Henry Stark and John W. Woods,
Probability, Statistics, and Random Variables for Engineers, 4th ed.,
Pearson Education Inc., 2012. ISBN: 978-0-13-231123-6

Chapter 1: Introduction to Probability

Sections

An understanding of Probability and Statistics is necessary in most if not all work related to science and engineering.

Statistics: the study of and the dealing with data.

Probability: the study of the likeliness of result, action or event occurring.

Often based on prior knowledge or the statistics of similar or past events!

Terms: *Random Variables, Random Processes or Stochastic Processes*

For any **measured phenomenon** there will be Uncertainty, Expected Variations, Randomness, or even Expected Errors included.

- when an outcome is non-deterministic
- where an exact value is subject to errors … e.g. noise, measurement

Easy examples of such phenomenon include all games of chance

- Flipping coins, rolling dice, dealing cards, etc.

Engineering Applications include

- Realistic signals – with noise or characteristic "unknown" parts
- Signal-to-noise Ratios, Noise-Power Measurements, Background Noise
- Expected Values, Variances, Distributions
- Thermal Motion, Electron Movement
- Reliability, Quality, Failure Rates, etc.
- Thermal Motion, Electron Movement

Probability theory is necessary for engineering system modeling and simulations.

- unknown initial conditions (random)
- noisy measurements, expected inaccuracies, etc. during operation

Textbook Intro Quote

By Eugen Merzbacher:

"The probability doctrine of quantum mechanics asserts that the *indetermination ... is a property inherent in nature and not merely a profession of our temporary ignorance* from which we expect to be relieved by a future better and more complete theory. "

"The conventional interpretation thus denies the possibility of an ideal theory that would encompass the innumerable experimentally verified predictions of quantum mechanics but would be free of its supposed defects, the most notorious "imperfection" of quantum mechanics being the abandonment of strict classical determinism. "

Different kinds of Probability

Suggested that there are essentially 4 types

- Probability by Intuition
    - "Lucky Numbers"
- Probability as the Ratio of Favorable to Total Outcomes (Classical Theory)
    - Measured Statistical Expectations
- Probability as a Measure of the Frequency of Outcomes
- Probability Based on Axiomatic Theory

Definitions of used in Probability

Experiment

- An experiment is some action that results in an outcome.
- A random experiment is one in which the outcome is uncertain before the experiment is performed.

Possible Outcomes

- A description of all possible experimental outcomes.
- The set of possible outcomes may be discrete or form a continuum.

Trials

- The single performance of a well-defined experiment.

Event

- An elementary event is one for which there is only one outcome.
- A composite event is one for which the desired result can be achieved in multiple ways. Multiple outcomes result in the event described.

Equally Likely Events/Outcomes

- When the set of events or each of the possible outcomes is equally likely to occur.
- A term that is used synonymously to equally likely outcomes is random.

Probability as the Ratio of Favorable to Total Outcomes (Classical Theory) – 2 Dice example

Table 1.2-1 Outcomes of Throwing Two Dice

	1st die					
2nd die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

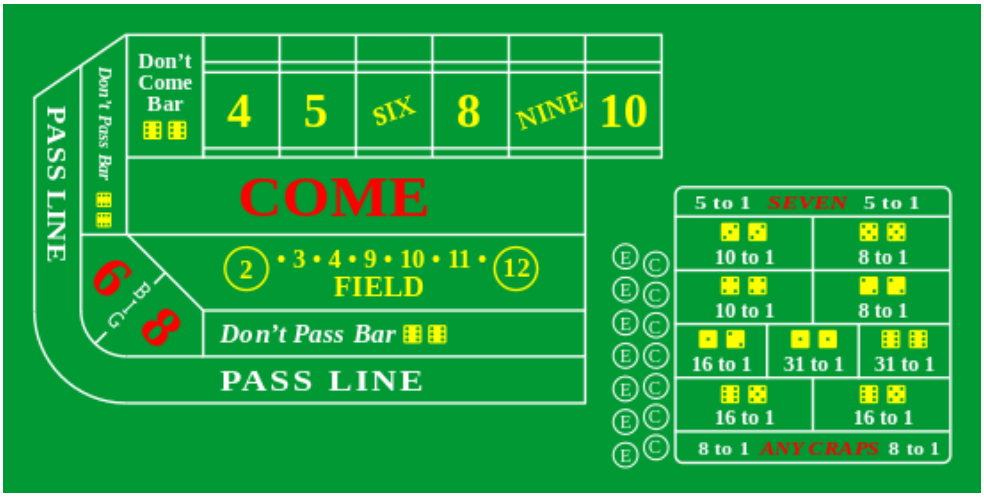
There are 36 possible outcomes

An elemental event can be defined as the total of the two die …

The total number of outcome resulting in each unique event is known.

The probability of each event can be computed and described
… if the die are "fair".

So … the "true odds" can be computed … and a gambling game with skewed odds in "the houses" favor can be created …



from: https://en.wikipedia.org/wiki/Craps

Probability as the Ratio of Favorable to Total Outcomes (Classical Theory) – flipping two coins example

Flip two coins:

What are the possible outcomes {HH, HT, TH, TT}

Define an event as the getting of getting at least one Tail.

Probability is the favorable outcomes/total outcomes, = 3/4

Possible Outcomes with Probabilities:

HH – probability 1/4

HT – probability 1/4

TH – probability 1/4

TT – probability ¼

Possible events: one head, one tail, at least one head, at least one tail, at most one head, at most one tail, two heads, two tails no heads, or no tails.

Probability as a Measure of the Frequency of Outcomes

Experiment: Selecting a sequence of random numbers.

- The random numbers are between 1 and 100.

Determining the relative frequency of a single number as from 01 to 10,000 numbers are selected.

- The statistics of "observed events" is approaching 1/100 …. (infinite trials?)

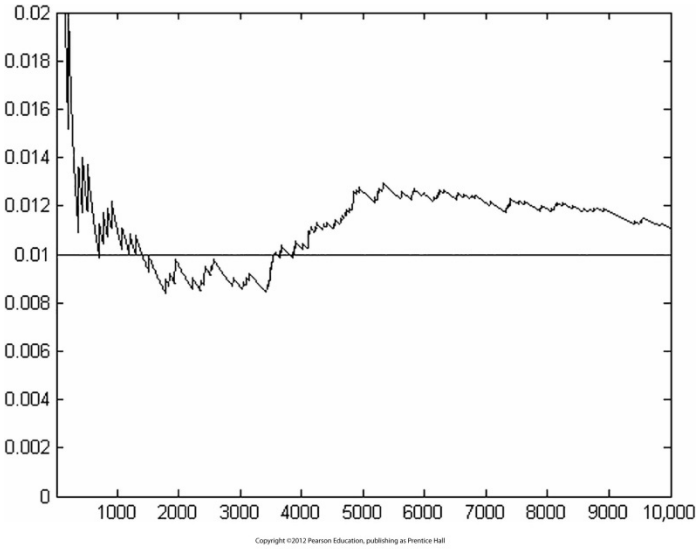


Figure 1.2-1 Event = {occurrence of number 5}

(Numbers derived from website RANDOM.ORG).

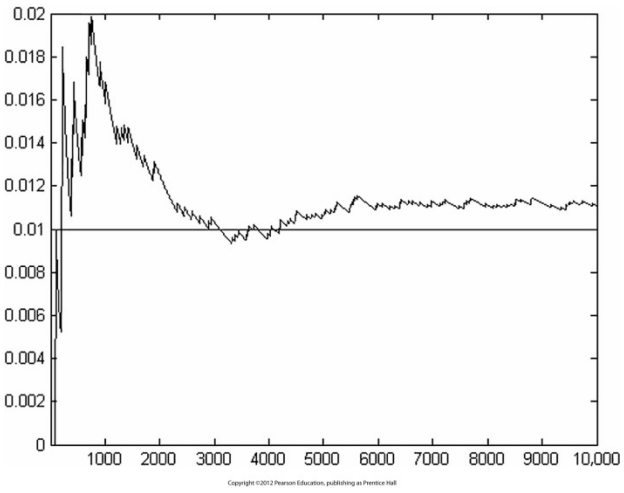


Figure 1.2-2 Event = {occurrence of number 23}

Probability Based on an Axiomatic Theory

Develop the coherent mathematical theory:

- Statistics collected data on random experiments
  - Possible outcomes, sample space, events, etc.
- From the statistics, probability structure can be observed and defined
  - Random processes follow defined probabilistic models of performance.
- Mathematical properties applied to probability derives derive new/alternate expectations
  - Probabilistic expectations can be verified by statistical measurement.

This can be considered as modeling a system prior to or instead of performing an experiment. Note that the results are only as good as the model or "theory" match the actual experiment.

Misuses, Miscalculations, and Paradoxes in Probability

Old time quotation … "There are three kinds of lies: lies, damned lies, and statistics!"
https://en.wikipedia.org/wiki/Lies,_damned_lies,_and_statistics

From the CNN headlines …

"Math is racist: How data is driving inequality", by Aimee Rawlins, September 6, 2016
http://money.cnn.com/2016/09/06/technology/weapons-of-math-destruction/index.html

Another example …

As a person, you are a unique individual and not a statistical probability …

but future "chances" may be based on others like you that have come before.

The class as a whole may exhibit statistical expectations … although it is made up of unique individuals.

For Sci-Fi readers ….

Issac Asimov's Foundation Trilogy – "psychohistory" used to predict the future ….

Sets, Fields and Events

Conceptually Defining a Problem

- Relative Frequency Approach (statistics)
- Set Theory Approach (formal math)
- Venn Diagrams (pictures based on set theory)

If you like "pictures" try to use Venn Diagrams to help understand the concepts.

