Lab 13

Exercise 1

Is there a linear relationship between the number of tree stumps left behind by beavers and the number of beetle larvae? Researchers laid out 10 circular plots, each 4 meters in diameter, in an area where beavers were cutting down cottonwood trees. The number of stumps and the number of clusters of beetle larvae were recorded in each plot with the following results:

Stumps	Beetle Larvae
2	10
2	30
1	12
3	24
4	40
1	11
5	56
3	40
1	8
2	14

Draw a scatterplot and determine whether there appears to be a linear relationship between the number of stumps and the number of clusters of beetle larvae. If so, describe the relationship and calculate Pearson correlation ρ.

Exercise 2

Johnson (1973) conducted a study to determine whether, in collegiate schools of nursing, relationships between certain variables could be identified. Two variables of interest for which indices were constructed were "extent of agreement (between the dean and the faculty) on the responsibilities for decision making" and "faculty satisfaction". The ranks on the two variables of the 12 institutions that participated in the study are shown in table below. Find Spearman's rank correlation and Kendall's tau from the data?

School	Rank on faculty satisfaction	Rank on decision-making agreement
A	1	12
B	2	9
C	3	1
D	4	7
E	5	3
F	6	10
G	7	11
H	8	8
I	9	2
J	10	6
K	11	4
L	12	5

Exercise 3

For each of the following situations, state the independent variable and the dependent variable.

a. A study is done to determine if elderly drivers are involved in more motor vehicle fatalities than other drivers. The number of fatalities per 100,000 drivers is compared to the age of drivers.

b. A study is done to determine if the weekly grocery bill changes based on the number of family members.

c. Insurance companies base life insurance premiums partially on the age of the applicant.

d. Utility bills vary according to power consumption.

e. A study is done to determine if a higher education reduces the crime rate in a population.

Exercise 4

Revision: **TRUE or FALSE**

1. Interval estimates are preferred over point estimates since a confidence level can be specified.

2. An estimator is consistent if as the sample size decreases, the value of the estimator approaches the value of the parameter estimated.

3. Interval estimate may or may not contain the value of the parameter being estimated.

4. A statistical test uses the data obtained from a sample to make a decision about whether the alternative hypothesis should be rejected.

5. The level of significance is the maximum probability of committing a type I error.

6. If the test is left-tailed, the critical region, with an area equal to α, will be on the left side of the mean.

7. No error is committed when the null hypothesis is rejected when it is false.

8. The test value separates the critical region from the noncritical region.

9. When the value of α is increased, the probability of committing a type I error is decreased

10. When you are testing the difference between two means, it is not important to distinguish whether the samples are independent of each other.

11. If the same diet is given to two groups of randomly selected individuals, the samples are considered to be dependent.

12. To test the equality of two variances, you would use a F-test.

[1]

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var\,X\;Var\,Y}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

2

$$d$$

$$(r_i - q_i)^2$$

$$D = 382$$

$$S = \frac{1}{3} n (n^2 - 1)$$

$$\rho = 1 - \frac{2D}{S}$$