STATISTICSWORKSHEET-4

Q1toQ15 aredescriptivetypes. Answerinbrief.

1. What is central limit theorem and why is it important?

   Answer.

   The central limit theorem states that the distribution of samples means.approximates a normal distribution as the sample size gets larger,regardless of the population's distribution. The central limit theorem is important for statistics because it allows us to safely assume that the sampling distribution of the mean will be normal in most cases.

2. What is sampling? How many sampling methods do you know?

   Answer.

   Sampling means selecting the group that you will actually collect data from in your research. Sampling is a technique of selecting individual members for a subset of the population to make statistical inferences from them and estimate the characteristics of the world population. There are two types of sampling methods. Probability sampling and non probability sampling.

3. Whatis the difference between type1and typeIIerror?

   Answer.

   Add type I error (false-positive) occurs if an investigator rejects a null hypothesis that is actually true in the population. A type II error (false-negative) occurs if the investigator fails to reject the null hypothesis that is actually false in the population.

4. Whatdoyouunderstandbytheterm Normaldistribution?

   Answer.

   A normal distribution is a type of continuous probability distribution in which most data points cluster toward the middle of the range. While the rest taper off symmetricallly toward either extreme.

5. Whatiscorrelationandcovarianceinstatistics?

   Answer.

   In statistics, correlation is a measure that determines the degree to which two or more random variables move in sequence. When an equivalent movement of another variable reciprocates the movement of one variable in some way or another during the study of two variables, the variables are said to be correlated.

   Covariance is a statistical term that refers to a systematic relationship between two random variables in which a change in other reflects a change in one variable.

6. Differentiatebetweenunivariate,Biavariate,andmultivariateanalysis.

   Answer.

Univariate statistics summarise only one variable at a time,Biavariate statistics compare two variables. Multivariate statistics compare more than two variables.

7. What do you understand by sensitivity and how would you calculate it?

Answer.

Sensitivity being the ability of screening test to detect true positive. Being based on the true positive rate reflecting a test's ability to correctly identify all the people who have a condition or If 100% identifying all people with a condition of industr by those people testing positive on the test.
Sensitivity = TP/TP+FN

8. What is hypothesis testing? What is H0 and H1? What is H0 and H1 for two-tail test?

Answer.

Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. H0 is Null hypothesis and H1 is Alternative Hypothesis. In a tooth tailed test. The generic null and alternative hypothesis are the following null: The effects equals zero alternative the effect does no equal zero.

9. What is quantitative data and qualitative data?

Answer.

Quantitative data measures of values or counts and are expressed as numbers. Quanlitative data are measures of types, and maybe represented by a name,symbol or a number code.

10. How to calculate range and interquartile range?

Answer.

Range is calculated by subtracting the lowest value from the highest value.
Interquartile range or IQR = Q3-Q1

11. What do you understand by bell curve distribution?

Answer.

A bell curve is a graph that is a normal distribution. The graph is a bell shaped line with
the curve's highest points show the most probable event in a number or series of data.

12. Mention one method to find outliers.

Answer.

Sorting method. We can sort quantitative variables from low to high and scan for extremely low or extremely high values. Flag any extreme values that we find.

13. What is p-value in hypothesis testing?

Answer.

The P value is a number calculated from a statistical test that describes how likely you are to have found a particular set of observations if the null hypothesis were true, P values are used in hypothesis testing to help decide whether to reject the null hypothesis.

14. What is the Binomial Probability Formula?

    Answer.

    Binomial probability refers to the probability of exactly x successes on a repeated trials in an experiment which has two possible outcomes. commonly called binomial experiment.

15. Explain ANOVA and it's applications.

    Answer.

    ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two or more categorical groups by testing for differences of means using variance. Another key part of ANOVA is that it is splits the independent variable into two or more groups. Applications of ANOVA is one-way ANOVA and two-way ANOVA

FLIP ROBO