STATISTICS

WORKSHEET-3

1. B) Total Variation = Residual Variation + Regression Variation

2. C) binomial

3. A) 2

4. A) Type-I error

5. A

6. A

7. B) Hypothesis

8. D) All of the mentioned

9. A) 0

10) The Bayes theorem is a mathematical formula for calculating conditional probability in probability and statistics. In other words, it's used to figure out how likely an event is based on its proximity to another. Bayes law or Bayes rule are other names for the theorem.

Bayes Theorem Formula : $P(A \mid B) = P(B \mid A)P(A) / P(B)$

where :

P(A | B) is the conditional probability of event A occurring, given that B is true.

P(B | A) is the conditional probability of event B occurring, given that A is true.

P(A) and P(B) are the probabilities of A and B occurring independently of one another.

11) z-score shows how much standard deviation of data away from mean .If z-score value is 0 means data point score is identical to mean and z-score value is 1 means 1 standard deviation from mean.
z-score basically tells there is outliers present or not . Normal data values range is standard deviation from -3 to +3 . If standard deviation is outside of -3 and + 3 means outliers or abnormal data , it is determined by z-score.
z-score may be positive indicates score is above the mean and can be negative also indicates score is below the mean

12) A T-test is a statistical method of comparing the means or proportions of two samples gathered from either the same group or different categories. It is aimed at hypothesis testing, which is used to test a hypothesis pertaining to a given population. It is the difference between population means and a hypothesized value.
T-Test used to compare 2 groups.
When we have least no of rows, we can do T-Test(datas are provided identical to mean value)

13) Percentile is used to indicate the value below which the group of percentage of data falls below so can understand and interpret data.
Ex : If anyone scores 75th percentile, means scored better than 75% people out of total people .

14) ANOVA (Analysis of Variance)
An ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two or more than two categorical groups by testing for differences of means using variance.

Anova Hypothesis :
Null Hypothesis (H_0) : Group means are equal (no variation in means of group)
Alternative Hypothesis(H_A/H_1): Atleast, one group mean is different from other group.

15) ANOVA helps in selecting the best features. This minimizes the number of input variables to reduce the complexity of that model.
ANOVA helps to determine if an independent variable is influencing a target variable.