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
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, then 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:
   ➡ (a) Decreases
   (b) Increases
   (c) Both of them
   (d) None of them
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) (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

Q10and 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 a probability when we know certain other probabilities.
The formula is:

$$P(A|B) = P(A)\,P(B|A)\mathbf{P(B)}$$

Which tells us: how often A happens *given that B happens*, written **P(A|B)**,

When we know: how often B happens *given that A happens*, written **P(B|A)**

and how likely A is on its own, written **P(A)**

and how likely B is on its own, written **P(B)**

11. What is z-score?

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

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.

A t-test looks at the t-statistic, the t-distribution values, and the degrees of freedom to determine the statistical significance. To conduct a test with three or more means, one must use an analysis of variance.

13. What is a percentile?

a **percentile** (or a **centile**) is a type of quantile which divides the given probability distribution, or sample, into 100 equal-sized intervals; this allows the data to be analyzed in terms of percentages. For example, the 20th percentile is the value (or score) below which 20% of the observations are found, and above which 80% are found.

14. What is ANOVA?

Analysis of variance (**ANOVA**) is a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyze the differences among group means in a sample.

15. How can ANOVA help?

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. Basically, you're testing groups **to** see if there's a difference between them.