

a systematic study of birth and death statistics. The computation of mortality tables and calculation of life expectancy at different stages led to the establishment of the life Insurance Institution in London in 1698. However, all these studies were carried out under the name of 'Political Arithmetic'.

The first study of the theory of probability was also done in the mid seventeenth century, inspired by an attempt to estimate the chances of winning or losing in a gamble. The study of J. Bernoulli, which contained and 'Law of Large numbers', was published in 1713. De-Moivre published his famous "Doctrine of Chance" in 1718 and also discovered the normal probability curve. The term 'statistics' is said to have been used as a 'subject matter' in 1749 by Gottfried Ache wall. The modern theory of statistics is said to have been formulated by L.A.J. Quelled who introduced the concept of 'average' and deviation from average. He discovered the principles of 'Constancy of Great 'Numbers' which forms the basis of sampling. There have been some grate contribution to the science of statistics in the nineteenth and twentieth centuries. Karl Pearson who conceptualized 'Chi-Square test.' founded the greatest statistical laboratory in England Sir Ronald Fisher pioneered the study of Estimation Theory' and also applied statistics to a variety of diversified fields such as genetics, biometry, psychology, agriculture and education. The application of scientific methods has considerably widened in the last 2 decades and there is hardly and field that does not make use of statistics. The use of information technology had made it easy and cost effective to use statistics.

Originally, statistics was restricted to recording and classification of data. Today, it has bloomed into a science and an analytical tool with significant applications in almost all disciplines ranging from Mathematics and Economics to areas such as Sociology, Psychology,

Definition of Statistics

The term 'Statistics' has been defined differently by different authors. Some have defined the word in the sense of numerical data, whereas others have defined in the sense of statistical methods. A few definitions have been discussed here.

Statistics as numerical data (Plural Sense)

Bowel defined Statistics as 'Numerical statement of facts in any department of enquiry placed in relation to each other' According to Yule and Kendall "By statistics we mean quantitative data affected to a marked extent by multiplicity of causes"

A comprehensive definition is given by Prof. Horace Secrist that points out all characteristics that numerical data must possess to be called statistics.

According to Horace Secrist "Statistics may be defined as the aggregates of facts affected to marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to reasonable standard of accuracy, collected in a systematic manner for a predetermined purpose and placed in relation to each other." According to above definition, Statistics should possess the following characteristics:

- 1. Statistics are aggregates of facts :** Single and unconnected figures are not statistics, for example, if it is stated that Raj, a student, secured 50 marks, it is not statistics. However, marks secured by the students of the class would constitute statistics. A single figure relating to production income, marks height, etc. cannot be regarded as statistics but aggregates of such facts would be regarded as statistics.
- 2. Statistics are affected to a marked extent by multiplicity of causes :** Numerical facts are affected by a multiplicity of factors. For example, the price, of a commodity is affected by number of factors such as supply, demand, imports, exports, money in circulation, competitive products in the market and so on. It is very difficult to study

separately the effect of these factors on the price of the commodity. In physical sciences, it is possible to isolate the effect of various factors on a single item. But statistics are commonly used in social sciences. In social sciences it is very difficult to study the effect of any one factor separately. In statistical methods, the effects of various factors affecting a particular phenomenon are generally studied in a combined form, though attempts are also made to study the effects of different factors separately as well.

3. **Statistics are numerically expressed** : Only numerical data constitutes statistics. Qualitative expressions like good, bad, young, old etc. cannot be regarded as statistics. Statement like the standard of living of people in India has improved or 'production of petroleum products has increased in India do not constitute statistics. But Statements like "Rice production in India in 2003-04 is expected to be 86.4 million tons as against 72.66 million tons in 2002-03" (Source Economic Survey 2003-2004) is statistics.
4. **Statistics are enumerated or estimated according to reasonable standards of Accuracy** : Numerical information can be either enumerated or estimated. If they are enumerated i.e actually counted or measured, they are supposed to be exact and accurate. If complete enumeration is not possible because of the large size or high cost, then data is estimated by using sampling technique. Estimated figures cannot be absolutely accurate and precise. The degree of accuracy depends to a large extent on the particular purpose for which the information is collected and the nature of the particular problem about which the data is collected. There cannot be uniform standard of accuracy for all types of data collection. For example, while calculating the marks of students in an entrance examination or the number of votes received by a candidate in an election, it is necessary that it should be absolutely accurate, but while calculating the number or persons watching a cricket match on Television, the numbers need not be accurate.
5. **Statistics are collected in a systematic manner** : The data should be collected in a very systematic manner. For any socioeconomic survey, a proper schedule depending on the object of enquiry should be prepared and trained investigators should be used to collect the data. An attempt should be made to reduce personal bias to the minimum. Data collected in haphazard manner may give inaccurate results.
6. **Statistics are collected for a pre-determined purpose** : The Purpose of collecting the should be pre-determined. Otherwise, the data collected may not serve any purpose and may become useless. One should not waste time and money in collecting information that is irrelevant for the enquiry. For example, if the purpose of enquiry is to measure the cost of living of low-income group people, we should collect information about the items that are generally consumed by them. Hence for such an index it is useless to collect information on items such as cars, and other luxury items.
7. **Statistics should be placed in relation to each other** : Statistical information is collected mostly for the purpose of comparison. If the data collected cannot be compared, then much of the purpose of collection will be lost. The information collected should be homogeneous and not heterogeneous in character. Statistical data are often compared period wise or region-wise. For example, the data relating to population of a country for different years or population of different countries in some fixed period will be regarded as statistics. But data relating to the size of shoe of an individual and his intelligence quotient do not constitute statistics. Statistics should contain the above characteristics. In the absence of such characteristics, numerical data cannot be called Statistics. Hence, "all statistics are numerical statements of facts, but all numerical statements of facts are not statistics."

