**Classwork 3**

*October 3rd 2018*

**Descriptive Statistics on the TI-83/TI-84**

Part 1 #1: Descriptive Statistics With Samples

Instructions: Before you begin the classwork, generate a sample of 36 observations drawn randomly from between 15 and 20 and paste it into the L1 list. Recall, in order to accomplish this, you will navigate to the EDIT submenu in the STAT menu

*STAT > 1 : Edit*

Then place your cursor over the title of the L1 column and bring up the List Operations submenu

*2ND > LIST > Ops > 5 : seq*

And then type the following in the **seq** program

***Expr****: MATH > PRB > 4 : RandInt > 15 , 20*

***Var:*** *X*

***start****:* *1*

***end:*** *36*

***step:*** *1*

After you have the random dataset entered into L1, continue on with the first tutorial.

Tutorial #1: 1 Variable Statistics With A Sample

TI calculators will summarize a great deal of information about a sample of data. To display all of the information available type in the following command

*STAT > CALC > 1 : 1-Var Stats*

Selecting this option will bring up the 1 Variable Statistics program. Enter in L1 Into the *List:* argument. Ignore the *FreqList:* argument for now (*we'll talk about that in the next part of this classwork!*) If you then tap calculate, the program output will display on screen.

**Problems**

1. List the following statistics outputted by the 1-Var Stats program.

a. Mean: \_\_\_\_\_\_\_\_\_

b. Median: \_\_\_\_\_\_\_\_\_

c. The sum of observations: \_\_\_\_\_\_\_\_\_

d. The sum of squared observations: \_\_\_\_\_\_\_\_\_

e. The sample standard deviation: \_\_\_\_\_\_\_\_\_

f. The 25th Percentile: \_\_\_\_\_\_\_\_\_

g. The 75th Percentile: \_\_\_\_\_\_\_\_\_

2. Apply the shortcut formula for the sample standard deviation,

And verify this formula agrees with the output from the calculator.

Part #2: Descriptive Statistics With Frequency Distributions

Instructions: For the next set of exercises, you are given the following frequency distribution of a sample,

|  |  |
| --- | --- |
| *x* | **f(** *x* **)** |
| 10 | 5 |
| 13 | 13 |
| 17 | 26 |
| 22 | 9 |

Enter the observation column into an empty list and then the frequency column into a *different* list. Read through the short tutorial and then answer the questions at the end of the section.

Tutorial #2: 1 Variable Statistics With A Frequency Distribution

Navigate back to the 1-Variable Stat program on the STAT menu

*STAT > CALC > 1 : 1-Var Stats*

The argument we ignored in the previous part, *FreqList:*, specifies the frequency of each observation contained in the sample. Note, in order for this program to work properly, each observation in the list containing the sample must be *unique;* its number of occurences in the sample is determined the frequency list.

**Problems**

3. List the following statistics outputted by the 1-Var Stats program.

a. Mean: \_\_\_\_\_\_\_\_\_

b. Median: \_\_\_\_\_\_\_\_\_

c. The sum of observations: \_\_\_\_\_\_\_\_\_

d. The sum of squared observations: \_\_\_\_\_\_\_\_\_

e. The sample standard deviation: \_\_\_\_\_\_\_\_\_

f. The 25th Percentile: \_\_\_\_\_\_\_\_\_

g. The 75th Percentile: \_\_\_\_\_\_\_\_\_

4. Apply the shortcut formula for the sample standard deviation,

And verify this formula agrees with the output from the calculator.

Part #3 : Z Scores and Outliers

Instructions: For the next set of problems, use the following data set that represents the final exam scores for a class of MAT 135 students,

S = { *10, 82, 77, 89, 90, 91, 95, 88, 75* }

Enter this sample into an empty list on your calculator and then follow the tutorial to calculate the Z score of each observation.

Note: If you run out of empty lists, you can clear a list with the **ClrList**

*STAT > 4 : ClrList*

You can then select the list you want to clear and press ENTER to execute. If you want to clear multiple lists, separate the lists with a comma before executing.

Tutorial #3: Editing Whole Lists at Once

A Z-score is a standardized measure of how far from the mean a particular observation in a sample is; Z-scores from different samples can be used to compare the relative standing of totally different observations with respect to their underlying samples. To see why, take a look at the formula for a Z-score

The numerator is the distance of an observation from the sample mean. The denominator is the sample standard deviation. Thus, a Z-score tells us the distance to the mean *per unit variation*. This is why we call a Z-score standardized; it takes a sample's variation into account.

Consider making an observation of 206 from sample with a mean of 200 and a standard deviation of 2. While the absolute distance to the mean is only 206 – 200 = 6, this is a misleading statistic; it doesn't take into account the standard deviation, which tells us most of the data is tightly is clustered around the mean. The Z-score for this particular would be , which indicates this is an outlying observation (Recall if Z > 3 or Z < -3, then we call those observation *outliers.* )

Compare this to making an observation of 206 from a sample with a mean of 200 and standard deviation of 10. In this case, the standard deviation tells us the data is more spread out around the mean. If we calculate the Z score, we get 0.6, which is well within the expected bounds!

It is, however, time consuming calculating the Z-score of each individual observation. In order to identify outliers in a sample, we employ our TI calculators!

Navigate to EDIT submenu in STAT Menu

*STAT > 1 : Edit*

In a blank list next to the list containing the data, highlight the name of the column itself with your cursor. This will allow you to enter in a formula for the entire column. Suppose our sample of data is located in list L4. Then, into blank column L5, you should enter the formula,

*L5 = (L4 – mean(L4))/stdDev(L4)*

And then execute the formula. Recall the **mean** and **stdDev** can be accessed through the MATH submenu in the LIST menu

*2ND > LIST > MATH > 3 : mean*

*2ND > LIST > MATH > 7 : stdDev*

After you execute the formula, the blank column should fill up with elements. Note there is one entry for ever entry in the column containing our dataset. This is because the formula operates on each entry in the former column to produce the latter. This new column represents the Z-scores of each observation!

**PROBLEM**

5. Calculate the Z-scores for the sample of data given in the instructions for Part 3. After you have them calculated, search through them for any potential outlying observations. List the value of the outlying observation: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_