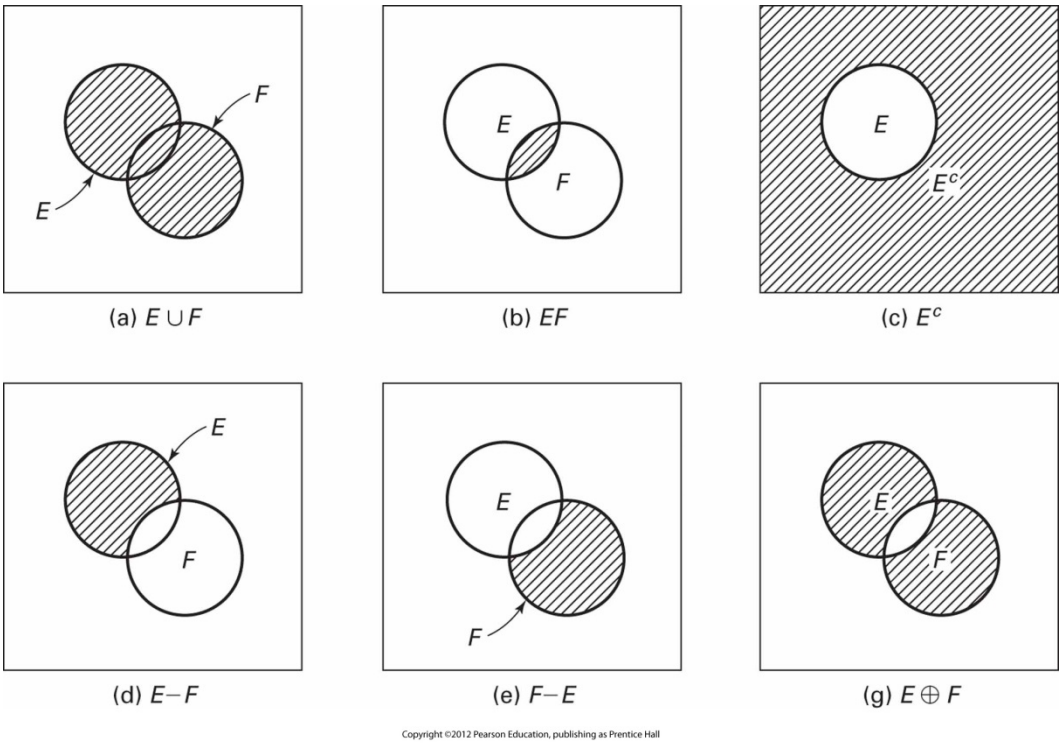


Figure 1.4-1 Venn diagrams for set operations.

Set Theory Definitions – (A review?!)

Set

- A collection of objects known as elements

$$A = \{a_1, a_2, \cdots, a_n\}$$

Subset

- The set whose elements are all members of another set (usually larger but possible the same size).

$$B = \{a_1, a_2, \cdots, a_{n-k}\} \quad \text{therefore} \quad B \subset A$$

Space

- The set containing the largest number of elements or all elements from all the subsets of interest. For probability, the set containing the event description of all possible experimental outcomes.

$$A_i \subset S \text{, for all i subsets}$$

Null Set or Empty Set

- The set containing no elements … $A \subset \varnothing$

Venn Diagrams can help when considering set theory …

- A graphical (geometric) representation of sets that can provide a way to visualize set theory and probability concepts and can lead to an understanding of the related mathematical concepts.

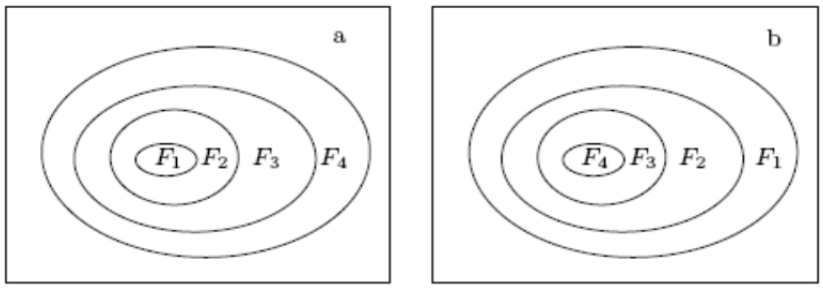


Figure 2.2 (a) Increasing sets, (b) decreasing sets

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. A pdf file version can be found at http://www-ee.stanford.edu/~gray/sp.html

Notes and figures are based on or taken from materials in the course textbook: Probability, Statistics and Random Processes for Engineers, 4th ed., Henry Stark and John W. Woods, Pearson Education, Inc., 2012.

More Set Theory Definitions

Equality

- Set A equals set B if and only if (iff) every element of A is an element of B AND every element of B is an element of A.

$$A = B \text{ iff } A \subset B \text{ and } B \subset A$$

Sum or Union (logic OR function)

- The sum or union of sets results in a set that contains all of the elements that are elements of every set being summed.

$$S = A_1 \cup A_2 \cup A_3 \cdots \cup A_N$$

- Laws for Unions

$$A \cup B = B \cup A$$

$$A \cup A = A$$

$$A \cup \varnothing = A$$

$$A \cup S = S$$

$$A \cup B = A, \text{ if } B \subset A$$

Products or Intersection (logic AND function)

- The product or intersection of sets results in a set that contains all of the elements that are present in every one of the sets.

$$S \cap \varnothing = \varnothing$$

- Laws for Intersections

$$A \cap B = B \cap A$$

$$A \cap A = A$$

$$A \cap \varnothing = \varnothing$$

$$A \cap S = A$$

$$A \cap B = B, \text{ if } B \subset A$$

Mutually Exclusive or Disjoint Sets

- Mutually exclusive or disjoint sets of no elements in common.

$$A \cap B = \varnothing$$

- NOTE: The intersection of two disjoint sets is a set … the null set!

Complement

- The complement of a set is the set containing all elements in the space that are not elements of the set.

$$A \cap \overline{A} = \varnothing \text{ and } A \cup \overline{A} = S$$

- Laws for Complement

$$\overline{\varnothing} = S$$

$$\overline{S} = \varnothing$$

$$\left(\overline{A}\right) = A$$

$$\overline{A} \subset \overline{B}, \text{ if } B \subset A$$

$$\overline{A} = \overline{B}, \text{ if } B = A$$

- DeMorgan's Law

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B}$$

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B}$$

Differences

- The difference of two sets, A-B, is the set containing the elements of A that are not elements of B.

$$A - B = A \cap \overline{B} = A - (A \cap B)$$

- Laws for Differences

$$(A - B) \cup B \neq B$$

$$(A \cup A) - A = \varnothing$$

$$(A - A) \cup A = A$$

$$A - \varnothing = A$$

$$A - S = \varnothing$$

$$S - A = \overline{A}$$

Venn Diagram set theory concepts
. (2D pictures that can help you understand set theory)

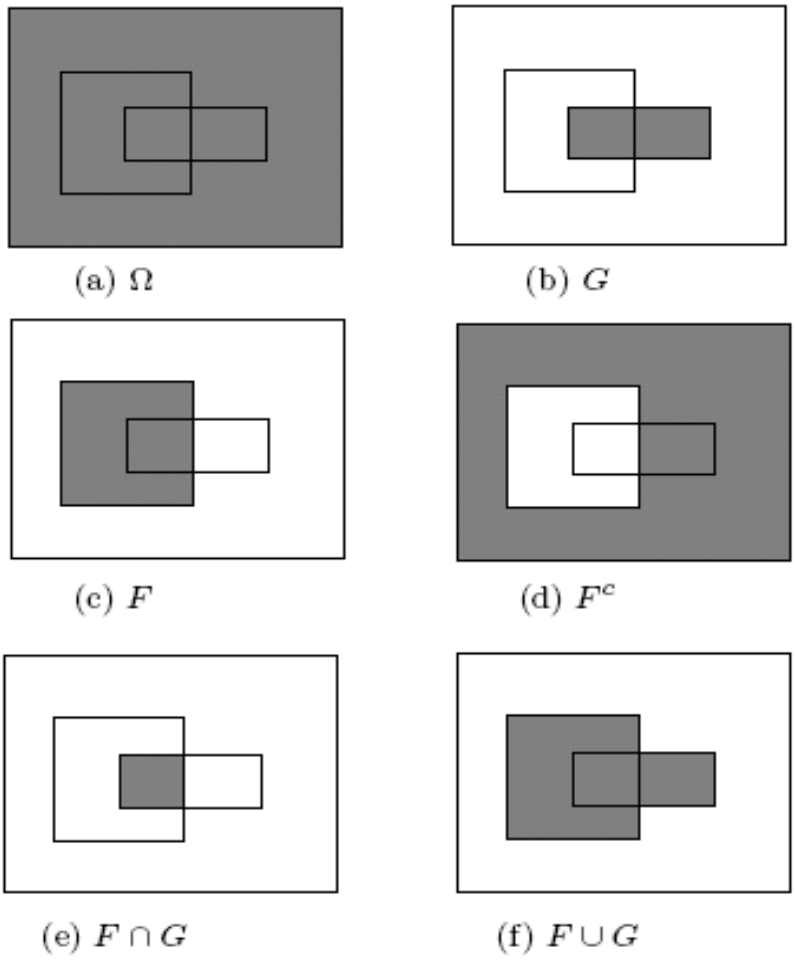
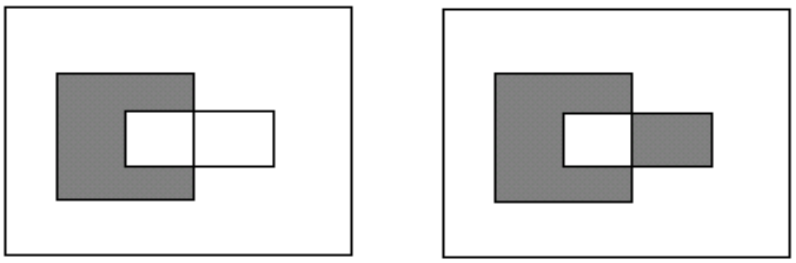


Figure A.1 Basic set operations

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

(a) The space
(b) Subset G
(c) Subset F
(d) The Complement of F
(e) Intersection of F and G
(f) Union of F and G

More Venn Diagrams



(a) $F - G$ (b) $F \Delta G$

Figure A.2 Set difference operations

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

(a) Difference F-G

(b) Difference F-G Union with Difference G-F

$$(F - G) \cup (G - F)$$

If events can be describe in set theory or Venn Diagrams, then probability can directly use the concepts and results of set theory!

