

A

Name of students:

Q1: For each of the following questions, Determine the appropriate type of statistical analysis and explain the reason (short answers) (10 degree)

1. A researcher is interested in understanding the relationships between two sets of variables, such as socioeconomic status and health outcomes?
2. A finance analyst wants to estimate the uncertainty around the mean return on investment using a small sample?
3. An Agriculture engineer is modeling the growth of a plant over time, and the relationship seems curvilinear?
4. A data scientist is predicting the sales of a product based on advertising expenditure and product features?
5. A researcher wants to explore the relationship between hours of study and exam scores among university students?
6. A sociologist is comparing the median income levels of individuals from three different cities?
7. A marketing researcher wants to compare the impact of three advertising strategies on multiple outcome variables, such as purchase intention, brand perception, and customer satisfaction?
8. An education researcher is comparing the academic performance of students from different schools, accounting for the effect of socioeconomic status?
9. An economist is examining the distribution of income in a population to assess if it follows a normal distribution?
10. Which statistical procedure in SPSS is appropriate for examining the differences between two or more related groups across multiple time points?

Q2 :From the following questions, chose the right answer?

(10 degree)

1- What does the R-squared value indicate in regression model?

- a) The correlation between the independent variables.
- b) The proportion of variance in the dependent variable explained by the independent variables.
- c) The number of predictors in the model.
- d) The coefficient of determination.
- e) Not of Them.

2- In multiple regression, what does the term "Mahalanobis Distance" assess in the context of multiple regression?

- a) Homoscedasticity.
- b) Independence of residuals.
- c) Presence of outliers.
- d) Linearity of the relationship.
- e) Not of Them.

3- When interpreting the p-value associated with a predictor variable's coefficient in SPSS multiple regression output, what does a (p-value= 0.05) generally suggest?

- a) The variable is not a significant predictor.
- b) The variable is not included in the model.
- c) The variable has a perfect correlation with the dependent variable.
- d) The variable is not included in the model.
- e) Not of Them.

4- What does the term "Multicollinearity" refer to?

- a) The assumption of normality in residuals.
- b) The presence of outliers in the data.
- c) The dependence of residuals.
- d) The independence of residuals.
- e) Not of Them.

5. When conducting hypothesis testing, what does the p-value indicate?

- a) The probability of rejecting a true null hypothesis.
- b) The probability of accepting a false null hypothesis.
- c) The level of confidence in the research findings.
- d) The significance of the research topic

6- The relation between (Tolerance & VIF) in Multiple Linear Regression?

- a) Positive.
- b) Negative.
- c) zero.
- d) No Relation.

7- Which type of analysis in SPSS is suitable for predicting a categorical outcome variable based on one or more predictor variables?

- a) Discriminant Analysis.
- b) Factor Analysis.
- c) Logistic Regression.
- d) Canonical Correlation.

8- What is the primary purpose of the "Cluster Analysis"?

- a) Identifying outliers in a dataset.
- b) Grouping similar cases based on selected variables.
- c) Examining the relationship between Clusters.
- d) Testing the mean differences between two independent groups

9- What is the basic building block of an Artificial Neural Network (ANN)?

- a) Neuron
- b) Node
- c) Link
- d) Edge

10- Which statement is true about a model with an AUC of 0.5 in a ROC curve?

- a) The model is perfect.
- b) The model is no better than random guessing.
- c) The model has high sensitivity.
- d) The model has high specificity.

مع تمنياتي لكم بالنجاح و الموفقية



Asst. Prof. Dr.Rizgar Maghded Ahmed

A

Name of students:

Q1// From the sample data (Employee data. sav) answer that: **(10 degree)**
 Note: used ($\alpha=0.05$) for all questions.

a) The researcher needs to estimate the relationship between the two variables (educ & salary) represented by a statistical model that can be predicted?

Name of Model	Write the estimate model?	F-value	R ²	p-value

b) Based on the model estimated in (a), how can its predictive values be found? Just write the steps on SPSS?

c) The researcher needs to answer the following question: Are the two variables (jobcat & minority) independent of each other or related?

Name of test	Value of statistic?	Hypothesis	p-value

Decision:

d) The researcher needs to test the comparison of the average (salary) based on the (gender) variable and whether the differences between them are significant or not?:

Name of test	Value of statistic?	Hypothesis	p-value

Decision:

- e) The researcher needs to classify a variable (minority) based on a group of variables including (salary, jobtime, educ). In how many methods can this be done? Determine them and then write a classification table for each method? What is the importance of each variable in the classification process?



Asst. Prof. Dr.Rizgar Maghded Ahmed