

PRESENTED BY
MS.MANIMOZHI
TUTOR
ICON

1. STATISTICS: TYPES, CONCEPTS AND SIGNIFICANCE

DEFINITION

- ★ Statistics is the science of making effective use of numerical data which is related to collection, analysis and interpretation of data.
- ★ Statistics is the study of how to collect, organizes, analyze, and Interpret data.
- ★ Statistics is a form of mathematical analysis that uses quantified models, representations and synopses for a given set of experimental data or real-life studies.

TYPES OF STATISTICS:

❖ **Descriptive statistics:**

It consists of methods for organizing and summarizing information, it Includes graphs. charts, tables & calculation of averages, percentiles

❖ **Inferential statistics:**

✚ It consists of methods for drawing and measuring the reliability of conclusions about population based on information obtained.

✚ It Includes point estimation, interval estimation, hypothesis testing.

Both are interrelated. Necessary to use methods of descriptive statistics to organize and summarize the information obtained before methods of inferential statistics can be used.

Concepts:

Measures of central tendency

Measures of dispersion and variability

i. **Measures of central tendency**

A researcher collects the values of any variables under study from a large population, and the observation values cluster around a central value. This features of concentration of the value around the central value is known as central tendency.

a. **Arithmetic mean:**

The mean or average is probably the most commonly used method of describing central tendency.

b. **Median:**

Median of a set of values is the middle most value when the data is arranged in ascending order of magnitude. The middle value will divide the number of observation in the data into two equal parts.

c. **Mode:**

It is the value which has the highest frequency in the data.

ii. **Measures of dispersion and variability**

The observation deviating from the central value is different in different set of values of character. In some distributions, the difference may be less whereas in some other, it may be more. This property of deviation of the values from the average is called variation or dispersion.

a. **Range:**

It is the difference between highest and lowest value in the data.

b. **Standard deviation:**

It is the positive square root of mean of the squared deviations of values from the arithmetic mean.

SIGNIFICANCE:

- ✚ Statistics plays a vitally important role in the research.
- ✚ It help to answer important research questions and field in study.
- ✚ Helps you understand how to apply statistical method
- ✚ Important to understand what tools are suitable for a particular research study
- ✚ Statistics enables to understand specified statistical concepts and procedures.
- ✚ It simplifies mass of data (condensation);
- ✚ Helps to get concrete information about any problem
- ✚ Helps for reliable and objective decision making:
- ✚ It presents facts in a precise & definite form;
- ✚ It facilitates comparison(Measures of central tendency and measures of dispersion);
- ✚ It facilitates Predictions (Time series and regression analysis are the most commonly used methods towards prediction.);
- ✚ Researchers link the statistical analyses they choose with the research question, design, and level of data collected.
- ✚ Allows us to critically analyze the results.
- ✚ Provide organization and meaning to data.

2. DIFFERENCE BETWEEN PARAMETRIC AND NON- PARAMETRIC TEST:

Parametric test:

The statistical methods of inference make certain assumptions about the population from which the samples are known.

non- parametric test:

- ★ A hypothesis test that does not require any specific conditions concerning the shape of the population or the value of any population parameters.
- ★ Generally easier to perform than parametric tests.
- ★ Usually less efficient than parametric tests (stronger evidence is required to reject the null hypothesis).

PARAMETRIC	NON- PARAMETRIC
Test statistics based on the distribution	Test statistics is arbitrary
Parametric tests are applicable only for variable	It is applied bot variable and artributes
No parametric test exist for nominal scale data	Non parametric test do exist for nominal scale and ordinary scale data
Parametric test is powerful, if it exist	It is not so powerful like parametric test
Information about the population is completely known	No Information about the population is available.
Specific assumption are made regarding the population	No assumption are made regarding the population
Null hypothesis is made on parameters of the population distribution	Null hypothesis Is free from parameters

3. PROBABILITY AND ITS CHARACTERISTICS:**Probability:**

The probability of a given event is an expression of likelihood of occurrence of an event. A probability is a number which ranges from 0 to 1.

Characteristics:**i. The normal curve is symmetrical:**

The Normal Probability Curve (N.P.C.) is symmetrical about the ordinate of the central point of the curve. It implies that the size, shape and slope of the curve on one side of the curve is identical to that of the other.

ii. The normal curve is unimodal:

Since there is only one point in the curve which has maximum frequency, the normal probability curve is unimodal, i.e. it has only one mode.

iii. Mean, median and mode coincide:

The mean, median and mode of the normal distribution are the same and they lie at the centre. They are represented by 0 (zero) along the base line. [Mean = Median = Mode]

iv. The maximum ordinate occurs at the centre:

The maximum height of the ordinate always occurs at the central point of the curve that is, at the mid-point. The ordinate at the mean is the highest ordinate and it is denoted by Y_0 . (Y_0 is the height of the curve at the mean or mid-point of the base line).

v. The normal curve is asymptotic to the X-axis:

The Normal Probability Curve approaches the horizontal axis asymptotically i.e., the curve continues to decrease in height on both ends away from the middle point (the maximum ordinate point); but it never touches the horizontal axis.

vi. The height of the curve declines symmetrically:

In the normal probability curve the height declines symmetrically in either direction from the maximum point. Hence the ordinates for values of $X = \mu \pm K$, where K is a real number, are equal.

vii. The points of Influx occur at point ± 1 Standard Deviation ($\pm 1 \sigma$):

The normal curve changes its direction from convex to concave at a point recognized as point of influx. If we draw the perpendiculars from these two points of influx of the curve on horizontal axis, these two will touch the axis at a distance one Standard Deviation unit above and below the mean ($\pm 1 \sigma$).

viii. The total percentage of area of the normal curve within two points of influxation is fixed:

Approximately 68.26% area of the curve falls within the limits of ± 1 standard deviation unit from the mean

ix. Normal curve is a smooth curve:

The normal curve is a smooth curve, not a histogram. It is moderately peaked.

x. The normal curve is bilateral:

The 50% area of the curve lies to the left side of the maximum central ordinate and 50% lies to the right side. Hence the curve is bilateral.:

xi. The normal curve is a mathematical model in behavioural sciences:

The curve is used as a measurement scale. The measurement unit of this scale is $\pm \sigma$ (the unit standard deviation).

xii. Greater percentage of cases at the middle of the distribution:

- ★ There is a greater percentage of cases at the middle of the distribution. In between -1σ and $+1\sigma$, 68.26% ($34.13 + 34.13$), nearly $2/3$ of cases lie.
- ★ To the right side of $+1\sigma$, 15.87% ($13.59 + 2.14 + .14$), and to the left of -1σ , 15.87% ($13.59 + 2.14 + .14$) of cases lie.
- ★ Beyond $+2\sigma$, 2.28% of cases lie and beyond -2σ also 2.28% of cases lie.
- ★ Thus, majority of cases lie at the middle of the distribution and gradually number of cases on either side decreases with certain proportions.

xiii. The scale of X-axis in normal curve is generalised by Z deviates

xiv. The equation of the normal probability curve reads

(equation of the normal probability curve) in which x = scores (expressed as deviations from the mean) laid off along the base line or X-axis.

xv. The normal curve is based on elementary principles of probability and the other name of the normal curve is the 'normal probability curve'.