

Statistics– WORKSHEET 6

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
b) Total Variation = Residual Variation + Regression Variation
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
c) binomial
3. How many outcomes are possible with bernoulli trial?
a) 2 b) 3 c) 4 d) None of the mentioned
a) 2
4. If H_0 is true and we reject it, then it is called:
(a) Type-I error (b) Type-II error (c) Standard error (d) Sampling error
(a) Type-I 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
c) Level of confidence
6. The chance of rejecting a true hypothesis decreases when sample size:
(a) Decreases (b) Increases (c) Both of them (d) None of them
(b) Increases
7. Which of the following testing is concerned with making decisions using data?
(a) Probability (b) Hypothesis (c) Causal (d) None of the mentioned
(b) Hypothesis
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
d) All of the mentioned
9. Normalized data is centered at and has unit equal to standard deviations of the original data.
(a) 0 (b) 5 (c) 1 (d) 10
a) 0

Q10 and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What Is Bayes' Theorem?
Bayes Theorem is a way of finding of probability when we know certain other probabilities.
The formula is :

$$P(A | B) = P(A) P(B | A)P(B)$$

11. What is z-score?

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, you need to know the mean and also the population standard deviation.

12. What is t-test?

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.

13. What is a percentile?

The most common definition of percentile is a number where a certain percentage of scores fall below that number. It can be used in calculating the IQR using boxplot.

14. What is ANOVA?

An ANOVA is a way to find out if survey or experiment results are significant. In other words, they help us to figure out if we need to reject the null hypothesis or accept the alternate hypothesis. Basically, we are testing groups to see if there is a difference between them.

15. How can ANOVA help?

The one-way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups and useful in hypothesis testing whether to accept null or not and used to calculate the significance level value.