Lecture 2: Discrete Distributions, Normal Distributions

Chapter 1

Reminders

- Course website:
  - www. stat.purdue.edu/~xuanyaoh/stat350
- Office Hour: Mon 3:30-4:30, Wed 4-5
- Bring a calculator, and copy Tables I – III.
- Start Hw#1 now.
  - Due by beginning of Next Fri class

Exercise 1

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Is this a density? Check the 3 properties

Exponential Distribution

- In fact, any variable *x* with density:

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

is said to have an exponential distribution.
Important!!!

- Exponentials model the "lifetime" or "lifespan" of many real life phenomena.

Terminology

- If variable x has an exponential density function $f(x)$, we say that x is exponentially distributed
- Or x has an exponential distribution

- What about a uniform density (see example (1) on Pg 29, Notes)? Normal? Etc.
  - *We say x is uniformly/normally/other-ly distributed*
  - *Or x has a normal/uniform/other distribution*

Discrete Distributions

- Discrete variables are treated similarly but are called mass functions instead of densities
- Example: toss a (fair) dice
  - X can take any discrete value 1, 2, 3, 4, 5, or 6
  - Suppose you can throw a dice forever, you can imagine that you will get each number 1/6 of the time
  - The mass function will be a table, instead of a curve.
- What is the mass function of tossing a single dice?

Answer:

x	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Mass functions

- Similar to density functions, the mass function follows 3 properties:

1. $p(x) \geq 0$
2. $\sum p(x) = 1$ – Summation over all possible x values
3. For any two numbers a and b with $a < b$,

   the proportion of values between a and b (inclusive)

   $= p(a) + \ldots\ldots + p(b)$

Another example—tossing a coin

- Suppose you toss a coin 10 times. Let x = the number of heads in 10 tosses.
  - What are the possible values of x?
  - What is the mass function? (We'll come back to this later)
  - Here x actually follows a Binomial Distribution
    - x has a Binomial mass function
    - x is Binomially distributed

Specific distributions

- We now look at several important distributions

- Continuous
  - Normal
- Discrete
  - Binomial
  - Poisson

1.4
Normal distribution

- Back to continuous distributions...
- A very special kind of continuous distribution is called a Normal distribution. It's density function is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{where } \sigma > 0, \ -\infty < x < \infty$$

- where μ and σ are specific parameters of the function.

Normal distribution

- Most widely encountered distribution: lots of real life phenomena such as errors, heights, weights, etc
- Chapter 5: how to use the normal distribution to approximate many other distributions (Central Limit Theorem)
  - Particularly useful when using sums or averages!

Normal Density Function

- Verifies 2 properties
  - *f(x)* is indeed nonnegative
  - Area under the curve is indeed 1 (can't integrate normally but it does integrate to 1)
- Bell-shaped and Unimodal
- Centered at μ
- σ controls the spread
  - Larger σ, wider distribution
  - Smaller σ, taller and narrower
  - Distance from μ to point of inflection

Finding probabilities for normal data

- Tables for a normal distribution with $\mu = 0$ and $\sigma = 1$ are available
- First learn how to find out different probabilities for the the standard normal
- Then we'll learn to convert **ANY** normal distribution to a standard normal and find the corresponding probability

Standard Normal Distribution

- Gets special "letter", *z* or *z*-score
- Always has $\mu = 0$ and $\sigma = 1$, so:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \text{where } -\infty < x < \infty$$

- Again, we can't integrate but we have the Z table that gives us probabilities for specific areas of the *z*-curve.
  - See table I or the front cover of the text.

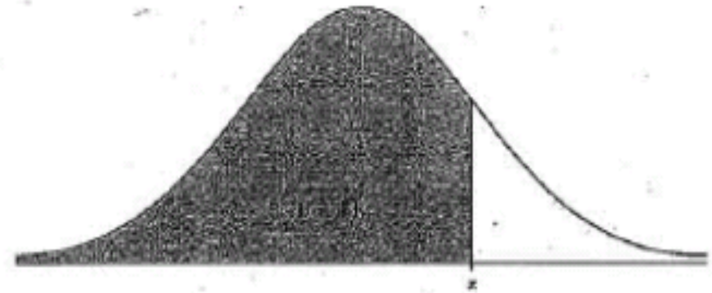

TABLE A: STANDARD NORMAL PROBABILITIES (CONTINUED)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Appendix I (page 560 - 561)

Examples

- What proportion of observations on a standard normal variable Z take values

  - less than 2.2 ?

  .9861, or say, 98.61%

  - greater than -2.05 ?

  97.98%

What about backwards?

- If I give you a probability, can you find the corresponding z value?

  → called percentiles

  - What is the *z*-score for the 25th percentile of the N(0,1) curve?

-0.67

  - 90th percentile?

1.28

Standardizing

- We can convert any normal to a standard normal distribution
- To do this, just subtract the mean and divide by the standard deviation
- z-score – standardized value of x (how many standard deviations from the mean)

$$z = \frac{x - \mu}{\sigma}$$

Standardizing

- Put differently…
- Suppose we want the area between a and b for *x*
- This is exactly the same area between a* and b* for *z*,
  - where a* is the a standardized and b* is b standardized

$$\int_{a}^{b} f(x)dx = \int_{a^*}^{b^*} f(z)dz$$

Standard Normal Distribution

- The **standardized values** for any distribution always have mean 0 and standard deviation 1.
- If the original distribution is normal, the standardized values have normal distribution with mean 0 and standard deviation 1
- Hence, the *standard normal distribution* is extremely important, especially it's corresponding Z table.
  - Remember we can do this forward or backward (using percentiles)

Practice

- In 2000 the scores of students taking SATs were approximately normal with mean 1019 and standard deviation 209. What percent of all students had the SAT scores of:
  - at least 820? (limit for Division I athletes to compete in their first college year)

    82.89%

  - between 720 and 820? (partial qualifiers)

    9.47%

  - How high must a student score in order to place in the top 20% of all students taking the SAT?

    1195

  - Berry's score was the 68th percentile, what score did he receive?

    1117

Connection between Normal Distribution and Discrete Populations …

- Self reading: page 40-41 in text
- Hw question in section 1.4

When you go home

- Review sections 1.3 (mass function) and 1.4, and the last part of section 1.4 “The normal Distribution and Discrete Populations”
- Self study: section 1.5 (not covered in exams)
- Hw#1 and Lab#1
  - due by the <u>beginning of next Friday</u>
- Read sections 1.6 and 2.1