What can be said about $\Pr(F \cup G)$? [read as the probability of event F union event G]

$$F \cup G = (F - G) \cup G = F \cup (G - F) = F + G - (F \cap G)$$

Therefore,

$$\Pr(F \cup G) = \Pr(F) + \Pr(G) - \Pr(F \cap G)$$

Set algebra is often used to help define probabilities …

Equalities in Set Algebra

$F \cup G = G \cup F$ commutative law	(A.1)
$F \cup (G \cup H) = (F \cup G) \cup H$ associative law	(A.2)
$F \cap (G \cup H) = (F \cap G) \cup (F \cap H)$ distributive law	(A.3)
$(F^c)^c = F$	(A.4)
$F \cap F^c = \emptyset$	(A.5)
$(F \cap G)^c = F^c \cup G^c$ DeMorgan's "law"	(A.6)
$F \cap \Omega = F$	(A.7)
$F \cap G = G \cap F$ commutative law	(A.8)
$F \cap (G \cap H) = (F \cap G) \cap H$ associative law	(A.9)
$(F \cup G)^c = F^c \cap G^c$ DeMorgan's other "law"	(A.10)
$F \cup F^c = \Omega$	(A.11)
$F \cup \emptyset = F$	(A.12)
$F \cup (F \cap G) = F = F \cap (F \cup G)$	(A.13)
$F \cup \Omega = \Omega$	(A.14)
$F \cap \emptyset = \emptyset$	(A.15)
$F \cup G = F \cup (F^c \cap G) = F \cup (G - F)$	(A.16)
$F \cup (G \cap H) = (F \cup G) \cap (F \cup H)$ distributive law	(A.17)
$\Omega^c = \emptyset$	(A.18)
$F \cup F = F$	(A.19)
$F \cap F = F$	(A.20)

Table A.1 *Set algebra*

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Appendix A, Set Theory. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

Axiomatic Definitions Using Sets

For event A

$$0 \le \Pr(A) \le 1$$

$$\Pr(S) = 1$$

$$\Pr(\varnothing) = 0$$

Disjoint Sets

$$\textit{If } A \cap B = \varnothing, \textit{ then } \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

Complement (complementary sets) (defining the complement may be easier sometimes)

$$\textit{If } A \cap \overline{A} = \varnothing, \textit{ then } \Pr(A \cup \overline{A}) = \Pr(A) + \Pr(\overline{A}) = \Pr(S) = 1$$

$$\Pr(A) = 1 - \Pr(\overline{A}) \le 1$$

Not a Disjoint Sets (solution)

$$\textit{If } A \cap B \neq \varnothing, \textit{ then } \Pr(A \cup B) = ???$$

- Manipulation (1)

$$A \cup B = A \cup (\overline{A} \cap B), \text{ the union of disjoint sets}$$

$$\Pr(A \cup B) = \Pr[A \cup (\overline{A} \cap B)] = \Pr(A) + \Pr(\overline{A} \cap B)$$

- Manipulation (2)

$$B = (A \cap B) \cup (\overline{A} \cap B), \text{ the union of disjoint sets}$$

$$\Pr(B) = \Pr[(A \cap B) \cup (\overline{A} \cap B)] = \Pr(A \cap B) + \Pr(\overline{A} \cap B)$$

- Manipulation (3)

$$\Pr(\overline{A} \cap B) = \Pr(B) - \Pr(A \cap B), \text{ rearranging from (2)}$$

- Substitution for (1)

$$\Pr(A \cup B) = \Pr(A) + \Pr(\overline{A} \cap B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Note that we can generally define a bound where

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

equality holds for A and B being disjoint sets!

Example: 6-sided die

$$\Pr(\alpha_i) = \frac{1}{6}$$

A: The probability of rolling a 1 or a 3, event $A = \{1,3\} = \{1 \cup 3\}$

$$\Pr(A) = \Pr(1) + \Pr(3) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

B: The probability of rolling a 3 or 5, event $B = \{3,5\} = \{3 \cup 5\}$

$$\Pr(B) = \Pr(3) + \Pr(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

C: The probability of event A or event B, event $C = \{A \cup B\}$

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

$$\Pr(C) = \frac{1}{3} + \frac{1}{3} - \frac{1}{6} = \frac{1}{2}$$

Note: $C = \{A \cup B\} = \{1,3,5\}$

$$\Pr(C) = \Pr(1) + \Pr(3) + \Pr(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

When in doubt, write it out to double check your results!

A Venn diagram may also help.

Probability of A union of events (from section 1.5)

An extension of the set theory for unions ….

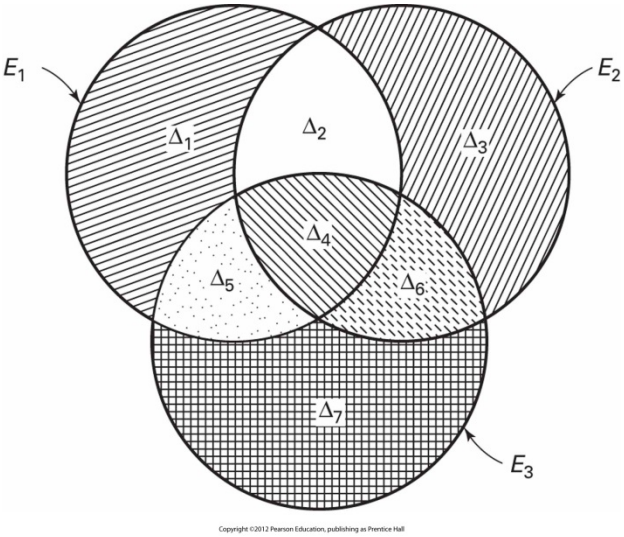


Figure 1.5-1 Partitioning $\bigcup_{i=1}^{3} E_i = \sum_{j=1}^{7} \Delta_j$ into seven disjoint regions Δ1,...,Δ7.)

If $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$, what about $\Pr(A \cup B \cup C) = ???$

$$\begin{aligned}\Pr(A \cup B \cup C) = {} & \Pr(A) + \Pr(B) + \Pr(C) \\ & - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) \\ & + \Pr(A \cap B \cap C)\end{aligned}$$

$$\begin{aligned}\Pr(E_1 \cup E_2 \cup E_3) = {} & \sum_{i=1}^{3} \Pr(E_i) \\ & - \sum_{i=1}^{2} \sum_{j=i+1}^{3} \Pr(E_i \cap E_j) \\ & + \sum_{i=1}^{1} \sum_{j=i+1}^{2} \sum_{k=j+1}^{3} \Pr(E_i \cap E_j \cap E_k)\end{aligned}$$

Can you recognize a pattern …

- "+"singles … "-"doubles … "+"triples … "-"quads … etc

What about $\Pr(A \cup B \cup C \cup D \cup E \cup F) = ???$

Section 1.6: More Definitions

Probability, the *relative frequency method*:

The number of trials and the number of times an event occurs can be described as

$$N = N_A + N_B + N_C + \cdots$$

the relative frequency is then

$$r(A) = \frac{N_A}{N}$$

note that

$$\frac{N}{N} = \frac{N_A + N_B + N_C + \cdots}{N} = r(A) + r(B) + r(C) + \cdots = 1$$

When experimental results appear with "statistical regularity", the relative frequency tends to approach the probability of the event.

$$\Pr(A) = \lim_{N \to \infty} r(A)$$

and

$$\Pr(A) + \Pr(B) + \Pr(C) + \cdots = 1$$

Where $\Pr(A)$ is defined as the probability of event A.

Mathematical definition of probability:

1. $0 \le \Pr(A) \le 1$
2. $\Pr(A) + \Pr(B) + \Pr(C) + \cdots = 1$, for mutually exclusive events
3. An impossible event, A, can be represented as $\Pr(A) = 0$.
4. A certain event, A, can be represented as $\Pr(A) = 1$.

Odds or probabilities can be assigned to every possible outcome of a "future" trial, experiment, contests, game that has some prior historical basis of events or outcomes.

Joint Probability

Defining probability based on multiple events … two classes for considerations.

- Independent experiments: The outcome of one experiment is not affected by past or future experiments.
  - flipping coins
  - repeating an experiment after initial conditions have been restored
  - Note: these problems are typically easier to solve

- Dependent experiments: The result of each subsequent experiment is affected by the results of previous experiments.
  - drawing cards from a deck of cards
  - drawing straws
  - selecting names from a hat
  - for each subsequent experiment, the previous results change the possible outcomes for the next event.
  - Note: these problems can be very difficult to solve (the "next experiment" changes based on previous outcomes!)

Conditional Probability

Defining the conditional probability of event A given that event B has occurred.

Using a Venn diagram, we know that B has occurred … then the probability that A has occurred given B must relate to the area of the intersection of A and B …

$$\Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B), \text{ for } \Pr(B) > 0$$

or

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \text{ for } \Pr(B) > 0$$

For elementary events,

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}, \text{ for } \Pr(B) > 0$$

Special cases for $A \subset B$, $B \subset A$, and $B \cap A = \varnothing$.

- If A is a subset of B, then the conditional probability must be

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)}, \text{ for } A \subset B$$

  Therefore, it can be said that

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)} \geq \Pr(A), \text{ for } A \subset B$$

- If B is a subset of A, then the conditional probability becomes

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1, \text{ for } B \subset A$$

- If A and B are mutually exclusive,

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0}{\Pr(B)} = 0, \text{ for } B \cap A = \varnothing$$

Independence

Two events, A and B, are independent if and only if

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

Independence is typically assumed when there is no apparent physical mechanism by which the two events could depend on each other. For events derived from independent elemental events, their independence may not be obvious but may be able to be derived.

Independence can be extended to more than two events, for example three, A, B, and C. The conditions for independence of three events is

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) \qquad \Pr(B \cap C) = \Pr(B) \cdot \Pr(C) \qquad \Pr(A \cap C) = \Pr(A) \cdot \Pr(C)$$

$$\Pr(A \cap B \cap C) = \Pr(A) \cdot \Pr(B) \cdot \Pr(C)$$

Note that it is not sufficient to establish pair-wise independence; the entire set of equations is required.

For multiple events, every set of events from n down must be verified. This implies that $2^n - (n+1)$ equations must be verified for n independent events.

Important Properties of Independence

Unions – help in simplifying the intersection term – if events are independent!

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

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) \cdot \Pr(B)$$

Independent intersection with a Union

$$\Pr(A \cap (B \cup C)) = \Pr(A) \cdot \Pr(B \cup C)$$

There will be some example problems where you must determine if events are independent in order to solve the problem.

- switch problems in homework and skills

Total Probability

For a space, S, that consists of multiple mutually exclusive events, the probability of a random event, B, occurring in space S, can be described based on the conditional probabilities associated with each of the possible events.

Proof:

$$S = A_1 \cup A_2 \cup A_3 \cdots \cup A_n$$

and

$$A_i \cap A_j = \varnothing, \text{ for } i \neq j$$

$$B = B \cap S = B \cap \left(A_1 \cup A_2 \cup A_3 \cdots \cup A_n\right) = \left(B \cap A_1\right) \cup \left(B \cap A_2\right) \cup \left(B \cap A_3\right) \cdots \cup \left(B \cap A_n\right)$$

$$\Pr(B) = \Pr\left(B \cap A_1\right) + \Pr\left(B \cap A_2\right) + \Pr\left(B \cap A_3\right) \cdots + \Pr\left(B \cap A_n\right)$$

But

$$\Pr\left(B \cap A_i\right) = \Pr\left(B \mid A_i\right) \cdot \Pr\left(A_i\right), \text{ for } \Pr\left(A_i\right) > 0$$

Therefore

$$\Pr(B) = \Pr\left(B \mid A_1\right) \cdot \Pr\left(A_1\right) + \Pr\left(B \mid A_2\right) \cdot \Pr\left(A_2\right) + \cdots + \Pr\left(B \mid A_n\right) \cdot \Pr\left(A_n\right)$$

Remember your math properties: distributive, associative, commutative etc. applied to set theory.

