Concept of Statistics in Social Sciences

Dr. P.K. Gupta*

Abstract

Statistics is a mathematical science involving the collection, interpretation, measurement, enumerations or estimation analysis, and presentation of natural or social phenomena, through application of various tools and technique the raw data becomes meaningful and generates the information's for decision making purpose. It is the systematic arrangement of data and information exhibits their inner relation between the things. Statistics plays a vital role in every fields of human activity and has important role in determining the existing position of per capita income, unemployment, population growth rate, housing, schooling medical facilities etc. in a country, by which the decision making and development plans of the government becomes concentric. Now statistics holds a central position in almost every field of research like Industry, Commerce, Trade, Physics, Chemistry, Economics, Mathematics, Biology, Botany, Psychology, Astronomy, management of decision making etc. Through this research article, I have tried to discuss about the role and Importance of Statistics in various sociological and applied fields.

Key Words: Statistics, Parametric & Non-Parametric Statistics, Social Sciences.

Introduction:

When tribal life came to an end and early cities began to be established rules required facts concerning their domain. They wanted to know the fighting capacity of their subjects and the possibility of living taxes. Herodotus gave details of one of our earliest statistical collection of data regarding wealth and population of Egypt about 3050 B.C. Remesus-II, the ruler of Egypt took a survey of the whole Egyptian land in 1400 B.C. in order to distribute it among his subjects. Moses numbered the tribal of Israel in order to determine their fighting capacity. David also did the same about 1018 B.C. at about the same time, in 1200 B.C., the Chinese Government compiled data regarding various provinces. In Rome also elaborate census were taken for the purpose of taxation.

Statistics began only in the sixteenth century in the real sense. Feudal Lords in Germany enumerated the population of their domain. Names of Charlesmagne, William and Fredrick-II may be mentioned in this regard. In 1554, Sebstian Muenster, a professor at Heidelberg, published a systematized treatise on ancient countries. In 1612, Prof. George Obrecht of Strassburg University asked the government to keep complete record of vital and commercial statistics. In 1661, John Graunt made the first analysis of vital statistics. The seventeenth century, however saw the beginnings of serious interest in the mathematical chance. Jacques Bernouilli (1654–1705), the well non-mathematician gave the theory of probability. De Moivre (1667–1827) had written what is considered the greatest single work on probability. It was Gauss (1777–1855) who demonstrated the great practical value of the normal curve, showing how it applied to great practical value of the normal curve, showing how it applied to the distribution of measurement and to errors made in scientific observations. It was he who devised the fundamental modes of computations of means, probable errors, and the like. To this day we often see the normal curve referred to as the Gaussian Curve. Quetlet (1796–1874) royal astronomer became the great promoter of statistical methods on the continent of Europe. He demonstrated the fact that normal law of distribution applies to various types of anthropometric measurements when unselected populations are used. Sir Francis Galton (1822-1911) evolved the concept of correlation. Karl persons and Spearman have developed the statistical methods of correlations. R.A. Fisher had contributed in the field of statistics. History of statistics is associated with the development of tests.

* **Department of Statistics, D.A.V. (P.G.) College, Muzaffarnagar (U.P.)**

Definitions of Statistics

The term statistics has been defined widely by many mathematician and Statistician. Some important definitions are:

According to **W.G.Sulcliffe** : ''Statistics comprises the collection, tabulation, presentation and analysis of an aggregate of facts collected in a methodical manner without basis and related to a predetermined purpose.''

According to **M.G. Kendal** : ''Statistics is the science of collecting, analyzing and interpreting numerical data.''

According to **George A. Freguson** : ''Statistics is a branch of scientific methodology. It deals with collection, classification, analysis and interpretation of data obtained by conducting of experiments, observations and surveys. It's essential purpose is to draw inferences about the properties of population.''

The definition of Ferguson indicates on one type of statistics i.e. inferential or parametric statistics. The main focus of the definition is to infer about the characteristics of population. It is known as the process of generalization. The group or sample observations are not confined to the group but are used to estimate properties of the population from which the sample has been drawn. It does not cover the whole concept of statistics.

According **G. Simpson and F. Kofika** (1969): ''The word statistics today refers either to quantitative information or to a method of dealing with quantitative information of qualitative information.

On the basis of above definitions we conclude that statistics is**:**

1. a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data
2. a collection of quantitative data

Now the 'Statistics' is the independent field of study or academic discipline. It has wide application in behavioural sciences and social science.

Characteristics of Statistics

The above definitions indicate the main characteristics of Statistics. These properties of statistics have been enumerated in the following paragraphs.

