MATH 3070 Introduction to Probability and Statistics Lecture notes The Standard Normal Distribution

Objectives:

- 1. Review concepts of mean, standard deviation, and standardization
- 2. Apply these concepts to the normal distribution
- 3. Find the area under the standard normal distribution

The Normal Probability Distribution

In most of statistics we refer to the "standard normal" curve, or distribution. This is the most useful and common distribution. From our previous discussions we were focusing on discrete random variables. The normal is a **continuous** distribution where the random variable can take on an infinite number of values. By applying the concepts of mean, standard deviation, and standardization we can covert and statistic to one that fits the normal distribution.

It is important to note that the terms **percentage**, **proportion**, and **probability** are interchangeable. For any of these concepts we can use the standard normal distribution to answer our questions.

Mean, Standard Deviation, and Standardization

We will forego a discussion of mean and standard deviation. Those concepts are assumed to be easily understood and recalled. (See Sections 2.4 through 2.6). We will, however, revist the concept of standardization.

The phrase, "Comparing apples to oranges," is used to imply that two things being compared really can't be. An example would be SAT to ACT scores. The tests have different measuring scales and it's hard to know if a score of 620 on the SAT verbal section is equal to a score of 18 on the ACT verbal. So how can we do this?

If the data being compared come from normal distributions, we can transform the data so that the two sets are equivalent. We call this transformation **standardization**. By doing this we change the values from their original units into standard deviation units. We also change the starting measuring point to be the mean of the distribution and convert that value to zero.

The result of this operation is called a **z**-score. The z-score tells us how far above or below the mean (in units of standard deviations) an observation is. It also allows us to compare the values since both now are measured the same (in terms of standard deviations) and measured from the same starting point (the mean, zero). Taking advantage of the fact that the area under the normal curve is equal to one (1), we can measure the percentage area for each value.

The formula for this transformation is

$$z = \frac{x - \bar{x}}{s}$$

Characteristics of the Standard Normal Curve

- 1. The total are under the normal curve is equal to 1.
- 2. The distribution is mounded and symmetrical and extends infinitely in both directions along the x axis.
- 3. The distribution has a mean of 0 and a standard deviation of 1. (This is sometimes written as N(0, 1) to indicate the standard normal.)
- 4. The mean divides the area in half.
- 5. Nearly all area is between z = -3.00 and z = 3.00.

Area Under the Curve

To compute the area, or percentage, at a given z-score, we must always remember that the mean (0) is at 50%. This means that any value to the left (below) the mean will have less than 50% area below it and any value to the right (above) the mean will have more than 50% area below it. The area above the z-score will be the inverse.

We find the area under the curve using a table. This table can be either one sided or two sided, depending upon the author. Ours is one sided. (Johnson and Kuby) In this table, we already know that one half (0.50) of the area is accounted for so the table shows us only the area for the upper half.

First we compute the z-score to two decimal places. Then we look at the table. The column lists the values for the gross measurement, the one's digit and the tenth's. The columns list the fine measurement, the hundredth's. Where the two intersect is the area under the curve for that z-score.