0244

7.16 0.9985

7.17 0.1056

7.18 0.8413

7.19 0.8201

7.20 0.224

Section 8

8.1 $E(X) = \frac{2}{\lambda^2}$ and $Var(X) = \frac{1}{\lambda^2}$

8.2 A normal random variable with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$

8.3 0.70

8.4 41.9

8.5 e^{13t^2+4t}

8.6 4

8.7 -28

8.8 2

8.9 5000

8.10 10560

8.11 $(tp + 1 - p)^n$

8.12 $\frac{tp}{1-t(1-p)}$ provided that $|t| < (1-p)^{-1}$

8.13 True

8.14 $t^a P_X(t^b)$

8.15 $E(X) = \frac{1}{p}$ and $\text{Var}(X) = \frac{1-p}{p^2}$

8.16 $0.4t^2 + 0.2t^3 + 0.2t^5 + 0.2t^8$

8.17 $\frac{1}{3}t^{-2} + \frac{1}{6}t^3 + \frac{1}{8}t^{\pi} + \frac{3}{8}t^{\frac{7}{2}}$

8.18 $\frac{t^3}{2-t}$

8.19

x	-1	0	1	2	3
$p(x)$	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$

8.20 $E(X) = 3.5$ and $\text{Var}(X) = 6.25$

Section 9

9.1 Let $m > 0$. Then there is $M > 0$ such that $e^{bx} \geq x^{m+1}$ which is equivalent to saying that $x^m e^{-bx} \leq \frac{1}{x}$ for $x \geq M$. By the comparison test of improper integrals we find that

$$\int_0^{\infty} x^m e^{-bx} < \infty.$$

Since $E(X^k)$ is an integral of the above form, we conclude that $E(X^k) < \infty$ for all $k > 0$. That is, the distribution of X is light-tailed.

9.2 Since X is heavy-tailed, we have $E(X^k) = \int_0^{\infty} x^k f_X(x)dx = \infty$ for some $k > 0$. Now, let $t > 0$. Let N be large enough so that $e^{tx} \geq x^k$ for all $x \geq N$. Hence,

$$\int_0^N x^k f_X(x)dx + \int_N^{\infty} e^{tx} f_X(x)dx \geq \int_0^N x^k f_X(x)dx + \int_N^{\infty} x^k f_X(x)dx = \int_0^{\infty} x^k f_X(x)dx = \infty.$$

Since $\int_0^N x^k f_X(x)dx < \infty$, we conclude that $\int_N^{\infty} e^{tx} f_X(x)dx = \infty$

9.3 We have

$$M_X(t) = \int_0^\infty e^{tx} f_X(x)dx \geq \int_N^\infty e^{tx} f_X(x)dx = \infty$$

9.4 From Table C, we have

$$E(X^k) = \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)}$$

for all $k > 0$. Hence, the Gamma distribution is light-tailed.

9.5 From Table C/ Exam 4, we have

$$E(X^k) = \theta^k \Gamma\left(1 - \frac{k}{\tau}\right)$$

provided that $k < \tau$. Since $E(X^k)$ is only valid for $k < \tau$, the distribution is heavy-tailed.

9.6 The Pareto distribution has a more heavy-tailed than the Gamma distribution,

9.7 The Weibull distribution has a lighter tail than the inverse Weibull distribution.

9.8 X is more heavy-tailed than Y

9.9 X and Y have similar or proportional tails.

9.10 X is more light-tailed than Y

9.11 X has a heavier tail than Y

9.12

Distribution	Heavy-Tail	Light-Tail
Weibull		✓
Inverse Pareto	✓	
Normal		✓
Loglogistic	✓	

9.13

Distribution	Heavy-Tail	Light-Tail
Paralogistic	✓	
Lognormal		✓
Inverse Gamma	✓	
Inverse Gaussian		✓

9.14

Distribution	Heavy-Tail	Light-Tail
Inverse Paralogistic ✓		
Inverse Exponential ✓		

9.15 $\lim_{x\to\infty} \frac{S_X(x)}{S_Y(x)} = \infty$

9.16 $\lim_{x\to\infty} \frac{S_X(x)}{S_Y(x)} = \infty$

9.17 $\lim_{x\to\infty} \frac{S_X(x)}{S_Y(x)} = \infty$

9.18 $c > 0$

9.19 X has a heavier tail than Y

9.20 The tail of X is heavier than that of Y which in turn is heavier than the tail of Z

Section 10

10.1 $\frac{d}{dx}\left[\frac{f(x+y)}{f(x)}\right] = \frac{y(1-\alpha)}{x^2}\left(1+\frac{y}{x}\right)^{\alpha-2} e^{-\frac{y}{\theta}} < 0$ for $\alpha > 1$.

10.2 For $\alpha = 1$, the Gamma distribution is just the exponential function which has a constant hazard rate

10.3 We have

$$h(x) = \frac{f(x)}{S(x)} = \frac{\tau x^{\tau-1}}{\theta^\tau}.$$

Hence,

$$h'(x) = \frac{\tau(\tau-1)x^{\tau-2}}{\theta^\tau}.$$

Thus, $h(x)$ is increasing (light-tailed distribution) for $\tau > 1$ and decreasing (heavy-tailed distribution) for $0 < \tau < 1$

10.4 X is light-tailed

10.5 We have

$$H(x) = \frac{\alpha\theta^\alpha}{(x+y+\theta)^{\alpha+1}} \cdot \frac{(x+\theta)^{\alpha+1}}{\alpha\theta^\alpha} = \left(\frac{x+\theta}{x+y+\theta}\right)^{\alpha+1}.$$

Hence,

$$H'(x) = (\alpha+1)\left(\frac{x+\theta}{x+y+\theta}\right)^\alpha \frac{y}{(x+y+\theta)^2} > 0.$$

Thus, $H(x)$ is increasing and by Theorem 10.1, $h(x)$ is decreasing which shows that the Pareto distribution is heavy-tailed

10.6 We have

$$\begin{aligned}\lim_{x\to\infty} h(x) &= \lim_{x\to\infty} \frac{f(x)}{S(x)} = \lim_{x\to\infty} \frac{f'(x)}{-f(x)} \\ &= -\lim_{x\to\infty} \frac{d}{dx}\left[\ln\left(f(x)\right)\right] = -\lim_{x\to\infty} \frac{d}{dx}\left[(\alpha-1)\ln x - \frac{x}{\theta}\right] \\ &= \lim_{x\to\infty}\left(\frac{1}{\theta} - \frac{\alpha-1}{x}\right) = \frac{1}{\theta}.\end{aligned}$$

10.7 We have

$$\lim_{x\to\infty} e(x) = \lim_{x\to\infty} \frac{\int_x^\infty S_X(t)dt}{S(x)} = \lim_{x\to\infty} \frac{-S_X(x)}{-f_X(x)} = \lim_{x\to\infty} \frac{1}{h(x)}.$$

10.8 θ

10.9 Since $0 < \alpha < 1$, the hazard rate function is decreasing and hence $e(x)$ is increasing. The result follows from the fact that $e(0) = \alpha\theta$ and $e(\infty) = \theta$.

10.10 Since $\alpha > 1$, the hazard rate function is inecreasing and hence $e(x)$ is decreasing. The result follows from the fact that $e(0) = \alpha\theta$ and $e(\infty) = \theta$

10.11 For this distribution we have $f_X(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}$ and $S_X(x) = \frac{\theta^\alpha}{(x+\theta)^\alpha}$. Hence, $h(x) = \frac{f_x(x)}{S_X(x)} = \frac{\alpha}{x+\theta}$. Thus,

$$\lim_{x\to\infty} e(x) = \frac{1}{\lim_{x\to\infty} h(x)} = \infty.$$

This shows that $e(x)$ is increasing and hence the distribution is heavy-tailed

10.12 $\lim_{x\to\infty} e(x) = \infty$

10.13 (a)$S(x) = \frac{1}{(x+1)^2},\ f(x) = \frac{2}{(x+1)^3}$, and $h(x) = \frac{2}{x+1}$.
(b) $E(X) = 1$ and $E(X^2) = \infty$ so that X is heavy-tailed.

10.14 Since the hazard rate is nonicreasing, X is a heavy-tailed

10.15 Since $e(x)$ is nondecreasing, X is a heavy-tailed

10.16 $\frac{S(x+y)}{S(x)}$ is nonincreasing so that $e(x)$ is nonincreasing and therefore X is light-tailed

10.17 $\frac{f(x+y)}{f(x)}$ is nondecreasing so that $h(x)$ is nonincreasing. Thus, X is heavy-tailed

Section 11

11.1 (a) $S(x) = \int_x^\infty 2te^{-t^2}dt = -e^{-t^2}\Big|_0^\infty = e^{-x^2}$.
(b) $f_e(x) = \frac{S(x)}{E(X)} = \frac{2}{\sqrt{\pi}}e^{-x^2}$ for $x > 0$ and 0 otherwise.

11.2 $h_e(x) = \frac{1}{e(x)} = \frac{2x+3}{2x+5}$

11.3 $S(x) = \frac{e(0)}{e(x)}e^{-\int_0^x \frac{1}{e(t)}dt} = (1+x)e^{-x-\frac{x^2}{2}},\ x > 0$

11.4 $S_e(x) = \frac{e(x)}{e(0)}S(x) = e^{-x-\frac{x^2}{2}},\ x > 0$

11.5 0.6559

11.6 (a) $E(X) = 3$ and $E(X^2) = \frac{24}{5}$ (b) $\frac{4}{5}$

11.7 $S(x) = \left(\frac{10}{10+9x}\right)^{\frac{10}{9}}$

11.8 λ

Section 12

12.1 By (P3), we have $\rho(0) = \rho(\alpha 0) = \alpha\rho(0)$. Choosing $\alpha \neq 1$, we conclude that $\rho(0) = 0$. Since there are no losses, no capital is required to support the risk

12.2 This follows from (P3) and (P4)

12.3 Check the properties (P1)-(P4)

12.4 (b)

12.5 (a)

12.6 (a) and (c)

12.7 All three are correct

12.8 Simple algebra

12.9 0

12.10 We have

$$\begin{aligned}
\rho(L+\alpha) =& E(L+\alpha) + \beta\sqrt{\text{Var}(L+\alpha)}\\
=& E(L) + \alpha + \beta\sqrt{\text{Var}(L)}\\
=& \rho(L) + \alpha\\
\rho(\alpha L) =& E(\alpha L) + \beta\sqrt{\text{Var}(\alpha L)}\\
=& \alpha E(L) + \alpha\beta\text{Var}(L)\\
=& \alpha\rho(L)
\end{aligned}$$

where $\alpha > 0$.

Section 13

13.1 $a(1-p)+pb$

13.2 1.8974

13.3 2.0227

13.4 60

13.5 $\theta(1-p)^{-\frac{1}{\alpha}}$

13.6 40

13.7 10,000

13.8 347.21

13.9 5

13.10 $\text{VaR}_{0.96} = 400$. This says that there is 4% chance the losses will exceed 400

Section 14

14.1 (a) $e(x) = \frac{1}{2}(b-x)$ (b) $\text{TVaR}_p(L) = \frac{1}{2}[(a+b)+p(b-a)]$

14.2 (a) $\pi_{0.90} =\approx 1.8974$ and $e(1.8974) = 0.0051$ (b) $\text{TVaR}_{0.90}(L) = 1.9025$

14.3 $\pi_{0.85} = 7$ and $\text{TVaR}_{0.85}(L) = 10.25$

14.4 (a) $\theta = 1000$ (b) $e(100) = 220$ (c) $\pi_{0.95} = 647.55$ (d) $\text{TRaV}_{0.95}(L) = 867.55$

14.5 $\text{TVaR}_{0.75}(L) = 756$

14.6 1.3

14.7 1.651

14.8 2.02

14.9 0.82

14.10 100,000

14.11 120.62

Section 15

15.1 We have

$$F_{cX}(x) = \Pr(cX \le x) = \Pr(X \le \frac{x}{c}) = F_X\left(\frac{x}{c}\right) = 1 - e^{-\lambda\frac{x}{c\theta}} = F_X(\frac{x}{c}).$$

This is an exponential distribution with parameter $c\theta$

15.2 $F_Y(y) = 1 - e^{-\left(\frac{y}{c}\right)^2}$

15.3 $F_Y(y) = \frac{y-c\theta}{c\theta}$

15.4 Let $Y = cX,\ c > 0$. We have

$$\begin{aligned} F_Y(y) =& \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right) \\ =& e^{-\left(\frac{y}{c\theta}\right)^{-\alpha}}. \end{aligned}$$

This is a Fréchet distribution with parameters $c\theta$ and α.

15.5 Let $Y = cX,\ c > 0$. We have

$$\begin{aligned} F_Y(y) =& \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right) \\ =& 1 - \frac{1}{\left[1 + \left(\frac{y}{c\theta}\right)^{\gamma}\right]^{\alpha}} \end{aligned}$$

This is a Burr distribution with parameters $\alpha, c\theta$, and γ

15.6 (a) θ (b) θ (c) θ (d) θ

15.7 (1) is the only correst answer

15.8 Let $Y = cX$. Then

$$F_Y(y) = \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right)$$
$$= \Phi\left(\frac{\ln\left(\frac{y}{c}\right) - \mu}{\sigma}\right)$$
$$= \Phi\left(\frac{\ln y - \ln c\mu}{\ln c\sigma}\right).$$

This is a lognormal distribution with parameters $\ln c\mu$ and $\ln c\sigma$ and consequently has no scale paramter

15.9 The Gamma distribution with parameters α and θ has a cdf $F_X(x) = \frac{1}{\Gamma(\alpha)}\int_0^{\frac{x}{\theta}} t^{\alpha-1}e^{-t}dt$. Let $Y = cX$. Then

$$F_Y(y) = \Pr(Y \le y) = \Pr\left(X \le \frac{y}{c}\right)$$
$$= \frac{1}{\Gamma(\alpha)}\int_0^{\frac{y}{c\theta}} t^{\alpha-1}e^{-t}dt.$$

This is a Gamma distribution with parameters α and $c\theta$

15.10 100

15.11 Letting $\alpha = 1$, we obtain $f_X(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}$ which is the pdf of an exponential distribution

15.12 0.0295

Section 16

16.1 0.0949

16.2 100

16.3 $E(Y_1) = 2$ and $E(Y_2) = 6$

16.4 $E(X) = 1300$ and $\mathrm{Var}(X) = 6935000$

16.5 0.0568

16.6 35

16.7 $F_X(x) = 1 - a_1\left(\frac{\theta_1}{x+\theta_1}\right)^{\alpha_1} - a_2\left(\frac{\theta_2}{x+\theta_2}\right)^{\alpha_2} - \cdots - a_N\left(\frac{\theta_N}{x+\theta_N}\right)^{\alpha_N}$ where $a_j > 0$ and $\sum_{i=1}^N a_i = 1$.
$f_X(x) = a_1\left[\frac{\alpha_1\theta_1^{\alpha_1}}{(x+\theta_1)^{\alpha_1+1}}\right] + a_2\left[\frac{\alpha_2\theta_2^{\alpha_2}}{(x+\theta_2)^{\alpha_2+1}}\right] + \cdots + a_N\left[\frac{\alpha_N\theta_N^{\alpha_N}}{(x+\theta_N)^{\alpha_N+1}}\right]$
$h_X(x) = \frac{a_1\left[\frac{\alpha_1\theta_1^{\alpha_1}}{(x+\theta_1)^{\alpha_1+1}}\right]+a_2\left[\frac{\alpha_2\theta_2^{\alpha_2}}{(x+\theta_2)^{\alpha_2+1}}\right]+\cdots+a_N\left[\frac{\alpha_N\theta_N^{\alpha_N}}{(x+\theta_N)^{\alpha_N+1}}\right]}{a_1\left(\frac{\theta_1}{x+\theta_1}\right)^{\alpha_1}+a_2\left(\frac{\theta_2}{x+\theta_2}\right)^{\alpha_2}+\cdots+a_N\left(\frac{\theta_N}{x+\theta_N}\right)^{\alpha_N}}$

16.8 15

16.9 400

16.10 0.146

16.11 0.7566

Section 17

17.1 $E(X) = \frac{2478}{13},\ \text{Var}(X) = \frac{2283152}{169}$ and the mode is 10

17.2 $E(X \wedge 105) = \frac{1351}{13}$ and $e_X(x) = \frac{1127}{9}$

17.3 0.1659

17.4

x	94	104	134	180	210	350	524
$p_X(x)$	$\frac{1}{13}$	$\frac{3}{13}$	$\frac{2}{13}$	$\frac{4}{13}$	$\frac{1}{13}$	$\frac{1}{13}$	$\frac{1}{13}$

The cdf is defined by

$$F_X(x) = \frac{1}{13}\text{number of elements in the sample that are} \le x$$

17.5 0.61

17.6 17,566,092.92

17.7

$$f_X(x) = \begin{cases} \frac{1}{8} \cdot \frac{1}{13} = \frac{1}{104}, & 90 \le x \le 98 \\ \frac{1}{8} \cdot \frac{3}{13} = \frac{3}{104}, & 100 \le x \le 108 \\ \frac{1}{8} \cdot \frac{2}{13} = \frac{2}{104}, & 130 \le x \le 138 \\ \frac{1}{8} \cdot \frac{4}{13} = \frac{4}{104}, & 176 \le x \le 184 \\ \frac{1}{8} \cdot \frac{1}{13} = \frac{1}{104}, & 206 \le x \le 214 \\ \frac{1}{8} \cdot \frac{1}{13} = \frac{1}{104}, & 346 \le x \le 354 \\ \frac{1}{8} \cdot \frac{1}{13} = \frac{1}{104}, & 520 \le x \le 528 \end{cases}$$

and 0 otherwise.

Section 18

18.1 0.75

18.2 We have that $Y = 0.5X$. Thus,

$$f_Y(y) = 2f_X(2y) = 3y^2$$

for $0 < y < 1$ and 0 otherwise.
Thus,

$$F_Y(y) = y^3, \ 0 \le y \le 1$$

and $F_Y(y) = 1$ for $y > 1$. Also,

$$S_Y(y) = 1 - y^3, 0 \le y < 1$$

and 0 for $y > 1$

18.3 $E(Y) = 0.75E(X) = \frac{0.75(2000)}{3-1} = 750$ and $\sigma_Y = 1299.04$

18.4 We have

$$F_Y(y) = F_X\left(\frac{y}{\theta}\right) = 1 - \left(1 + \frac{y}{\theta}\right)^{-\alpha} = 1 - \left(\frac{\theta}{y+\theta}\right)^{\alpha}.$$

This is the cdf of a Pareto distribution with parameters α and θ. The pdf is

$$f_Y(y) = \frac{\alpha\theta^\alpha}{(y+\theta)^{\alpha+1}}.$$

18.5 We have

$$F_Y(y) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right).$$

Thus,

$$F_Z(z) = F_Y\left(\frac{z}{\theta}\right) = \Phi\left(\frac{\ln\left(\frac{z}{\theta}\right) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln z - (\mu + \ln\theta)}{\sigma}\right).$$

Hence, Z is a lognormal distribution with parameters $\mu + \ln\theta$ and σ.

