**Study Guide**

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PART I. HYPOTHESIS TESTING**

**PROBLEM 1 (30 points): Construct a 95 per cent confident interval given the information below.**

A certain brand of fluorescent light tube was advertised as having an average illumination life-span of **XXXXX** hours. A random sample of 64 bulbs burned out with a mean life-span of **XXXXX** hours and a sample standard deviation of XX hours. Construct a 95 percent confident interval based on this sample. Be sure to interpret the meaning of this interval.

**The XXXX’s will be replaced with numerical values on my quiz. Please edit the problem to include values and then solve the problem showing in detail how you accomplished your answer.**

**It would be helpful for me to understand the formulas used prior to inserting values into said formula. If you have seen this problem previously, please utilize the prior values that may have been presented.**

**PROBLEM 2 (40 points):** Given the following data from two independent samples from which the population standard deviation is known, conduct a *two-tailed* hypothesis test to determine if the first sample mean is smaller than the second sample mean, given a 0.10 level of significance.

n1 = 42 n2 = **XX**

xbar1= **XX** xbar2 = **XX**

sigma1=**XX** sigma2 = **XX**

**The XXXX’s will be replaced with numerical values on my quiz. Please edit the problem to include values and then solve the problem showing in detail how you accomplished your answer.**

**It would be helpful for me to understand the formulas used prior to inserting values into said formula. If you have seen this problem previously, please utilize the prior values that may have been presented.**

**PROBLEM 3 (44 points): Conduct a *one-tailed* hypothesis test given the information below.**

A test was conducted to determine whether gender of a spokesperson affected the likelihood that consumers would prefer a new product. A survey of consumers at a trade show employing a female spokesperson determined that **XX** out of 150 customers preferred the product, while **XX** of 180 customers preferred the product when a male spokesperson was employed. At the 0.01 level of significance, do the samples provide sufficient evidence to indicate that on the average, fewer consumers prefer a new product when the spokesperson is female?

**The XXXX’s will be replaced with numerical values on my quiz. Please edit the problem to include values and then solve the problem showing in detail how you accomplished your answer.**

**It would be helpful for me to understand the formulas used prior to inserting values into said formula. If you have seen this problem previously, please utilize the prior values that may have been presented.**

**PROBLEM 4 (52 points): Conduct a *two-tailed* hypothesis test given the information below.**

**Assuming that the population standard deviations are unknown, but equal** for male and female Grade Point Averages (GPAs), use the following sample data to test whether the averages are different at the 0.05 level of significance.

|  |  |  |
| --- | --- | --- |
|  | **Male GPA’s** | **Female GPA’s** |
| **Sample Size** | **XX** | **XX** |
| **Sample Mean** | **XXX** | **XXX** |
| **Sample Standard Dev** | **XX** | **XX** |

**The XXXX’s will be replaced with numerical values on my quiz. Please edit the problem to include values and then solve the problem showing in detail how you accomplished your answer.**

**It would be helpful for me to understand the formulas used prior to inserting values into said formula. If you have seen this problem previously, please utilize the prior values that may have been presented.**

**PART II REGRESSION ANALYSIS**

**PROBLEM 5 (54 points): Answer questions (a) through (e) using the following information and output for multiple regression.**

**A real estate investor has devised a model to estimate home prices in a new suburban development. Data for a random sample of 30 homes were gathered on the selling price of the home ($ thousands), the home size (square feet), the lot size (thousands of square feet), and the number of bedrooms.**

**The following multiple regression output was generated:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Regression Statistics* | |  |  |  |
| Multiple R |  |  |  |  |
| R Square |  |  |  |  |
| Adjusted R Square |  |  |  |  |
| Standard Error |  |  |  |  |
| Observations |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept |  |  |  |  |
| X1 (Employ) |  |  |  |  |
| X2 (competitors) |  |  |  |  |
| X3 (Mkt Share) |  |  |  |  |

**Your complete Megastas Matrix will be provided.**

1. Why is the coefficient for lot employment a positive number?
2. Which is the most statistically significant variable? What evidence shows this?
3. Which is the least statistically significant variable? What evidence shows this?
4. For a 0.05 level of significance, should any variable be dropped from this model? Why or why not?
5. Predict the annual sales for a representative with 58 months of employment whose competitors earn $34.5 million ($34,500 thousands) in sales and with a 12 percent product market share.

**PART III SPECIFIC KNOWLEDGE SHORT-ANSWER QUESTIONS**

**Problem 6 (40 points):** Define autocorrelation in the following terms:

a. In which type of regression is it likely to occur?

b. What is the negative impact of autocorrelation in a regression?

c. Which method is used to determine if it exists?

d. If found in a regression, how is autocorrelation eliminated?

**Problem 7 (40 points):** Define multicollinearity in the following terms:

a. In which type of regression is it likely to occur?

b. What is the negative impact of multicollinearity in a regression?

c. Which method is used to determine if it exists?

d. If multicollinearity is found in a regression, how is it eliminated?