



Introduction to Statistics

INTRODUCTION TO STATISTICS

Modern age is the age of statistics. H.G. Wells' prediction that, "*Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write*," has become true. Every citizen finds statistics in newspapers, magazines, advertisements in T.V. and radio etc. The figures relating to the various aspects of his life - social, political and economic - represent and support the observed facts or situations. The reader analyses the figures and arrives at certain conclusions.

Moreover, many have seen statistics as a device to achieve the degree of precision in the concept and theories of social sciences. In a nut shell, if we analyze the way in which statistics is looked at, *we broadly find two categories, one refers statistics as a set of figures and the other connotes it as a set of techniques.*

Origin and growth of Statistics

The origin of statistics is revealed by the word itself which is said to have been **derived** either from the Latin Word 'Status' or the Italian word 'statista' or the German word 'Statistik' which means political state.

Statistics was used as a by-product of administrative activity. Govt. maintains records of various types of numerical data on population, births, deaths, literates, illiterates, employment, unemployment, Income, Taxes, Imports, exports etc.

Statistics was used as a technique to collect periodical data to ascertain the manpower and material strength for military and fiscal purposes.

The theoretical development of statistics has its origin in the mid-seventeenth century when many gamblers and mathematicians of France, Germany and England are credited for its development. Pascal and P. Fermat, the two great French mathematicians made innovative efforts to solve the famous 'Problem of point' which was posed by the famous French gambler Chevalier De Mere. Their contribution became the foundation stone of the Science of Probability.

James Bernoulli (1654-1705) developed the 'Normal Curve'. The use of 'Statistics' was popularized by Sir John Sinclair in his work Statistical Account of Scotland (1791-1799). Modern Theory of Statistics was gradually developed during the 18th, 19th and 20th centuries mathematicians.

Laplace (1749-1827) gave the principles of 'Least squares' and the 'Normal Law of Errors'. The famous statisticians **Sir Francis Galton (1822-1911)**, **Karl Pearson (1857-1936)** and **W.S. Gosset** contributed to the study of Regression Analysis, Correlation Analysis as well as Chi-square test of Goodness of Fit, and t-test respectively. **A. Fisher, who is called "Father of Statistics"**, has developed statistics for use in genetics, biometry, agriculture, psychology and education.

Need or Importance of Statistics

Statistical techniques do have universal applicability. These techniques are used in almost all fields of knowledge e.g. social science, medical science, physical science, natural science and so on. So far as the field of economics is concerned, this technique has become so important that even the understanding of Elementary Economics requires knowledge of statistics. In this regard [Marshall opined, "Statistics are the straw out of which I like every other economist have to make bricks."](#)

However, the significance of statistical analysis is on the basis of the following grounds:

- It is a tool collecting information and processing the data.
- It provides a tool for scientific analysis.
- It provides solution for various business problems.
- It enables proper allocation of resources.
- It helps in minimizing waiting and servicing cost.
- It enables the management to decide when to buy and how to buy.
- It helps in choosing an optimum strategy.
- It renders great help in the optimum allocation of resources.
- It facilitates the process of decision making.
- Management can know the reactions of the integrated business system through quantitative analysis
- It helps in planning for the future and forecasting.

MEANING AND DEFINITION OF STATISTICS

Meaning of Statistics

According to Tate, "You can compute statistics by statistics from statistics."

Therefore the word statistics has three aspects;

- (a) Statistics
- (b) Statistical science
- (c) Statistical measurement.

According to Oxford Dictionary, "Statistics has two meanings, as in plural sense and as in singular sense.

In plural sense, it means a systematic collection of numerical facts. (Data)

In singular sense, it is the science or the methods of collecting, classifying and using statistics. (Method)

Statistics in the Plural Sense (Statistical Data):

In plural sense, it means a collection of numerical facts. It is in this sense that the public usually think of statistics, say figures concerned with population or production of wheat in India in different years or number of man-hours lost in industry in a specific year. Secondary Statistics e.g., percentages, averages and coefficients derived from numerical facts, are also included in the term statistics in this sense.

A.L. Bowley has given a series of definitions. Some definitions give emphasis on one aspect and others on some other aspect. At one place Bowley says "**Statistics, may be called the science of counting**". This view is not perfect and correct. Statistics is not concerned with counting only. It deals more with estimates. At another place, he says that "**Statistics may rightly be called the science of averages**". But calling statistics as a science of counting or averages, confines the scope of statistics. Bowley himself realized this drawback and stated that statistics cannot be confined to any one sense.

