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

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

a) Decrease

b) Increase

c) Both of them

d) None

Ans 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

Ans 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

Ans 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

Ans 9 a) 0

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

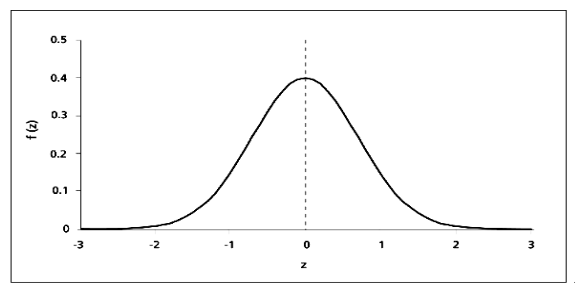
10. What Is Bayes' Theorem?

According to Bayes theorem, if an event A occurs based on the occurrence of another event B, then the conditional probability is equal to the likelihood of the second event given the first event multiplied by the probability of the first event. For example, we are to calculate the probability of a man having stroke given that they high hypertension. In such a case, if we know the probability of male having stroke and the probability of men who had stroke and had high hypertension, then

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

P(Event of man having stroke, given that he has hyper tension) = P(Had stoke and had high hypertension) \* P(Probability of having stroke)/ P(Probability of having high hypertension)

11. What is z-score?



Given a standard normal distribution, the centre depicts the mean given by zero, such that it divides the curve symmetrically. Then if we have a raw score, then standardizing it to z-score will describe the number of standard deviations away it is from the mean.

Calculation of z-score:

z = x- *μ/σ*

Where x is the raw score, μ is the mean and σ the standard deviation of the population.

12. What is t-test?

A t-test, also called student’s t- test is a method used for comparing the means of two groups. It is often used during hypothetical testing where the t-values helps us determine if there is statistically significant difference between the samples, on the basis of which we either reject or accept a Null Hypothesis. A t-test can be classified into i) independent test and ii) paired t-test. In independent t-test we take two separate samples and in paired t-test, we take the same sample for testing. A t-test can either be one-tailed or two-tailed, depending on the direction of the variation. For applying a t-test, the population we are dealing with should be normally distributed, have similar variance in each of the groups, and must have the same number of test size preferably between 20-30.

13. What is percentile?

When we have a continuous probability distribution, percentile will let us know the percentage of score that fall below a particular value. For example, if math test score 80 is at 80th percentile, we can assess that that their test score is higher than 80 percentage of the people who took the math test.

14. What is ANOVA?

Analysis of Variance (ANOVA) is a statistical test that is used for evaluating the variations found in more than two groups to see how dependent variable is impacted by the independent. variable/variables. There are two main aspects we compute during an ANOVA test which is i) Variance within a group and ii) variance between the groups. The ratio, which is the F value indicates if the means of the groups are related or not. If F-value is high, then the groups may have different means.

**F= Variance between the groups/ Variance within a group**

Depending on the result, we decide whether or not to reject the null hypothesis. ANOVA can be classified into one-way ANOVA and two-way ANOVA. One way ANOVA has only one factor effect the dependent variable where as two-way ANOVA has more than one factor.

15. How can ANOVA help?

While t-test is a simple test for comparison of two variables, the calculation becomes longer when it is more that two. This is where ANOVA mostly used. Hence, with the help of ANOVA we can test hypothesis to see if the independent variable have significant effect on the dependent variable.