Experiment 1: A bag of marbles, draw 1

A bag of marbles: 3-blue, 2-red, one-yellow

- Objects: Marbles
- Attributes: Color (Blue, Red, Yellow)
- Experiment: Draw one marble, with replacement
- Sample Space: {B, R, Y}
- Probability (relative frequency method)

The probability for each possible event in the sample space is ….

Event	Probability
Blue	3/6
Red	2/6
Yellow	1/6
Total	6/6

This experiment would be easy to run and verify … after lots of trials.

see Matlab Sec1_Marble1.m
ntrials = 6 vs. 600 vs. 6000 (repeat execution a few times)

(Another problem: if we ran 6 trials, what is the probability that we get events that exactly match the probability? 3-Blue, 2-Red, 1 Yellow - a much harder problem)

Experiment 2: A bag of marbles, draw 2

- Experiment: Draw one marble, replace, draw a second marble.
  "with replacement"
- Sample Space: {BB, BR, BY, RR, RB, RY, YB, YR, YY}

Define the probability of each event in the sample space ….

Joint Probability

- When a desired outcome consists of multiple events. (Read the joint probability of events A and B).

$$\Pr(A, B)$$

Statistically Independent Events

- When the probability of an event does not depend upon any other prior events.
  If trials are performed with replacement and/or the initial conditions are restored, you expect trial outcomes to be independent.

$$\Pr(A, B) = \Pr(B, A) = \Pr(A) \cdot \Pr(B)$$

- The marginal probability of each event is not affected by prior/other events.
  The probability of event A given event B occurred is the same as the probability of event A and vice versa.

$$\Pr(A \mid B) = \Pr(A) \quad \text{and} \quad \Pr(B \mid A) = \Pr(B)$$

- Applicable for multiple objects with single attributes and with replacement.

Therefore

1st-rows\2nd-col	Blue	Red	Yellow
Blue	$\left(\frac{3}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{9}{36}$	$\left(\frac{3}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{6}{36}$	$\left(\frac{3}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{3}{36}$
Red	$\left(\frac{2}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{6}{36}$	$\left(\frac{2}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{4}{36}$	$\left(\frac{2}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{2}{36}$
Yellow	$\left(\frac{1}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{3}{36}$	$\left(\frac{1}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{2}{36}$	$\left(\frac{1}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{1}{36}$

Next Concept

Conditional Probability

- When the probability of an event depends upon prior events.
  If trials are performed without replacement and/or the initial conditions are not restored, you expect trial outcomes to be dependent on prior results or conditions.

$$\Pr(A \mid B) \neq \Pr(A) \quad \text{when A follows B}$$

- The joint probability is.

$$\Pr(A, B) = \Pr(B, A) = \Pr(A \mid B) \cdot \Pr(B) = \Pr(B \mid A) \cdot \Pr(A)$$

- Applicable for objects that have multiple attributes and/or for trials performed without replacement.

Experiment 3: A bag of marbles, draw 2 without replacement

- Experiment: Draw two marbles, without replacement
- Sample Space: {BB, BR, BY, RR, RB, RY, YB, YR}

Therefore

1st-rows \ 2nd-col	Blue	Red	Yellow	1st Marble
Blue	$\left(\frac{3}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{6}{30}$	$\left(\frac{3}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{6}{30}$	$\left(\frac{3}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{3}{30}$	$\frac{3}{6}$
Red	$\left(\frac{2}{6}\right)\cdot\left(\frac{3}{5}\right)=\frac{6}{30}$	$\left(\frac{2}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{2}{30}$	$\left(\frac{2}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{2}{30}$	$\frac{2}{6}$
Yellow	$\left(\frac{1}{6}\right)\cdot\left(\frac{3}{5}\right)=\frac{3}{30}$	$\left(\frac{1}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{2}{30}$	$\left(\frac{1}{6}\right)\cdot\left(\frac{0}{5}\right)=\frac{0}{30}$	$\frac{1}{6}$
2nd Marble	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{6}{6}$

Matlab Marble Simulation Examples:

Sec1_Marble1.m

- example to show small versus large number of sample statistics vs. probability

Sec1_Marble2.m

- example to validate probability and/or small versus large number of trials

Sec1_Marble3.m

- example to validate probability and/or small versus large number of trials

Resistor Example: Joint and Conditional Probability

Assume we have a bunch of resistors (150) of various impedances and powers…
Similar to old textbook problems (more realistic resistor values)

	50 ohms	100 ohms	200 ohms	Subtotal
¼ watt	40	20	10	70
½ watt	30	20	5	55
1 watt	10	10	5	25
Subtotal	80	50	20	150

Each object has two attributes: impedance (ohms) and power rating (watts)

Marginal Probabilities: (uses subtotals)

Pr(¼ watt) = 70/150	Pr(½ watt) = 55/150	Pr(1 watt) = 25/150
Pr(50 ohms) = 80/150	Pr(100 ohms) = 50/150	Pr(200 ohms) = 20/150

These are called the marginal probabilities … when fewer than all the attributes are considered (or don't matter).

Joint Probabilities: divided each member of the table by 150!

	50 ohms	100 ohms	200 ohms	Subtotal
¼ watt	40/150=0.266	20/150=0.133	10/150=0.066	70/150=0.466
½ watt	30/150=0.20	20/150=0.133	5/150=0.033	55/150=0.366
1 watt	10/150=0.066	10/150=0.066	5/150=0.033	25/150=0.166
Subtotal	80/150=0.533	50/150=0.333	20/150=0.133	150/150=1.0

These are called the joint probabilities … when all unique attributes must be considered.

(Concept of *total probability* … things that sum to 1.0)

Conditional Probabilities:

When one attributes probability is determined based on the existence (or non-existence) of another attribute. Therefore,

The probability of a ¼ watt resistor given that the impedance is 50 ohm.

Pr(¼ watt given that the impedance is 50 ohms) = Pr(¼ watt | 50 ohms) = 40/80 = 0.50

	50 ohms
¼ watt	40/80=0.50
½ watt	30/80=0.375
1 watt	10/80=0.125
Total	80/80=1.0

Simple math that does not work to find the solution: (they are not independent)

Pr(¼ watt) = 70/150 and Pr(50 ohms) = 80/150

Pr(¼ watt) x Pr(50 ohms) = 70/150 x 80/150 = 56/225 = 0.249 NO!!! Not independent!!

Math that does work

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)} = \frac{40/150}{80/150} = 40/80 = 0.50$$

What about Pr(50 ohms given the power is ¼ watt)

	50 ohms	100 ohms	200 ohms	Total
¼ watt	40/70=0.571	20/70=0.286	10/70=0.143	70/70=1.0

Pr(50 ohms | ¼ watt) = Pr(50 | ¼) = 40/70 = 0.571

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)} = \frac{40/150}{70/150} = 40/70 = 0.571$$

Can you determine?

Pr(100, ½) = Pr(100) =

Pr(50, ½) = Pr(½ | 50) =

Pr(50 | ½) = Pr(1) =

Using the "table" it is rather straight forward …

	50 ohms	100 ohms	200 ohms	Subtotal
¼ watt	40	20	10	70
½ watt	30	20	5	55
1 watt	10	10	5	25
Subtotal	80	50	20	150

Joint Probabilities $\Pr(A \cap B) = \Pr(A, B)$

Pr(100, ½) = Pr(50, ½) =

Conditional Probabilities $\Pr(A \mid B) = \dfrac{\Pr(A \cap B)}{\Pr(B)} = \dfrac{\Pr(A, B)}{\Pr(B)}$

Pr(½ | 100) = Pr(200 | ½) =

Marginal Probability $\Pr(B) = \Pr(B \mid A_1) \cdot \Pr(A_1) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)$

Pr(1) = Pr(100) =

Are there multiple ways to conceptually define such problems? … Yes

- Relative Frequency Approach (statistics)
- Set Theory Approach (formal math)
- Venn Diagrams (pictures based on set theory)

All ways to derive equations that form desired probabilities ….

- The Relative Frequency Approach is the slowest and requires the most work!

A Priori and A Posteriori Probability (Sec. 1.7 Bayes Theorem)

The probabilities defined for the expected outcomes, $\Pr(A_i)$, are referred to as *a priori* probabilities (before the event). They describe the probability before the actual experiment or experimental results are known.

After an event has occurred, the outcome B is known. The probability of the event belonging to one of the expected outcomes can be defined as

$$\Pr(A_i \mid B)$$

or from before

$$\Pr(A_i \cap B) = \Pr(B \mid A_i) \cdot \Pr(A_i) = \Pr(A_i \mid B) \cdot \Pr(B)$$

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B)}, \text{ for } \Pr(B) > 0$$

Using the concept of total probability

$$\Pr(B) = \Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)$$

We also have the following forms

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)}$$

or

$$\Pr(A_j \mid B) = \frac{\Pr(B \mid A_j) \cdot \Pr(A_j)}{\sum_{i=1}^{n} \Pr(B \mid A_1) \cdot \Pr(A_1)} = \frac{\Pr(B \mid A_j) \cdot \Pr(A_j)}{\Pr(B)}$$

This probability is referred to as the *a posteriori* probability (after the event).

It is also referred to as **Bayes Theorem**.