1. Basically the term 'statistics' is used for the indexes or measures of a sample or a group. Whereas data or scores are the measures of an individual. The score is the numerical description of an individual and statistics are the quantitative description of a group or a sample.
2. 'Statistics' is the independent field of study or an academic discipline which deals with collection, classification, analysis and interpretation of data to draw the inferences.
3. 'Statistics' is an empirical science. The data are collected by conducting an experiments, observation and survey.
4. The foundation of statistical method is provided by mathematics, but now it has achieved recognition as an area of specialization in the general academic field.
5. Statistical methods are the techniques used to facilitate in the analysis and the interpretation of numerical data. It makes the raw data meaningful.
6. Statistics or statistical methods are used to draw inferences about the properties of population on the basis of sample statistics or observations.
7. Statistical techniques are used to estimate the sampling error or chance error for accurate or valid conclusions.
8. In social sciences and behavioural sciences large sample is selected for the study. A large sample reduces the error of measurement and the statistics of a large sample are the good estimates of parameters.

Kinds of Statistics:

There are two kinds of statistics

1. Descriptive or non parametric or Non-sampling and
2. Inferential or parametric statistics of the sampling statistics.

Statistical method and indexes are of two kinds: The Descriptive and the sampling. The first are measures, or indexes, the show the principal characteristics of a set of data. When it is found, for example, that 65 percent of entering college freshman class complete their course of study; this is a simple descriptive statistic for the group. When it is shown that the individuals in the lowest decide group (lowest to percent) on a scholastic aptitude test only 10 percent complete the course of study, we have a somewhat more analytical, descriptive between ranks on the test and achievement in college course for the entire group. For this purpose, several different statistical than is a percent; but is still in index descriptive of the observed data for this one group.

These indexes of not necessarily hold for groups or individuals other than the ones actually studied. General predictions cannot be made from these data and their indexes; nor can they be applied to results that might be obtained from other groups, unless the statistical techniques of predicting from a sample are used. In the case, the characteristics of the groups from which, and for which, the predictions are being made must be considered. That is to say, not only are sampling errors calculated, but comparability of groups must be established.

Analysis of the sampling error of obtained data is the second type of statistical method. The purpose of determining sampling errors is to indicate the limits with in which predictions can be made from one particular group to another and comparable group, or to predict from past to future performance of the one particular group. In other words, an answer is sought to this questions : ‘‘How dependable are these obtained data for predictive purposes ?’ Analysis of sampling errors provides a statistical method of finding an answer.

The basic problem in sampling statistics is to estimate the unknown values for the population from the known values obtained on sample observations. The statistical techniques are used for estimating parameters on the basis of statistics. It is known as statistical generalization.

In the literal meaning of the terms, a **parametric** statistical test is one that makes assumptions about the parameters (defining properties) of the population distribution(s) from which one's data are drawn, while a **non-parametric** test is one that makes no such assumptions. In this strict sense, "non-parametric" is essentially a null category, since virtually all statistical tests assume one thing or another about the properties of the source population(s).the difference between parametric and non-parametric statistics is shown in the following table:

Parametric and Non-parametric Statistics

S.No.	Parametric Statistics	Non-Parametric Statistics
1.	Information about population is completely known.	No information about the population is available.
2.	Specific assumptions are made regarding the population.	No assumptions are made regarding the population.
3.	Null hypothesis is made on parameters of population distribution.	The null hypothesis is free from parameters.

4.	Test statistics is based on normal distribution	Test statistics is arbitrary. Can be any.
5.	Assumed variance is homogeneous	Assume variance can be any-homogeneous or heterogeneous.
6.	Parametric test are applicable only for variables.	It is applied for both variables and attributes.
7.	Parametric exist for Ratio and Interval scale data	Non-parametric exist for Nominal and ordinal scale data
8.	Parametric test are powerful. Can draw more conclusions	It is not so powerful. Simplicity; less affected by outliers.
9.	Usual central measure is mean.	Usual central measure is median.
10.	Correlation method – Pearson	Correlation method – Spearman

Types of Data

Continuous and Discrete Series are the two major types of data:

1. Continuous Series or data and
2. Discrete Series or data.

1. Continuous Series or Data

A Continuous Series or data is one which is capable of any degree of subdivision. Though in practice divisions smaller than same unit are gravely met.

For example, I Q.s. achievement scores, mental test scores fall into continuous series; within the given range and score integral or fractional, may exist and meaningful. The mental and physical measurement yield data on continuous series. The weight of a person is 50.5 kg and 1.Q. is 110.5. Gaps in a truly continuous series can be attributable to get enough data. The precision of measurement fulfills the gaps.

2. Discrete Series of Data

A Discrete Series exhibit or indicate the real gaps. The data on nominal scale are on discrete series. For example, a class consists of 35 boys and 25 girls. There is real gap between the two groups.

In the field of Social Sciences and Education discrete data are obtained in terms of frequency on nominal scale. The interval scale or measurement provides continuous series.