18.6 The cdf of X is $F_X(x) = \int_1^x 3t^{-4}dt = 1 - x^{-3}$ and $\Pr(Y > 2.2) = 1 - (1 - 2^{-3}) = \frac{1}{8} = 0.125$

18.7 We have

$$\begin{aligned} F_W(w) =& F_Z\left(\frac{w}{1+r}\right) \\ =& \frac{1}{2}\frac{w^2}{w^2 + 1000(1+r)^2} + \frac{1}{2}\frac{w}{w + 1000(1+r)}. \end{aligned}$$

Thus, W is an equal mixture of a loglogistic distribution with parameters $\gamma = 2$ and $\theta = 10\sqrt{10}(1+r)$ and a Pareto distribution with parameters $\alpha = 1$ and $\theta = 1000(1+r)$

Section 19

19.1 In the transformed case, we have

$$F_Y(y) = 1 - e^{-\left(\frac{y^\tau}{\theta}\right)} \text{ and } f_Y(y) = \frac{\tau}{\theta}y^{\tau-1}e^{-\left(\frac{y^\tau}{\theta}\right)}.$$

In the inverse transformed case, we have

$$F_Y(y) = e^{-\left(\frac{y^{-\tau}}{\theta}\right)} \text{ and } f_Y(y) = \frac{\tau}{\theta}y^{-\tau-1}e^{-\left(\frac{y^{-\tau}}{\theta}\right)}.$$

In the inverse case, we have

$$F_Y(y) = e^{-\left(\frac{1}{y\theta}\right)} \text{ and } f_Y(y) = \frac{1}{y^2\theta}e^{-\left(\frac{1}{y\theta}\right)}$$

19.2 We have

$$F_Y(y) = 1 - F_X(y^{-1}) = \left(\frac{\theta}{\frac{1}{y} + \theta}\right)^\alpha = \left(\frac{y}{y + \frac{1}{\theta}}\right)^\alpha.$$

Y has the inverse Pareto distribution with parameters α and $\frac{1}{\theta}$

19.3 $f_Y(y) = y^{-2}f_X(y^{-1}) = \frac{1}{y^2} \cdot \frac{2}{y} = \frac{2}{y^3}$ for $y > 1$ and 0 otherwise

19.4 $f_Y(y) = \frac{1}{y^2}f_X(y^{-1}) = \frac{1}{y^2}\frac{y^{1-\alpha}e^{-\frac{1}{y}}}{\Gamma(\alpha)}$ and 0 otherwise

19.5 $f_Y(y) = \tau y^{\tau-1}f_X(y^\tau) = \frac{\tau y^{\tau-1}}{b}$ for $0 \le y \le b^{\frac{1}{\tau}}$ and 0 otherwise

19.6 Y has an exponential distribution with parameter α

19.7 We have

$$F_Y(y) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right)$$

and

$$f_Y(y) = \frac{1}{y}f_Z\left(\frac{\ln y - \mu}{\sigma}\right) = \frac{1}{y\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2}$$

where Z is the standard normal distribution

19.8 $f_Y(y) = \frac{1}{y}f_X(\ln y) = \frac{1}{2by}$ for $1 < t < e^b$ and 0 otherwise

19.9 We have

$$f_Y(y) = \frac{1}{y}f_X(\ln y) = \frac{1}{y}\frac{1}{\theta}e^{-\frac{\ln y}{\theta}} = \frac{y^{-\left(\frac{1}{\theta}+1\right)}}{\theta}$$

for $y > 1$ and 0 otherwise

19.10 0.25

Section 20

20.1 $\text{Var}(X) = \frac{5\alpha-2}{12(\alpha-1)^2(\alpha-2)}b^2$

20.2 $f_X(x) = \frac{4}{(4+x)^2}$ for $x > 0$ and 0 otherwise

20.3 1.7975

20.4 $E(\Lambda) = \alpha\beta = E(X)$

20.5 $F_X(x) = \frac{\theta x^\gamma}{1+\theta x^\gamma}$

20.6 0.6094

20.7 14

20.8 0.61

20.9 0.75

Section 21

21.1 $S_X(x) = M_\lambda(-x) = (1+\theta x)^{-\alpha}$

21.2 $S_X(x) = \frac{x}{(1+x)\ln(1+x)},\ x > 0$

21.3 $M_\Lambda(x) = e^{\frac{x}{\theta}},\ x \geq 0$

21.4 $a(x) = -1$

21.5 $S_X(x) = M_\Lambda(-x) = e^{e^{-x}-1}$

21.6 $S_X(x) = \frac{1}{10x}(e^{-x} - e^{-11x})$

21.7 $S_X(x) = M_\Lambda(-A(x)) = (1+\theta x^\gamma)^{-1}$

21.8 $S_{X|\Lambda}(x|\lambda) = e^{-\lambda(\sqrt{1+\theta x}-1)}$

21.9 $S_X(x) = M_\Lambda(-A(x)) = (1+\theta x)^{-\alpha}$

21.10 $f_X(x) = -S'_X(x) = -M'_\Lambda(-A(X))(-A(x))' = a(x)M'_\Lambda[-A(x)]$

21.11 $h_X(x) = \frac{f_X(x)}{S_X(x)} = \frac{a(x)M'_\Lambda[-A(x)]}{M_\Lambda(-A(x))}$

Section 22

22.1

$$f(x) = \begin{cases} \frac{1}{\alpha}, & 0 < x < c \\ (1-\alpha)\theta e^{-\theta x}, & x > c. \end{cases}$$

22.2 0.9252

22.3 $3 + \ln 5$

22.4 5.61

22.5

$$f(x) = \begin{cases} \frac{1}{1000+\theta}, & 0 < x \leq 1000 \\ \frac{1}{1000+\theta} e^{\frac{1000}{\theta} - \frac{x}{\theta}}, & x > 1000 \end{cases}$$

and 0 otherwise.

22.6 461.78

Section 23

23.1 We have

$$\begin{aligned}
\lim_{\tau\to\infty} \ln w_1 =& \lim_{\tau\to\infty} \frac{\ln\left(1+\frac{\alpha}{\tau}\right)}{(\alpha+\tau-\frac{1}{2})^{-1}} \\
=& \lim_{\tau\to\infty} \frac{\left(1+\frac{\alpha}{\tau}\right)^{-1}\left((-\frac{\alpha}{\tau^2}\right)}{-(\alpha+\tau-1/2)^{-2}} \\
=& \lim_{\tau\to\infty} \left(1+\frac{\alpha}{\tau}\right)^{-1} \alpha \left(\frac{\alpha}{\tau}+1-\frac{1}{2\tau}\right)^2 \\
=&\alpha.
\end{aligned}$$

Let

$$w_2 = \left[1 + \frac{(\xi/x)^\gamma}{\tau}\right]^{\alpha+\tau}.$$

We have

$$\begin{aligned}
\lim_{\tau\to\infty} \ln w_1 =& \lim_{\tau\to\infty} (\alpha+\tau) \ln\left[1+\frac{(\xi/x)^\gamma}{\tau}\right] \\
=& \lim_{\tau\to\infty} \frac{\left[1+\frac{(\xi/x)^\gamma}{\tau}\right]^{-1}(-\tau^{-2})(\xi/x)^\gamma}{-(\alpha+\tau)^{-2}} \\
=& \lim_{\tau\to\infty} \left[1+\frac{(\xi/x)^\gamma}{\tau}\right]^{-1} \left(\frac{\xi}{x}\right)^\gamma \left(1+\frac{\alpha}{\tau}\right)^2 \\
=& \left(\frac{\xi}{x}\right)^\gamma.
\end{aligned}$$

23.2 For large α, Stirling's formula gives

$$\Gamma(\alpha) \approx e^{-\alpha} \alpha^{\alpha-\frac{1}{2}} (2\pi)^{\frac{1}{2}}.$$

Also, we let $\xi = \theta\tau^{\frac{1}{\gamma}}$ so that $\theta = \xi\tau^{-\frac{1}{\gamma}}$.
Using this and Stirling's formula in the pdf of a transformed beta distribution, we find

$$\begin{aligned}
f_X(x) =& \frac{\Gamma(\alpha+\tau)\gamma x^{\gamma\tau-1}}{\Gamma(\alpha)\Gamma(\tau)\theta^{\gamma\tau}(1+x^\gamma\theta^{-\gamma})^{\gamma+\tau}} \\
\approx& \frac{e^{-\alpha-\tau}(\alpha+\tau)^{\alpha+\tau-\frac{1}{2}}(2\pi)^{\frac{1}{2}}\gamma x^{\gamma\tau-1}}{\Gamma(\alpha)e^{-\tau}(\tau)^{\tau-\frac{1}{2}}(2\pi)^{\frac{1}{2}}\xi^{\gamma\tau}\tau^{-\tau}(1+x^\gamma\xi^{-\gamma}\tau)^{\gamma+\tau}} \\
=& \frac{e^{-\alpha}\left(1+\frac{\alpha}{\tau}\right)^{\alpha+\tau-\frac{1}{2}}\gamma x^{-\gamma\alpha-1}}{\Gamma(\alpha)\tau^{-\alpha-\tau}\xi^{\gamma(\tau+\alpha)}\xi^{-\gamma\alpha}x^{-\gamma(\tau+\alpha)}(1+x^\gamma\xi^{-\gamma}\tau)^{\gamma+\tau}} \\
=& \frac{e^{-\alpha}\left(1+\frac{\alpha}{\tau}\right)^{\alpha+\tau-\frac{1}{2}}\gamma x^{-\gamma\alpha-1}}{\Gamma(\alpha)\xi^{-\gamma\alpha}\left[1+\frac{(\xi/x)^\gamma}{\tau}\right]^{\alpha+\tau}}
\end{aligned}$$

Let

$$w_1 = \left(1+\frac{\alpha}{\tau}\right)^{\alpha+\tau-\frac{1}{2}}.$$

From the previous problem, $\lim_{\tau\infty} w_1 = e^\alpha$. Now, let

$$w_2 = \left[1+\frac{(\xi/x)^\gamma}{\tau}\right]^{\alpha+\tau}.$$

Then

$$\lim_{\tau\to\infty} w_2 = e^{\left(\frac{\xi}{x}\right)^\gamma}.$$

Hence,

$$\lim_{\tau\to\infty} f_X(x) = \frac{\gamma\xi^{\gamma\alpha}}{\Gamma(\alpha)x^{\gamma\alpha+1}e^{\left(\frac{\xi}{x}\right)^\gamma}}$$

which is the pdf of an inverse transformed Gamma distribution

23.3 Let $\xi = \tau\theta$. Then

$$\begin{aligned}
f_{\text{inverser Pareto}}(x) =& \frac{\tau\theta x^{\tau-1}}{(x+\theta)^{\tau+1}} \\
=& \frac{\xi}{x^2}\left(1+\frac{\theta}{x}\right)^{-1}\left(1+\frac{\theta}{x}\right)^{-\frac{\xi}{\theta}} \\
\to& \frac{\xi e^{-\frac{\xi}{x}}}{x^2}
\end{aligned}$$

which is the pdf of an inverse exponential distribution with parameter ξ

Section 24

24.1 The pdf of the Gamma distribution can be written as

$$f(x,\theta) = \frac{1}{\Gamma(\alpha)\theta^{-\alpha}} x^{\alpha-1} e^{-\frac{x}{\theta}}.$$

We have: $q(\theta) = \theta^{-\alpha},\ p(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}$, and $r(\theta) = -\frac{1}{\theta}$

24.2 The pdf of X is

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-\lambda x} e^{x\ln\lambda}}{x!}.$$

Thus, $p(x) = \frac{1}{x!},\ q(\lambda) = e^{\lambda}$, and $r(\lambda) = \ln|lambda$

24.3 The pdf of X is

$$f(x,p) = C(m,x)p^x(1-p)^{m-x} = C(m,x)\left(\frac{p}{1-p}\right)^x (1-p)^m = \frac{C(m,x)e^{\ln\left(\frac{p}{1-p}\right)}}{(1-p)^{-m}}.$$

Thus, $p(x) = C(m.x),\ q(p) = (1-p)^{-m}$, and $r(p) = \ln\left(\frac{p}{1-p}\right)$

24.4 $E(X) = \lambda$ and $\text{var}(X) = \lambda$

24.5 $E(X) = mp$ and $\text{var}(X) = mp(1-p)$

Section 25

25.1 0.91873

25.2 $5e^{-5}$

25.3 0.1412

25.4 5 cars per week

Section 26

26.1 For the Poisson distribution the variance is equal to the mean. For the negative binomial and geometric the variance exceeds the mean. So the answer is (d)

26.2 0.75

26.3 $E(N) = 3$ and $\text{Var}(N) = 12$

26.4 $E(N) = 6$ and $\text{Var}(N) = 18$

26.5 192

26.6 $r = 2$ and $\beta = 4$

26.7 $P_N(z) = [1 - \beta(z-1)]^{-1}$

26.8 $CV(N) = \sqrt{\frac{1+\beta}{r\beta}}$

Section 27

27.1 The Poisson distribution has a variance equal to the mean. The negative binomial and geometric distributions have a varaince exceeding the mean. The binomial distribution has a variance less than the mean. Thus, the answer is (a)

27.2 38.34

27.3 $E(N) = 1.4$ and $\text{Var}(N) = 0.42$

27.4 0.0057

27.5 6.2784

27.6 0.172

27.7

m	0	1	2	3	4
p_m	0.1074	0.2684	0.3020	0.2013	0,0881
$F(m)$	0.1074	0.3758	0.6778	0.8791	0.9672

27.8 N_1 represents the number of successes in m_1 independent trials, each of which results in a success with probability q. Similarly, N_2 represents the number of successes in m_2 independent trials, each of which results in a success with probability q. Hence, as N_1 and N_2 are assumed to be independent, it follows that $N_1 + N_2$ represents the number of successes in $m_1 + m_2$ independent trials, each of which results in a success with probability q. So $N_1 + N_2$ is a binomial random variable with parameters $(m_1 + m_2, q)$

Section 28

28.1 For the given recursice equation, we find $a = -\frac{1}{3} < 0$ and $b = 4$. Thus, N is a binomial distribution. From $-\frac{q}{1-q} = -\frac{1}{3}$ we find $q = \frac{1}{4}$. From $(m+1)\frac{q}{1-q} = 4$ we find $m = 11$

28.2 $E(N) = 2.75$ and $\text{Var}(N) = 2.065$

28.3 0.125

28.4 3

28.5 0.0118

28.6 8

28.7 $a = 0.2$ and $b = 0.8$

28.8 (III)

28.9 0.3012

28.10 0.8

28.11 0.09

Section 29

29.1 We have

$$\frac{p_k^M}{p_k^T} = \frac{\left(\frac{1-p_0^M}{1-p_0}\right)p_k}{\frac{1}{1-p_0}p_k} = 1 - p_0^M$$

29.2 We have

$$E(N^M) = [P_N^M]'(1) = \left(\frac{1-p_0^M}{1-p_0}\right)P_N'(1) = \left(\frac{1-p_0^M}{1-p_0}\right)E(N)$$

29.3 $E[N^M(N^M-1)] = \left(\frac{1-p_0^M}{1-p_0}\right)E[N(N-1)]$ and

$$\mathrm{Var}(N^M) = \left(\frac{1-p_0^M}{1-p_0}\right)E[N(N-1)] + \left(\frac{1-p_0^M}{1-p_0}\right)E(N) - \left[\left(\frac{1-p_0^M}{1-p_0}\right)E(N)\right]^2$$

29.4

$$\begin{aligned}
p_0 =& e^{-\lambda} = e^{-1} = 0.3679\\
a =& 0\\
b =& 1\\
p_1 =& p_0\frac{1}{1} = 0.3679\\
p_2 =& p_1\frac{1}{2} = 0.1839\\
E(N^T) =& \frac{E(N)}{1-p_0} = \frac{1}{1-0.3679} = 1.582\\
E[N^T(N^T-1)] =& \left(\frac{1}{1-p_0}\right)E[N(N-1)] = \frac{1}{1-0.3679} = 1.582\\
\mathrm{Var}(N^T) =& 1.582 + 1.582 - 1.582^2 = 0.661276
\end{aligned}$$

29.5 (a) $P_N^M(z) = \frac{1}{2}\left(1 + \frac{z}{3-2z}\right)$ (b) $E(N^M) = 1.5$ and $\mathrm{Var}(N^M) = 2.25$

29.6 0.5

Section 30

30.1

$$\begin{aligned}
a =& \frac{\beta}{1+\beta} = \frac{1}{1+1} = 0.5 \\
b =& (r-1)\frac{\beta}{1+\beta} = -0.75 \\
p_2^T =& p_1^T\left(0.5 - \frac{0.75}{2}\right) = 0.106694 \\
p_3^T =& p_2^T\left(0.5 - \frac{0.75}{3}\right) = 0.026674 \\
p_1^M =& (1-p_0M)p_1^T = 0.341421 \\
p_2^M =& (1-p_0M)p_2^T = 0.042678 \\
p_3^M =& (1-p_0M)p_3^T = 0.010670
\end{aligned}$$

30.2 $E(N) = \frac{\beta}{\ln(1+\beta)}$

30.3 $E[N(N-1)] = P_N''(1) = \frac{\beta^2}{\ln(1+\beta)}$

30.4 $\mathrm{Var}(N) = \frac{\beta}{\ln(1+\beta)}\left[1+\beta-\frac{\beta}{\ln(1+\beta)}\right]$

30.5 We have

$$\begin{aligned}
P_N^T(z) = \sum_{n=0}^{\infty} p_n^T z^n = \sum_{n=1}^{\infty} p_n^T z^n \\
=& \frac{1}{1-p_0}\sum_{n=1}^{\infty} p_n z^n \\
=& \frac{1}{1-p_0}\sum_{n=0}^{\infty} p_n z^n - \frac{p_0}{1-p_0} \\
=& \frac{P_N(z)-p_0}{1-p_0}
\end{aligned}$$

30.6 $P_N^T(z) = \frac{[1-\beta(z-1)]^{-r}-(1+\beta)^{-r}}{1-(1+\beta)^{-r}}$

30.7 $E(N^T) = [P_N^T]'(1) = \frac{P_N'(1)}{1-p_0} = \frac{r\beta}{1-(1+\beta)^{-r}}$

30.8 $E[N^T(N^T-1)] = [P_N^T]''(1) = \frac{P_N''(1)}{1-p_0} = \frac{r(r+1)\beta^2}{1-(1+\beta)^{-r}}$

30.9 $\text{Var}(N^T) = \frac{r(r+1)\beta^2}{1-(1+\beta)^{-r}} + \frac{r\beta}{1-(1+\beta)^{-r}} - \left[\frac{r\beta}{1-(1+\beta)^{-r}}\right]^2$

Section 31

31.1

$$f_{Y^L}(y) = \begin{cases} 1-e^{-0.25} & y=0 \\ 0.0002(y+50)e^{-\left(\frac{y+50}{100}\right)^2} & y>0. \end{cases}$$

$$F_{Y^L}(y) = \begin{cases} 1-e^{-0.25} & y=0 \\ 1-e^{-\left(\frac{y+50}{100}\right)^2} & y>0. \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} e^{-0.25} & y=0 \\ e^{-\left(\frac{y+50}{100}\right)^2} & y>0. \end{cases}$$

31.2

$$\begin{aligned} f_{Y^P}(y) =& 0.0002(y+50)e^{-0.0001y^2-0.01y} \\ F_{Y^P}(y) =& 1-e^{-0.0001y^2-0.01y} \\ S_{Y^P}(y) =& e^{-0.0001y^2-0.01y} \\ h_{Y^P}(y) =& 0.0002(y+50). \end{aligned}$$