Example

More Resistors

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Subtotal
10 ohm	500	0	200	800	1200	1000	3700
100 ohm	300	400	600	200	800	0	2300
1000 ohm	200	600	200	600	0	1000	2600
Subtotal	1000	1000	1000	1600	2000	2000	8600

What is the probability of selecting a 10 ohm resistor from a random bin?

$$\text{Given Bin marginal probability } \Pr(Bin\#) = \frac{1}{6}$$

$$\Pr(10\Omega \mid Bin1) = \frac{500}{1000} \qquad \Pr(10\Omega \mid Bin2) = \frac{0}{1000} \qquad \Pr(10\Omega \mid Bin3) = \frac{200}{1000}$$

$$\Pr(10\Omega \mid Bin4) = \frac{800}{1600} \qquad \Pr(10\Omega \mid Bin5) = \frac{1200}{2000} \qquad \Pr(10\Omega \mid Bin6) = \frac{1000}{2000}$$

$$\Pr(B) = \Pr(B \mid A_1)\cdot\Pr(A_1) + \Pr(B \mid A_2)\cdot\Pr(A_2) + \cdots + \Pr(B \mid A_n)\cdot\Pr(A_n)$$

$$\Pr(B) = \frac{500}{1000}\cdot\frac{1}{6} + \frac{0}{1000}\cdot\frac{1}{6} + \frac{200}{1000}\cdot\frac{1}{6} + \frac{800}{1600}\cdot\frac{1}{6} + \frac{1200}{2000}\cdot\frac{1}{6} + \frac{1000}{2000}\cdot\frac{1}{6}$$

$$\Pr(B) = \frac{5}{10}\cdot\frac{1}{6} + \frac{0}{10}\cdot\frac{1}{6} + \frac{2}{10}\cdot\frac{1}{6} + \frac{5}{10}\cdot\frac{1}{6} + \frac{6}{10}\cdot\frac{1}{6} + \frac{5}{10}\cdot\frac{1}{6} = \frac{23}{10}\cdot\frac{1}{6} = 0.3833$$

Assuming a 10 ohm resistor is selected, what is the probability it came from bin 3?

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i)\cdot\Pr(A_i)}{\Pr(B \mid A_1)\cdot\Pr(A_1) + \Pr(B \mid A_2)\cdot\Pr(A_2) + \cdots + \Pr(B \mid A_n)\cdot\Pr(A_n)}$$

$$\Pr(Bin3 \mid 10\Omega) = \frac{\Pr(10\Omega \mid Bin3)\cdot\Pr(Bin3)}{\Pr(10\Omega \mid Bin1)\cdot\Pr(Bin1) + \cdots + \Pr(10\Omega \mid Bin6)\cdot\Pr(Bin6)}$$

$$\Pr(Bin3 \mid 10\Omega) = \frac{\frac{2}{10}\cdot\frac{1}{6}}{0.3833} = 0.08696$$

Digital Transmissions

A digital communication system sends a sequence of 0 and 1, each of which are received at the other end of a link. Assume that the probability that 0 is received correctly is 0.90 and that a 1 is received correctly is 0.90. Alternately, the probability that a 0 or 1 is not received correctly is 0.10 (the cross-over probability, β). Within the sequence, the probability that a 0 is sent is 60% and that a one is sent is 40%. [S is Send and R is Receive}

$$\Pr(S_0) = 0.60 \qquad \Pr(S_1) = 0.40$$

$$\Pr(R_0 \mid S_0) = 0.90 = 1-\beta \qquad \Pr(R_1 \mid S_1) = 0.90 = 1-\beta$$

$$\Pr(R_1 \mid S_0) = 0.10 = \beta \qquad \Pr(R_0 \mid S_1) = 0.10 = \beta$$

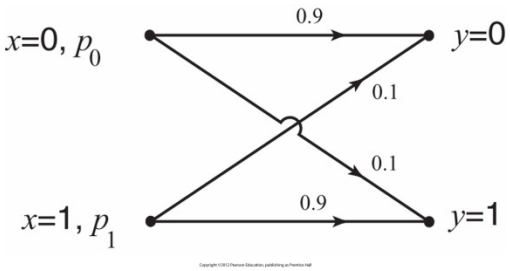


Figure 1.7-1

a) What is the probability that a zero is received?

Total Probability: $\Pr(B) = \Pr(B \mid A_1)\cdot\Pr(A_1) + \Pr(B \mid A_2)\cdot\Pr(A_2) + \cdots + \Pr(B \mid A_n)\cdot\Pr(A_n)$

$$\Pr(R_0) = \Pr(R_0 \mid S_0)\cdot\Pr(S_0) + \Pr(R_0 \mid S_1)\cdot\Pr(S_1)$$

$$\Pr(R_0) = 0.90\cdot 0.60 + 0.10\cdot 0.40$$

$$\Pr(R_0) = 0.54 + 0.04 = 0.58$$

b) What is the probability that a one is received?

$$\Pr(R_1) = \Pr(R_1 \mid S_0)\cdot\Pr(S_0) + \Pr(R_1 \mid S_1)\cdot\Pr(S_1)$$

$$\Pr(R_1) = 0.10\cdot 0.60 + 0.90\cdot 0.40$$

$$\Pr(R_1) = 0.06 + 0.36 = 0.42$$

Digital Communications (continued)

$$\Pr(S_0) = 0.60 \qquad \Pr(S_1) = 0.40$$

$$\Pr(R_0 \mid S_0) = 0.90 \qquad \Pr(R_1 \mid S_1) = 0.90$$

$$\Pr(R_1 \mid S_0) = 0.10 \qquad \Pr(R_0 \mid S_1) = 0.10$$

c) What is the probability that a received zero was transmitted as a 0?

Bayes Theorem

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)}$$

$$\Pr(S_0 \mid R_0) = \frac{\Pr(R_0 \mid S_0) \cdot \Pr(S_0)}{\Pr(R_0)} = \frac{\Pr(R_0 \mid S_0) \cdot \Pr(S_0)}{\Pr(R_0 \mid S_0) \cdot \Pr(S_0) + \Pr(R_0 \mid S_1) \cdot \Pr(S_1)}$$

$$\Pr(S_0 \mid R_0) = \frac{\Pr(R_0 \mid S_0) \cdot \Pr(S_0)}{0.58} = \frac{0.90 \cdot 0.60}{0.58} = \frac{0.54}{0.58} = 0.931$$

d) What is the probability that a received one was transmitted as a 1?

Bayes Theorem

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)}$$

$$\Pr(S_1 \mid R_1) = \frac{\Pr(R_1 \mid S_1) \cdot \Pr(S_1)}{\Pr(R_1)} = \frac{\Pr(R_1 \mid S_1) \cdot \Pr(S_1)}{\Pr(R_1 \mid S_0) \cdot \Pr(S_0) + \Pr(R_1 \mid S_1) \cdot \Pr(S_1)}$$

$$\Pr(S_1 \mid R_1) = \frac{\Pr(R_1 \mid S_1) \cdot \Pr(S_1)}{0.42} = \frac{0.90 \cdot 0.40}{0.42} = \frac{0.36}{0.42} = 0.857$$

Digital Communications (continued)

$$\Pr(S_0) = 0.60 \qquad \Pr(S_1) = 0.40$$

$$\Pr(R_0 \mid S_0) = 0.90 \qquad \Pr(R_1 \mid S_1) = 0.90$$

$$\Pr(R_1 \mid S_0) = 0.10 \qquad \Pr(R_0 \mid S_1) = 0.10$$

e) What is the probability that a received zero was transmitted as a 1?

Bayes Theorem

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)}$$

$$\Pr(S_1 \mid R_0) = \frac{\Pr(R_0 \mid S_1) \cdot \Pr(S_1)}{\Pr(R_0)} = \frac{\Pr(R_0 \mid S_1) \cdot \Pr(S_1)}{\Pr(R_0 \mid S_0) \cdot \Pr(S_0) + \Pr(R_0 \mid S_1) \cdot \Pr(S_1)}$$

$$\Pr(S_1 \mid R_0) = \frac{\Pr(R_0 \mid S_1) \cdot \Pr(S_1)}{0.54 + 0.04} = \frac{0.10 \cdot 0.40}{0.58} = \frac{0.04}{0.58} = 0.069$$

Note: $\Pr(S_1 \mid R_0) = 1 - \Pr(S_0 \mid R_0) = 1 - 0.931 = 0.069$

f) What is the probability that a received one was transmitted as a 0?

Bayes Theorem

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)}$$

$$\Pr(S_0 \mid R_1) = \frac{\Pr(R_1 \mid S_0) \cdot \Pr(S_0)}{\Pr(R_1)} = \frac{\Pr(R_1 \mid S_0) \cdot \Pr(S_0)}{\Pr(R_1 \mid S_0) \cdot \Pr(S_0) + \Pr(R_1 \mid S_1) \cdot \Pr(S_1)}$$

$$\Pr(S_0 \mid R_1) = \frac{\Pr(R_1 \mid S_0) \cdot \Pr(S_0)}{0.42} = \frac{0.10 \cdot 0.60}{0.42} = \frac{0.06}{0.42} = 0.143$$

Note: $\Pr(S_0 \mid R_1) = 1 - \Pr(S_1 \mid R_1) = 1 - 0.857 = 0.143$

Digital Communications (continued)

$$\Pr(S_0) = 0.60 \qquad \Pr(S_1) = 0.40$$

$$\Pr(R_0 \mid S_0) = 0.90 \qquad \Pr(R_1 \mid S_1) = 0.90$$

$$\Pr(R_1 \mid S_0) = 0.10 \qquad \Pr(R_0 \mid S_1) = 0.10$$

e) What is the probability that a symbol is received in error?

$$\Pr(Error) = \Pr(R_0 \mid S_1) \cdot \Pr(S_1) + \Pr(R_1 \mid S_0) \cdot \Pr(S_0)$$

$$\Pr(Error) = 0.10 \cdot 0.40 + 0.10 \cdot 0.60 = 0.04 + 0.06 = 0.10$$

Alternately,

$$\Pr(Error) = \Pr(S_0 \mid R_1) \cdot \Pr(R_1) + \Pr(S_1 \mid R_0) \cdot \Pr(R_0)$$

$$\Pr(Error) = 0.143 \cdot 0.42 + 0.069 \cdot 0.58 = 0.060 + 0.040 = 0.100$$

Which way is easier?

Notice that you were told originally that there was a 0.10 chance of receiving a symbol in error!

Summary:

A-priori Probabilities

$$\Pr(S_0) = 0.60 \qquad \Pr(S_1) = 0.40$$

$$\Pr(R_0 \mid S_0) = 0.90 \qquad \Pr(R_1 \mid S_1) = 0.90$$

$$\Pr(R_1 \mid S_0) = 0.10 \qquad \Pr(R_0 \mid S_1) = 0.10$$

Computed Total Probability

$$\Pr(R_0) = 0.58 \qquad \Pr(R_1) = 0.42$$

Bayes Theorem (A-posteriori Probabilities)

$$\Pr(S_0 \mid R_0) = 0.931 \qquad \Pr(S_1 \mid R_1) = 0.857$$

$$\Pr(S_1 \mid R_0) = 0.069 \qquad \Pr(S_0 \mid R_1) = 0.143$$

Example 1.7-2: Amyloid test: is it a good test for Alzheimer's?

An amyloid test for Alzheimer's disease had reported results/information for people 65 and older.

