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

**ANS: 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

**ANS: c) binomial**

3**. How many outcomes are possible with Bernoulli trial?**

a) 2

b) 3

c) 4

d) None of the mentioned

**ANS: a) 2**

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

**ANS: 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

**ANS: b) Size of the test**

6. **The chance of rejecting a true hypothesis decreases when sample size is:**

a) Decrease

b) Increase

c) Both of them

d) None

**ANS: c) Both 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

**ANS: a) Probability**

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 mentione

**ANS: a) Minimize errors**

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

**ANS: a) 0**

10**. What Is Bayes' Theorem**?

ANS: In statistics and probability theory, the Bayes’ theorem is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes’ theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.

Formula for Bayes’ Theorem:

The Bayes’ theorem is expressed in the following formula:

P(A|B)=P(B|A)P(A)\P(B)

Where:

* P(A|B) – the probability of event A occurring, given event B has occurred
* P(B|A) – the probability of event B occurring, given event A has occurred
* P(A) – the probability of event A
* P(B) – the probability of event B

11. What is z-score?

**Ans:** Z-score is also known as**standard score** gives us an idea of how far a data point is from the mean. It indicates how many standard deviations an element is from the mean. Hence, Z-Score is measured in terms of standard deviation from the mean.

**The Formula for Z-Score:**

*A z-score can be calculated using the following formula.*

**z = (X – μ) / σ**

*where,   
z = Z-Score,   
X = The value of the element,   
μ = The population mean, and   
σ = The population standard deviation.*

12. **What is t-test?**

Ans: 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. The t-test is a test used for hypothesis testing in statistics and uses the t-statistic, the distribution values, and the degrees of freedom to determine statistical significance.

13. **What is percentile**?

ANS: A percentile is a measurement that tells us what percent of the total frequency of a data set was at or below that measure. As an example, let us consider a student’s percentile in some exams.

If on this test, a given student scored in the 60th percentile on the quantitative section, she scored at or better than 60% of the other students.

Percentile is calculated by the ratio of the number of values below ‘x’ to the total number of values.

The Percentile Formula is given as,

|  |
| --- |
| Percentile (P)=(Number of Values Below “x”\Total Number of Values)×100 |

14. **What is ANOVA**?

ANS: Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not.

* Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components to use for additional tests.
* A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables.
* If no true variance exists between the groups, the ANOVA's F-ratio should equal close to 1.

**15. How can ANOVA help?**

**Ans:** ANOVA can help to identify the sources of variation in a data set. This can help to improve the accuracy of data predictions and analyses. Additionally, ANOVA can help to identify relationships between different variables in a data set. This information can be used to improve data models and predictions. Overall, ANOVA can greatly improve the quality of data science research and results by allowing researchers to focus on the areas that need improvement.