31.3 $E(Y^L) = \frac{1}{1000}e^{-500{,}000}$

31.4 $E(Y^P) = 0.001$

31.5 $2000e^{-400\theta}$

31.6 1708.70

31.7 $\frac{(b-d)^2}{12}$

31.8 (a) 188.75 (b) 269.64

31.9 (a) $\frac{\theta^\alpha}{(\alpha-1)(\theta+\alpha)^{\alpha-1}}$ (b) ∞

31.10 30

Section 32

32.1

$$f_{Y^L}(y) = \begin{cases} \frac{d}{\theta}, & y = 0 \\ \frac{1}{\theta}, & y > d. \end{cases}$$

$$F_{Y^L}(y) = \begin{cases} \frac{d}{\theta}, & 0 \leq y \leq d \\ \frac{y}{\theta}, & y > d. \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} 1 - \frac{d}{\theta}, & 0 \leq y \leq d \\ 1 - \frac{y}{\theta}, & y > d. \end{cases}$$

$$h_{Y^L}(y) = \begin{cases} 0, & 0 < y < d \\ \frac{1}{\theta - y}, & y > d. \end{cases}$$

32.2

$$f_{Y^P}(y) = \frac{1}{\theta - d},\ y > d.$$

$$F_{Y^P}(y) = \begin{cases} 0, & 0 \leq y \leq d \\ \frac{y-d}{\theta-d}, & y > d. \end{cases}$$

$$S_{Y^P}(y) = \begin{cases} 1, & 0 \leq y \leq d \\ \frac{\theta-y}{\theta-d} & y > d. \end{cases}$$

$$h_{Y^P}(y) = \begin{cases} 0, & 0 < y < d \\ \frac{1}{\theta - y}, & y > d. \end{cases}$$

32.3 $E(Y^L) = \frac{\theta^2 - d^2}{2\theta}$ and $E(Y^P) = \frac{\theta + d}{2}$

32.4 340.83

32.5 900

32.6 $d = \frac{\ln 0.40}{-\theta}$

32.7 (a) 310 (b) 387.5

32.8 320.83

32.9 456

32.10 6400

Section 33

33.1 LER = 79.2%. The is the percentage of savings in claim payments due to the presence of the deductible 30.

33.2 86.6%

33.3 1546

33.4 500

33.5 333.33

33.6 0.07418

33.7 0.5

33.8 510.16

33.9 0.625

Section 34

34.1 We have

$$f_Y(y) = \begin{cases} e^{-\theta u}, & y = u \\ \theta e^{-\theta y}, & y < u \\ 0, & y > u \end{cases}$$

and

$$F_Y(y) = \begin{cases} 1 - e^{-\theta y}, & y < u \\ 1 & y \geq u \end{cases}$$

34.2 $E(X \wedge u) = \int_0^u e^{-\theta x} dx = \frac{1}{\theta}\left(1 - e^{-\theta u}\right)$

34.3 $E((1+r)X \wedge u) = \frac{(1+r)}{\theta}\left(1 - e^{-\frac{\theta u}{1+r}}\right)$

34.4 48

34.5 182.18

34.6 2011.80

34.7 5176.78

Section 35

35.1 1.115

35.2 990,938.89

35.3 0.4163

35.4 133

35.5 353.55

35.6 3031.06

35.7 85%

35.8 29.93

35.9 109.4

35.10 0.583

Section 36

36.1 8.0925

36.2 0.5553

36.3 The pgf of N^L is $P_{N^L}(z) = e^{0.5(z-1)}$. Hence,

$$P_{N^P}(z) = P_{N^L}[1 + v(z-1)] = e^{0.5(0.5553)(z-1)} = e^{0.27756(z-1)}.$$

Note that N^P is a Poisson distribution with mean $\lambda - 0.27756$

36.4 0.242444

36.5 $P_{N^P}(z) = (1 - 0.11852(z-1))^{-1}$

36.6 0.6304

36.7 0.4424

Section 37

37.1 (a) This is false, In a collective risk model, all the loss amounts are identically distributed.
(b) This is true since the loss amounts need not all have the same distribution.
(c) This is true. In the collective risk model, N (the frequency random variable) is determined independently of each of the subscripted Xs, the severity random variables.
(d) This is false. If frequency is independent of severity, as it is in the collective risk model, then this implies that the number of payments does not affect the size of each individual payment

37.2 0

37.3 $E(S) = 70$ and $\text{Var}(S) = 7900$

37.4 This is a collective loss model

37.5 The probability generating function of N is given by

$$P_N(z) = [1 - \beta(z-1)]^{-r}$$

so that the pgf can be expressed in the form

$$P_N(z;\alpha) = Q(z)^\alpha$$

where $\alpha = r$ and $Q(z) = [1 - \beta(z-1)]^{-1}$

Section 38

38.1 $E(S^3) = E(N^3)E(X)^2 - 3E(N^2)E(X)^2 + 2E(N)E(X)^2 + 3E(N^2)E(X)E(X^2) - 3E(N)E(X)E(X^2) + E(N)E(X^3)$

38.2 $E[(S-E(S))^3] = E(N)E[(X-E(X))^3]+3E[(N-E(N))^2]E(X)E[(X-E(X))^2] + E[(N-E(N))^3]E(X)^3$

38.3 $F_X^{*2}(x) = \int_0^x \Phi\left(\frac{\ln(x-y)-\mu}{\sigma}\right)\frac{\phi\left(\frac{\ln y-\mu}{\sigma}\right)}{\sigma y}dy$

38.4 We have

$$E(S) = 0.18 + 0.182(2) + 0.0909(3) + 0.0182(4) = 0.8895$$
$$E(S^2) = 0.18 + 0.182(2^2) + 0.0909(3^2) + 0.0182(4^2) = 2.0173$$

38.5 1.226

38.6 $E(Y) = 2.5$ and $\mathrm{Var}(Y) = 23.75$

38.7 0.1587

38.8 0.1637

38.9 0.0681

38.10 $E(S) = 5600$ and $\mathrm{Var}(S) = 9,710,400$

38.11 100

38.12 0.1003

38.13 0.242

38.14 0.1230

38.15 24

38.16 518

38.17 40

38.18 0.2483

38.19 0.0039

38.20 0.37

38.21 65.3

38.22 0.4207

38.23 0.0233

Section 39

39.1 1.014

39.2 2.25

39.3 1/3

39.4 2.064

39.5 25/16

39.6 2.3608

39.7 2.064

39.8 18.15

39.9 18.81

Section 40

40.1 The mgf of S is

$$\begin{aligned}
M_S(z) =& P_N[M_X(z)] = P_N[(1-\theta z)^{-1}]\\
=& \left\{1-\beta[(1-\theta z)^{-1}-1]\right\}^{-r}\\
=& \left(\frac{1-\theta z}{[1-\theta(1+\beta)z]}\right)^r\\
=& \left\{\frac{1}{1+\beta}\left[\frac{1-\theta(1+\beta)z+\beta}{1-\theta(1+\beta)z}\right]\right\}^r\\
=& \left\{\frac{1}{1+\beta}\left[1+\frac{\beta}{1-\theta(1+\beta)z}\right]\right\}^r\\
=& \left[1+\frac{\beta}{(1+\beta)[1-\theta(1+\beta)z]}-\frac{\beta}{1+\beta}\right]^r\\
=& \left(1+\frac{\beta}{1+\beta}\left\{[1-\theta(1+\beta)z]^{-1}-1\right\}\right)^r\\
=& P_N^*[M_X^*(z)]
\end{aligned}$$

where

$$P_N^*(z) = \left[1+\frac{\beta}{1+\beta}(z-1)\right]^r$$

is the pgf of the binomial distribution with parameters r and $\frac{\beta}{1+\beta}$ and $M_X^*(z) = [1-\theta(1+\beta)z]^{-1}$ is the mgf of the exponential distribution with mean $\theta(1+\beta)$. Thus, the negative binomial-exponential model is equivalent to the binomial exponential model.

40.2 This follows from Exercise 40.1 and

$$F_S(x) = 1-\sum_{n=1}^{\infty}\Pr(N=n)\sum_{j=0}^{n-1}\frac{(x/\theta)^j e^{-x/\theta}}{j!}$$

40.3 $F_S(x) = 1-\frac{\beta}{1+\beta}e^{-\frac{x}{\theta(1+\beta)}},\ x\geq 0$

40.4 $f_S(x) = \frac{\beta}{\theta(1+\beta)^2}e^{-\frac{x}{\theta(1+\beta)}},\ x\geq 0$. Note that S is a mixed distribution with the discrete part $\Pr(S=0) = F_S(0) = (1+\beta)^{-1}$ and the continuous part is the exponential distribution with mean $\theta(1+\beta)$

40.5 Suppose that $X_1, X_2, \cdots, X_N$ are independent Poisson random variables with parameters $\lambda_1, \cdots, \lambda_N$. Then

$$M_S(t) = e^{\lambda_1(e^t-1)}\cdots e^{\lambda_N(e^t-1)} = e^{(\lambda_1+\cdots+\lambda_N)(e^t-1)}.$$

Hence, S is a Poisson random variable with parameter $\lambda_1 + \cdots + \lambda_N$

40.6 Suppose that $X_1, X_2, \cdots, X_N$ are independent binomial random variables with parameters $(n_1, p), (n_2, p), \cdots, (n_N, p)$. Then

$$M_S(t) = (q + pe^t)^{n_1} \cdots (q + pe^t)^{n_N} = (q + pe^t)^{n_1 + n_2 + \cdots n_N}.$$

Hence, S is a binomial random variable with parameters $(n_1 + \cdots + n_N, p)$.

40.7 Suppose that $X_1, X_2, \cdots, X_N$ are independent negative binomial random variables with parameters $(r_1, q), (r_2, q), \cdots, (r_N, q)$. Then

$$P_S(t) = [1-\beta(z-1)]^{-r_1}[1-\beta(z-1)]^{-r_2} \cdots [1-\beta(z-1)]^{-r_N} = [1-\beta(z-1)]^{-(r_1+r_2+\cdots+r_N)}.$$

Hence, S is a negative binomial random variable with parameters $(r_1 + \cdots + r_N, \beta)$. In particular, the family of geometric random variables is closed under convolution.

40.8 Suppose that $X_1, X_2, \cdots, X_N$ are independent gamma random variables with parameters $(\alpha_1, \theta), (\alpha_2, \theta), \cdots, (\alpha_N, \theta)$. Then

$$M_S(t) = (1 - \theta t)^{-\alpha_1}(1 - \theta t)^{-\alpha_2} \cdots (1 - \theta t)^{-\alpha_N} = (1 - \theta t)^{-(\alpha_1 + \alpha_2 + \cdots + \alpha_N)}.$$

Hence, S is a gamma random variable with parameters $(\alpha_1 + \cdots + \alpha_N, \theta)$. In particular, the family of exponential random variables is closed under convolution.

40.9 We have

$$\begin{aligned}F_S(x) =& 1 - e^{-\frac{x}{2}} \sum_{j=0}^{\infty} \frac{(x/2)^j}{j!} \sum_{n=j+1}^{\infty} \Pr(N = n) \\ =& 1 - \frac{1}{2} e^{-\frac{x}{2}} \left(1 + \frac{x}{10}\right).\end{aligned}$$

40.10 $M_S(2) = 3.6 \times 10^{48}$

Section 41

41.1 $f_S(0) = f_N(0) = e^{-0.04} = 0.9608$ and

$$\begin{aligned}
f_S(1) =&0.04f_X(1)f_S(0) = 0.04(0.5)(0.9608) = 0.019216\\
f_S(2) =&0.02[f_X(1)f_S(1) + 2f_X(2)f_S(0)]\\
=&0.02[0.5(0.019216) + 2(0.4)(0.9608)] = 0.01556496\\
f_S(3) =&\frac{0.04}{3}[f_X(1)f_S(2) + 2f_X(2)f_S(1) + 3f_X(3)f_S(0)]\\
=&0.0041487371\\
f_S(4) =&0.01[f_X(1)f_S(3) + 2f_X(2)f_S(2) + 3f_X(3)f_S(1) + 4f_X(4)f_S(0)]\\
=&0.00037586.
\end{aligned}$$

41.2 We have

$$\begin{aligned}
f_S(n) =&\frac{0.04}{n}[f_X(1)f_S(n-1) + f_X(2)f_S(n-2) + f_X(3)f_S(n-3)]\\
=&\frac{0.04}{n}[0.5f_S(n-1) + 0.4f_S(n-2) + 0.1f_S(n-3)]\\
=&\frac{1}{n}[0.02f_S(n-1) + 0.016f_S(n-2) + 0.004f_S(n-3)].
\end{aligned}$$

41.3 0.15172

41.4 12

41.5 76

41.6 165

41.7 1.0001

41.8 0.3336

41.9 $f_S(1) = 0.3687$ and $f_S(2) = 0.2055$

41.10 0.0921

41.11 0.2883

Section 42

42.1 (a)

$$\begin{aligned}
f_0 =&F_X(5) = \frac{5}{50} = 0.1\\
f_1 =&F_X(15) - F_X(5) = \frac{15}{50} - \frac{5}{50} = 0.2\\
f_2 =&F_X(25) - F_X(15) = \frac{25}{50} - \frac{15}{50} = 0.2\\
f_3 =&1 - F_X(25) = 1 - \frac{25}{50} = 0.5.
\end{aligned}$$

(b) 0.1935

42.2 0.0368

42.3 0.0404

42.4 We have

$$m_0^1 + m_1^1 = F_X(6) - F_X(3) = \left(\frac{5}{8}\right)^3 - \left(\frac{5}{11}\right)^3 = 0.150226$$

and

$$3m_0^1 + 6m_1^1 = \int_3^6 \frac{3(5)^3 x}{(x+5)^4} dx = 0.62897.$$

Solving this system we find $m_0^1 = 0.090796$ and $m_1^1 = 0.05943$.

42.5 We have $f_0 = m_0^0 = 0.4922$ and $f_3 = m_1^0 + m_0^1 = 0.2637 + 0.090796 = 0.3545$.

42.6 0.03682

42.7 0.03236

Section 43

43.1

$$\begin{aligned}F_{Y^L}(y) =&1-v+vF_{Y^P}(y)\\=&\begin{cases}1-v+v\left[\frac{\Pr(X>6)-\Pr\left(X>6+\frac{y}{0.75}\right)}{\Pr(X>6)}\right], & y<13.5\\ 2-v, & y\geq 13.5\end{cases}\end{aligned}$$

where $v=0.15259$

43.2 $E(S)=262.4621$ and $\text{Var}(S)=487,269,766.1$

43.3 $F_{Y^P}(y)=1-\Pr(0.53Z>y)=\frac{\Pr(X>30)-\Pr\left(X>30+\frac{y}{0.53}\right)}{\Pr(X>30)}$ and $F_{Y^P}(y)=1$ for $y\geq 164.30$

43.4

$$\begin{aligned}f_0 =&F_{Y^P}(15)=0.2465\\f_1 =&F_{Y^P}(45)-F_{Y^P}(15)=0.3257\\f_2 =&F_{Y^P}(75)-F_{Y^P}(45)=0.1849\\f_3 =&F_{Y^P}(105)-F_{Y^P}(75)=0.0778\\f_4 =&F_{Y^P}(135)-F_{Y^P}(105)=0.0322\\f_5 =&F_{Y^P}(165)-F_{Y^P}(135)=0.058\\f_n =&1-1=0,\ n=6,7,\cdots\end{aligned}$$

43.5

$$\begin{aligned}f_S(0) =&e^{7(1-0.3257)}=112.179\\f_S(n) =&\frac{5.18574}{n}\sum_{j=1}^{n}f_nf_S(n-j)\end{aligned}$$

43.6 $M_{Y^L}(t)=0.25918+0.74082((1-100t)^{-1}$

43.7 $P_{N^P}(z)=e^{1.06813(z-1)}$

Section 44

44.1 $E(S) = 100,000$ and $\text{Var}(S) = 1.6002 \times 10^{11}$

44.2 $M_S(t) = \left(0.80 + 0.20(1 - 0.001t)^{-1}\right)^{100}$

44.3 $P_S(z) = M_S[\ln z] = \left(0.80 + 0.20(1 - 0.001 \ln z)^{-1}\right)^{100}$

44.4 (a) The mean is $E(S) = 163$ and the variance is $\text{Var}(S) = 1107.77$. (b) 217.75

44.5 $E(S) = 352n$ and $\text{Var}(S) = 88856n$

Section 45

45.1 53

45.2 0.29

45.3 0.18

45.4 1975

45.5 $\lambda = 1.52$ or $\lambda = 5.45$

Section 46

46.1 5

46.2 0.2

46.3 151.52

46.4 $\frac{n+8}{18(n-1)^2}$

46.5 $20\sqrt{10}$

46.6 $\frac{1}{n-1}\theta$

46.7 12

46.8 (D)

Section 47

47.1 1.64

47.2 2.58

47.3 0.2

47.4 $78.14 \le \mu \le 81.86$

47.5 $0.1825 \le p \le 0.2175$

Section 48

48.1 (a) $H_0 : \mu = 18,000$ and $H_1 : \mu < 18,000$
(b) $H_0 : \mu = 18,000$ and $H_1 : \mu \ne 18,000$
(c) $H_0 : \mu = 18,000$ and $H_1 : \mu > 18,000$

48.2 (i) Two-tailed (ii) Left-tailed (iii) Right-tailed

48.3 The null and alternative hypotheses are

$$\begin{aligned} H_0 :& \mu \ge 30 \\ H_1 :& \mu < 30. \end{aligned}$$

The test statistic for the given sample is

$$z = \frac{20 - 30}{6/\sqrt{5}} = -3.727.$$

The rejection region is $Z < -1.28$. Since $-3.727 < -1.28$, so we reject the null hypothesis in favor of the alternative.Thus, the mean time to find a parking space is less than 30 minutes.

48.4 We reject the null hypothesis when the level of confidence is greater than or equal to the $p-$value. Thus, the answer to the question is:(ii), (iii), and (iv).

48.5 The null and alternative hypotheses are

$$H_0 : \mu = 20$$
$$H_1 : \mu > 20.$$

The test statistic is

$$z = \frac{22.60 - 20}{2.50/\sqrt{49}} = 7.28.$$

The rejection region corresponding to $\alpha = 0.02$ is $Z > 2.06$. Since $7.28 > 2.06$, we reject H_0 and conclude that the typical amount spent per customer is more than \$20.

48.6 The null and alternative hypotheses are

$$H_0 : \mu = 16$$
$$H_1 : \mu \neq 16.$$

The test statistic is

$$z = \frac{16.32 - 16}{0.8/\sqrt{30}} = 2.19.$$

The rejection region corresponding to $\alpha = 0.10$ is $|Z| > 1.645$. Since $2.19 > 1.645$, we reject H_0 and conclude that the process is out of control.

48.7 The null and alternative hypotheses are

$$H_0 : \mu = 10$$
$$H_1 : \mu \neq 10.$$

We have $z_{\frac{\alpha}{2}} = z_{0.01} = 2.33$. Thus, the critical values are -2.33 and 2.33.