- Alzheimer's patients with disease = 90% had amyloid protein
- Alzheimer's free patients = 36% had amyloid protein

General population facts for Alzheimer's
- Total Alzheimer's probability = 10%
- Total non-Alzheimer's probability = 1-10% = 90%

The setup – a-priori probabilities (given)

$$\Pr(am \mid Alz) = 0.90 \quad \text{and} \quad \Pr(am \mid nonAlz) = 0.36$$

$$\Pr(Alz) = 0.10 \quad \text{and} \quad \Pr(nonAlz) = 0.90$$

What we want to know – if someone had the amyloid protein, what is the probability they have Alzheimer's?

$$\Pr(Alz \mid am) = ???$$

Using Bayes Theorem

$$\Pr(Alz \mid am) = \frac{\Pr(am \mid Alz) \cdot \Pr(Alz)}{\Pr(am)}$$

But we need to know $\Pr(am)$ … determine the total probability

$$\Pr(am) = \Pr(am \mid Alz) \cdot \Pr(Alz) + \Pr(am \mid nonAlz) \cdot \Pr(nonAlz)$$

$$\Pr(am) = 0.90 \cdot 0.10 + 0.36 \cdot 0.90 = 0.414$$

Therefore

$$\Pr(Alz \mid am) = \frac{0.90 \cdot 0.10}{0.414} = 0.2174$$

The diagnosis is better than 10%, but for completeness … what about the non-Alzheimer's population … too many for a good test

$$\Pr(nonAlz \mid am) = \frac{0.36 \cdot 0.90}{0.414} = 0.7826$$

… too high a probability for a good test

1.8 Combinatorics

Some important math before doing more probability …

From Merriam Webster's Dictionary.
http://www.merriam-webster.com/dictionary/combinatorial

> *Combinatorial - of or relating to the arrangement of, operation on, and selection of discrete mathematical elements belonging to finite sets or making up geometric configurations.*

For a population of size n … the set contains n elements (a deck of 52 playing cards)

A subpopulation of size r can be defined (draw 5 cards at random from the deck)

How many unique subpopulations of r can we expects (notice that the same r elements can be selected in numerous ways).

(i) Sampling with replacement is the easy way …

Possible combinations → $n \cdot n \cdot n \cdot \ldots \cdot n = n^r$

(ii) Sampling without replacement

Possible combinations → $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-r+1) = \dfrac{n!}{(n-r)!}$

Next considerations … how many ways can the r things be selected …

Possible selections → $r \cdot (r-1) \cdot (r-2) \cdot \ldots \cdot 1 = r!$

Now we can consider the "unique" combinations

Unique combinations → $\dfrac{PossibleCombination}{PossibleSelection} = \dfrac{n!}{(n-r)!\,r!}$

We have now defined an operator to determine unique values for "n choose r"

$$C_r^n = \binom{n}{r} = \frac{n!}{(n-r)!\,r!}$$ also sometimes shown as $${}_nC_r = C(n,r) = \frac{n!}{(n-r)!\,r!}$$

This is also called a binomial coefficient. There is also some important "definitions"

$$C_r^n = \binom{n}{0} = \frac{n!}{(n)!\,0!} = 1 \quad \text{and} \quad C_r^n = \binom{n}{n} = \frac{n!}{(n)!\,0!} = 1$$

Why are they called binomial coefficients?

Binomial Powers $(1+x)^n$ or $(x+y)^n$

- the coefficients for the various power resulting from 2 summed elements to the nth power.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^{n-k} \cdot y^k$$

The coefficients can also be selected using Pascal's Triangle also called Binomial Expansion.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1

Each row starts (and ends) with 1 and then sums the adjacent coefficients from the next higher row.

So, by inspection $(a+b)^4 = 1 \cdot a^4 + 4 \cdot a^3 \cdot b^1 + 6 \cdot a^2 \cdot b^2 + 4 \cdot a^1 \cdot b^3 + 1 \cdot b^4$.

Now, if a flip a coin 4 times, what are the possible combinations and how many times do they occur?

What if a said that: a=Heads and b=Tails

1 – H^4 4 – H^3xT^1 6 – H^2xT^2 4 – H^1xT^3 1 – T^4

Also of note $$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Letting x=y=1 $$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^{n-k} \cdot y^k$$

Multinomial Coefficients

Theorem 1.8-1: Let n consists of multiple subsets, each of r_i elements such that

$$n = \sum_{i=1}^{K} r_i$$

The number of ways in which the population of n elements can be partitioned into K subpopulations of which each contains r_i element is

$$\frac{n!}{r_1! \cdot r_2! \cdot r_3! \cdot \ldots \cdot r_K!}$$

5-Card Draw Combinatorial

How many ways can 5 cards be drawn from a deck of 52 playing cards?

$$C_r^n = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\binom{52}{5} = \frac{52!}{(52-5)!5!} = 2{,}598{,}960$$

If you are poker player … see

https://en.wikipedia.org/wiki/Poker_probability

1.9 Bernoulli Trials

A repeated trial can take the form of:

1. Repeated experiments where the relative frequency of occurrence is of interest
2. The creation of a new experiment that consists of a defined number of elementary events

Bernoulli Trials: **Determining the probability that an event occurs *k* times in *n* independent trials of an experiment.**

For some experiment let: $\Pr(A) = p$ and $\Pr(\overline{A}) = q$

where $p + q = 1$

Then for an experiment where we get 2 event "A"s followed by 2 "not A" (i.e., $B = \{AA\overline{A}\overline{A}\}$) …

$$\Pr(B) = \Pr(A)\cdot\Pr(A)\cdot\Pr(\overline{A})\cdot\Pr(\overline{A}) = p^k \cdot q^{n-k}$$

But what about the other ways to have 2 event A's in 4 trials? Note that for each instance, the probability of occurring will be the same as just defined … so how many of them are there?

$$\{AA\overline{A}\overline{A}, A\overline{A}A\overline{A}, A\overline{A}\overline{A}A, \overline{A}AA\overline{A}, \overline{A}A\overline{A}A, \overline{A}\overline{A}AA\}$$

The number of occurrences can be defined using binomial coefficients and the *Binomial Theorem*.

The number of instances is defined by the binomial coefficient, ${}_nC_k$ or $\binom{n}{k}$.

- the number of ways to select k elements out of a set of n elements ...

Where $$\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$$

Therefore, to describe the desired outcome of 2 A's in 4 trials, the probability is

$$\Pr(A\,occuring\,2\,times\,in\,4\,trials) = p_4(2) = \binom{4}{2}\cdot p^2 \cdot q^{4-2} = \frac{4}{2!\cdot(4-2)!}\cdot p^2 \cdot q^{4-2}$$

Therefore …

Bernoulli Trials

The probability that an event occurs k times in n independent trials of an experiment can be defined as

$$\Pr\left(A\,occuring\,k\,times\,in\,n\,trials\right) = p_n(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

Example Flipping Coins

The probability for each outcome of flipping a coin 4 times, where Pr(H)= p and Pr(T)=q with

$$p = q = \frac{1}{2}$$

4 H : $$\Pr(HHHH) = p_4(4) = \binom{4}{4} \cdot p^4 \cdot q^{4-4} = \binom{4}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^0 = 1 \cdot \frac{1}{16} = \frac{1}{16}$$

3 H & 1 T: $$\Pr(HHHT) = p_4(3) = \binom{4}{3} \cdot p^3 \cdot q^{4-3} = \binom{4}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^1 = 4 \cdot \frac{1}{16} = \frac{4}{16}$$

2 H & 2 T: $$\Pr(HHTT) = p_4(2) = \binom{4}{2} \cdot p^2 \cdot q^{4-2} = \binom{4}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^2 = 6 \cdot \frac{1}{16} = \frac{6}{16}$$

1 H & 3 T: $$\Pr(HTTT) = p_4(1) = \binom{4}{1} \cdot p^1 \cdot q^{4-1} = \binom{4}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^3 = 4 \cdot \frac{1}{16} = \frac{4}{16}$$

4 T: $$\Pr(TTTT) = p_4(0) = \binom{4}{0} \cdot p^0 \cdot q^{4-0} = \binom{4}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^4 = 1 \cdot \frac{1}{16} = \frac{1}{16}$$

What if p = 0.6 and q=0.4? An “unfair coin”!!

Example Binary Communications

Example 1: For a bit-error-rate (BER) of 10^{-3} in a binary data stream, what is the probability of 1 error in a 32-bit word?

$$p_{32}(1) = \binom{32}{1} \cdot \left(10^{-3}\right)^1 \cdot \left(1-10^{-3}\right)^{31}$$

$$p_{32}(1) = 32 \cdot 10^{-3} \cdot \left(1-10^{-3}\right)^{31}$$

$$p_{32}(1) \approx 32 \cdot 10^{-3} \cdot 0.9695 \approx 0.0310$$

Example 2: For a bit-error-rate (BER) of 10^{-3} in a binary data stream, what is the probability of 0 errors in a 32-bit word?

$$p_{32}(0) = \binom{32}{0} \cdot \left(10^{-3}\right)^0 \cdot \left(1-10^{-3}\right)^{32}$$

$$p_{32}(0) = 1 \cdot 1 \cdot \left(1-10^{-3}\right)^{32}$$

$$p_{32}(1) \approx 0.9685$$

Example 3: What is that probability of having one or more errors in 32 bits?

$$\sum_{i=1}^{32} p_{32}(i) = \sum_{i=1}^{32} \left[\binom{32}{i} \cdot \left(10^{-3}\right)^i \cdot \left(1-10^{-3}\right)^{32-i} \right]$$

or

$$\sum_{i=1}^{32} p_{32}(i) = 1 - p_{32}(0) = 1 - 0.9685 = 0.0315$$

Power Ball Lottery

The lottery is a 69 choose 5 game combined with a 26 choose 1 game.

see MATLAB code

Power Ball total combinations = 292201338.

Prob of 0+0 balls is 1 in 1.53296 min payout is $0.
Prob of 0+1 balls is 1 in 38.3239 min payout is $4.
Prob of 1+0 balls is 1 in 3.6791 min payout is $0.
Prob of 1+1 balls is 1 in 91.9775 min payout is $4.
Prob of 2+0 balls is 1 in 28.0531 min payout is $0.
Prob of 2+1 balls is 1 in 701.328 min payout is $7.
Prob of 3+0 balls is 1 in 579.765 min payout is $7.
Prob of 3+1 balls is 1 in 14494.1 min payout is $100.
Prob of 4+0 balls is 1 in 36525.2 min payout is $100.
Prob of 4+1 balls is 1 in 913129 min payout is $50000.
Prob of 5+0 balls is 1 in 1.16881e+07 min payout is $1000000.
Prob of 5+1 balls is 1 in 2.92201e+08 min payout is $40000000.

