Homework assignment 04

Use black text (if possible) for everything you include in this document. Keep both your answers and the original questions. Save this document in PDF format and submit it on Canvas. Include your last name, the course number and the module number in the name of your file.

1. Show a documentation header. The documentation header is a description of who wrote the program, when it was written, what the purpose of the program is (briefly), and what restrictions (if any) that you may place on the program. For SPSS, you can just type the documentation as free format text. For other programs, you might use the comment feature (such as /\* and \*/ in SAS).

Consider the following problem and then follow the steps of Hypothesis testing as presented in class and answer the questions. **Show all work.** Please write legibly. Typed answers should be in a font color that is NOT black. An answer should be entered into the document immediately following the question asked. This HW assignment is worth 20 points. **Rounding only occurs with the final answer. Rounding within the calculation reduces precision. Statistical analysis requires precision.**

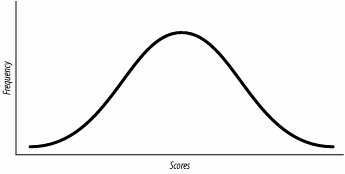
The equations that you will use are:

**For a more in-depth explanation of the above equations, including examples of application, please see the Zoom Class Video for this week – To post on Tuesday following the class meeting.**

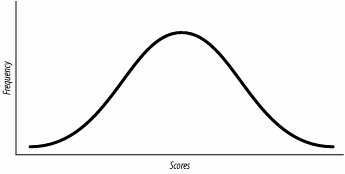
**Question Set 1:**

**General Knowledge:** Show your answers to the first two questions in this section using the following normal curves: Please include the *location of the mean*, *standard deviation marks*, z-score marks, the z-score for Q1 and Q2, the percent answer to each question, and the shaded areas depicting the null hypothesis rejection areas. Be as precise as possible. You can do this “by Hand” or computer.

Q1. You are conducting a **Two-tailed test**. You have decided to use an alpha of .05 for significance. What is the appropriate Z-score(s) that would mark the null hypothesis rejection area(s) for this test? What percent of the data would be included in the null hypothesis rejection area(s)?



Q2. You are conducting a **One-tailed test and assume a decrease in the measured variable**. You have decided to use an alpha of .05 for significance. What is the appropriate Z-score(s) that would mark the null hypothesis rejection area(s) for this test? What percent of the data would be included in the null hypothesis rejection area(s)?



Q3. Considering a One-tailed or a Two-tailed test:

1. Which test is more powerful? \_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. Which test would more likely result in a Type I Error? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. Which test would more likely result in a Type II Error? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Question Set 2:**

**Presenting Problem:** Systolic Arterial Pressure is a normally distributed variable that has **a population mean of 120 mmHg and a population Standard Deviation of 10mmHg**. Use this information to answer the following questions.

Use: and the z-table to answer the following questions. Show all work.

1. What is the z-score for a BP value of 130 mmHg?
2. What percent of values (area of the curve) are located above this z-score for 130 mmHg?
3. What is the z-score for a BP value of 145 mmHg?
4. What percent of values (area of the curve) are located above this z-score for 145 mmHg?
5. Identify the area of the curve that lies between +30 mmHg and -30 mmHg from the mean.
6. What is the value of BP that is found at the 95th percentile?
7. What is the value of BP that is found at the 5th percentile?
8. What is the value of BP that is found at the 97.5th percentile?
9. What is the value of BP that is found at the 2.5th percentile?

**Question Set 3:** Recalling that the purpose of inferential statistics is to infer that the sample is representative to the population:

1. Population Data: Systolic Arterial Pressure is a normally distributed variable that has a **population mean of 120 mmHg**.

The Problem: A Physician randomly selects a group of patients to assess their systolic arterial pressure. They selected **36 patients** and find that their **systolic arterial** **pressure averaged 124 mmHg +/- 10 mmHg SD**. The data was amazingly normally distributed, so the z distribution was deemed appropriate for this assessment. How often would a sample of 36 patients have a mean systolic arterial pressure greater than 124 mmHg? This is the same as asking: Given repeated sampling, what proportion of samples would have mean values of systolic arterial pressure above 124 mmHg?