48.8 (d)

Section 49

49.1

x	1	2	3	4	5	6	7
$p_{12}(x)$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$

$$F_{12}(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{6} & 1 \le x < 2 \\ \frac{1}{4} & 2 \le x < 3 \\ \frac{5}{12} & 3 \le x < 4 \\ \frac{7}{12} & 4 \le x < 5 \\ \frac{2}{3} & 5 \le x < 6 \\ \frac{5}{6} & 6 \le x < 7 \\ 1, & x \ge 7. \end{cases}$$

49.2 The empirical mean is $\overline{X} = \frac{41}{12}$ and the empirical variance is $\frac{1331}{144}$.

49.3 (a) We have

$$\hat{H}(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{6}, & 1 \le x < 2 \\ \frac{4}{15}, & 2 \le x < 3 \\ \frac{22}{45}, & 3 \le x < 4 \\ \frac{244}{315}, & 4 \le x < 5 \\ \frac{307}{315}, & 5 \le x < 6 \\ \frac{929}{630}, & 6 \le x < 7 \\ \frac{1559}{630}, & x \ge 7. \end{cases}$$

(b) We have

$$\hat{S}(x) = \begin{cases} 1, & x < 1 \\ 0.8465, & 1 \le x < 2 \\ 0.7659, & 2 \le x < 3 \\ 0.6133, & 3 \le x < 4 \\ 0.4609, & 4 \le x < 5 \\ 0.3773, & 5 \le x < 6 \\ 0.2289, & 6 \le x < 7 \\ 0.0842, & x \ge 7. \end{cases}$$

49.4

$$S_n(x) = \begin{cases} 1, & x < 49 \\ \frac{8}{9} = \frac{1}{9}, & 49 \le x < 50 \\ \frac{5}{9} = \frac{4}{9}, & 50 \le x < 60 \\ \frac{4}{9} = \frac{5}{9}, & 60 \le x < 75 \\ \frac{3}{9} = \frac{2}{3}, & 75 \le x < 80 \\ \frac{2}{9} = \frac{7}{9}, & 80 \le x < 120 \\ \frac{1}{9} = \frac{8}{9}, & 120 \le x < 130 \\ 0, & x \ge 130. \end{cases}$$

49.5 6

49.6 1.291

49.7 12

Section 50

50.1 (a) 47.50 (b) 3958.33

50.2 81

50.3 120

50.4 0.396

50.5 20,750

Section 51

51.1 (A) and (D) are false. (B) and (C) are true

51.2 Losses above a policy limit are right-censored and losses below a policy deductible are left-truncated. The answer is (D)

51.3

Life	d_i	x_i	u_i
1	0	–	0.2
2	0	–	0.3
3	0	–	0.5
4	0	0.5	–
5	1	–	0.7
6	1.2	1.0	–
7	1.5	–	2.0
8	2	2.5	–
9	2.5	–	3.0
10	3.1	3.5	–
11	0	–	4.0
12	0	–	4.0
13	0	–	4.0
14	0	–	4.0
15	0	–	4.0
16	0	–	4.0
17	0	–	4.0
18	0	–	4.0
19	0.5	–	4.0
20	0.7	1.0	–
21	1.0	3.0	–
22	1.0	–	4.0
23	2.0	2.5	–
24	2.0	–	2.5
25	3.0	3.5	–

51.4

j	y_j	s_j	r_j
1	4	1	$2+2-1=3$
2	8	1	$1+0-0=1$

51.5

j	y_j	s_j	r_j
1	0.9	1	$5+5-3=7$
2	1.5	1	$4+4-2=6$
3	1.7	1	$3+2-0=5$
4	2.1	2	$2+1-0=3$

Section 52

52.1 0.52

52.2 0.583

52.3 0.067

52.4 0.7143

52.5 2

52.6 0.385

52.7 0.112

52.8 100

Section 53

53.1 We have

$$E[p_n(x_j)] = E\left(\frac{N_j}{n}\right) = \frac{E(N_j)}{n} = \frac{np(x_j)}{n} = p(x_j).$$

This shows that the estimator is unbiased. Finding the variance of $p_n(x_j)$ we have

$$\mathrm{Var}[p_n(x_j)] = \frac{np(x_j)[1-p(x_j)]}{n^2} = \frac{p(x_j)[1-p(x_j)]}{n} \to 0$$

as $n \to \infty$. This shows that the estimator is consistent

53.2 3.951×10^{-7}

53.3 We have

$$\begin{aligned}p_{386}(2) =& \frac{64}{386} = 0.1658\\ \widehat{\mathrm{Var}}[p_{386}(2)] =& \frac{p_{386}(2)[1-p_{386}(2)]}{n}\\ =& \frac{\frac{64}{386}\frac{322}{386}}{386} = 3.58 \times 10^{-4}\end{aligned}$$

53.4 The endpoints of the interval are:

$$0.1658 \pm 1.96\sqrt{\frac{0.1658(1-0.1658)}{386}} \Longrightarrow (0.1287, 0.2029)$$

Section 54

54.1 Since $X_1, X_2, \cdots, X_n$ are independent so are $X_1^2, X_2^2, \cdots, X_n^2$. We have

$$\begin{aligned}\mathrm{Var}(X_1X_2\cdots X_n) =& E[(X_1X_2\cdots X_n)^2] - E(X_1X_2\cdots X_n)^2\\ =& E(X_1^2)E(X_2^2)\cdots E(X_n^2) - E(X_1)^2E(X_2)^2\cdots E(X_n)^2\\ =& \prod_{i=1}^n(\mu_i^2+\sigma_i^2) - \prod_{i=1}^n \mu_i^2\end{aligned}$$

54.2 We have

$$(1+a_1)(1+a_2)\cdots(1+a_n) = 1 + a_1 + a_2 + \cdots + a_n + \text{ prodcuts of } a_is.$$

But the a_is are given small so that a product of a_is is even smaller. Ignoring all the product terms, we obtain the desired result

54.3 0.0148

54.4 0.03086

54.5 10

Section 55

55.1 $\hat{H}(3) = 0.6875$ and $\widehat{\text{Var}}(\hat{H}(3)) = 0.02376$

55.2 The 95% linear confidence interval is

$$(0.6875 - 1.96\sqrt{0.02376}, 0.6875 + 1.96\sqrt{0.02376}) = (0.3854, 0.9896).$$

The 95% log-transformed confidence interval is

$$(0.6875e^{-1.96\frac{\sqrt{0.02376}}{0.6875}}, 0.6875e^{1.96\frac{\sqrt{0.02376}}{0.6875}}) = (0.4431, 1.0669)$$

55.3 0.607

55.4 $(0.189, 1.361)$

55.5 0.779

55.6 $(0.443, 1.067)$

55.7 0.2341

Section 56

56.1 0.485

56.2 0.53125

56.3 0.026

56.4 0.3

56.5 $1 \leq x \leq 2$

56.6 0.3

Section 57

57.1 (a) 990 (b) 0.0080

57.2 We have

$$\begin{aligned}r_j =&\sum_{i=0}^{j} d_j - \sum_{i=1}^{j-1}(x_j + u_j)\\ =&[\sum_{i=0}^{j-1} d_j - \sum_{i=1}^{j-2}(x_j + u_j)] + d_j - (x_{j-1} + u_{j-1})\\ =&r_{j-1} + d_j - (x_{j-1} + u_{j-1})\end{aligned}$$

57.3 0.75

57.4 0.6

Section 58

58.1 0.52490

58.2 0.52

58.3 369

58.4 26,400

58.5 13.75

58.6 107.8

58.7 20

58.8 384

58.9 −0.24

58.10 4.468

58.11 246.6

58.12 17.55

58.13 208.3

58.14 1.614

58.15 296.21

58.16 224

58.17 104.4

58.18 118.32

58.19 0.983

58.20 0.345

Section 59

59.1 $\tilde{\theta} = \frac{x_1+x_2+\cdots+x_n}{2n}$

59.2 $\hat{\theta} = \max\{x_1, x_2, \cdots, x_n\}$

59.3 4.3275

59.4 3.97

59.5 $\hat{\theta} = \min\{x_1, x_2, \cdots, x_n\}$

59.6 0.2507

59.7 2

59.8 0.6798

59.9 1996.90

59.10 0.447

59.11 $L(p) = \left(\frac{pe^{-1}}{100} + \frac{1-p}{10{,}000}e^{-0.01}\right)\left(\frac{p}{100}e^{-2} + \frac{1-p}{10{,}000}e^{-0.2}\right)$

59.12
$$\ell'(\alpha_1) = \frac{n}{\alpha_1} - \sum_{i=1}^{n} \ln x_i + \frac{2m}{\alpha_1(2+\alpha_1)} - \frac{6}{(2+\alpha_1)^2}\sum_{i=1}^{m} \ln y_i$$

59.13 16.74

59.14 916.7

Section 60

60.1 73

60.2 233.333

60.3 703

60.4 3000

60.5 2.41877

60.6 $L(\theta) = \frac{e^{-\frac{1100}{\theta}}}{\theta^3}$

60.7 471

60.8 $f(50)f(15)f(60)f(500)[1 - F(100)][1 - F(500)]$

60.9 3.089

60.10 3,325.67

60.11 0.09

60.12 1067

60.13 3/8

60.14 52.68

Section 61

61.1 $\ell'(\alpha)^2 = \frac{16}{\alpha^2} - \frac{18.501}{\alpha} + 5.3481$

61.2 $\ell''(\alpha) = -\frac{4}{\alpha^2}$

61.3 $I(\alpha) = \frac{16}{1.73^2} = 5.346$

61.4 0.1871

61.5 $[0.8822, 2.5778]$

61.6 $\frac{1}{n}$

61.7 0.447

61.8 $\frac{3\theta^2}{n}$

61.9 0.97

Section 62

62.1

$$\ell(\alpha,\theta) = \sum_{i=1}^{n}[\ln\alpha + \alpha\ln\theta - (\alpha+1)\ln(x_i+\theta)]$$

62.2

$$\begin{aligned}\frac{\partial^2\ell}{\partial\alpha^2} &= -\frac{n}{\alpha^2}\\ \frac{\partial^2\ell}{\partial\theta\partial\alpha} &= \frac{n}{\theta} - \sum_{i=1}^{n}\frac{1}{x_i+\theta}\\ \frac{\partial^2\ell}{\partial\theta^2} &= -\frac{n\alpha}{\theta^2} + \sum_{i=1}^{n}\frac{(\alpha+1)}{(x_i+\theta)^2}\end{aligned}$$

62.3

$$I(\theta) = \begin{bmatrix} -\frac{n}{\alpha^2} & \frac{n}{\theta} - \frac{n\alpha}{(\alpha+1)\theta} \\ \frac{n}{\theta} - \frac{n\alpha}{(\alpha+1)\theta} & -\frac{n\alpha}{\theta^2} + \frac{n\alpha(\alpha+1)}{(\alpha+2)\theta^2} \end{bmatrix}$$

62.4

$$I(\theta)^{-1} = \frac{1}{det[I(\theta)]}\begin{bmatrix} -\frac{n\alpha}{\theta^2} + \frac{n\alpha(\alpha+1)}{(\alpha+2)\theta^2} & -\frac{n}{\theta} + \frac{n\alpha}{(\alpha+1)\theta} \\ -\frac{n}{\theta} + \frac{n\alpha}{(\alpha+1)\theta} & -\frac{n}{\alpha^2} \end{bmatrix}$$

62.5

$$I(\alpha,\theta) = \begin{bmatrix} 5.0391 & -0.4115 \\ -0.4115 & -0.0524 \end{bmatrix}$$

62.6

$$\widehat{\text{Var}}(\hat{\alpha},\theta) = \frac{1}{(5.0391)(-0.0524) - 0.4115^2}\begin{bmatrix} 0.0524 & -0.4115 \\ -0.4115 & -5.0391 \end{bmatrix} = \begin{bmatrix} -0.1209 & 0.9495 \\ 0.9495 & 11.6274 \end{bmatrix}$$

62.7 0.0187

62.8 (a) We have

$$M_Y(e^{tY}) = M_Y(e^{tX_1}e^{tX_2}) = (1-\theta t)^{-2}.$$

Thus, Y is a Gamma distribution with parameters $\alpha = 2$ and θ.
(b) 0.15
(c) 0.079

62.9 $(0.23, 0.69)$

62.10 0.02345

62.11

$$\begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$$

Section 63

63.1 -1.00774

63.2 $[0.70206, 4.20726]$

63.3 $[2.641591352, 8.358408648]$

Section 64

64.1 This follows from

$$\int_0^\infty \frac{d\theta}{\theta} = \ln\theta|_0^\infty = \infty$$

64.2 (a) The model distribution is

$$f_{\mathbf{X}|Q}(\mathbf{x}|q) = (q)(q) = q^2.$$

(b) The joint distribution is

$$f_{\mathbf{X},Q}(\mathbf{x},q) = q^2\frac{q^2}{0.039} = \frac{q^4}{0.039}.$$

(c) The marginal distribution in X is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{0.2}^{0.5} \frac{q^4}{0.039}dq = 0.15862.$$

(d) The posterior distribution is

$$\pi_{Q|\mathbf{X}}(q|\mathbf{x}) = \frac{q^4}{0.006186}$$

64.3 $\pi_{\Lambda|\mathbf{X}}(\lambda|\mathbf{x}) = \frac{4^{13}}{12!}e^{-4\lambda}\lambda^{12}$

64.4 $\pi_{\Lambda|\mathbf{X}}(\lambda|\mathbf{x}) = \frac{\lambda^{10}(0.8e^{-\frac{7\lambda}{6}}+0.6e^{-\frac{13\lambda}{12}})}{0.395536(10!)}$

64.5 0.5572

64.6 0.721

64.7 0.64

64.8 1.90

64.9 $\frac{x+c}{2}$

64.10 $\frac{27}{16}$

64.11 1.9899

64.12 0.622

64.13 0.148

Section 65

65.1 2

65.2 0.45

65.3 0.000398

65.4 450

65.5 1.319

65.6 0.8148

Section 66

66.1 The model distribution is

$$f_{\mathbf{X}|Q}(\mathbf{x}|q) = \prod_{i=1}^{n} \binom{m}{x_i} q^{x_i}(1-q)^{m-x_i}.$$

The joint distribution is

$$f_{\mathbf{X},Q}(\mathbf{x},q) = \prod_{i=1}^{n} \binom{m}{x_i} q^{x_i}(1-q)^{m-x_i}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1}.$$

The marginal distribution is

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) =& \prod_{i=1}^{n} \binom{m}{x_i} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 q^{a+\sum_{i=1}^n x_i-1}(1-q)^{b+nm-\sum_{i=1}^n x_i-1}dq \\ =& \prod_{i=1}^{n} \binom{m}{x_i} \frac{\Gamma(a+b)\Gamma(a+\sum_{i=1}^n x_i)\Gamma(b+nm-\sum_{i=1}^n x_i)}{\Gamma(a)\Gamma(b)\Gamma(a+b+nm)}. \end{aligned}$$

The posterior distribution is

$$\pi_{Q|\mathbf{X}}(q,\mathbf{x}) = \frac{\Gamma(a+b+nm)}{\Gamma(a+\sum_{i=1}^n x_i)\Gamma(b+nm-\sum_{i=1}^n x_i)}q^{a+\sum_{i=1}^n x_i-1}(1-q)^{b+nm-\sum_{i=1}^n x_i-1}.$$

Hence, $Q|\mathbf{X}$ has a beta distribution with parameters $a+\sum_{i=1}^n x_i, b+nm-\sum_{i=1}^n x_i$, and 1.

66.2 The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda)=\frac{\lambda^n e^{-\lambda\sum_{i=1}^n\frac{1}{x_i}}}{\prod_{i=1}^n x_i^2}.$$

The joint distribution is

$$f_{\mathbf{X},\Lambda}(\mathbf{x},\lambda)=\frac{\lambda^{n+\alpha-1}e^{-\lambda\left(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i}\right)}}{\theta^\alpha\Gamma(\alpha)\prod_{i=1}^n x_i^2}.$$

The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x})=\int_0^\infty\frac{\lambda^{n+\alpha-1}e^{-\lambda\left(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i}\right)}}{\theta^\alpha\Gamma(\alpha)\prod_{i=1}^n x_i^2}d\lambda=\frac{\Gamma(n+\alpha)}{\theta^\alpha(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i})^{n+\alpha}\Gamma(\alpha)\prod_{i=1}^n x_i^2}.$$

The posterior distribution is

$$\pi_{\Lambda|\mathbf{X}}(\theta,\mathbf{x})=\frac{\lambda^{n+\alpha-1}e^{-\lambda\left(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i}\right)}}{\left(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i}\right)^{-(n+\alpha)}\Gamma(n+\alpha)}.$$

Hence, $\Lambda|\mathbf{X}$ has a Gamma distribution with parameters $\alpha+n$ and $\left(\frac{1}{\theta}+\sum_{i=1}^n\frac{1}{x_i}\right)^{-1}$

66.3 The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda)=\frac{1}{\lambda^n}e^{-\frac{\sum_{i=1}^n x_i}{\lambda}}.$$

The joint distribution is

$$f_{\mathbf{X},\Lambda}(\mathbf{x},\lambda)=\frac{1}{\lambda^n}e^{-\frac{\sum_{i=1}^n x_i}{\lambda}}\frac{\theta^\alpha e^{-\frac{\theta}{\lambda}}}{\lambda^{\alpha+1}\Gamma(\alpha)}.$$

The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x})=\int_0^\infty\frac{\theta^\alpha e^{-\frac{1}{\lambda}(\sum_{i=1}^n x_i+\theta)}}{\lambda^{\alpha+n+1}\Gamma(\alpha)}d\lambda=\frac{\theta^\alpha\Gamma(\alpha+n)}{(\theta+\sum_{i=1}^n x_i)^{\alpha+n}\Gamma(\alpha)}.$$

The posterior distribution is

$$\pi_{\Lambda|\mathbf{X}}(\lambda,\mathbf{x})=\frac{(\theta+\sum_{i=1}^n x_i)^{\alpha+n}e^{-\frac{1}{\lambda}(\sum_{i=1}^n x_i+\theta)}}{\lambda^{\alpha+n+1}\Gamma(\alpha+n)}.$$

Hence, $\Lambda|\mathbf{X}$ has an inverse Gamma distribution with parameters $\alpha + n$ and $\sum_{i=1}^n x_i + \theta$

66.4 The model distribution is

$$f_{\mathbf{X}|\Lambda}(\mathbf{x}|\lambda) = \frac{1}{\lambda^n}.$$

The joint distribution is

$$f_{\mathbf{X},\Lambda}(\mathbf{x}, \lambda) = \frac{\alpha\theta^\alpha}{\lambda^{n+\alpha+1}}, \ \lambda > M$$

where $M = \max\{x_1, x_2, \cdots, x_n, \theta\}$. The marginal distribution is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_M^\infty \frac{\alpha\theta^\alpha}{\lambda^{n+\alpha+1}} d\lambda = \frac{\alpha\theta^\alpha}{(n+\alpha)M^{n+\alpha}}.$$

The posterior distribution is

$$\pi_{\Lambda|\mathbf{X}}(\lambda, \mathbf{x}) = \frac{(n+\alpha)M^{n+\alpha}}{\lambda^{n+\alpha-1}}.$$

Hence, $\Lambda|\mathbf{X}$ has a single-parameter Pareto distribution with $\alpha' = n + \alpha$ and $\theta' = M$

Section 67

67.1 (a) Let n_k be the number of policies un which frequency of exactly k accidrnts occurred. The Poisson distribution parameter estimate by the method of moments is

$$\hat{\lambda} = \overline{x} = \frac{\sum_{k=1}^6 kn_k}{\sum_{k=0}^6 n_k} = \frac{103}{84} = 1.2262.$$

(b) The likelihood function is

$$L(\lambda) = \prod_{k=0}^6 \left(\frac{e^{-\lambda}\lambda^k}{k!}\right)^{n_k}.$$

The loglikelihood function is

$$\ell(\lambda) = \sum_{k=0}^6 n_k \ln\left(\frac{e^{-\lambda}\lambda^k}{k!}\right) = -\lambda\sum_{k=0}^6 n_k + \sum_{k=1}^6 kn_k \ln\lambda - \sum_{k=0}^6 n_k \ln(k!).$$

The MLE of λ is found from

$$\frac{d}{d\lambda}\ell(\lambda) = -\sum_{k=0}^{6} n_k + \frac{1}{\lambda}\sum_{k=1}^{6} kn_k = 0 \Longrightarrow \hat{\lambda} = \frac{\sum_{k=1}^{6} kn_k}{\sum_{k=0}^{6} n_k} = 1.2262.$$

