Statistics, Day 3

Objectives: To find measures of variation in a set of data; To identify outliers; To use the graphing calculator to find summary statistics and make Histograms and Box Plots.

We have already talked about measures of central tendency in a data set. What are they?

Today we will look at measures of variation (or spread). Let’s look at the test score data again:

 55, 60, 73, 76, 84, 91, 95, 97, 99, 99

The simplest measure of spread is the range, the difference between the maximum and minimum value in the data set. It is strongly influenced by outliers.

 Range = Max – Min =

The Interquartile Range (IQR) is another measure of variation that is not as influenced by outliers. The IQR is the difference between Q3 and Q1. It represents the middle 50% of the data.

 IQR = Q3 – Q1 =

A data point is considered an outlier if it lies more than 1.5\*IQR away from Q3 or Q1. Outliers are represented as individual data points on a box plot.

Example: Use the 1.5 IQR rule to determine if any of the test scores are outliers.

The IQR is used as a measure of spread when the median is used as a measure of center. When the mean is used as a measure of center, the standard deviation is used as the measure of spread. It measures an average distance of each data point in a sample from the mean. It is calculated by

Example: Find the standard deviation of the data set:

0, 5, 5, 10.

The graphing calculator can be used to find summary statistics, such as the mean or standard deviation. The calculator can also be used to make plots of the distribution.

55, 84, 99, 60, 73, 99, 76, 95, 91, 97

Example: Enter the test score data into your calculator by selecting: STAT: EDIT. If you already have data in List 1, you can clear it by placing your cursor on L1, and then selecting Clear, Enter.

To find the mean, standard deviation and five number summary:

 STAT CALC 1:1-Var Stats (Enter) L1 (Enter)

To make a histogram:

 2nd STAT PLOT 1 (Enter)

 Turn PLOT 1 On, Type: histogram, Xlist: L1

 Use WINDOW to set the intervals. Then GRAPH.

 TRACE gives the number of data points in each interval. (You could use Trace to help make the frequency table)

Example: Use several different intervals to change the look of your graph. Which one seems the most useful?

To make a box plot: Return to 2nd STATPLOT 1 and select the first box plot icon.

Example: Does the data have any outliers? How can you tell?