

STATISTICS Assignment 4

Answer 1. if you take a sufficiently large sample size from a population with a finite level of variance, the mean of all samples from that population will be roughly equal to the population mean.

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.

Answer 2. sampling is a method when researchers determine a representative segment of a larger population that is then used to conduct a study. Sampling generally comes in two forms probability sampling and non-probability sampling.

simple random sampling, systematic sampling, stratified sampling, and cluster sampling

Answer 3. Type 1 error is known as false positive, i.e when we reject the correct null hypothesis, whereas type-2 error is also known as a false negative, i.e when we fail to reject the false null hypothesis

Answer 4. 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 symmetrically toward either extreme.

Answer 5. in statistical terms we use correlation to denote association between two quantitative variables.

Covariance is a measure of the relationship between two random variables and to what extent, they change together.

Answer 6. Univariate statistics summarize only one variable at a time. Bivariate statistics compare two variables. Multivariate statistics compare more than two variables.

Answer 7. Sensitivity analysis determines how different values of an independent variable affect a particular dependent variable under a given set of assumptions. The sensitivity is calculated by dividing the percentage change in output by the percentage change in input

Answer 8. Hypothesis Testing is a type of statistic analysis in which you put your assumptions about a population parameter to the test. it is used to estimate the relationship between 2 statistical variable. What does H0 and H1 mean?

In hypothesis testing there are two mutually exclusive hypotheses; the Null Hypothesis (H0) and the Alternative Hypothesis (H1).

Our null hypothesis is that the mean is equal to μ . A two-tailed test will test both if the mean is significantly greater than μ and if the mean significantly less than μ .

Answer 9. Quantitative data is data expressing a certain quantity, amount or range. Qualitative data is information that cannot be counted, measured or easily expressed using numbers.

Answer10. The range is calculated by subtracting the lowest value from the highest value. To find the interquartile range (IQR), first find the median (middle value) of the lower and upper half of the data.

Answer11. The term "bell curve" is used to describe a graphical depiction of a normal probability distribution, whose underlying standard deviations from the mean create the curved bell shape.

Answer12. Sorting method.,Data visualization method,Statistical tests (z scores),Interquartile range method.

Answer13. 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

Answer14. Binomial probability refers to the probability of exactly x successes on n repeated trials in an experiment which has two possible outcomes (commonly called a binomial experiment). If the probability of success on an individual trial is p , then the binomial probability is $nCx \cdot p^x \cdot (1-p)^{n-x}$.

Answer15. Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components to use for additional tests