Probability

Chapter 6

Class Outline – 7-14-08

- Turn in Problem Set #1
- Exam Review
- Probability – Chap. 6
- Workshop #2

Exam Tomorrow!

- Chapters 1-5
- Chapter 6 (today's material) will not be on this exam!
- I will not give you the definitional formula for standard deviation (make sure you learn it!)
- Bring a calculator – show all work.
- Questions now and after class

Chap. 6 - Probability

- Inferential statistics (using a sample to infer something about a population) use the concept of probability.
- Probability is used to predict what kind of samples are likely to be obtained from a population
  - In this population, have a 50% chance of drawing a black or white marble
  - But, if most of the marbles were black, have a much greater probability of drawing a black marble

Probability

- Definition of probability
  - In a situation where several different outcomes are possible, we define the probability for any particular outcome as a fraction or proportion of all the possible outcomes
  - Probability of A occurring = $\frac{\text{\# of outcomes classified as A}}{\text{total \# of possible outcomes}}$
  - Probability values can be expressed as a decimal or fraction (same thing)

Probability

- Example – probability of a coin landing on tails when flipped
  - Probability of tails = $\frac{\text{\# classified as tails (1)}}{\text{total \# of possible outcomes (2)}}$
  - We would say the probability of getting tails is 1 over 2 which is equal to .50 or 50%
- Example – rolling dice
  - Probability of rolling a 2 = $\frac{\text{\# classified as a 2 (1)}}{\text{total \# possible outcomes (6)}}$
  - Probability of rolling an even # = $\frac{\text{\# classified as even (3)}}{\text{total \# possible outcomes (6)}}$

Random Sampling

- It is necessary to use random sampling for this definition of probability to hold for samples from populations (and for inferential statistics to be effective)
- Random Sampling:
  - 1) Each individual in the population must have an *equal chance* of being selected
  - 2) If more than one individual is to be selected for a sample, there must be *constant probability* for each and every selection

Random Sampling

- Every individual must have an equal chance or be equally likely to be chosen (if this is not true, probability will not be accurate)
- Constant probability states that you must replace each sample before you select from the population again (sampling with replacement)
  - Otherwise 1/100 becomes 1/ 99, 1/98, etc.
- It is, however, common to define random sampling without having to replace samples (in some experiments it is impossible to replace samples)
  - True random sampling difficult in research context

Probability Example

- Population = 3, 3, 5, 6, 7, 8, 9, 9, 9, 10
- p(X=3)?
- p(X<9)?
- p(X= odd #)?
- Can visualize this by drawing a frequency distribution histogram and defining probability as *area*

The Normal Distribution

- Because the normal distribution is ALWAYS the same, we can apply probability theory to it.
- 68.26% of all scores in a normal distribution are located between – 1 and +1 standard deviations above and below the mean (or have a z-score ranging from -1 to +1)
- These proportions apply to any normal distribution regardless of the mean or standard deviation

The Normal Distribution

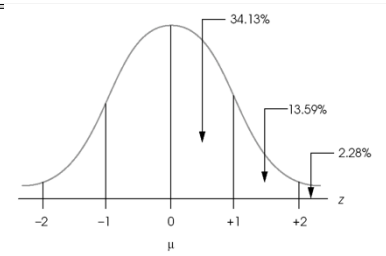


Figure 6.4 (p. 135)
The normal distribution following a *z*-score transformation

Example 6.2, p. 136

- *p*(X>80)?

The Unit Normal Table

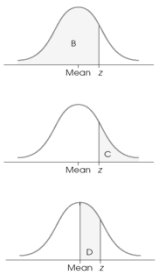
- The score you're calculating won't always fall neatly at SD = +/– 1 or 2
- The unit normal table (z table) provides a complete listing of z-scores and tells you what proportion of a normal distribution lies to the right and to the left of each z-score
- Appendix B table B.1

The Unit Normal Table

- Using the table:
  - The larger portion of the distribution is known as the **body** while the smaller portion is known as the **tail**
    - (If the z-score is positive, the body will be to the left of the z-score and the tail to the right. If the z-score is negative, the body will be to the right of the z-score and the tail to the left)
  - Example and closer look at unit normal table (1st column is z-score, 2nd is proportion in body, 3rd is proportion in tail, and 4th is proportion between mean and z-score)

The Unit Normal Table

(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
0.00	.5000	.5000	.0000
0.01	.5040	.4960	.0040
0.02	.5080	.4920	.0080
0.03	.5120	.4880	.0120
...	...	...	...
0.21	.5832	.4168	.0832
0.22	.5871	.4129	.0871
0.23	.5910	.4090	.0910
0.24	.5948	.4052	.0948
0.25	.5987	.4013	.0987
0.26	.6026	.3974	.1026
0.27	.6064	.3936	.1064
0.28	.6103	.3897	.1103
0.29	.6141	.3859	.1141
0.30	.6179	.3821	.1179
0.31	.6217	.3783	.1217
0.32	.6255	.3745	.1255
0.33	.6293	.3707	.1293
0.34	.6331	.3669	.1331



The Unit Normal Table

- Because normal distribution is symmetrical, the proportions on the right side are the same as the proportions on the left side (same proportions for z-score of +2.00 as -2.00)
- So, although z can be + or – , proportion will always be positive, and column C will be the proportion in the tail regardless if it's on the right or left

The Unit Normal Table

- In addition to finding a proportion if you know a z-score, you can also find a z-score if you know a proportion
- Example – using Appendix B.1, what z-score separates the top 5% from the bottom 95%?

The Unit Normal Table

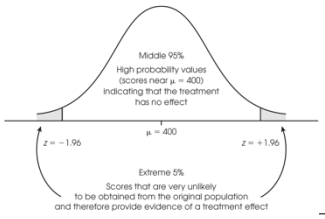
- What if we wanted to know the probability of finding a score that was between two specific scores?
  - Find z-score for each boundary
  - GRE example
- What if we had a probability and we wanted to find a specific X value associated with that probability?
  - work backwards to find z-score associated with probability
  - then get X value from z-score)

The Unit Normal Table

- Let's say I want to join MENSA – what IQ score would I need to achieve in order to score at or above the 98th percentile?
- $\mu = 100,\ \sigma = 15$
- $z = \frac{X - \mu}{\sigma}$

Looking Ahead

- In inferential statistics, we want to compare a treatment group to a control group (representative of the g



Questions?

Exam questions?
Workshop #2