Prob of winning something is 1 in 24.8671.

Expected Winnings per $2 without Jackpot = $0.32
Expected Winnings per $2 min $40M Jackpot = $0.46

Single Winner Break Even Jackpot = $490,936,628.00
Total computed without considering taxes.

Max US tax rate 39.6%, MI tax rate 4.25%)!!
Single Winner Break Even Jackpot with taxes= $848,886,670.24

Example Baseball/Softball Statistics

Example 1: A batter has a 0.250 batting average. What is the probability that the batter gets 1 hit in 4 at bats?

$$\Pr(A\,occuring\,k\,times\,in\,n\,trials) = p_n(k) = \binom{n}{k} p^k \cdot q^{n-k}$$

$$p_4(1) = \binom{4}{1} \cdot (0.25)^1 \cdot (1-0.25)^3$$

$$p_4(1) = \frac{4!}{1!\cdot(4-1)!} \cdot (0.25)^1 \cdot (0.75)^3 = 4 \cdot \frac{1}{4} \cdot \frac{3\cdot 3\cdot 3}{4\cdot 4\cdot 4} = \frac{27}{64} = 0.422$$

Example 2: A batter has a 0.250 batting average. What is the probability that the batter gets 2 hit in 4 at bats?

$$p_4(2) = \binom{4}{2} \cdot (0.25)^2 \cdot (1-0.25)^2$$

$$p_4(2) = \frac{4!}{2!\cdot(4-2)!} \cdot (0.25)^2 \cdot (0.75)^2 = \frac{4\cdot 3}{2} \cdot \frac{1\cdot 1}{4\cdot 4} \cdot \frac{3\cdot 3}{4\cdot 4} = \frac{27}{128} = 0.211$$

Example 3: A batter has a 0.250 batting average. What is the probability that the batter gets at least 1 hit in 4 at bats?

$$p_4(1) + p_4(2) + p_4(3) + p_4(4) = 1 - p_4(0)$$

$$1 - p_4(0) = 1 - \frac{4!}{0!\cdot(4-0)!} \cdot (0.25)^0 \cdot (0.75)^4 = 1 - 1 \cdot \frac{3\cdot 3\cdot 3\cdot 3}{4\cdot 4\cdot 4\cdot 4} = 1 - \frac{81}{256} = \frac{175}{256} = 0.684$$

Example 4: A batter has a 0.250 batting average. What is the probability that the batter gets at most 1 hit in 4 at bats?

$$p_4(0) + p_4(1) = \frac{4!}{0!\cdot(4-0)!} \cdot (0.25)^0 \cdot (0.75)^4 + \frac{4!}{1!\cdot(4-1)!} \cdot (0.25)^1 \cdot (0.75)^3$$

$$p_4(0) + p_4(1) = 1 \cdot \frac{3\cdot 3\cdot 3\cdot 3}{4\cdot 4\cdot 4\cdot 4} + 4 \cdot \frac{1}{4} \cdot \frac{3\cdot 3\cdot 3}{4\cdot 4\cdot 4} = \frac{81}{256} + \frac{108}{256} = \frac{189}{256} = 0.738$$

Defining a player having a hitting slump … how many at bats until it is a slump?

How many at bats would the batter need to take …
to have a 90% (or 99%) probability of getting at least one hit.
Cabrera's average in 2014 ….was 0.313 ? (see Excel Spread Sheet

$$1 - p_m(0) \geq 0.900 \qquad 1 - p_m(0) \geq 0.990$$
$$1 - p_7(0) = 0.9278 \qquad 1 - p_{13}(0) = 0.9924$$

Example 1-10.2 from Cooper-McGillem

In playing an opponent of equal ability, which is more probable:

$$\Pr(A\,occuring\,k\,times\,in\,n\,trials) = p_n(k) = \binom{n}{k} p^k \cdot q^{n-k}$$

a) To win 4 games out of 7, or to win 5 games out of 9?

$$p_7(4) = \binom{7}{4} p^4 \cdot q^{7-4} = \frac{7!}{4!\cdot 3!} \cdot (0.5)^4 \cdot (0.5)^{7-3}$$

$$p_7(4) = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2} \cdot (0.5)^7 = \frac{210}{6} \cdot \frac{1}{128} = 0.2734$$

$$p_9(5) = \binom{9}{5} p^5 \cdot q^{9-5} = \frac{9!}{5!\cdot 4!} \cdot (0.5)^5 \cdot (0.5)^{9-5}$$

$$p_9(5) = \frac{9 \cdot 8 \cdot 7 \cdot 6}{4 \cdot 3 \cdot 2} \cdot (0.5)^9 = \frac{3024}{24} \cdot \frac{1}{512} = 0.2461$$

Therefore, winning 4 out of 7 is more probable.

b) To win at least 4 games out of 7, or to win at least 5 games out of 9.

$$p_7(4) + p_7(5) + p_7(6) + p_7(7) = \left[\binom{7}{4} + \binom{7}{5} + \binom{7}{6} + \binom{7}{7}\right] \cdot (0.5)^7$$

$$p_7(4) + p_7(5) + p_7(6) + p_7(7) = \left[\frac{210}{6} + \frac{42}{2} + \frac{7}{1} + \frac{1}{1}\right] \cdot \frac{1}{128}$$

$$p_7(4) + p_7(5) + p_7(6) + p_7(7) = [35 + 21 + 7 + 1] \cdot \frac{1}{128} = \frac{64}{128} = 0.50$$

$$p_9(5) + p_9(6) + p_9(7) + p_9(8) + p_9(9) = \left[\binom{9}{5} + \binom{9}{6} + \binom{9}{7} + \binom{9}{8} + \binom{9}{9}\right] \cdot (0.5)^9$$

$$p_9(5) + p_9(6) + p_9(7) + p_9(8) + p_9(9) = \left[\frac{3024}{24} + \frac{504}{6} + \frac{72}{2} + \frac{9}{1} + 1\right] \cdot \frac{1}{512}$$

$$p_9(5) + p_9(6) + p_9(7) + p_9(8) + p_9(9) = [126 + 84 + 36 + 9 + 1] \cdot \frac{1}{512} = \frac{256}{512} = 0.50$$

The probabilities are the same! (You should have a 50-50 chance of winning or losing) !

1.10 Asymptotic Behavior of the Binomial Law

For "Bernoulli Trials" or "Binomial Law" ➔ exactly k successes in n trials

$$\Pr(A\,occuring\,k\,times\,in\,n\,trials) = p_n(k) \equiv b(k;n,p) \equiv \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

For "Summation Binomial Law" ➔ k or fewer successes in n trials

$$B(k;n,p) \equiv \sum_{i=0}^{k} b(i;n,p) \equiv \sum_{i=0}^{k} \binom{n}{i} \cdot p^i \cdot q^{n-i}$$

When n gets large, there are some approximations that can be used …

Why approximate, the combinatorial function can cause calculators and computers to loose numerical precision … and produce incorrect results if they produce results at all.

Poisson probability mass function (pmf) and approximation from future chapters …

Conditions: $n >> 1$, $p << 1$, $k << n$, but $n \cdot p = \mu$ is a constant term, then

$$b(k;n,p) = p_n(k) = \binom{n}{k} p^k \cdot q^{n-k} \approx \frac{1}{k!} \cdot \mu^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k}$$

Then, if we consider an infinite number of trials … that is $n \to \infty$

$$b(k;n,p) \approx \frac{1}{k!} \cdot \mu^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} \approx \frac{\mu^k}{k!} \cdot \exp(-\mu)$$

Example 1.10-1. Computer Component Failure

n = number of components = 10,000

p = component failure rate per year $= 10^{-4}$

Assuming the computer fails if 1 or more component fails.

… using the approximation …

What is the probability the computer will still be working one year from now ?

$$b(k;n,p) \approx \frac{1}{k!}\cdot \mu^k \cdot \left(1-\frac{\mu}{n}\right)^{n-k} \approx \frac{\mu^k}{k!}\cdot \exp(-\mu)$$

The probability of 0 failures is …

$$b\left(0;10^4,10^{-4}\right) \approx \frac{1}{0!}\cdot \exp\left(-10^4\cdot 10^{-4}\right) = \frac{1}{1}\cdot \exp(-1)$$

$$b\left(0;10^4,10^{-4}\right) \approx \exp(-1) = 0.368$$

Example 1.10-2. Random points in time

Suppose n independent points (events) are placed randomly in time from 0 to T.

We want to observe the interval for a short period. $0 < t_1 < t_2 < T$. What is the probability of observing exactly k points (events) in the interval?

Gut reactions to the exercise:

For $\tau = t_2 - t_1$ you would expect the number of points = $n \cdot \frac{\tau}{T}$

Setup … binomial $b(k;n,p) \approx \frac{1}{k!}\cdot \mu^k \cdot \left(1-\frac{\mu}{n}\right)^{n-k} \approx \frac{\mu^k}{k!}\cdot \exp(-\mu)$

Letting $p = \cdot\frac{\tau}{T}$ and $n \cdot p = \mu$

$$b(k;n,p) \approx \frac{\mu^k}{k!}\cdot \exp(-\mu) = \left(\frac{n\cdot\tau}{T}\right)\cdot\frac{1}{k!}\cdot\exp\left(-\frac{n\cdot\tau}{T}\right)$$

Application of Poisson

Physics: radioactive decay

Telecommunications: planning the size of a telephone call center or server farm for the internet

Biology: water pollution or organism monitoring

Optics: designing optimal receivers based on photons received per second

Example 1.10-4. Web Server

On the average, assume there are 16 access request per minute. If the server can handle at most 24 accesses per minute, what is the probability that in any one minute interval that the web site would be saturated?

$$b(k;n,p) \approx \frac{(\lambda \cdot \tau)^k}{k!} \cdot \exp(-\lambda \cdot \tau)$$

$$n \cdot p = \mu = \lambda \cdot \tau = 16 \quad \text{and} \quad \text{k} = 0 \text{ to } 24$$

$$B(nosaturation) \approx \sum_{k=0}^{24} \frac{(\lambda \cdot \tau)^k}{k!} \cdot \exp(-\lambda \cdot \tau) = \sum_{k=0}^{24} \frac{(16)^k}{k!} \cdot \exp(-16)$$

$$B(saturation) \approx \sum_{k=25}^{\infty} \frac{(\lambda \cdot \tau)^k}{k!} \cdot \exp(-\lambda \cdot \tau) = \sum_{k=25}^{\infty} \frac{(16)^k}{k!} \cdot \exp(-16)$$

See Matlab solution … it does not agree with the textbook result?!