(c) We have

$$\begin{aligned} E(\hat{\lambda}) =& E(N) = \lambda \\ \mathrm{Var}(\hat{\lambda}) =& \frac{\mathrm{Var}(N)}{n} = \frac{\lambda}{n}, \end{aligned}$$

where $n = \sum_{k=0}^{6} n_k$. Thus, $\hat{\lambda}$ is unbiased and consistent.
(d) The asymptotic variance is found as follows:

$$\begin{aligned} I(\lambda) =& nI(N|\lambda) = -nE\left[\frac{\partial^2}{\partial\lambda^2}\ln\left(\frac{e^{-\lambda}\lambda^N}{N!}\right)\right] \\ =& nE\left(\frac{N}{\lambda^2}\right) = \frac{n}{\lambda} \end{aligned}$$
$$\mathrm{Var}(\hat{\lambda}) = \frac{\lambda}{n}.$$

(e) The confidence interval is

$$\left(1.2262 - 1.96\sqrt{1.2262/84}, 1.2262 + 1.96\sqrt{1.2262/84}\right) = (0.9894, 1.463)$$

67.2 1.6438

67.3 $p_k = \frac{e^{-\lambda}\lambda^k}{k!}$

67.4 $\hat{\beta} = 0.03$ and $\hat{r} = 41.6667$

67.5

$$\begin{aligned} H(\hat{r}) =& 100\ln\left(1 + \frac{1.25}{\hat{r}}\right) - \sum_{k=1}^{4} n_k \sum_{m=0}^{3} (\hat{r} + m)^{-1} \\ =& 100\ln\left(1 + \frac{1.25}{\hat{r}}\right) - \left(\frac{70}{\hat{r}} + \frac{35}{\hat{r}+1} + \frac{15}{\hat{r}+2} + \frac{5}{\hat{r}+3}\right) \end{aligned}$$

67.6 $\hat{\beta} = \frac{\overline{x}}{r}$

67.7 $\hat{q} = 0.1472$ and $\hat{m} = 18.27$

67.8 0.06

Section 68

68.1 We have

$$\hat{p}_0^M = \frac{n_0}{n} = \frac{9048}{10000} = 0.9048$$

and

$$\hat{\beta} = \frac{\sum_{k=0}^{\infty} kn_k}{n - n_0} - 1 = \frac{n\overline{x}}{n - n_0} - 1 = 0.05147$$

68.2 We have

$$\hat{p}_0^M = \frac{n_0}{n} = \frac{9048}{10000} = 0.9048$$

and

$$\overline{x}(1 - e^{-\lambda}) = \frac{n - n_0}{n}\lambda \Longrightarrow 0.1001(1 - e^{-\lambda}) = 0.0952\lambda \Longrightarrow \hat{\lambda} = 0.1012$$

68.3 We have

$$\hat{p}_0^M = \frac{n_0}{n} = \frac{10}{20} = 0.5$$

and

$$\overline{x} = \frac{1 - \hat{p}_0^M}{1 - p_0} mq \Longrightarrow 0.7 = 3q \Longrightarrow \hat{q} = 0.2333$$

68.4 -22.5547

Section 69

69.1 -0.0714

69.2 For $x < 0.3$, we have $F^*(x) > F_n(x)$ so that the fitted distribution is thicker on the left than the empirical distribution. For $x > 0.85$, $F^*(x) > F_n(x)$ which implies $S^*(x) < S_n(x)$. That is, the fitted distribution is thinner on the right than the empirical distribution. Also, note that near the median $x = 0.5$, the slope is less than 1. That is, less probability on the fitted than the empirical. Hence, the answer is (E)

69.3 Let's choose $x_5 = 30$. Than $F(30) \approx 0.6$ but smaller than 0.6. If X is uniform in $[1, 100]$ then its cdf is $F_u(x) = \frac{x-1}{99}$ and $F_u(30) \approx 0.29$ so that (C) is eliminated. If X is exponential with mean 10 then $F_e = 1 - e^{-0.1x}$ and $F_e(30) \approx 0.95$. Thus, (D) is eliminated. If $F(x) = \frac{x}{x+1}$ then $F(30)\frac{30}{31} \approx 0.97$ so that (B) is eliminated. IF X is normal with mean 40 and standard deviation 40 then $F_n(30) = \Phi\left(\frac{30-40}{40}\right) = \Phi(-0.25) = 0.40$ so that (E) is eliminated. Note that with the function in (A) we have $F(30) = 1 - 30^{-0.25} \approx 0.57$. Hence, the answer is (A)

Section 70

70.1 0.2727

70.2 0.4025

70.3

Level of comfidence	0.10	0.05	0.025	0.01
Critical value	0.5456	0.6082	0.6619	0.729
Test Result	Reject	Reject	Reject	Fail to reject

70.4 0.1679

70.5 Fail to reject the null hypothesis

70.6

Level of comfidence	0.10	0.05	0.02	0.01
Critical value	0.0863	0.0962	0.1075	0.1152
Test Result	Reject	Reject	Reject	Fail to reject

Section 71

71.1 0.252

71.2

Level of confidence α	0.10	0.05	0.01
Critical value	1.933	2.492	3.857
Test Result	Fail to reject	Fail to reject	Fail to reject

71.3 (A) Using sample data gives a better than expected fit and therefore a test statistic that favors the null hypothesis, thus increasing the Type II error probability.
(B) The K-S test works only on individual data and so B is false.
(C) The Anderson-Darling test emphasizes the tails so that (C) is false.
Hence, the answer is (D)

71.4 (A), (B), and (C) are all correct. Thus, the answer is (D)

71.5 (A)

Section 72

72.1 9.151

72.2 (a) $\hat{\lambda} = 1.6438$
(b) The Chi-square statistic is

$$\chi^2 = \frac{421.4809}{70.53} + \frac{36.7236}{115.94} + \frac{32.49}{95.30} + \frac{76.9129}{83.23} = 7.558.$$

We have

Level of Significance	$\chi^2_{k-r-1,1-\alpha}$	Test result
10%	4.605	Reject
5%	5.991	Reject
2.5$	7.378	Reject
1%	9.210	Do not reject

where the degrees of freedom is $k - r - 1 = 4 - 1 - 1 = 2$

72.3 The Chi-square statistic is

$$\chi^2 = \frac{225}{250} + \frac{225}{350} + \frac{100}{240} + \frac{1}{110} + \frac{49}{40} + \frac{144}{10} = 17.594.$$

We have

Level of Significance	$\chi^2_{k-r-1,1-\alpha}$	Test result
10%	9.236	Reject
5%	11.070	Reject
2.5$	12.833	Reject
1%	15.086	Reject

The degrees of freedom is $6 - 1 = 5$

72.4 $\chi^2 = 6.659$

72.5 A is false. Using sample data gives a better than expected fit and therefore a test statistic that favors the null hypothesis, thus increasing the Type II error probability. The K-S test works only on individual data and so B is false. The A-D test emphasizes the tails, thus C is false. D is false because the critical value depends on the degrees of freedom which in turn depends on the number of cells, not the sample size. So the answer is (E)

72.6 (A)

Section 73

73.1 We have 0 degrees of freedom in the null hypothesis, since both parameters are specified, and 2 degrees of freedom in the alternative hypothesis, since both parameters are freely chosen in maximizing $L(\alpha, \theta)$. We thus have 2 degrees of freedom overall

73.2

α	5%	2.5%	1%	0.5%
c_α	5.991	7.378	9.210	10.597

73.3

α	5%	2.5%	1%	0.5%
c_α	5.991	7.378	9.210	10.597
Test Result	Reject	Reject	Do not reject	Do not reject

73.4

α	10%	5%	2.5%	1%
c_α	2.706	3.841	5.024	6.635
Test Result	Reject	Reject	Do not reject	Do not reject

73.5 7

Section 74

74.1 (i)

74.2 (A)

74.3 (I)

74.4 Generalized Pareto

Section 75

75.1 16,913

75.2 2,381

75.3 0.10

75.4 $384.16\left(\frac{\alpha+1}{\alpha}\right)$

75.5 960

Section 76

76.1 0.47

76.2 (E)

76.3 0.3723

76.4 5,446,250

76.5 138

76.6 0.8

Section 77

77.1 For the risk parameter, we have

$$\pi(\lambda) = \frac{\lambda^{\alpha-1}e^{-\frac{\lambda}{\beta}}}{\beta^\alpha\Gamma(\alpha)}.$$

For the claims, we have

$$f_{X|\Lambda}(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

77.2 For the risk parameter, we have

$$\pi(\theta) = \frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(\theta-\mu)^2}{2\sigma_2^2}}.$$

For the claims, we have

$$f_{X|\Theta}(x|\theta) = \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2\sigma_1^2}}$$

77.3 For the risk parameter, we have

$$\pi(q) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1},\ 0 < q < 1.$$

For the claims, we have

$$f_{X|\mathbf{Q}}(x|q) = \left(\begin{array}{c} m \\ x \end{array}\right) q^x(1-q)^{m-x},\ x = 0, 1, \cdots, m$$

77.4 For the risk parameter, we have

$$\pi(\lambda) = \frac{\theta^\alpha e^{-\frac{\theta}{\lambda}}}{\lambda^{\alpha+1}\Gamma(\alpha)}.$$

For the claims, we have

$$f_{X|\Lambda}(x|\lambda) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}$$

77.5 For the risk parameter, we have

$$\pi(\lambda) = \frac{\alpha\theta^\alpha}{\lambda^{\alpha+1}},\ \lambda > \theta.$$

For the claims, we have

$$f_{X|\Lambda}(x|\lambda) = \frac{1}{\lambda},\ 0 \le x \le \lambda$$

Section 78

78.1 (a) $f_{XY}(x,y) = f_X(x)f_{Y|X}(y|x) = \frac{1}{x}, 0 < x < 1, 1-x < y < 1$ (b) $\frac{1+\ln 2}{2}$

78.2 $f_{X|Y}(x|y) = \frac{6x(2-x-y)}{4-3y}$

78.3 $e^{-\frac{1}{y}}$

78.4 (a) Observe that X only takes positive values, thus $f_X(x) = 0,\ x \le 0$. For $0 < x < 1$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_1^\infty f_{XY}(x,y)dy = \frac{\alpha-1}{\alpha}$$

For $x \ge 1$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_x^\infty f_{XY}(x,y)dy = \frac{\alpha-1}{\alpha x^\alpha}$$

(b) For $0 < x < 1$ we have

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\alpha}{y^{\alpha+1}},\quad y > 1.$$

Hence,

$$E(Y|X=x) = \int_1^\infty \frac{y\alpha}{y^{\alpha+1}}dy = \alpha\int_1^\infty \frac{dy}{y^\alpha} = \frac{\alpha}{\alpha-1}.$$

If $x \ge 1$ then

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\alpha x^\alpha}{y^{\alpha+1}},\quad y > x.$$

Hence,

$$E(Y|X=x) = \int_x^\infty y\frac{\alpha x^\alpha}{y^{\alpha+1}}dy = \frac{\alpha x}{\alpha-1}$$

78.5 $E(X|Y=y) = \frac{2}{3}y$ and $E(Y|X=x) = \frac{2}{3}\left(\frac{1-y^3}{1-y^2}\right)$

78.6 The marginal density functions are

$$f_X(x) = \int_{x^2}^1 \frac{21}{4}x^2 y dy = \frac{21}{8}x^2(1-x^4), \quad -1 < x < 1$$

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2 y dx = \frac{7}{2}y^{\frac{5}{2}}, \quad 0 < y < 1.$$

Thus, a first way for finding $E(Y)$ is

$$E(Y) = \int_0^1 y\frac{7}{2}y^{\frac{5}{2}}dy = \int_0^1 y\frac{7}{2}y^{\frac{7}{2}}dy = \frac{7}{9}.$$

For the second way, we use the double expectation result

$$E(Y) = E(E(Y|X)) = \int_{-1}^1 E(Y|X)f_X(x)dx = \int_{-1}^1 \frac{2}{3}\left(\frac{1-x^6}{1-x^4}\right)\frac{21}{8}x^2(1-x^6) = \frac{7}{9}$$

78.7 $\frac{1}{12}$

78.8 $\text{Var}(Y) = 13$

78.9 We have $E(X) = \lambda = \text{Var}(X)$, $E(Y|X=x) = \alpha x$, and $\text{Var}(Y|X = x) = \beta^2 x^2$. Thus,

$$E(Y) = E[E(Y|X)] = E(\alpha X) = \alpha E(X) = \alpha\lambda$$

and

$$\begin{aligned}
\text{Var}(Y) =& E[\text{Var}(Y|X)] + \text{Var}[E(Y|X)] \\
=& E(\beta^2 X^2) + \text{Var}(\alpha X) \\
=& \beta^2 E(X^2) + \alpha^2 \text{Var}(X) \\
=& \beta^2(\lambda + \lambda^2) + \alpha^2\lambda
\end{aligned}$$

Section 79

79.1 (a) The prior distribution is

$$\begin{aligned}\pi(G) =&0.70\\ \pi(A) =&0.20\\ \pi(B) =& 0.10.\end{aligned}$$

(b) The model distribution is

$$\begin{aligned}f_{\mathbf{X}|\Theta}(\mathbf{x}|G) =&(0.25)(0.10) = 0.025\\ f_{\mathbf{X}|\Theta}(\mathbf{x}|A) =&(0.40)(0.20) = 0.08\\ f_{\mathbf{X}|\Theta}(\mathbf{x}|B) =&(0.30)(0.20) = 0.06\end{aligned}$$

79.2 (a) The marginal probability is

$$\begin{aligned}f_{\mathbf{X}}(1,2) =&\sum_{\theta} f_{X_1|\Theta}(1|\theta)f_{X_2|\Theta}(2|\theta)\pi(\theta)\\ =&(0.25)(0.10)(0.7) + (0.4)(0.2)(0.2) + (0.3)(0.2)(0.10)\\ =&0.0395.\end{aligned}$$

(b) The joint distribution is

$$\begin{aligned}f_{\mathbf{X},X_3}(1,2,0) =&\sum_{\theta} f_{X_1|\Theta}(1|\theta)f_{X_2|\Theta}(2|\theta)f_{X_3|\Theta}(0|\theta)\pi(\theta)\\ =&(0.25)(0.10)(0.65)(0.70) + (0.4)(0.2)(0.4)(0.2) + (0.3)(0.2)(0.5)(0.10)\\ =&0.020775\\ f_{\mathbf{X},X_3}(1,2,1) =&\sum_{\theta} f_{X_1|\Theta}(1|\theta)f_{X_2|\Theta}(2|\theta)f_{X_3|\Theta}(1|\theta)\pi(\theta)\\ =&(0.25)(0.10)(0.25)(0.70) + (0.4)(0.2)(0.4)(0.2) + (0.3)(0.2)(0.3)(0.10)\\ =&0.012575\\ f_{\mathbf{X},X_3}(1,2,2) =&\sum_{\theta} f_{X_1|\Theta}(1|\theta)f_{X_2|\Theta}(2|\theta)f_{X_3|\Theta}(2|\theta)\pi(\theta)\\ =&(0.25)(0.10)(0.10)(0.70) + (0.4)(0.2)(0.2)(0.2) + (0.3)(0.2)(0.2)(0.10)\\ =&0.00615\end{aligned}$$

79.3 (a) The predictive distribution is

$$\begin{aligned}
f_{X_3|\mathbf{X}}(0|1,2) =& \frac{0.020775}{0.0395} = 0.5259\\
f_{X_3|\mathbf{X}}(1|1,2) =& \frac{0.012575}{0.0395} = 0.3183\\
f_{X_3|\mathbf{X}}(2|1,2) =& \frac{0.00615}{0.0395} = 0.1557.
\end{aligned}$$

(b) The posterior probabilities are

$$\begin{aligned}
\pi(G|1,2) =& \frac{f(1|G)f(2|G)\pi(G)}{f(1,2)} = \frac{(0.25)(0.10)(0.70)}{0.0395} = 0.4430\\
\pi(A|1,2) =& \frac{f(1|A)f(2|A)\pi(A)}{f(1,2)} = \frac{(0.40)(0.20)(0.20)}{0.0395} = 0.4051\\
\pi(B|1,2) =& \frac{f(1|B)f(2|B)\pi(B)}{f(1,2)} = \frac{(0.30)(0.20)(0.10)}{0.0395} = 0.1519
\end{aligned}$$

79.4 (a) The hypothetical means are

$$\begin{aligned}
\mu_3(G) =& 0(0.65) + 1(0.25) + 2(0.10) = 0.45\\
\mu_3(A) =& 0(0.40) + 1(0.40) + 2(0.20) = 0.80\\
\mu_3(B) =& 0(0.50) + 1(0.30) + 2(0.20) = 0.70.
\end{aligned}$$

(b) The pure premium is

$$\mu_3 = E(X_3) = 0.45(0.70) + 0.80(0.20) + 0.70(0.10) = 0.545$$

79.5 (a) Without using the hypothetical means, we have

$$E(X_3|\mathbf{X}] = 0(0.5259) + 1(0.3183) + 2(0.1557) = 0.6297.$$

(b) The Bayesian premium, using hypothetical means is

$$E(X_3|\mathbf{X}] = (0.45)(0.4430) + (0.80)(0.4051) + (0.70)(0.1519) = 0.62976$$

79.6 0.158

79.7 3.83

79.8 0.6794

79.9 7.202

79.10 10,322

79.11 0.278

Section 80

80.1 (a) The model distribution is

$$f(x|\lambda) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}.$$

The prior distribution is

$$\pi(\lambda) = \frac{225e^{-\frac{15}{\lambda}}}{\lambda^3},\ \lambda > 0.$$

(b) The joint density of x and λ is

$$f(x,\lambda) = f(x|\lambda)\pi(\lambda) = \frac{225e^{-\frac{(x+15)}{\lambda}}}{\lambda^4},\ \lambda > 0.$$

For $x = 12$, the joint density is

$$f(12,\lambda) = \frac{225e^{-\frac{27}{\lambda}}}{\lambda^4},\ \lambda > 0.$$

The marginal density of x is

$$f(x) = \int_0^\infty \frac{225e^{-\frac{(x+15)}{\lambda}}}{\lambda^4}d\lambda$$

and

$$\begin{aligned}
f(12) =& \int_0^\infty \frac{225e^{-\frac{27}{\lambda}}}{\lambda^4} d\lambda \\
=& \frac{225}{27^3}\Gamma(3) \underbrace{\int_0^\infty \frac{27^3 e^{-\frac{27}{\lambda}}}{\lambda^4 \Gamma(3)} d\lambda}_{1} \\
=& \frac{450}{(27^3)}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
f(12, x_2) =& \int_0^\infty \frac{1}{\lambda^2} e^{-\frac{(12+x_2)}{\lambda}} \frac{225e^{-\frac{15}{\lambda}}}{\lambda^3} d\lambda \\
=& 225 \int_0^\infty \frac{1}{\lambda^5} e^{-\frac{(27+x_2)}{\lambda}} d\lambda \\
=& \frac{225}{(27+x_2)^4}\Gamma(4) \underbrace{\int_0^\infty \frac{(27+x_2)^4 e^{-\frac{(27+x_2)}{\lambda}}}{\lambda^5 \Gamma(4)} d\lambda}_{1} \\
=& \frac{1350}{(27+x_2)^4}.
\end{aligned}$$