Use: and the z-table to answer the following questions. Show all work.

1. Calculate the z-score for a mean systolic arterial pressure of **124 mmHg**?
2. What is the percentile for this mean score?
3. What percent or proportion of samples are expected to have a mean above 124 mmHg?
4. Population Data: Systolic Arterial Pressure is a normally distributed variable that has a **population mean of 120 mmHg**.

The above Physician randomly repeated this study with a new group of patients to assess their systolic arterial pressure. The physician again selected different **36 patients** (Amazingly normally distributed again) and found this time that patient **systolic arterial pressure averaged 123 mmHg +/- 10 mmHg**. The physician wants to know the probability of obtaining this mean value and if this mean value “differs significantly from the population mean”. They are concerned that they are beginning to see more patients with hypertension (high blood pressure).

1. First, the physician decides to perform a One-tailed test. After all, they are examining possible “hypertension”.
   1. State the Null Hypothesis
   2. Calculate the z-score for this sample mean. Remember, you are assessing where in the proposed population distribution the mean falls. Show all work. Circle the z-score that you calculated.
   3. What is the z-score (alpha = .05) standard that marks the beginning of the null hypothesis rejection area for a One-tailed test?. (Again – This is a standard)
   4. Interpret the result as a One-tailed test. (Alpha = .05)
   5. What conclusion does the physician make? (Are the patients in this sample hypertensive or not based on the results of this test?)
2. You read this physician’s study in the Questionable Statistics Medical Journal. Since you are taking Biostatistics, you learned that it is better to consider a Two-tailed test because usually physiologic variables can fall above or below the mean. This means that the sample systolic arterial pressure could fall lower than the population mean in some patients.
   1. What are the z-score (alpha = .05) standards that mark the beginning of the null hypothesis rejection areas for a Two-tailed test? (Again –standards)
   2. Interpret the z-score result from the question above, using a Two-tailed test.
   3. What conclusion do you make? (Are the patients in this sample hypertensive or not based on the results of this test?)

**Question Set 4:**

It is well know that childhood obesity is a continuing health problem in the United States. The number of children with the disease of obesity has increased 10-fold since the 1970’s. In the 70’s the average female weighed 140 lbs. and in 2015, that average weight had climbed to 166 lbs. Knowing this, a teacher at a local school and a parent who is a medical researcher (who were alarmed that their students appeared increasingly larger over the past few years) wanted to assess if the students attending their school followed the national trends of obesity disease. The researcher decided that since their child was a girl, that the first assessment would be the weight of females. The researcher conducted a literature search and found public health statistics that reported that the average weight of 17-year-old adolescent females in 1970 was 140 lbs.

The study received IRB approval. A random sample of 32 female students in the Senior class was consented and obtained. Student weight was recorded on a morning when all students arrived at school 12 hours fasted as instructed. All students weighed on the same scale and wore a robe provided by the school. The scale was calibrated (zeroed) for each subject. All data was de-identified. Average weight was obtained and compared against the 1970 population standard of 140 lbs.

1. The data can be found in the Homework Data set provided in the module. You do not need to create a data set! **The variable name should be renamed to include your initials.** There is one variable and one column of data. Each cell of the data set is the weight in pounds of a single student.

2. Using SPSS, run a one sample test of the hypothesis to determine if this current day sample is “the same or different than” the 1970s standard of 140lbs. (SPSS Directions can be found in the Canvas Module for this week).

Answer the following questions and provide the requested output.

1. Write the Null and the Alternative Hypotheses for this study question.
2. What are the Mean and the Standard Deviation of the variable in this set of data? (Look in the output table)
3. State the calculated p-value for the hypothesis test that was run in this analysis.
4. What is the probability of making a Type one error for this analysis? (State as a percent)
5. Using all of the above information, what conclusion can you draw about the mean weight of this current day sample and the possibility of obesity disease?

SPSS Output to include for Question Set 4:

1. *One-Sample Statistics* table from SPSS output

2. *One-Sample Test table* from SPSS output