Webster defined Statistics as "**The classified facts respecting the condition of the people in a state- especially those facts which can be stated in numbers or in tables of numbers or in any tabular or classified arrangement**".

This definition has limited the scope of statistics. It relates statistics only to those facts which are concerned with the condition of the people in a state. This concept does not suit the modern world. Furthermore, this definition is not exhaustive because it does not take into account all aspects of human activity.

Characteristics or Features of Statistics in the Plural sense

The basic feature of statistics as a quantitative/numerical data run as:

Aggregate of Facts

Statistics does not refer to a single figure but it refers to a series of figures. A single weight of 50 kg is not statistics but a series relating to the weight of a group of persons is called statistics. It means, all those figures which relate to the totality of facts are called statistics.

Affected by Multiplicity of Causes

Statistics are not affected by one factor only, rather they are affected by a large number of factors. It is because statistics are commonly used in social sciences. It is not an easy job to study the effects of any one factor on a phenomenon and effects of different sets of factors separately.

In nutshell, we can say that statistics are affected considerably by multiply causes e.g. prices are affected by conditions of demand, supply, money supply, imports, exports and various other factors.

Numerically expressed

Another characteristic of statistics is that qualitative expressions like young, old, good, bad etc. are not statistics. For example, the statement like "The average score in the last ABC Exam was 640". Statements must contain figures so that they are called numerical statements of facts. Numerical expressions are precise, meaningful and convenient form of communication.

- Continued in the next slide

Characteristics or Features of Statistics in the Plural sense

- Continued from previous slide

Enumerated or Estimated according to Reasonable Standards of Accuracy

When the numerical statements are precise and accurate, they can be enumerated. But in case the number of observations is very large, in that case the figures are estimated. It is obvious that the estimated figures cannot be absolutely accurate and precise. The accuracy, of course, depends on the purpose for which statistics are collected. There cannot be uniform standard of accuracy for all types of enquiries.

Collected in a Systematic Manner

For accuracy or reliability of data, the figures should be collected in a systematic manner. Thus for reasonable standard of accuracy, the data should be collected in a systematic manner, otherwise the results would be erroneous.

Collected for a Pre-determined Purpose

The usefulness of the data collected would be negligible if the data are not collected with some pre-determined purpose. The purpose of collecting data must be decided well in advance. Besides, the objective should be concrete and specific. For example, if we want to collect data on prices, then we must be clear whether we have to collect whole-sale or retail prices. If we want data on retail prices, then we have to see the number of goods required to serve the objective.

Placed in Relation to each other

The collection of data is generally done with the motive to compare. If the figures collected are not comparable in that case, they lose a large part of their significance.

Singular Definitions or Statistical Methods

In singular sense, statistics implies statistics methods. It is a body or technique of methods relating to the collection, classification, presentation, analysis and interpretation of information.

According to Boddington: "Statistics is the science of estimates and probabilities."

According to W.I. King, "Science of statistics is the method of judging collective natural or social phenomena from the results obtained by the analysis of an enumeration or collection of estimates."

According to P.H. Karmel, "The subject statistics is concerned with the collection, presentation, description and analysis of data which are measurable in numerical terms."

According to Seligman, "Statistics is the science which deals with the method of collecting presenting, comparing and interpreting numerical data-collected to throw some light on any sphere of inquiry."

Stages in Statistical Investigations

Collection

It is the first step in a statistical inquiry. The collection part is the backbone of the inquiry. If the collection of data is not in proper form, in that case the conclusions drawn can never be reliable. The source of data may be primary i.e., collected directly, or the data may be secondary i.e., available from existing published sources. The first hand collection of data is one of the most difficult and important tasks faced by the investigator.

Organisation

The data collected from published sources are generally in organized form. But the figures which are collected from a survey, need organisation. The most important point in organizing a group of data is editing. This is done to correct omission, inconsistencies and wrong calculations in the survey. The classification is done to arrange the data according to some common characteristics possessed by the items. The last step in organisation is tabulation. The object of tabulation is to arrange the data in columns and rows for complete clarity as far as presentation of data is concerned. Thus organisation can be classified into three stages, (i) Editing (ii) Classification (iii) Tabulation.

Presentation

After collection and organisation, the data should be presented. If the data are presented in an orderly manner, the statistical analysis gets facilitated. As far as the presentation of data is concerned. The classified data are to be presented in such a way that it becomes easily understandable.

