<u>Normal Distribution and Z-Scores Introduction</u>

In life, researchers collect data sets to examine probabilities of all sorts of topics such as height, weight, and even shoe size. When they collect these samples they form what is known as a normal distribution.

A normal distribution is when the values of that data set are centered around the median, forming what appears to be a bell, thus giving it the name of a “bell curve”. The curve is completely symmetrical and has the mean of the data set in the middle. From that mean point follows a number of standard deviations, a given number of units that the data is separated by. A standard deviation normally is marked as 1. 1 unit away from the mean, both in the negative and positive direction.

One way to remember this layout is through the “Empirical Rule”. We must think of a normal distribution as ranging from 1% to 100%, so this rule says that 68% of data points will be 1 standard deviation away from the mean, these are the points mainly concentrated in the center. Then, 95% are 2 standard deviations away, and later 99.7% are 3 standard deviations away. Showing that the farther you get from the mean, the less likely that data point is likely to occur.

Now, when researchers need to look at a specific data point on the distribution, they need to calculate what is known as a z-score. The z-score displays the relationship between that point of interest and the mean of the entire data set. This comes in handy when looking at two different sets of data that do not line up perfectly. Say a teacher needed to compare the midterm grades to the final grades even though they are on a different grading scale, calculating a z-score would allow them to compare a student’s midterm grade to their final grade without worrying about the difference in the total test points.

For example, a college entrance example has a mean of 85 points, and a standard deviation of 3. If a student receives a 92 on the exam, we could calculate their z-score to see how they did in comparison to all other scores. If they were to get a z-score of +2.3, we could use the z-table to see how much better that student did compare to all other scores.

A negative z-score would mean that a specific data point falls below the mean, while a positive would mean it is above the mean, and a zero would mean it is equal to the mean.

This then allows researchers to see how one data point can compare to the rest, or even compare data points from different data sets.

Pretest/Posttest Possibilities:

Suppose a college professor wants to compare midterm grades between two different intro to Psychology classes. Section A has a mean of 88 with a standard deviation of 3. Section B has a mean of 80 with a standard deviation of 4. There are 50 students in each class.

1.) In Section A, what would two standard deviations be from the mean?
    a.) 3
    b.) 5
    c.) 6
    d.) 9

2.) Student Y in section A has a z-score of -0.67. Student Z in the same section has a z-score of +1.2. Which student did better?
    a.) Student Y
    b.) Student Z
    c.) They did the same
    d.) Not enough information

3.) From the Empirical Rule, what percentage of kids would have scored 6 points away from the mean?
    a.) 50%
    b.) 95%
    c.) 20%
    d.) 68%

The average running time of a movie is 96.5 minutes with a standard deviation of 10 minutes.

1.) If you calculated a z-score for the newest released movie and got a +0.6 what does that tell you?
    a.) It is shorter than the average running time
    b.) It is longer than the average running time
    c.) It is equal to the average running time
    d.) Not enough information