The predictive distribution is

$$f(x_2|12) = \frac{\frac{1350}{(27+x_2)^4}}{\frac{450}{(27^3)}} = \frac{3(27^3)}{(27+x_2)^4}$$

which is a type 2 Pareto distribution with parameters $\alpha = 3$ and $\theta = 27$.
(c) The posterior distribution of Λ is

$$\pi(\lambda|12) = \frac{27^3}{2} \frac{e^{-\frac{27}{\lambda}}}{\lambda^4}.$$

(d) $E(X_2|12) = 13.5$

80.2 3.25

80.3 The posterior distribution of P is

$$\pi(p|4) = \frac{2\left(\frac{10!}{4!6!}\right) p^5(1-p)^6}{2\frac{\Gamma(10)\Gamma(7)}{\Gamma(4)\Gamma(13)}} = \frac{\Gamma(13)}{\Gamma(6)\Gamma(7)} p^5(1-p)^6$$

which is a beta distribution with $a = 6, b = 7$, and $\theta = 1$

80.4 (i) Letting $q = \frac{1}{1+\beta}$ we can write

$$\begin{aligned}
p_k =& \frac{r(r+1)\cdots(r+k-1)}{k!}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^k\\
=&\frac{(r-1)!r(r+1)\cdots(r+k-1)}{k!(r-1)!}\left(\frac{1}{1+\beta}\right)^r\left(1-\frac{1}{1+\beta}\right)^k\\
=&\frac{\Gamma(r+k)}{\Gamma(r)\Gamma(k+1)}q^r(1-q)^k.
\end{aligned}$$

(ii) The model distribution is

$$f(x|q) = \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}q^r(1-q)^x.$$

The prior distribution is

$$\pi(q) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1}.$$

The joint distribution of X and Q is

$$\begin{aligned}
f(x,q) =& f(x|q)\pi(q) = \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}q^r(1-q)^x\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1}\\
=&\frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a+r-1}(1-q)^{b+k-1}.
\end{aligned}$$

The marginal distribution is

$$\begin{aligned}
f(x) =& \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 q^{a+r-1}(1-q)^{b+k-1}dq\\
=&\frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+r)\Gamma(b+k)}{\Gamma(a+b+k+r)}\underbrace{\int_0^1\frac{\Gamma(a+b+k+r)}{\Gamma(a+r)\Gamma(b+k)}q^{a+r-1}(1-q)^{b+k-1}dq}_{1}\\
=&\frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+r)\Gamma(b+k)}{\Gamma(a+b+k+r)}.
\end{aligned}$$

The posterior distribution is

$$\begin{aligned}
\pi(q|x) =& \frac{\frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a+r-1}(1-q)^{b+k-1}}{\frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+r)\Gamma(b+k)}{\Gamma(a+b+k+r)}}\\
=&\frac{\Gamma(a+b+k+r)}{\Gamma(a+r)\Gamma(b+k)}q^{a+r-1}(1-q)^{b+k-1}
\end{aligned}$$

which is a beta distribution with $a' = a + r$ and $b' = b + k$

80.6 15

Section 81

81.1 The unbiasedness equation yields

$$\hat{\alpha}_0 + \mu \sum_{j=1}^{n} \hat{\alpha}_j = \mu$$

which implies

$$\sum_{j=1}^{n} \hat{\alpha}_i = 1 - \frac{\hat{\alpha}_0}{2}$$

81.2 For $i = 1, 2, \cdots, n$, we have

$$\sum_{\substack{j = 1 \\ j \neq i}}^{n} \hat{\alpha}_j \rho \sigma^2 + \sigma^2 \hat{\alpha}_i = \rho \sigma^2$$

or equivalently

$$\sum_{j=1}^{n} \hat{\alpha}_j + \hat{\alpha}_i (1 - \rho) = \rho$$

81.3 Problem 81.2 followed by Problem 81.1, we find

$$\hat{\alpha}_i = \frac{\rho(1 - \sum_{j=1}^{n} \hat{\alpha}_j)}{1 - \rho} = \frac{\rho \hat{\alpha}_0}{\mu(1 - \rho)}$$

81.4 From Problem 81.3, we find

$$\sum_{j=1}^{n} \hat{\alpha}_j = \frac{n \rho \hat{\alpha}_0}{\mu(1 - \rho)}.$$

This combined with Problem 81.1 yield the equation

$$1 - \frac{\hat{\alpha}_0}{\mu} = \frac{n \rho \hat{\alpha}_0}{\mu(1 - \rho)}.$$

Solving this equation, we find

$$\hat{\alpha}_0 = \frac{(1-\rho)\mu}{1-\rho+n\rho}.$$

Plugging this into Problem 81.3, we find

$$\hat{\alpha}_i = \frac{\rho}{1-\rho+n\rho}$$

81.5 The credibility premium is

$$\begin{aligned}\hat{\alpha}_0 + \sum_{j=1}^{n}\hat{\alpha}_j X_j =& \frac{(1-\rho)\mu}{1-\rho+n\rho} + \sum_{j=1}^{n}\frac{\rho X_j}{1-\rho+n\rho}\\ =&(1-Z)\mu + Z\overline{X}\end{aligned}$$

where

$$Z = \frac{n\rho}{1-\rho+n\rho}$$

Section 82

82.1 10,622

82.2 0.85651

82.3 0.22

82.4 1063.47

82.5 3

82.6 14

82.7 $\frac{\theta-\theta^2}{1.5-\theta^2}$

82.8 8.33

82.9 $\frac{1}{9}$

82.10 3.27

Section 83

83.1 $\frac{n\beta}{n\beta+1}\overline{X} + \frac{1}{n\beta+1}(\alpha\beta)$

83.2 0.9375

83.3 0.905

83.4 1

83.5 0.93

83.6 8.69

83.7 0.428

83.8 0.8

Section 84

84.1 (a) We have

$$\begin{aligned}\operatorname{Var}\left(\left.\frac{m_iX_i+m_jX_j}{m_i+m_j}\right|\Theta\right) =&\left(\frac{m_i}{m_i+m_j}\right)^2\operatorname{Var}(X_i|\Theta)+\left(\frac{m_j}{m_i+m_j}\right)^2\operatorname{Var}(X_j|\Theta)\\ =&\left(\frac{m_i}{m_i+m_j}\right)^2\left(w(\theta)+\frac{v(\theta)}{m_i}\right)+\left(\frac{m_j}{m_i+m_j}\right)^2\left(w(\theta)+\frac{v(\theta)}{m_j}\right)\\ =&\frac{m_i^2+m_j^2}{(m_i+m_j)^2}w(\theta)+\frac{v(\theta)}{m_i+m_j}.\end{aligned}$$

(b) We have

$$\begin{aligned}
E(X_i) =& E[E(X_i|\Theta)] = E[\mu(\Theta)] = \mu \\
\mathrm{Cov}(X_i, X_j) =& E(X_i, X_j) - E(X_i)E(X_j) \\
=& E[E(X_iX_j|\Theta)] - E[\mu(\Theta)]^2 \\
=& E[E(X_i|\Theta)E(X_j|\Theta)] - E[\mu(\Theta)]^2 \quad \text{(by independence)} \\
=& E[\mu^2(\Theta)] - E[\mu(\Theta)]^2 \\
=& \mathrm{Var}[\mu(\Theta)] = a \\
\mathrm{Var}(X_i) =& E[\mathrm{Var}(X_i|\Theta)] + \mathrm{Var}[E(X_i|\Theta)] \\
=& E\left[w(\theta) + \frac{v(\Theta)}{m_i}\right] + \mathrm{Var}[\mu(\Theta)] \\
=& w + \frac{v}{m_i} + a
\end{aligned}$$

84.2 The unbiasedness equation is

$$E(X_{n+1}) = \mu = \hat{\alpha}_0 + \sum_{i=1}^{n} \hat{\alpha}_i \mu \Longrightarrow \sum_{i=1}^{n} \hat{\alpha}_i = 1 - \frac{\hat{\alpha}_0}{\mu}.$$

For $i = 1, 2, \cdots, n$, (81.4) becomes

$$\begin{aligned}
a =& \sum_{\substack{j=1 \\ j \neq i}}^{n} \hat{\alpha}_j a + \hat{\alpha}_i \left(a + \frac{v}{m_i} + w\right) \\
=& \sum_{j=1}^{n} \hat{\alpha}_j a + \hat{\alpha}_i \left(w + \frac{v}{m_i}\right) \\
=& a\left(1 - \frac{\hat{\alpha}_0}{\mu}\right) + \hat{\alpha}_i \left(w + \frac{v}{m_i}\right).
\end{aligned}$$

Solving this equation for $\hat{\alpha}_i$, we find

$$\hat{\alpha}_i = \frac{a\hat{\alpha}_0/\mu}{w + v/m_i}.$$

Summing both sides from 1 to n, we find

$$\frac{a\hat{\alpha}_0}{\mu} \sum_{j=1}^{n} \frac{m_j}{v + wm_j} = 1 - \frac{\hat{\alpha}_0}{\mu}.$$

Solving this equation, we find

$$\hat{\alpha}_0 = \frac{1}{(a/\mu)\sum_{j=1}^{n}\frac{m_j}{v+wm_j}+\frac{1}{\mu}} = \hat{\alpha}_0 = \frac{\mu}{1+am^*}$$

84.3 We have

$$\begin{aligned}\hat{\alpha}_0+\sum_{j=1}^{n}\hat{\alpha}_jX_j =&\frac{\mu}{1+am^*}+\sum_{j=1}^{n}\frac{am_j}{v+wm_j}\frac{1}{1+am^*}X_j\\ =&\frac{\mu}{1+am^*}+\frac{a}{1+am^*}\sum_{j=1}^{n}\frac{m_j}{v+wm_j}X_j\\ =&Z\overline{X}+(1-Z)\mu.\end{aligned}$$

84.4 2.4

84.5 12

84.6 11.13

84.7 257.11

84.8 $\frac{4}{3}$

84.9 (A) is false. This is true for Bhlmann. The Bhlmann-Straub model allows the variation in size and exposure.
(B) is false. The model is valid for any type of distributions.
(C) is false. There is no cap on the number of exposure.
Thus, the answer to the problem is (E)

84.10 $\frac{n}{n+\frac{w}{a}}$

Section 85

85.1 Problem 66.3 shows that the posterior distribution is an inverse Gamma distribution with parameters $\alpha' = \alpha + n$ and $\theta' = \sum_{i=1}^{n} x_i + \theta$. The hypothetical mean is

$$\mu(\lambda) = E(X_i|\Lambda) = \lambda$$

and the Bayesian premium is

$$\begin{aligned}
E(X_{n+1}|\mathbf{X}) =& \int_0^\infty \lambda \frac{(\theta + \sum_{i=1}^n x_i)^{\alpha+n} e^{-\frac{1}{\lambda}(\sum_{i=1}^n x_i + \theta)}}{\lambda^{\alpha+n+1}\Gamma(\alpha+n)} d\lambda \\
=& \frac{(n\overline{X}+\theta)}{\alpha+n-1} \int_0^\infty \frac{(\theta + \sum_{i=1}^n x_i)^{\alpha+n-1} e^{-\frac{1}{\lambda}(\sum_{i=1}^n x_i + \theta)}}{\lambda^{\alpha+n-1+1}\Gamma(\alpha+n-1)} d\lambda \\
=& \frac{(n\overline{X}+\theta)}{\alpha+n-1}.
\end{aligned}$$

Note that the Bayesian premium is a linear function of $X_1, X_2, \cdots, X_n$. Next, we find the Bühlmann credibility. We have

$$\begin{aligned}
\mu(\lambda) =& E(X_i|\Lambda) = \lambda \\
\mu =& E(\Lambda) = \frac{\theta}{\alpha - 1} \\
v(\lambda) =& \text{Var}(X_i|\Lambda) = \lambda^2 v = \qquad E(\Lambda^2) = \frac{\theta^2}{(\alpha-1)(\alpha-2)} \\
a =& \text{Var}(\Lambda) = \frac{\theta^2}{(\alpha-1)^2(\alpha-2)} \\
k =& \frac{v}{a} = \alpha - 1 \\
Z =& \frac{n}{n+k} = \frac{n}{n+\alpha-1} \\
P_c =& Z\overline{X} + (1-Z)\mu \\
=& \frac{n}{n+\alpha-1}\overline{X} + \frac{\alpha-1}{n+\alpha-1}\frac{\theta}{\alpha-1} \\
=& (n\overline{X}+\theta)n + \alpha - 1.
\end{aligned}$$

Thus, the Bühlmann credibility premium equals the Bayesian premium

85.2 Problem 66.1 shows that the posterior distribution has a beta distribution with parameters $a' = a + \sum_{i=1}^n x_i, b' = b + nm - \sum_{i=1}^n x_i$. The hypothetical mean is

$$\mu(q) = E(X_i|Q) = mq$$

and the Bayesian premium is

$$
\begin{aligned}
E(X_{n+1}|\mathbf{X}) =& \int_0^\infty mq \frac{\Gamma(a+b+nm)}{\Gamma(a+\sum_{i=1}^n x_i)\Gamma(b+nm-\sum_{i=1}^n x_i)} q^{a+\sum_{i=1}^n x_i-1}(1-q)^{b+nm-\sum_{i=1}^n x_i-1}dq\\
=& \frac{m(a+n\overline{x})}{a+b+mn}\int_0^\infty \frac{\Gamma(a+b+nm+1)}{\Gamma(a+\sum_{i=1}^n x_i+1)\Gamma(b+nm-\sum_{i=1}^n x_i)} q^{a+\sum_{i=1}^n x_i+1-1}(1-q)^{b+nm-\sum_{i=1}^n x_i-1}\\
=& \frac{m(a+n\overline{x})}{a+b+mn}.
\end{aligned}
$$

Note that the Bayesian premium is a linear function of $X_1, X_2, \cdots, X_n$. Next, we find the Bühlmann credibility. We have

$$
\begin{aligned}
\mu(q) =& E(X_i|Q) = mq\\
\mu =& E(mQ) = mE(Q) = \frac{ma}{a+b}\\
v(q) =& \mathrm{Var}(X_i|Q) = mq(1-q)\\
v =& E[mQ(1-Q)] = \frac{mab}{(a+b)(a+b+1)}\\
a =& \mathrm{Var}[mQ] = m^2\mathrm{Var}(Q) = \frac{m^2ab}{(a+b)^2(a+b+1)}\\
k =& \frac{v}{a} = \frac{a+b}{m}\\
Z =& \frac{n}{n+k} = \frac{nm}{nm+a+b}\\
P_c =& Z\overline{X} + (1-Z)\mu\\
=& \frac{m(a+n\overline{x})}{a+b+nm}.
\end{aligned}
$$

Thus, the Bühlmann credibility premium equals the Bayesian premium

85.3 By Example 66.3 the posterior distribution has normal distribution with mean $\left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}$ and variance $\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}$. The hypothetical mean is

$$\mu(\lambda) = E(X_i|\Lambda) = \lambda$$

and the Bayesian premium is

$$E(X_{n+1}|\mathbf{X}) = E[\Lambda|\mathbf{X}] = \left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}.$$

Note that the Bayesian premium is a linear function of $X_1, X_2, \cdots, X_n$. Next, we find the Bühlmann credibility. We have

$$\begin{aligned}
\mu(\lambda) =&\lambda \\
\mu =&E(\Lambda) \\
v(\lambda) =&\mathrm{Var}(X_i|\Lambda) = \sigma^2 \\
v =&E(\sigma^2) = \sigma^2 \\
\mathrm{Var}(\Lambda) =&a^2 \\
k =&\frac{v}{a} = \frac{sigma^2}{a^2} \\
Z =&\frac{n}{n+k} = \frac{na^2}{na^2+\sigma^2} \\
P_c =&Z\overline{X} + (1-Z)\mu \\
=&\frac{na^2\overline{x}}{na^2+\sigma^2} + \frac{\mu\sigma^2}{na^2+\sigma^2} \\
=&\left(\frac{\sum x_i}{\sigma^2} + \frac{\mu}{a^2}\right)\left(\frac{n}{\sigma^2} + \frac{1}{a^2}\right)^{-1}.
\end{aligned}$$

Thus, the Bühlmann credibility premium equals the Bayesian premium

85.4 0.0182

Section 86

86.1 0.818

86.2 687.375

86.3 0.78

86.4 0.8718

Section 87

87.1 0.3682

87.2 0.4987

87.3 0.852

87.4 1.351

87.5 98.26

87.6 0.323

87.7 7.56

Section 88

88.1 $a = \mu - v - \mu^2$

88.2 0.221

88.3 0.3928

88.4 0.6333

88.5 0.5747

88.6 0.2659

88.7 0.023209

Section 89

89.1 1000

89.2 1

89.3 (D)

89.4 2212.76

89.5 3477.81

89.6 7

Section 90

90.1 522.13

90.2 228,503

90.3 224.44

90.4 88.75

90.5 41.897

90.6 35.7

90.7 630.79

Section 91

91.1 We have the following sequence of calculation

$$\begin{aligned}
k =&[pn]+1=[3]+1=4\\
\widehat{\mathrm{VaR}}_p(X) =&123\\
\widehat{\mathrm{TVaR}}_p(X) =&\frac{1}{n-k+1}\sum_{i=k}^{n}y_i\\
=&\frac{1}{10-4+1}(123+150+153+189+190+195+200)=171.43\\
s_p^2 =&\frac{1}{n-k}\sum_{i=k}^{n}(y_i-\widehat{\mathrm{TVaR}}_p(X))^2\\
=&\frac{1}{10-4}[(123-171.43)^2+(150-171.43)^2+(153-171.43)^2+(189-171.43)^2+\\
\widehat{\mathrm{Var}}(\widehat{\mathrm{TVaR}}_p(X)) =&\frac{s_p^2+p[\widehat{\mathrm{TVaR}}_p(X)-\widehat{\mathrm{VaR}}_p(X)]^2}{n-k+1}\\
=&\frac{861.61+0.3(123-171.43)^2}{10-4+1}=223.52
\end{aligned}$$

Section 92

92.1 1

92.2 (A)

92.3 $\frac{44}{9}$

92.4 0.0131

92.5 $21\theta^4$

Exam C Tables

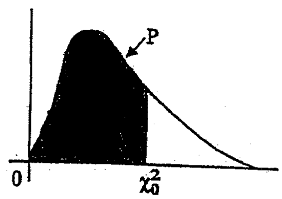
NORMAL DISTRIBUTION TABLE

Entries represent the area under the standardized normal distribution from -∞ to z, Pr(Z<z)
The value of z to the first decimal is given in the left column. The second decimal place is given in the top row.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Values of z for selected values of Pr(Z<z)							
z	0.842	1.036	1.282	1.645	1.960	2.326	2.576
Pr(Z<z)	0.800	0.850	0.900	0.950	0.975	0.990	0.995

χ^2 Distribution



The table below gives the value x_0^2 for which $P[x^2 < x_0^2] = P$ for a given number of degrees of freedom and a given value of P.

