

Assignment 1 Statistics

What exactly is the difference between descriptive and inferential statistics?

Descriptive statistics is analysis of data that helps to describe, show and summarize data in a meaningful way. This type of statistics is applied to already known data. Histograms, Bar graph, Pie-chart, scatterplot are some common technique to visualize the data.

In **Inferential statistics**, predictions are made by taking any group of data in which you are interested using different experiments. It can be defined as a random sample of data taken from a population to describe and make inferences about the population. Ex- Hypothesis testing, Null hypothesis, P-value , t-test, F-test and z-test etc.

I'm not sure what is the difference between a sample and a population?

A sample is a subset of individuals selected from a larger group (the population) to represent the population. The population is the entire group of individuals, objects or elements that have a common characteristic or interest. The sample is used to draw inferences and make conclusions about the population based on the data collected from the sample. The sample is usually smaller than the population and selected in such a way that it is representative of the population.

What distinguishes descriptive statistics from other types of statistics?

Descriptive statistics is a branch of statistics that deals with summarizing, organizing and presenting data in a meaningful way. It is used to describe the main features of a dataset, such as the mean, median, mode, range, and standard deviation.

On the other hand, inferential statistics uses sample data to make predictions or inferences about a larger population. In inferential statistics, we use statistical methods and models to test hypotheses, estimate population parameters, and make decisions based on the results.

So, the main difference between descriptive and inferential statistics is that descriptive statistics focuses on summarizing and describing the data, while inferential statistics focuses on using the sample data to make inferences about the population.

What is the difference between quantitative and qualitative data?

Quantitative data refers to data that is numerical and can be measured and quantified, such as age, height, weight, income, and scores on a test. These data can be analyzed using statistical methods and techniques to draw conclusions and make predictions.

Qualitative data, on the other hand, refers to non-numerical data, such as opinions, attitudes, beliefs, feelings, and categorical variables like gender or occupation. Qualitative data are often descriptive in nature and cannot be expressed as a numerical value. Qualitative data analysis often involves categorizing and organizing the data, and drawing conclusions based on patterns and themes within the data.

So, the main difference between quantitative and qualitative data is that quantitative data is numerical and can be measured and quantified, while qualitative data is non-numerical and descriptive in nature.

What is the definition of a percentile?

A percentile is a statistical measure that represents the value below which a certain percentage of observations in a dataset fall. In other words, it indicates the point at which a certain percentage of the data lies below that value.

For example, the 50th percentile, also known as the median, is the value that separates the lower half of a dataset from the upper half. If the median of a dataset is 30, then 50% of the observations in the dataset are less than or equal to 30, and 50% are greater than 30.

Percentiles are commonly used in data analysis to understand the distribution of a dataset and to identify values that fall outside the typical range. Percentiles can be calculated for any percentage, but the most commonly used percentiles are the 25th (lower quartile), 50th (median), and 75th (upper quartile).