Stages in Statistical Investigations

- Continued from

previous slide

Analysis

Once the data are collected, organized and presented, the next step is that of analysis. The main objective of analysis is to prepare data in such a fashion so as to arrive at certain definite conclusions. The methods most commonly used are measures of Central Tendency and are called measures of the first order. Measures of Dispersion are called measures of the second order. Skewness, Correlation, Regression, Interpolation etc. are called measures of the third order. The analysis of facts based on observation is termed as (a) Scientific analysis, (b) Numerical analysis, (c) empirical analysis.

Interpretation

The last stage of statistical investigation is to derive the results and give comments on the inquiry in question. Interpretation means to draw conclusions from the data collected and analyzed. The interpretation of data is not an easy job and requires a high degree of skill and experience. If the analyzed data are not properly interpreted, the whole object of the inquiry may be erroneous. It is only correct interpretation which may lead to reliable conclusions. Thus it is clear that statistics is a science of taking decisions on the basis of numerical data properly collected, organized

Functions / Applications of Statistics

The various applications or functions performed by statistics in modern times are discussed as under:

Simplification of Complex Facts: The foremost purpose of statistics is to simplify huge collection of numerical data. It is beyond the reach of human mind to remember and recollect the huge facts and figures. Statistical method makes it possible to understand the whole in the short span of time and in a better way.

Comparison: Comparison of data is another function of statistics. The object of statistics is to enable comparison with a view to ascertain the reasons for changes.

Relationship between Facts: Statistical methods are used to investigate the cause and effect relationship between two or more facts. The relationship between demand and supply, money-supply and price level are

Functions / Applications of Statistics

Formulation and Testing of Hypothesis: The most theoretical function of statistics is to test the various types of hypothesis and discover a new theory.

Measurement of Effects: Statistical methods act as a guide to measure the effect of a policy. For example, the effect of a change in bank rate or a change in incomes tax etc. can best be judged by the statistical methods.

Forecasting: Statistical methods are of great use to predict the future course of action of the phenomenon. It is only on the basis of statistical techniques that the planners in India prepare future estimates for production, consumption, investment etc.

Enlarges Individual Knowledge: Statistical methods sharpen the faculty of rational thinking and reasoning of an individual. It is a master-key that solves the problems of mankind in every sphere of life. Thus, Whipple has rightly opined that statistics enables one to enlarge his horizon.

Limitations of Statistics

Although, statistics is a very useful science yet it suffers from certain limitation. According to Newsholme, "It must be regarded as an instrument of research of great value but having several limitations which are not possible to overcome and as such they need a careful attention."

Statistics does not deal with individuals: Statistics deals only with the aggregates rather than individual items. In statistical methods, we deal with aggregates and not with a single figure. When we say that average height of a class is 5' 8', this individual figure refers to the aggregate of individuals. Thus it is clear that statistics is concerned with aggregates and not with individual items.

Qualitative Aspect Ignored: The statistical method cannot study the nature of phenomena which cannot be expressed in quantitative form. The phenomena which cannot be expressed quantitatively, cannot be a part of the study of statistics. These characteristics include health, intelligence etc. There is no doubt that the data which cannot be quantitatively expressed, needs conversion of qualitative data into quantitative data. Thus experiments are being conducted to measure the reactions of human mind statistically. One of the branch of statistics, the "Theory of Attributes" deals with qualitative data.

Limitations of Statistics

Statistics deals with Average: Statistical findings are true only on an average. According to W.I. King, "Statistics largely deals with averages and these averages may be made up of individual items radically different from each other." For instance, if we may say that the average production of wheat in the last ten years is 250 quintals, it does not mean that the production of every year is equal. Production of a particular year may be less or more than the other years. Thus, statistical information is true only on an average.

Statistics can be Misused: Statistics can be misused by ignorant or wrongly motivated persons. The data used by untrained people can lead to misleading results. The statistics can be handled correctly only by those who have sufficient knowledge in statistics. W.I. King points out, "One of the short-comings of statistics is that they do not bear on their face the label of their quality."

Results true only on an Average: We know that statistics is not as accurate science as other science. Similarly the statistical methods are not very precise and correct. In the same fashion, the laws of statistics are not universal like the laws of physics, Chemistry or Astronomy. The statistical laws are true only on average. Statistics are concerned with those phenomena which are affected by multiplicity of causes. In this way