Degrees of Freedom	Values of P									
	0.005	**0.010**	**0.025**	**0.050**	**0.100**	**0.900**	**0.950**	**0.975**	**0.990**	**0.995**
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.01	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997

Appendix A

An Inventory of Continuous Distributions

A.1 Introduction

The incomplete gamma function is given by

$$\Gamma(\alpha; x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t}\, dt, \quad \alpha > 0,\ x > 0$$

$$\text{with } \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt, \quad \alpha > 0.$$

Also, define

$$G(\alpha; x) = \int_x^\infty t^{\alpha-1} e^{-t}\, dt, \quad x > 0.$$

At times we will need this integral for nonpositive values of α. Integration by parts produces the relationship

$$G(\alpha; x) = -\frac{x^\alpha e^{-x}}{\alpha} + \frac{1}{\alpha} G(\alpha + 1; x)$$

This can be repeated until the first argument of G is $\alpha + k$, a positive number. Then it can be evaluated from

$$G(\alpha + k; x) = \Gamma(\alpha + k)[1 - \Gamma(\alpha + k; x)].$$

The incomplete beta function is given by

$$\beta(a, b; x) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1}(1 - t)^{b-1}\, dt, \quad a > 0,\ b > 0,\ 0 < x < 1.$$

A.2 Transformed beta family

A.2.2 Three-parameter distributions

A.2.2.1 Generalized Pareto (beta of the second kind)—α, θ, τ

$$\begin{aligned}
f(x) &= \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\theta^\alpha x^{\tau-1}}{(x+\theta)^{\alpha+\tau}} \qquad F(x) = \beta(\tau, \alpha; u), \quad u = \frac{x}{x+\theta} \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\tau+k)\Gamma(\alpha-k)}{\Gamma(\alpha)\Gamma(\tau)}, \quad -\tau < k < \alpha \\
\mathrm{E}[X^k] &= \frac{\theta^k \tau(\tau+1)\cdots(\tau+k-1)}{(\alpha-1)\cdots(\alpha-k)}, \quad \text{if } k \text{ is an integer} \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\tau+k)\Gamma(\alpha-k)}{\Gamma(\alpha)\Gamma(\tau)} \beta(\tau+k, \alpha-k; u) + x^k[1-F(x)], \quad k > -\tau \\
\text{mode} &= \theta \frac{\tau-1}{\alpha+1}, \quad \tau > 1, \text{ else } 0
\end{aligned}$$

A.2.2.2 Burr (Burr Type XII, Singh-Maddala)—α, θ, γ

$$\begin{aligned}
f(x) &= \frac{\alpha\gamma(x/\theta)^\gamma}{x[1+(x/\theta)^\gamma]^{\alpha+1}} \qquad F(x) = 1-u^\alpha, \quad u = \frac{1}{1+(x/\theta)^\gamma} \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(1+k/\gamma)\Gamma(\alpha-k/\gamma)}{\Gamma(\alpha)}, \quad -\gamma < k < \alpha\gamma \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha}-1]^{1/\gamma} \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(1+k/\gamma)\Gamma(\alpha-k/\gamma)}{\Gamma(\alpha)} \beta(1+k/\gamma, \alpha-k/\gamma; 1-u) + x^k u^\alpha, \quad k > -\gamma \\
\text{mode} &= \theta \left(\frac{\gamma-1}{\alpha\gamma+1}\right)^{1/\gamma}, \quad \gamma > 1, \text{ else } 0
\end{aligned}$$

A.2.2.3 Inverse Burr (Dagum)—τ, θ, γ

$$\begin{aligned}
f(x) &= \frac{\tau\gamma(x/\theta)^{\gamma\tau}}{x[1+(x/\theta)^\gamma]^{\tau+1}} \qquad F(x) = u^\tau, \quad u = \frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma} \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\tau+k/\gamma)\Gamma(1-k/\gamma)}{\Gamma(\tau)}, \quad -\tau\gamma < k < \gamma \\
\mathrm{VaR}_p(X) &= \theta(p^{-1/\tau}-1)^{-1/\gamma} \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\tau+k/\gamma)\Gamma(1-k/\gamma)}{\Gamma(\tau)} \beta(\tau+k/\gamma, 1-k/\gamma; u) + x^k[1-u^\tau], \quad k > -\tau\gamma \\
\text{mode} &= \theta \left(\frac{\tau\gamma-1}{\gamma+1}\right)^{1/\gamma}, \quad \tau\gamma > 1, \text{ else } 0
\end{aligned}$$

A.2.3 Two-parameter distributions

A.2.3.1 Pareto (Pareto Type II, Lomax)—α, θ

$$
\begin{aligned}
f(x) &= \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}} \qquad F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad -1 < k < \alpha \\
\mathrm{E}[X^k] &= \frac{\theta^k k!}{(\alpha-1)\cdots(\alpha-k)}, \quad \text{if } k \text{ is an integer} \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha} - 1] \\
\mathrm{TVaR}_p(X) &= \mathrm{VaR}_p(X) + \frac{\theta(1-p)^{-1/\alpha}}{\alpha-1}, \quad \alpha > 1 \\
\mathrm{E}[X \wedge x] &= \frac{\theta}{\alpha-1}\left[1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha-1}\right], \quad \alpha \neq 1 \\
\mathrm{E}[X \wedge x] &= -\theta \ln\left(\frac{\theta}{x+\theta}\right), \quad \alpha = 1 \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha)}\beta[k+1, \alpha-k; x/(x+\theta)] + x^k \left(\frac{\theta}{x+\theta}\right)^\alpha, \quad \text{all } k \\
\text{mode} &= 0
\end{aligned}
$$

A.2.3.2 Inverse Pareto—τ, θ

$$
\begin{aligned}
f(x) &= \frac{\tau\theta x^{\tau-1}}{(x+\theta)^{\tau+1}} \qquad F(x) = \left(\frac{x}{x+\theta}\right)^\tau \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\tau+k)\Gamma(1-k)}{\Gamma(\tau)}, \quad -\tau < k < 1 \\
\mathrm{E}[X^k] &= \frac{\theta^k (-k)!}{(\tau-1)\cdots(\tau+k)}, \quad \text{if } k \text{ is a negative integer} \\
\mathrm{VaR}_p(X) &= \theta[p^{-1/\tau} - 1]^{-1} \\
\mathrm{E}[(X \wedge x)^k] &= \theta^k \tau \int_0^{x/(x+\theta)} y^{\tau+k-1}(1-y)^{-k} dy + x^k \left[1 - \left(\frac{x}{x+\theta}\right)^\tau\right], \quad k > -\tau \\
\text{mode} &= \theta\frac{\tau-1}{2}, \quad \tau > 1, \text{ else } 0
\end{aligned}
$$

A.2.3.3 Loglogistic (Fisk)—γ, θ

$$
\begin{aligned}
f(x) &= \frac{\gamma(x/\theta)^\gamma}{x[1+(x/\theta)^\gamma]^2} \qquad F(x) = u, \quad u = \frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma} \\
\mathrm{E}[X^k] &= \theta^k \Gamma(1+k/\gamma)\Gamma(1-k/\gamma), \quad -\gamma < k < \gamma \\
\mathrm{VaR}_p(X) &= \theta(p^{-1} - 1)^{-1/\gamma} \\
\mathrm{E}[(X \wedge x)^k] &= \theta^k \Gamma(1+k/\gamma)\Gamma(1-k/\gamma)\beta(1+k/\gamma, 1-k/\gamma; u) + x^k(1-u), \quad k > -\gamma \\
\text{mode} &= \theta\left(\frac{\gamma-1}{\gamma+1}\right)^{1/\gamma}, \quad \gamma > 1, \text{ else } 0
\end{aligned}
$$

A.2.3.4 Paralogistic—α, θ

This is a Burr distribution with $\gamma = \alpha$.

$$
\begin{aligned}
f(x) &= \frac{\alpha^2(x/\theta)^\alpha}{x[1+(x/\theta)^\alpha]^{\alpha+1}} \qquad F(x) = 1-u^\alpha, \quad u = \frac{1}{1+(x/\theta)^\alpha} \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(1+k/\alpha)\Gamma(\alpha-k/\alpha)}{\Gamma(\alpha)}, \quad -\alpha < k < \alpha^2 \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha}-1]^{1/\alpha} \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(1+k/\alpha)\Gamma(\alpha-k/\alpha)}{\Gamma(\alpha)}\beta(1+k/\alpha, \alpha-k/\alpha; 1-u) + x^k u^\alpha, \quad k > -\alpha \\
\text{mode} &= \theta\left(\frac{\alpha-1}{\alpha^2+1}\right)^{1/\alpha}, \quad \alpha > 1, \text{ else } 0
\end{aligned}
$$

A.2.3.5 Inverse paralogistic—τ, θ

This is an inverse Burr distribution with $\gamma = \tau$.

$$
\begin{aligned}
f(x) &= \frac{\tau^2(x/\theta)^{\tau^2}}{x[1+(x/\theta)^\tau]^{\tau+1}} \qquad F(x) = u^\tau, \quad u = \frac{(x/\theta)^\tau}{1+(x/\theta)^\tau} \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\tau+k/\tau)\Gamma(1-k/\tau)}{\Gamma(\tau)}, \quad -\tau^2 < k < \tau \\
\mathrm{VaR}_p(X) &= \theta(p^{-1/\tau}-1)^{-1/\tau} \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\tau+k/\tau)\Gamma(1-k/\tau)}{\Gamma(\tau)}\beta(\tau+k/\tau, 1-k/\tau; u) + x^k[1-u^\tau], \quad k > -\tau^2 \\
\text{mode} &= \theta\,(\tau-1)^{1/\tau}, \quad \tau > 1, \text{ else } 0
\end{aligned}
$$

A.3 Transformed gamma family

A.3.2 Two-parameter distributions

A.3.2.1 Gamma—α, θ

$$
\begin{aligned}
f(x) &= \frac{(x/\theta)^\alpha e^{-x/\theta}}{x\Gamma(\alpha)} \qquad F(x) = \Gamma(\alpha; x/\theta) \\
M(t) &= (1-\theta t)^{-\alpha}, \quad t < 1/\theta \qquad \mathrm{E}[X^k] = \frac{\theta^k\Gamma(\alpha+k)}{\Gamma(\alpha)}, \quad k > -\alpha \\
\mathrm{E}[X^k] &= \theta^k(\alpha+k-1)\cdots\alpha, \quad \text{if } k \text{ is an integer}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\alpha+k)}{\Gamma(\alpha)}\Gamma(\alpha+k; x/\theta) + x^k[1-\Gamma(\alpha; x/\theta)], \quad k > -\alpha \\
&= \alpha(\alpha+1)\cdots(\alpha+k-1)\theta^k\Gamma(\alpha+k; x/\theta) + x^k[1-\Gamma(\alpha; x/\theta)], \quad k \text{ an integer} \\
\text{mode} &= \theta(\alpha-1), \quad \alpha > 1, \text{ else } 0
\end{aligned}
$$

A.3.2.2 Inverse gamma (Vinci)—α, θ

$$
\begin{aligned}
f(x) &= \frac{(\theta/x)^\alpha e^{-\theta/x}}{x\Gamma(\alpha)} \qquad\qquad F(x) = 1 - \Gamma(\alpha;\theta/x) \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad k<\alpha \qquad\qquad \mathrm{E}[X^k] = \frac{\theta^k}{(\alpha-1)\cdots(\alpha-k)}, \quad \text{if } k \text{ is an integer} \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)}[1-\Gamma(\alpha-k;\theta/x)] + x^k\Gamma(\alpha;\theta/x) \\
&= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)}G(\alpha-k;\theta/x) + x^k\Gamma(\alpha;\theta/x), \text{ all } k \\
\text{mode} &= \theta/(\alpha+1)
\end{aligned}
$$

A.3.2.3 Weibull—θ, τ

$$
\begin{aligned}
f(x) &= \frac{\tau(x/\theta)^\tau e^{-(x/\theta)^\tau}}{x} \qquad\qquad F(x) = 1 - e^{-(x/\theta)^\tau} \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1+k/\tau), \quad k > -\tau \\
\mathrm{VaR}_p(X) &= \theta[-\ln(1-p)]^{1/\tau} \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(1+k/\tau)\Gamma[1+k/\tau;(x/\theta)^\tau] + x^k e^{-(x/\theta)^\tau}, \quad k>-\tau \\
\text{mode} &= \theta\left(\frac{\tau-1}{\tau}\right)^{1/\tau}, \quad \tau>1, \text{ else } 0
\end{aligned}
$$

A.3.2.4 Inverse Weibull (log Gompertz)—θ, τ

$$
\begin{aligned}
f(x) &= \frac{\tau(\theta/x)^\tau e^{-(\theta/x)^\tau}}{x} \qquad\qquad F(x) = e^{-(\theta/x)^\tau} \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1-k/\tau), \quad k<\tau \\
\mathrm{VaR}_p(X) &= \theta(-\ln p)^{-1/\tau} \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(1-k/\tau)\{1-\Gamma[1-k/\tau;(\theta/x)^\tau]\} + x^k\left[1-e^{-(\theta/x)^\tau}\right], \quad \text{all } k \\
&= \theta^k\Gamma(1-k/\tau)G[1-k/\tau;(\theta/x)^\tau] + x^k\left[1-e^{-(\theta/x)^\tau}\right] \\
\text{mode} &= \theta\left(\frac{\tau}{\tau+1}\right)^{1/\tau}
\end{aligned}
$$

A.3.3 One-parameter distributions

A.3.3.1 Exponential—θ

$$
\begin{aligned}
f(x) &= \frac{e^{-x/\theta}}{\theta} \qquad F(x) = 1 - e^{-x/\theta} \\
M(t) &= (1-\theta t)^{-1} \qquad \mathrm{E}[X^k] = \theta^k\Gamma(k+1), \quad k > -1 \\
\mathrm{E}[X^k] &= \theta^k k!, \quad \text{if } k \text{ is an integer} \\
\mathrm{VaR}_p(X) &= -\theta\ln(1-p) \\
\mathrm{TVaR}_p(X) &= -\theta\ln(1-p) + \theta \\
\mathrm{E}[X\wedge x] &= \theta(1-e^{-x/\theta}) \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(k+1)\Gamma(k+1;x/\theta) + x^k e^{-x/\theta}, \quad k > -1 \\
&= \theta^k k!\Gamma(k+1;x/\theta) + x^k e^{-x/\theta}, \quad k \text{ an integer} \\
\text{mode} &= 0
\end{aligned}
$$

A.3.3.2 Inverse exponential—θ

$$
\begin{aligned}
f(x) &= \frac{\theta e^{-\theta/x}}{x^2} \qquad F(x) = e^{-\theta/x} \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1-k), \quad k < 1 \\
\mathrm{VaR}_p(X) &= \theta(-\ln p)^{-1} \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k G(1-k;\theta/x) + x^k(1-e^{-\theta/x}), \quad \text{all } k \\
\text{mode} &= \theta/2
\end{aligned}
$$

A.5 Other distributions

A.5.1.1 Lognormal—μ,σ (μ can be negative)

$$
\begin{aligned}
f(x) &= \frac{1}{x\sigma\sqrt{2\pi}}\exp(-z^2/2) = \phi(z)/(\sigma x), \quad z = \frac{\ln x - \mu}{\sigma} \qquad F(x) = \Phi(z) \\
\mathrm{E}[X^k] &= \exp(k\mu + k^2\sigma^2/2) \\
\mathrm{E}[(X\wedge x)^k] &= \exp(k\mu + k^2\sigma^2/2)\Phi\left(\frac{\ln x - \mu - k\sigma^2}{\sigma}\right) + x^k[1-F(x)] \\
\text{mode} &= \exp(\mu - \sigma^2)
\end{aligned}
$$

A.5.1.2 Inverse Gaussian—μ, θ

$$\begin{aligned}
f(x) &= \left(\frac{\theta}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{\theta z^2}{2x}\right), \quad z = \frac{x-\mu}{\mu} \\
F(x) &= \Phi\left[z\left(\frac{\theta}{x}\right)^{1/2}\right] + \exp\left(\frac{2\theta}{\mu}\right)\Phi\left[-y\left(\frac{\theta}{x}\right)^{1/2}\right], \quad y = \frac{x+\mu}{\mu} \\
M(t) &= \exp\left[\frac{\theta}{\mu}\left(1-\sqrt{1-\frac{2t\mu^2}{\theta}}\right)\right], \quad t < \frac{\theta}{2\mu^2}, \qquad \mathrm{E}[X] = \mu, \quad \mathrm{Var}[X] = \mu^3/\theta \\
\mathrm{E}[X \wedge x] &= x - \mu z \Phi\left[z\left(\frac{\theta}{x}\right)^{1/2}\right] - \mu y \exp\left(\frac{2\theta}{\mu}\right)\Phi\left[-y\left(\frac{\theta}{x}\right)^{1/2}\right]
\end{aligned}$$

A.5.1.3 log-t—r, μ, σ (μ *can be negative*)

Let Y have a t distribution with r degrees of freedom. Then $X = \exp(\sigma Y + \mu)$ has the log-t distribution. Positive moments do not exist for this distribution. Just as the t distribution has a heavier tail than the normal distribution, this distribution has a heavier tail than the lognormal distribution.

$$\begin{aligned}
f(x) &= \frac{\Gamma\left(\frac{r+1}{2}\right)}{x\sigma\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)\left[1+\frac{1}{r}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right]^{(r+1)/2}}, \\
F(x) &= F_r\left(\frac{\ln x - \mu}{\sigma}\right) \text{ with } F_r(t) \text{ the cdf of a } t \text{ distribution with } r \text{ d.f.,}
\end{aligned}$$

$$F(x) = \begin{cases} \dfrac{1}{2}\beta\left[\dfrac{r}{2}, \dfrac{1}{2}; \dfrac{r}{r+\left(\dfrac{\ln x - \mu}{\sigma}\right)^2}\right], & 0 < x \le e^{\mu}, \\[2ex] 1 - \dfrac{1}{2}\beta\left[\dfrac{r}{2}, \dfrac{1}{2}; \dfrac{r}{r+\left(\dfrac{\ln x - \mu}{\sigma}\right)^2}\right], & x \ge e^{\mu}. \end{cases}$$

A.5.1.4 Single-parameter Pareto—α, θ

$$\begin{aligned}
f(x) &= \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \quad x > \theta \qquad F(x) = 1 - (\theta/x)^\alpha, \quad x > \theta \\
\mathrm{VaR}_p(X) &= \theta(1-p)^{-1/\alpha} \qquad \mathrm{TVaR}_p(X) = \frac{\alpha\theta(1-p)^{-1/\alpha}}{\alpha - 1}, \quad \alpha > 1 \\
\mathrm{E}[X^k] &= \frac{\alpha\theta^k}{\alpha - k}, \quad k < \alpha \qquad \mathrm{E}[(X \wedge x)^k] = \frac{\alpha\theta^k}{\alpha - k} - \frac{k\theta^\alpha}{(\alpha - k)x^{\alpha - k}}, \; x \ge \theta \\
\text{mode} &= \theta
\end{aligned}$$

Note: Although there appears to be two parameters, only α is a true parameter. The value of θ must be set in advance.

A.6 Distributions with finite support

For these two distributions, the scale parameter θ is assumed known.

A.6.1.1 Generalized beta—a, b, θ, τ

$$
\begin{aligned}
f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^a (1-u)^{b-1} \frac{\tau}{x}, \quad 0 < x < \theta, \quad u = (x/\theta)^\tau \\
F(x) &= \beta(a,b;u) \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(a+b)\Gamma(a+k/\tau)}{\Gamma(a)\Gamma(a+b+k/\tau)}, \quad k > -a\tau \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(a+b)\Gamma(a+k/\tau)}{\Gamma(a)\Gamma(a+b+k/\tau)} \beta(a+k/\tau, b; u) + x^k [1 - \beta(a,b;u)]
\end{aligned}
$$

A.6.1.2 beta—a, b, θ

$$
\begin{aligned}
f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^a (1-u)^{b-1} \frac{1}{x}, \quad 0 < x < \theta, \quad u = x/\theta \\
F(x) &= \beta(a,b;u) \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(a+b)\Gamma(a+k)}{\Gamma(a)\Gamma(a+b+k)}, \quad k > -a \\
\mathrm{E}[X^k] &= \frac{\theta^k a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)}, \quad \text{if } k \text{ is an integer} \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)} \beta(a+k, b; u) \\
&\quad + x^k [1 - \beta(a,b;u)]
\end{aligned}
$$

Appendix B

An Inventory of Discrete Distributions

B.1 Introduction

The 16 models fall into three classes. The divisions are based on the algorithm by which the probabilities are computed. For some of the more familiar distributions these formulas will look different from the ones you may have learned, but they produce the same probabilities. After each name, the parameters are given. All parameters are positive unless otherwise indicated. In all cases, p_k is the probability of observing k losses.