The prob of satuaration is 0.0223155.

The prob of no satuaration is 0.977685.

1.11 Normal Approximation to the Binomial Law (DeMoivre-Laplace Theorem)

This is an approximation of the binomial distribution when the number of trials (n of n choose k) is large and other assumptions are met.

Assumptions: $n \cdot p \cdot q >> 1$ and $|k - n \cdot p| <= \sqrt{n \cdot p \cdot q}$

$$p_n(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k} \cong \frac{1}{\sqrt{2 \cdot \pi \cdot n \cdot p \cdot q}} \cdot \exp\left(\frac{-(k - n \cdot p)^2}{2 \cdot n \cdot p \cdot q} \right)$$

[Aside: we will be discussing the law of large numbers and that sums of larger numbers of events appear as a Gaussian distribution. This is the first example … and you haven't been told what a Gaussian distribution is yet.]

Text example: 100 coin tosses, equally likely head or tail, probability of k heads.

$$p_{100}(k) = \binom{100}{k} \cdot 0.5^{100} \cong \frac{1}{\sqrt{50 \cdot \pi}} \cdot \exp\left(\frac{-(k - 50)^2}{50} \right)$$

Note: assumptions valid for 40 <= k <= 60.

The probability for exactly 50 heads would be …

$$p_{100}(k = 50) = \binom{100}{50} \cdot 0.5^{100} \cong \frac{1}{\sqrt{50 \cdot \pi}} \cdot \exp\left(\frac{-(50 - 50)^2}{50} \right) = \frac{1}{\sqrt{50 \cdot \pi}}$$

ECE Applications of Bernoulli Trials

(1) Bit errors in binary transmissions:

Degree of error detection and correction needed. The theoretical validation of performance of the system after "extra bits" for error correction have been added.

- bit-error-rate may also increase if a greater bandwidth is needed because of the "extra bits"

(2) Radar (or similar) signal detection:

After setting a signal detection threshold, the expected signal should be above the threshold when being received for a fixed number of sample times. If the signal is above the threshold for m (or more) of n sample periods, one may also say the signal has been detected.

$$\Pr(Detection) = \sum_{k=m}^{n} p_s(n,k) = \sum_{k=m}^{n} \binom{n}{k} \cdot {p_s}^k \cdot (1-p_s)^{n-k}$$

One can also define a noise threshold where the noise should not be above a particularly level more than m (or more) of n time samples.

$$\Pr(False\_Alarm) = \sum_{k=m}^{n} p_a(n,k) = \sum_{k=m}^{n} \binom{n}{k} \cdot {p_a}^k \cdot (1-p_a)^{n-k}$$

(3) System reliability improvement using redundancy.

If a unit has a known failure rate, by incorporating redundant units, the system will have a longer expected lifetime.

Important when dealing with systems that cannot be serviced, systems that may be very expensive to service, systems that require very high reliability, system with components with high failure rates, etc. . (e.g. satellites, computer hard-disk farms, internet order entry servers).

Defining the probability that one of the redundant elements is still working …

$$\Pr(Functional) = 1 - \Pr((All\_Failed)$$

Hypergeometric Distribution – related information

From: http://en.wikipedia.org/wiki/Hypergeometric_distribution

In probability theory and statistics, the hypergeometric distribution is a discrete probability distribution (*probability mass function*) that describes the number of successes in a sequence of n draws from a finite population without replacement.

A typical example is the following: There is a shipment of N objects in which D are defective. The hypergeometric distribution describes the probability that in a sample of n distinctive objects drawn from the shipment exactly x objects are defective.

$$\Pr(x = X, N, D, n) = \frac{\binom{D}{x} \cdot \binom{N-D}{n-x}}{\binom{N}{n}}$$

for $$\max(0, D+n-N) \le x \le \min(n, D)$$

The equation is derived based on a non-replacement Bernoulli Trials …

Where the denominator term defines the number of trial possibilities, the 1st numerator term defines the number of ways to achieve the desired x, and the 2nd numerator term defines the filling of the remainder of the set.

Quality Control Example

A batch of 50 items contains 10 defective items. Suppose 10 items are selected at random and tested. What is the probability that exactly 5 of the items tested are defective?

The number of ways of selecting 10 items out of a batch of 50 is the number of combinations of size 10 from a set of 50 objects:

$$C_{10}^{50} = \binom{50}{10} = \frac{50!}{10!\cdot 40!}$$

The number of ways of selecting 5 defective and 5 nondefective items from the batch of 50 is the product N_1 x N_2 where N_1 is the number of ways of selecting the 5 items from the set of 10 defective items, and N_2 is the number of ways of selecting 5 items from the 40 nondefective items.

$$C_5^{10} \cdot C_5^{40} = \left(\frac{10!}{5!\cdot 5!}\right)\cdot\left(\frac{40!}{5!\cdot 35!}\right)$$

Thus the probability that exactly 5 tested items are defective is the desired ways the selection can be made divided by the total number of ways selection can be made, or

$$\frac{C_5^{10} \cdot C_5^{40}}{C_{10}^{50}} = \left(\frac{10!}{5!\cdot 5!}\right)\cdot\left(\frac{40!}{5!\cdot 35!}\right) \Big/ \left(\frac{50!}{10!\cdot 40!}\right) = \frac{252\cdot 658008}{10272278170} = 0.01614$$

Another Use: From "The Minnesota State Lottery" – a better description than Michigan

There are *N* objects in which *D* are of interest. The hypergeometric distribution describes the probability that in a sample of *n* distinctive objects drawn from the total set exactly *x* objects are of interest.

Lotteries …

N= number of balls to be selected at random
D = the balls that you want selected
n = the number of balls drawn
x = the number of desired balls in the set that is drawn

https://www.mnlottery.com/games/figuring_the_odds/hypergeometric_distribution/

Example: Michigan's Classic Lotto 47

Prize Structure For Classic Lotto 47 (web site data)

Match	Prize	Odds of Winning
6 of 6	Jackpot	1 in 10,737,573
5 of 6	$2,500 (guaranteed)	1 in 43,649
4 of 6	$100 (guaranteed)	1 in 873
3 of 6	$5 (guaranteed)	1 in 50

Overall Odds: 1 in 47

Matlab Odds

Match	Odds of Winning 1 in	Percent Probability
6 of 6	**10737573**	**$<1x10^{-5}$%**
5 of 6	**43648.7**	**0.0023%**
4 of 6	**872.97**	**0.1146%**
3 of 6	**50.36**	**1.9856%**
2 of 6	7.07	14.1471%
1 of 6	2.39	41.8753%
0 of 6	2	41.8753%

Chance of winning 2.1024%
ROI per dollar without jackpot ~ $0.2711

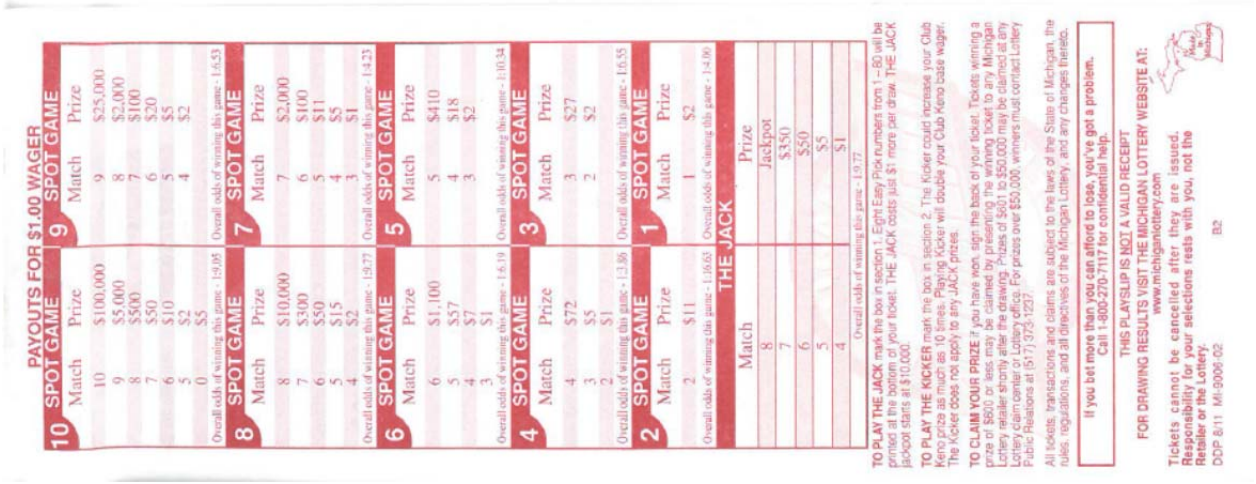
see Matlab simulation "MI_Lotto.m"

Matlab Note: binomial coefficient = nchoosek(n,k)

Keno anyone? Another Michigan gambling game 80 balls, 20 drawn, you need to match k of n selected for n=2 to 20.

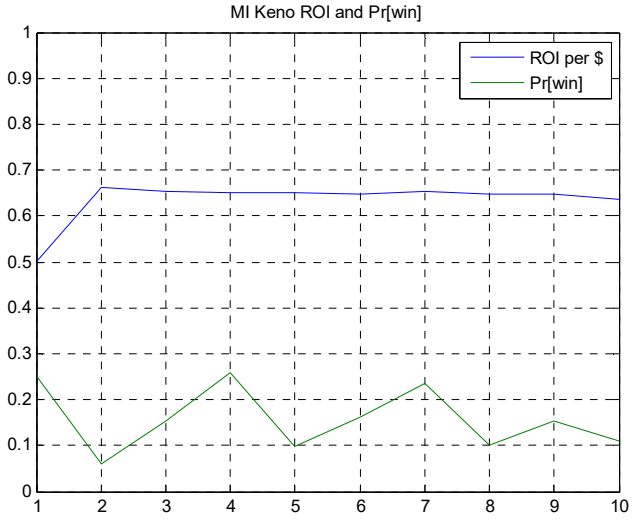
MI Keno

A Keno ticket with the payouts is shown!



Another hypergeometric density function

N= number of balls to be selected at random (80)
D = the balls that you want selected (D)
n = the number of balls drawn (20)
x = the number of desired balls in the set that is drawn (0:D)



In general, you get $0.65 back for every $1 played. I did not include a "kicker" bet. The overall odds of a Kicker (1, 2, 3, 4, 5, 10) number being 2 or higher are 1:1.67. see MI_Keno.m on the web site for more information and results.

Examples for Chapter 1 … more notes on line.