



Introduction to Statistics

SGK GOVT DEGREE COLLEGE,VINUKONDA

CH.VIJAYA KALPANA

LECTURER IN COMMERCE



Descriptive vs. Inferential Statistics

1 Descriptive Statistics

Describes the main features of a dataset, such as mean, median, mode, and standard deviation.

2 Inferential Statistics

Uses sample data to make inferences about a larger population. This includes tests of significance and confidence intervals.

3 Use Case

To compare the average salaries of two different departments, we would use descriptive statistics. However, if we wanted to know whether gender affects salary, we would use inferential statistics.

Data Types and Measurement Scales

Data Types

- Nominal
- Ordinal
- Interval
- Ratio

Measurement Scales

- Nominal
- Ordinal
- Interval
- Ratio

Use Case

When analyzing customer feedback, we may use nominal data to categorize positive and negative comments, and ordinal data to rank the severity of each comment.

Sampling Techniques

1

Simple Random Sampling

Each member of the population has the same chance of being selected as a part of the sample.

2

Stratified Sampling

The population is divided into strata based on particular characteristics, and samples are taken from each stratum.

3

Cluster Sampling

The population is divided into clusters based on geographic or other criteria, and samples are taken from each cluster.

Central Tendency and Dispersion

same thing or not.

MEAN	AVERAGE
1. MEAN refers to the central point of a specific list of values. In order to find it, you need to add all of the values together and then divide the result by the number of values. - If the set of values is 3, 7, 8: Mean = $(3 + 7 + 8) / 3 = 18 / 3 = 6$	1. You can find yourself talking about AVERAGE, and this is the middle point of all the numbers that you have. - If the numbers that you have are 3, 7, 8: Average = $(3 + 7 + 8) = 18 / 3 = 6$
2. You can always say that "average" is a synonym to "mean".	2. You can't always say that "mean" is a synonym to "average".

• The method used and the result found are the same, so what's the difference? The answer is very simple: only terminology is different. The number that statisticians call mean is the same with the number that mathematicians call average.



Mean

The average of all numbers in a dataset; calculated by adding all values and dividing by the number of values.

1, 3, 3, **6**, 7, 8, 9

Median = **6**

1, 2, 3, **4**, **5**, 6, 8, 9

Median = $(4 + 5) \div 2$
= **4.5**

Median

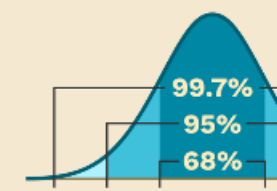
The middle value in a dataset; the value that separates the lower 50% of values from the upper 50%.

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

The number of data points

Each of the values of the data

The mean of x_i



Standard Deviation

A measure of the spread of values in a dataset; calculated by taking the square root of the variance.

Probability Concepts

Independent Events

Two events are independent if the outcome of one event does not affect the outcome of the other.

Conditional Probability

The probability of one event given that another event has occurred.

Bayes' Theorem

A formula for updating the probability of a hypothesis based on new evidence.

Use Case

To predict whether a customer will churn, we can use conditional probability to calculate the likelihood of that event given certain characteristics of the customer.

Hypothesis Testing

