FLIP ROBO

STATISTICS WORKSHEET-3

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

1. Which of the following is the correct formula for total variation?
   a) Total Variation = Residual Variation – Regression Variation
   b) **Total Variation = Residual Variation + Regression Variation**
   c) Total Variation = Residual Variation * Regression Variation
   d) All of the mentioned

2. Collection of exchangeable binary outcomes for the same covariate data are called_____outcomes.
   a) random
   b) direct
   c) **binomial**
   d) none of the mentioned

3. How many outcomes are possible with Bernoulli trial?
   a) **2**
   b) 3
   c) 4
   d) None of the mentioned

4. If Ho is true and we reject it is called
   a) **Type-I error**
   b) Type-II error
   c) Standard error
   d) Sampling error

5. Level of significance is also called:
   a) **Power of the test**
   b) Size of the test
   c) Level of confidence
   d) Confidence coefficient

6. The chance of rejecting a true hypothesis decreases when sample size is:
   a) Decrease
   b) **Increase**
   c) Both of them
   d) None

7. Which of the following testing is concerned with making decisions using data?
   a) Probability
   b) **Hypothesis**
   c) Causal
   d) None of the mentioned

8. What is the purpose of multiple testing in statistical inference?
   a) Minimize errors
   b) Minimize false positives
   c) Minimize false negatives
   d) **All of the mentioned**

9. Normalized data are centred at____and have units equal to standard deviations of the original data
   a) **0**
   b) 5
   c) 1
   d) 10

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What Is Bayes' Theorem?

Ans.10) Bayes' theorem describes the probability of occurrence of an event related to any

condition. It is also considered for the case of conditional probability. Bayes theorem is

also known as the formula for the probability of "causes".

Let E1, E2…, En be a set of events associated with a sample space S, where all the

events E1, E2…, En have a nonzero probability of occurrence and they form a partition

of S. Let A be any event associated with S, then according to Bayes theorem,

$P(Ei│A) = P(Ei)P(A│Ei)Σk=1nP(Ek)P(A|Ek)$

for any k = 1, 2, 3, …., n

11. What is z-score?

Ans.11) A z-score (also called a standard score) gives you an idea of how far from
the mean a data point is. But more technically it's a measure of how many standard
deviations below or above the population mean a raw score is.
A z-score can be placed on a normal distribution curve.
Z-scores range from -3 standard deviations (which would fall to the far left of the normal
distribution curve) up to +3 standard deviations (which would fall to the far right of the
normal distribution curve).
In order to use a z-score, we need to know the mean μ and also the population standard
deviation σ.
The basic z score formula for a sample is:
$z = (x − μ) / σ$

12. What is t-test?

Ans.12) A t-test is a statistical test that is used to compare the means of two groups. It is
often used in hypothesis testing to determine whether a process or treatment actually has
an effect on the population of interest, or whether two groups are different from one
another.
   o The t-test is a test used for hypothesis testing in statistics.
   o Calculating a t-test requires three fundamental data values including the
   o difference between the mean values from each data set, the standard deviation
   o of each group, and the number of data values.
   o T-tests can be dependent or independent.

13. What is percentile?

Ans.13) A percentile (or a centile) is a measure used in statistics indicating the value below
which a given percentage of observations in a group of observations fall. For example,
the 20th percentile is the value (or score) below which 20% of the observations may be
found.
The term percentile and the related term percentile rank are often used in the reporting
of scores from norm-referenced tests. Every score is in the 100th percentile.
The 25th percentile is also known as the first quartile (Q1), the 50th percentile as the
median or second quartile (Q2), and the 75th percentile as the third quartile (Q3). In
general, percentiles and quartiles are specific types of quantiles.

14. What is ANOVA?

Ans.14) ANOVA, which stands for Analysis of Variance, is a statistical test used to analyze the difference between the means of more than two groups. A one-way ANOVA uses one independent variable, while a two-way ANOVA uses two independent variables. An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis.

15. How can ANOVA help?

Ans.15) ANOVA is helpful for testing three or more variables. It is similar to multiple twosample t-tests. However, it results in fewer type I errors and is appropriate for a range of issues. ANOVA groups difference by comparing the means of each group and includes spreading out the variance into diverse sources. The one-way ANOVA can help you know whether or not there are significant differences between the means of your independent variables (such as the first example: age, sex, income). When you understand how each independent variable's mean is different from the others, you can begin to understand which of them has a connection to your dependent variable (landing page clicks), and begin to learn what is driving that behavior.