For finding moments, the most convenient form is to give the factorial moments. The jth factorial moment is $\mu_{(j)} = \mathrm{E}[N(N-1)\cdots(N-j+1)]$. We have $\mathrm{E}[N] = \mu_{(1)}$ and $\mathrm{Var}(N) = \mu_{(2)} + \mu_{(1)} - \mu_{(1)}^2$.

The estimators which are presented are not intended to be useful estimators but rather for providing starting values for maximizing the likelihood (or other) function. For determining starting values, the following quantities are used [where n_k is the observed frequency at k (if, for the last entry, n_k represents the number of observations at k or more, assume it was at exactly k) and n is the sample size]:

$$\hat{\mu} = \frac{1}{n}\sum_{k=1}^{\infty} k n_k, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{\infty} k^2 n_k - \hat{\mu}^2.$$

When the method of moments is used to determine the starting value, a circumflex (e.g., $\hat{\lambda}$) is used. For any other method, a tilde (e.g., $\tilde{\lambda}$) is used. When the starting value formulas do not provide admissible parameter values, a truly crude guess is to set the product of all λ and β parameters equal to the sample mean and set all other parameters equal to 1. If there are two λ and/or β parameters, an easy choice is to set each to the square root of the sample mean.

The last item presented is the probability generating function,

$$P(z) = \mathrm{E}[z^N].$$

B.2 The $(a, b, 0)$ class

B.2.1.1 Poisson—λ

$$\begin{aligned} p_0 &= e^{-\lambda}, \quad a = 0, \quad b = \lambda \qquad p_k = \frac{e^{-\lambda}\lambda^k}{k!} \\ \mathrm{E}[N] &= \lambda, \quad \mathrm{Var}[N] = \lambda \qquad P(z) = e^{\lambda(z-1)} \end{aligned}$$

B.2.1.2 Geometric—β

$$
\begin{aligned}
p_0 &= \frac{1}{1+\beta}, \quad a = \frac{\beta}{1+\beta}, \quad b = 0 \qquad p_k = \frac{\beta^k}{(1+\beta)^{k+1}} \\
\mathrm{E}[N] &= \beta, \quad \mathrm{Var}[N] = \beta(1+\beta) \qquad P(z) = [1-\beta(z-1)]^{-1}.
\end{aligned}
$$

This is a special case of the negative binomial with $r = 1$.

B.2.1.3 Binomial—$q, m, (0 < q < 1,\ m$ an integer$)$

$$
\begin{aligned}
p_0 &= (1-q)^m, \quad a = -\frac{q}{1-q}, \quad b = \frac{(m+1)q}{1-q} \\
p_k &= \binom{m}{k} q^k (1-q)^{m-k}, \quad k = 0, 1, \ldots, m \\
\mathrm{E}[N] &= mq, \quad \mathrm{Var}[N] = mq(1-q) \qquad P(z) = [1+q(z-1)]^m.
\end{aligned}
$$

B.2.1.4 Negative binomial—β, r

$$
\begin{aligned}
p_0 &= (1+\beta)^{-r}, \quad a = \frac{\beta}{1+\beta}, \quad b = \frac{(r-1)\beta}{1+\beta} \\
p_k &= \frac{r(r+1)\cdots(r+k-1)\beta^k}{k!(1+\beta)^{r+k}} \\
\mathrm{E}[N] &= r\beta, \quad \mathrm{Var}[N] = r\beta(1+\beta) \qquad P(z) = [1-\beta(z-1)]^{-r}.
\end{aligned}
$$

B.3 The $(a, b, 1)$ class

To distinguish this class from the $(a, b, 0)$ class, the probabilities are denoted $\Pr(N = k) = p_k^M$ or $\Pr(N = k) = p_k^T$ depending on which subclass is being represented. For this class, p_0^M is arbitrary (that is, it is a parameter) and then p_1^M or p_1^T is a specified function of the parameters a and b. Subsequent probabilities are obtained recursively as in the $(a, b, 0)$ class: $p_k^M = (a + b/k)p_{k-1}^M$, $k = 2, 3, \ldots$, with the same recursion for p_k^T There are two sub-classes of this class. When discussing their members, we often refer to the "corresponding" member of the $(a, b, 0)$ class. This refers to the member of that class with the same values for a and b. The notation p_k will continue to be used for probabilities for the corresponding $(a, b, 0)$ distribution.

B.3.1 The zero-truncated subclass

The members of this class have $p_0^T = 0$ and therefore it need not be estimated. These distributions should only be used when a value of zero is impossible. The first factorial moment is $\mu_{(1)} = (a+b)/[(1-a)(1-p_0)]$, where p_0 is the value for the corresponding member of the $(a, b, 0)$ class. For the logarithmic distribution (which has no corresponding member), $\mu_{(1)} = \beta/\ln(1+\beta)$. Higher factorial moments are obtained recursively with the same formula as with the $(a, b, 0)$ class. The variance is $(a+b)[1-(a+b+1)p_0]/[(1-a)(1-p_0)]^2$.For those members of the subclass which have corresponding $(a, b, 0)$ distributions, $p_k^T = p_k/(1-p_0)$.

B.3.1.1 Zero-truncated Poisson—λ

$$
\begin{aligned}
p_1^T &= \frac{\lambda}{e^\lambda - 1}, \quad a = 0, \quad b = \lambda, \\
p_k^T &= \frac{\lambda^k}{k!(e^\lambda - 1)}, \\
\mathrm{E}[N] &= \lambda/(1 - e^{-\lambda}), \quad \mathrm{Var}[N] = \lambda[1 - (\lambda + 1)e^{-\lambda}]/(1 - e^{-\lambda})^2, \\
\tilde{\lambda} &= \ln(n\hat{\mu}/n_1), \\
P(z) &= \frac{e^{\lambda z} - 1}{e^\lambda - 1}.
\end{aligned}
$$

B.3.1.2 Zero-truncated geometric—β

$$
\begin{aligned}
p_1^T &= \frac{1}{1+\beta}, \quad a = \frac{\beta}{1+\beta}, \quad b = 0, \\
p_k^T &= \frac{\beta^{k-1}}{(1+\beta)^k}, \\
\mathrm{E}[N] &= 1 + \beta, \quad \mathrm{Var}[N] = \beta(1+\beta), \\
\hat{\beta} &= \hat{\mu} - 1, \\
P(z) &= \frac{[1 - \beta(z-1)]^{-1} - (1+\beta)^{-1}}{1 - (1+\beta)^{-1}}.
\end{aligned}
$$

This is a special case of the zero-truncated negative binomial with $r = 1$.

B.3.1.3 Logarithmic—β

$$
\begin{aligned}
p_1^T &= \frac{\beta}{(1+\beta)\ln(1+\beta)}, \quad a = \frac{\beta}{1+\beta}, \quad b = -\frac{\beta}{1+\beta}, \\
p_k^T &= \frac{\beta^k}{k(1+\beta)^k \ln(1+\beta)}, \\
\mathrm{E}[N] &= \beta/\ln(1+\beta), \quad \mathrm{Var}[N] = \frac{\beta[1 + \beta - \beta/\ln(1+\beta)]}{\ln(1+\beta)}, \\
\tilde{\beta} &= \frac{n\hat{\mu}}{n_1} - 1 \quad \text{or} \quad \frac{2(\hat{\mu} - 1)}{\hat{\mu}}, \\
P(z) &= 1 - \frac{\ln[1 - \beta(z-1)]}{\ln(1+\beta)}.
\end{aligned}
$$

This is a limiting case of the zero-truncated negative binomial as $r \to 0$.

B.3.1.4 Zero-truncated binomial—$q, m, (0 < q < 1,\ m$ an integer)

$$
\begin{aligned}
p_1^T &= \frac{m(1-q)^{m-1}q}{1-(1-q)^m}, \quad a = -\frac{q}{1-q}, \quad b = \frac{(m+1)q}{1-q}, \\
p_k^T &= \frac{\binom{m}{k} q^k (1-q)^{m-k}}{1-(1-q)^m}, \quad k = 1, 2, \ldots, m, \\
\mathrm{E}[N] &= \frac{mq}{1-(1-q)^m}, \\
\mathrm{Var}[N] &= \frac{mq[(1-q)-(1-q+mq)(1-q)^m]}{[1-(1-q)^m]^2}, \\
\tilde{q} &= \frac{\hat{\mu}}{m}, \\
P(z) &= \frac{[1+q(z-1)]^m - (1-q)^m}{1-(1-q)^m}.
\end{aligned}
$$

B.3.1.5 Zero-truncated negative binomial—$\beta, r, (r > -1, r \neq 0)$

$$
\begin{aligned}
p_1^T &= \frac{r\beta}{(1+\beta)^{r+1}-(1+\beta)}, \quad a = \frac{\beta}{1+\beta}, \quad b = \frac{(r-1)\beta}{1+\beta}, \\
p_k^T &= \frac{r(r+1)\cdots(r+k-1)}{k![(1+\beta)^r - 1]} \left(\frac{\beta}{1+\beta}\right)^k, \\
\mathrm{E}[N] &= \frac{r\beta}{1-(1+\beta)^{-r}}, \\
Var[N] &= \frac{r\beta[(1+\beta)-(1+\beta+r\beta)(1+\beta)^{-r}]}{[1-(1+\beta)^{-r}]^2}, \\
\tilde{\beta} &= \frac{\hat{\sigma}^2}{\hat{\mu}} - 1, \quad \tilde{r} = \frac{\hat{\mu}^2}{\hat{\sigma}^2 - \hat{\mu}}, \\
P(z) &= \frac{[1-\beta(z-1)]^{-r} - (1+\beta)^{-r}}{1-(1+\beta)^{-r}}.
\end{aligned}
$$

This distribution is sometimes called the extended truncated negative binomial distribution because the parameter r can extend below 0.

B.3.2 The zero-modified subclass

A zero-modified distribution is created by starting with a truncated distribution and then placing an arbitrary amount of probability at zero. This probability, p_0^M, is a parameter. The remaining probabilities are adjusted accordingly. Values of p_k^M can be determined from the corresponding zero-truncated distribution as $p_k^M = (1-p_0^M)p_k^T$ or from the corresponding $(a, b, 0)$ distribution as $p_k^M = (1-p_0^M)p_k/(1-p_0)$. The same recursion used for the zero-truncated subclass applies.

The mean is $1-p_0^M$ times the mean for the corresponding zero-truncated distribution. The variance is $1-p_0^M$ times the zero-truncated variance plus $p_0^M(1-p_0^M)$ times the square of the zero-truncated mean. The probability generating function is $P^M(z) = p_0^M + (1-p_0^M)P(z)$, where $P(z)$ is the probability generating function for the corresponding zero-truncated distribution.

The maximum likelihood estimator of p_0^M is always the sample relative frequency at 0.

Bibliography

[1] S.A. Klugman , H.H. Panjer, and G.E. Willmot *Loss Models from Data to Decisions*, 3rd Edition (2008), Wiley.

[2] M.B. Finan, *A Probability Course for the Actuaries: A Preparation for Exam P/1*, 2007, Arkansas Tech University.

[3] M.B. Finan, *A Reading in the Theory of Life Contingency Models: A preparation for Exam MLC/3L*, 2011, Arkansas Tech University.

[4] SOA/CAS, *Exam C Sample Questions.*

[5] IAA, *Stochastic Modeling: Theory and Reality from an Actuarial Perspective*, 2010.

Index

$C(a, b, 0)$, 195
$C(a, b, 1)$, 201
$D(x)-$plot, 514
$p - p$ plot, 514

Absolute loss, 485
Aggregate claims model, 250
Aggregate loss, 257
Alternative hypothesis, 344
Anderson-Darling test, 527
Arithmetic distribution, 308
Arithmetization, 308
Asymptotic variance of the MLE, 459
Asymptotically unbiased, 335, 459

Bühlmann credibility factor, 605
Bühlmann credibility premium, 606
Bayes estimate, 485
Bayes Theorem, 475
Bayes' formula, 10
Bayesian approach, 474
Bayesian Central Limit Theorem, 489
Bayesian premium, 582
Bernoulli r.v., 189
Biased, 334
Binomial random variable, 40, 189
Bootstrap method, 686

Central limit theorem, 70
Central moment, 34
Chi square distribution, 471
Chi-square goodness-of-fit test, 532
Claim amount, 48, 211
Claim count distribution, 258
Claim count random variable, 258
Claim experience, 551
Classical credibility, 551
Coefficient of kurtosis, 37
Coefficient of skewness, 35
Coefficient of variation, 38
Coherent risk measure, 110
Coinsurance, 237
Collective premium, 582
collective premium, 605
Complement, 6
Complete data, 351, 363
Complete expectation of life, 49, 98
Compound distribution, 250, 258
Compound frequency model, 250
Compound Poisson distribution, 296
Conditional hazard function, 160
Conditional probability, 9
Confidence interval, 341
Confidence region, 471
Conjugate prior distribution, 493
Consistent estimator, 336
Continuous r.v., 18
Convex cone, 110
Convolutions, 264
Cost per loss, 52
Cost per-payment, 213
Cost-per loss, 212
Countable additivity, 7
Counting distribution, 177
Covariance matrix, 464

Cramér-Rao lower bound, 458
Credibility factor, 561
Credibility interval, 488
Credibility premium, 561
Credibility theory, 551
credibility weighted average method, 657
Critical values, 345
Cumulative distribution function, 22
Cumulative hazard function, 26
Cumulative hazard rate function, 353

Data-dependent distribution, 138
Decomposition property, 179
Delta method, 465
Deterministic model, 2
Discrete r.v., 18

Empirical Bayes estimation, 643
Empirical distribution, 138, 352
Empirical model, 45
Equilibrium distribution, 102
Estimate, 334
Estimator, 334
Event, 6
Exact credibility, 637
Excess loss variable, 48, 213
Expected shortfall, 119
Expected value of the hypothetical mean, 605
Expected value of the process variance, 605
Expected value premium principle, 113
Experiment, 6
Exponential premium principle, 112
Exponential random variable, 41
Exposure, 63, 552
Extended truncated negative binomial distribution, 206

Failure rate, 26
First quartile, 63
Fisher information, 455
Fisher information matrix, 463
force of mortality, 26
Frailty model, 160
Franchise deductible, 220
Frequency distribution, 258
Frequentest approach, 474

Gamma kernel, 402
Gamma random variable, 41
Geometric, 185
Geometric random variable, 40
Greatest accuracy credibility theory, 567
Greenwood's approximation, 389

Hazard rate function, 26
Heavy-tailed distribution, 87, 88
Hewitt's model, 633
Histogram, 358
Hypothesis testing, 344
Hypothetical mean, 582, 605

Improper prior distribution, 474
Incomplete data, 351, 363
incomplete Gamma function, 294
Independent random variables, 69
Individual risk model, 258, 321
Interquartile range, 64
Intersection, 6
interval estimation, 341
Inverse distribution, 148
Inverse transformed distribution, 148
Inversion method, 668

Joint density function, 153
Joint distribution, 581

k-point mixture, 131
Kapler-Meier estimator, 370

Kernel density estimator, 400
Kernel function, 400
Kernel smoothed density function, 139
Kernel smoothed distribution, 138
Kernel smoothing, 400
Kolmogorov axioms, 7
Kolmogorov-Smirnov statistic, 521
Kurtosis, 37

Left censored, 364
Left truncated, 364
Left truncated and shifted variable, 213
left truncated and shifted variable, 48
Left-censored and shifted, 52
Level of confidence, 341, 345
Life table approach, 415
Light tailed distribution, 88
Light-tailed distribution, 87
Likelihood function, 434
Likelihood ratio test, 540
Limited expected value, 53
Limited loss random variable, 233
Limited loss variable, 53
linear confidence intervals, 394
Linear exponential family, 172
Local moment matching method, 309
Log-transformed confidence intervals, 394
Logarithmic distribution, 206, 207
Loglikelihood function, 435
Loss Elimination Ratio, 226
Loss frequency, 257
Loss random variable, 48
Loss severity, 257

Manual rate, 551
Marginal density, 153
Marginal distribution, 581
Mass dispersal method, 308
Maximum covered loss, 237
Maximum likelihood estimate, 434
mean excess loss function, 49
mean residual life function, 49
Mean square error, 686
Mean-squared error, 337
Median, 62
Medium-tailed distribution, 96
Memoryless property, 185
Method of Rounding, 308
Method-of-moments estimate, 421
Mixed random variable, 18
Mixing weight, 131
Mixture distribution, 131
Mode, 63
Model distribution, 475, 581
Modified data, 363
Moment, 34
Moment generating function, 78
Monotonicity, 110
Mutually exclusive, 6

Negative binomial distribution, 183
Nelson-Åalen, 373
Nelson-Åalen estimate, 353
Net stop-loss premium, 284
Non-parametric distributions, 138
Nonparametric Estimation, 643
Normal equations, 601
Normal random variable, 40
Null hypothesis, 344

Observed information, 464
Observed pure premium, 552
Ogive, 357
Ordinary deductible, 212
Outcome, 6

p-value, 347
Panjer Recursive formula, 300

Parameters, 126
Parametric distribution, 126
Parametric distribution family, 128
Pareto random variable, 43, 44
Partial credibility, 561
Percentile, 63
Percentile matching estimate, 423
Point estimation, 334
Poisson r.v., 178
Poisson random variable, 40
Policy limit, 53
Positive homogeneity, 110
Posterior distribution, 581
Predictive distribution, 475, 581
Primary distribution, 250, 258
Principle of Parsimony, 546
Prior distribution, 474, 581
Probability, 7
Probability axioms, 7
Probability density function, 21
Probability function, 19, 177
Probability generating function, 82
Probability mass function, 19
Probability measure, 7
Probability plot, 514
Probability tree, 8
Process variance, 605
Pure premium, 552

Quantile, 63

Railty random variable, 160
Random variable, 18
Raw moment, 34
Rejection region, 345
Relative tail weight, 89
Reliability function,, 25
Residual mean lifetime, 98
Right censored, 364
Right tail, 87
Right truncated, 364
Right-censored r.v., 53
Risk group, 551
Risk management, 110
Risk measure, 110
Risk set, 352, 365

Sample space, 6
Scale distribution, 126
Scale parameter, 127
Schwarz Bayesian adjustment, 546
Schwarz Bayesian Criterion, 546
Secondary distribution, 250, 258
Semi-parametric distributions, 138
semiparametric, 661
Severity distribution, 258
Severity random variable, 48, 258
Simple event, 6
Single-decrement, 413
Skewness, 35
Smoothed empirical estimate, 423
Span, 308
Spliced distribution, 165
Squared-error loss, 485
standard deviation principle, 114
Standard of full credibiltiy, 554
Statistic, 334
Statistical hypothesis, 344
Stochastic model, 2
Stop loss premium, 52
Stop-loss insurance, 284
structural parameters, 643
Subadditivity, 110
Support, 18
Survival function, 25

Tail conditional expectation, 119
Tail-Value-at-Risk, 119
Test statistic, 345
Transformed distribution, 148

Translation invariance, 110
Triangular kernel, 401
Type I error, 345
Type II error, 346

Unbiased, 334
Unbiased equation, 600
Uniform kernel, 139, 400
Uniformly minimum variance unbiased estimator, 338
Union, 6

Value at risk, 115
Variable-component mixture distribution, 133
Variance of the hypothetical mean, 605
Variance premium principle, 111

Zero-modified, 201
Zero-one loss, 485
Zero-truncated, 201