Assignment



Surname 1

Student's Name Professor's Name Course Number Date



- To study about the correlation between height and shoe size, you need to collect a sample of nine
 (9) people using a Systematic Sampling method.
 - a. What is the population of people? Where and how are you going to collect your sample? Does your sample accurately represent your population? Why or why not? *The population is the campus where I study. I will select a sample from the college using a systematic sampling method to minimize the risk of data manipulation. The data will be selected at intervals of 30 students entering the campus using the main gate. This will be a good representation of the population.*
 - **b.** Collect the sample and record the data. Use a single unit for height. Do not use a mixed unit like feet and inches.
- 2. Construct a confidence interval to estimate the mean height and the mean shoe size: you must complete the following questions by first choosing a Confidence Level. You may choose from the familiar 90%, 95%, or 99% level of confidence. Denote this by choosing $\alpha = 0.05$
 - **a.** Find the sample mean and sample standard deviation of the height. Denote them as x-bar and sx respectively.

$$\overline{x} = \frac{\sum x_i}{n}$$

$$\overline{x} = \frac{63 + 60 + 77 + 66 + 70 + 65 + 65 + 67 + 74}{2}$$

Mean: 67.444

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(N-1)}}$$
$$s = \sqrt{\frac{19.749 + 55.413 + 91.317 + 2.085 + 6.533 + 5.973 + 5.973 + 0.197 + 42.981}{8}}$$

- Std Dev: 5.364
- **b.** Find the sample mean and sample standard deviation of the shoe sizes. Denote them as y-bar and sy respectively.



$$\overline{y} = \frac{\Sigma y_i}{n}$$

$$\overline{y} = \frac{5+4+12+8+9+7.5+6.5+11.5+10.5}{9}$$

Mean: 8.222

$$s = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{(N-1)}}$$
$$s = \sqrt{\frac{10.381 + 17.825 + 14.273 + 0.049 + 0.605 + 0.521 + 2.965 + 10.745 + 5.189}{8}}$$

Std Dev: 2.863

c. Construct and interpret a confidence interval to estimate the mean height of the population. You must first write the formula for the confidence interval and then substitute your appropriate numbers.

$$x - bar \pm t_{\alpha/2} \left(\frac{s_x}{\sqrt{n}} \right)$$

Substituting the figures into the equation gives t-value for 0.05 with 8 degrees of freedom gives: 2.306 $67.444 \pm 2.306(1.788)$ = (63.321, 71.567)

d. Construct and interpret a confidence interval to estimate the mean shoe size of the population. You must first write the formula for the confidence interval and then substitute your appropriate numbers.

$$y - bar \pm t_{\alpha/2} \left(\frac{s_y}{\sqrt{n}} \right)$$

t-value for 0.05 with 8 degrees of freedom gives 2.306 8.222 ± 2.306(0.932) (6.073, 10.371)

3. Test a claim that the mean height of your population is different from 64 inches. Use the appropriate significance level α you fixed earlier.



a. State the initial and alternative hypothesis.

$$H_0: \mu = 64$$
$$H_A: \mu \neq 64$$

b. Find the test statistic and the P-value. You must first write the formula for the test statistic and then substitute your appropriate numbers.

This test will use the t-test statistic. The formular is as follows:

$$t = \frac{x-\mu}{\frac{s}{\sqrt{n}}}$$

Upon substituting, we get,

$$t = \frac{\frac{67.444-64}{\frac{5.364}{\sqrt{9}}}}{$$

The t-stat is 1.9262 The p-value is 0.09

c. Draw a conclusion in context of the situation. Your conclusion should include both the formal language as well as an informal explanation.

We fail to reject the null hypothesis. There is insufficient evidence to conclude that the mean height of my population is different from 64 inches since the p-value is greater than the alpha level of significance.

- 4. Find a correlation between height and shoe size.
 - **a.** Create a scatterplot of the data. Height is x-axis and Shoe size is y-axis. Attach your scatterplot to the end of this document.
 - b. Find the linear correlation coefficient. What does this tell you about your data?
 Correlation coefficient = 0.863 There is a strong positive correlation between the height and shoe size.



c. Write the equation of the regression line and use it to predict the shoe size of a person that is 68 inches tall.

Regression line y = 0.45x - 22.131

Predicted shoe size of a person that is 68 inches tall gives: y = 0.45(68) - 22.131y = 8.469

5. Write a paragraph or two about what you have learned from this process. When you read, see, or hear a statistic in the future, what skills will you apply to know whether you can trust the result? *This process has enabled me to learn about the t-test and the requirements to make a t-test. For instance, a sample less than 30 will require a t-test. Also, if the population standard deviation is unknown, we can use the t-test statistic. Furthermore, understanding hypothesis testing helps in making proper and accurate inferences in the future.*



Figure 1: Scatter Plot