Statistical methods and techniques are used for analyzing and interpreting data. The data is master while statistics is assists to make the raw data meaningful. Therefore, it is essential to understand source of data and ways of collecting data.

Functions of Statistics in Social Sciences:

The main functions of statistical techniques were to analysis the data and interpret the results. Thus, statistics makes the raw-data meaningful of measurement. The object and person cannot be measured but its characteristics or variables are quantified. Measurement of individual performance by means of test is usually expressed in terms of scores. The score is the numerical description of an individual with regard to his some variable. The statistics are the numerical indices of a group or a sample with regard to its some variable. The process of measurement converts the variable into variate or data. The 'variate' is the quantified variable.

Intelligence is a variable when the intelligence is measured in terms of I Q.s. The I.Q. is the data or variate. Thus, statistical methods attempt to analyse the variates. The statistics involves three types of analysis.

1. **Uni-variate Analysis** : When one variable is quantified, measures of central tendency, measures of variability, normal distribution and frequency distribution are employed.
2. **Bi-variate Analysis** : When two variables are quantified Pearson's product moment correlation, Spearman Rank difference, and other techniques of correlation are used.
3. **Multi-Variate Analysis** : When more than two variables are quantified. Multiple correlation, partial correlation, factor analysis and analysis of variance techniques are used.

Kinds of Variables:

The variable is the characteristics or trait of an individual which reveals the individual difference. The main feature of a variable is that it is quantifiable and converted into variate. The variables or variates are of two kinds:

1. Continuous Variable and
2. Discontinuous or Discrete Variable.

1. Continuous Variable :The continuous variates may take any value within a defined range of score between any two values of the variable an indefinitely large number of values may occur.

When we say that a student I.Q. is 120, it means that the range of value is 119.5 to 120.5. Similarly, the score in mathematics 60, means that the value ranges from 59.5 to 60.5. There is several values in between 19.5 to 60.5 e.g., 59.6, 59.7, 60.3, 60.4, 60.49, etc. this type of variate is obtained on interval scale of measurement.

2. Discontinuous or Discrete Variable : The discrete or discontinuous variates may take the specific value. It has the true dichotomy.

When we say that a class-group of students consists of 40 boys and 20 girls, in the staff of a college have 25 male teachers and 15 female teachers; there is no continuity in these data it has true dichotomy. These types of data are usually on nominal scale of measurement.

Main Function of Statistical Methods in Social Sciences:

There are three main functions of statistical methods which are as follows:

1. Organization of data, and presentation of data.
2. Analysis of data or treatment of data and
3. Interpretation of data

Conclusion: Statistics plays a vital role in every fields of human activity and has important role in determining the existing position of per capita income, unemployment, population growth rate, housing, schooling medical facilities etc. in a country, by which the decision making and development plans of the government becomes concentric. Now statistics holds a central position in almost every field of research like Industry, Commerce, Trade, Physics, Chemistry, Economics, Mathematics, Biology, Botany, Psychology, Astronomy, management of decision making etc.

Statistics is a mathematical science involving the collection, interpretation, measurement, enumerations or estimation analysis, and presentation of natural or social phenomena, through application of various tools and technique the raw data becomes meaningful and generates the information's for decision making purpose. It is the systematic arrangement of data and information exhibits their inner relation between the things, with their limitations also discusses the integration and co-ordination of statistical tools with the computer technology, which makes the calculation and interpretation data in very efficient and effective manner. Now a day's Computer technology has becomes as lifeline of statistical tools and software's.

References :

- Asthana, Bipin, 'Measurement and Evaluation in Psychology and Education' Vinod Pustak Mandir, Agra.
- Downie, N.M & R.W. Health (1970), 'Basic Statistical Methods', Happer and Row, Publishers, New York.
- Garrett, H. E(1947), 'Statistics in Psychology and Education' Longmans Green & Co., New York.
- Guilford, J.P (1950), 'Fundamental Statistics in Psychology and Education' McGraw-Hill Book Co., New York..
- McNemar, Q., 'Psychological Statistics' John Wiley & Sons, New York, Chapman & Hall Ltd., London.
- Pal, Surendra, (2017) 'Assessment for Learning' Arjun Publishing House, New Delhi
- Siegel, Sidney (1956), 'Non-Parametric Statistics', McGraw Hill Book Co., New York.
- Sharma, R.A.(2001), 'Essentials of Measurement in Education and Psychology', R. Lall Book Depot, Meerut.
- Walker, H.W(1943), 'Elementary Statistical Methods' Henry Holt & Co., New York
- https://www.researchgate.net/publication/269699659_Importance_of_Statistics_in_Social_Sciences_Research [accessed Jun 9, 2017].

●