

Assignment 2

Score: 55 / 55



What is the difference between population and sample in statistics?

In statistics, the population and sample are different. Population is the total set of people or objects that the researcher is interested in examining.

A sample is a portion of the population chosen for research or observation. It is employed to derive conclusions or infer things about the population.

Feedback

Great answer! The explanation is clear and demonstrates a good understanding of the difference between population and sample in statistics.



Define the term 'mean' in statistics and provide the formula for its calculation.

The sum of all the values in a dataset divided by the total number of values yields the mean, sometimes referred to as the average, which is a measure of central tendency.



Formula:

Mean=Sum of all values/Number of values

Feedback

Great answer! The definition of mean in statistics and the formula for calculation are accurately provided.



Explain the concept of standard deviation and its significance in statistics.

A dataset's dispersion, or spread around its mean, is measured by the standard deviation.

When the standard deviation is low, it means that the data points are often near the mean; when it is large, it means that the data points are dispersed.

Because it gives an indication of the variability or uncertainty in the data, it is important in statistics.

Feedback

Great answer! Well explained and covers all aspects of standard deviation in statistics.





In a given data set, what is the interquartile range (IQR) and how is it calculated?

A measure of statistical dispersion, or more precisely, the dispersion of data points around the median, is the interquartile range (IQR).

It is computed as the difference between the dataset's first (Q1) and third (Q3) quartiles:

$$\text{IQR} = \text{Q3} - \text{Q1}.$$

Feedback

Great answer! Well explained and correct.



What is the purpose of conducting a hypothesis test in statistics?

Using sample data, hypothesis testing is used to make judgements or infer inferences about a population parameter.

It assists in establishing if an effect or difference in the sample that is noticed is due to chance or is statistically significant.



Feedback

Great Answer! Well explained and comprehensive.



What is the difference between correlation and causation in statistics? Provide an example to illustrate.

A statistical metric that characterises the degree to which two variables change together is called correlation. It doesn't suggest a connection.

Contrarily, causation suggests that a change in one variable causes a change in another.

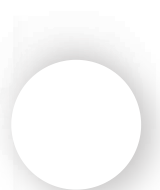
Example: The number of drowning events that occur in a city during the summer is highly correlated with ice cream sales. But drownings are not caused by ice cream sales; rather, hot weather is a common element that affects both variables.

Feedback

Great answer! Clear explanation distinguishing correlation and causation. Well-done!



Explain the concept of p-value in statistical hypothesis testing and its significance



If the null hypothesis is correct, the p-value represents the likelihood of receiving test findings that are as severe as the observed results.

It displays the degree of evidence that refutes the null hypothesis. A low p-value (usually less than 0.05) indicates that the observed results are not likely to be the result of chance, which means the null hypothesis is rejected.

Feedback

Great Answer! Well explained the concept and significance of p-value in hypothesis testing.



Provide a detailed explanation of the Central Limit Theorem and its importance in statistics.

Regardless of the form of the population distribution, the Central Limit Theorem asserts that as sample size rises, the sampling distribution of the sample mean approaches a normal distribution.

It is significant because it permits the use of inferential statistics such as hypothesis testing and confidence intervals even in cases where the population distribution is non-normal or unknown.



Feedback

Great answer! Explained the Central Limit Theorem and its importance clearly and concisely.



Define Type I and Type II errors in the context of hypothesis testing

When the null hypothesis is disregarded when it is true, this is known as a type I error. It stands for a fictitious positive.

When the null hypothesis is accepted even when it is incorrect, this is known as a type II error. It stands for a fabricated negative.

Feedback

Great answer! Well explained and understood.



What is the difference between inferential and descriptive statistics?

Methods for summarising and characterising a dataset's characteristics, such as its mean, median, mode, standard deviation, etc., are referred to as descriptive statistics.

Making assumptions or forecasts about a population from sample data involves inferential statistics, which



makes use of methods like regression analysis, confidence intervals, hypothesis testing, etc.

Feedback

Great answer! You have clearly explained the key differences between descriptive and inferential statistics.



Explain the concept of z-score in statistics and its applications.

The amount of standard deviations a data point deviates from the dataset mean is shown by its z-score, also known as its standard score.

The computation involves the following formula: $z = \frac{x - \mu}{\sigma}$, where x represents the individual data point, μ is the mean, and σ is the standard deviation.

Z-scores are useful for evaluating a data point's relative position within the dataset and comparing it to the mean. They are also employed in the normalisation and identification of outliers.

Feedback

Great Answer! Well explained z-score concept and its applications.