Statistics as Statistical Methods (Singular Sense)

Statisticians' defining statistics as statistical methods emphasize on the usage and interpretation. A.L Bowley defines statistics as science of counting, science of average and science of measurement.... at various points of time. Boddingtons defines statistics as a "science of estimates and probabilities." Harlow defines statistics as the science and art of handling aggregate of facts-observing, enumeration, recording, classifying and otherwise systematically treating them. According to Crouton and Cowden, "Statistics may be defined as the science of collection, presentation, analysis and interpretation of numerical data. On the basis of the above definitions, we can identify the following characters ties.

1. **Observation and Collection** : A Strong foundation to statistical analysis can be made only through minute observation of various elements (parts) of any study, such as impact of price reduction on consumption of soft drinks, etc, Observation should result in filtering redundant and irrelevant data. While observation will result in first hand collection of data, one also needs to collect existing published information with regard to subject of study. Maximum care is required at this stage, as faulty data will lead to faulty analysis.
2. **Organization of Data** : If the data collected is voluminous, it has to be organized through careful editing and appropriate classification. Data might also need Tabulation to ensure clarity at the time of presentation.
3. **Presentation** : An orderly presentation of data is invaluable for proper analysis. Data can be presented in the form of simple numbers (e.g. GDP of India is US\$ 2.07 trillion), Tables, charts, Diagrams etc. Presentation of data should ensure that data collected is not cluttered and that data required for specific analysis is readily available.
4. **Analysis** : Raw data collected and presented is converted into information through meaningful statistical analysis. The discipline of statistics offers a variety of methods for analysis, ranging from simple computation of Arithmetic Mean, to sophisticated computer assisted techniques such as Discriminate Analysis. Analysis of data presents facts which might not be obvious otherwise.
5. **Interpretation** : Interpretation refers to drawing conclusions from data analysis. Interpretation is subjective and hence, this stage is most crucial in terms of efficacy of statistical research. Objective interpretation requires skill, experience and courage to accept facts that might not be to one's liking. An objective interpretation is an invaluable aid in decision making.

Difference between the two definitions

Adapting a particular definition of Statistics would mean making some inferences about the characteristics of following are a few differences in terms of what can be inferred from the two definitions.

Statistics as numeric data

It is quantitative
 It is often in the raw state
 It is descriptive in nature
 It Provides material for processing Unprocessed data. It does not help in decision- making.
 As it is, it would not make much sense without application of the tools of analysis
 Surely the choice of tools will depend on the nature of data

Statistics as a method

It is an operational technique
 It helps in processing the raw data
 It is basically, a tool of analysis
 The processing is done by the Scientific methods of analysis and interpretation.
 Tools of analysis will be idle without facts available for making use of such tools.
 The nature of the data to be collected will also depend on the tools sought to be used for processing.

Scope of Statistics

There is hardly any field that has remained untouched by Statistics. Statistics is viewed not just as a device for collecting numerical data, but as a means of sound techniques for their handling, analysis and for drawing valid inference from them. From this perspective, the scope or subject matter to statistics can be broadly studied under 2 heads namely.

(A) Statistical Methods and (B) Applied Statistics

Statistical Methods

Statistical methods are the tools that are in the hands of the statistician. They include all the general principles and techniques that are commonly used in the collection, analysis and interpretation of data. These methods are applicable to all kinds of data. The stage involved in study of any kind of data are (i) Observation and Collection (ii) Organization (iii) Presentation (iv) Analysis and (v) Interpretation. These have been explained under the topic 'Definition of Statistics as Statistical Methods.'

Applied Statistics

Applied Statistics deals with application of Statistical methods of specific problems or concrete forms. To illustrate, if a software services firm is experiencing attrition (loss of personnel), it may be worthwhile to investigate the reasons for the same. Special techniques can be employed to understand the underlying trends. For example, a correlation analysis could probably indicate that increasing opportunities (growth in the industry) and attrition are highly correlated. All such techniques and the results obtained by employing such techniques form part of applied Statistics.

Applied Statistics can be further classified into different categories. These categories are not water-tight. These categories are

- (i) **Descriptive Statistics** : Descriptive Statistics deals with data that is known. The describe the main features of such data. They are 'basic' to any statistical analysis and bring out such characteristics of data which could have escaped attention of the statistician. Tabulation, Averages and trends are examples of such descriptive statistics.
- (ii) **Scientific Statistics** : Scientific statistics deals with formulation of statistical laws. These laws are based on quantitative data and generally hold true. For example, when we refer to the properties of standard normal distribution (to be studied under "Probability in Quantitative Techniques -II), we are referring to Scientific statistics. Scientific statistics are heavily used for the purpose of business forecasting.
- (iii) **Analytical Statistics** : This includes methods such as correlation, regression, etc that help in establishing functional relationship between variables. Two or more sets of data are compared and analyzed to arrive at the relationship between them. Such relationship could be continuous or at a defined period of time.
- (iv) **Inferential Statistics** : Statistical methods that help us arrive at certain conclusion based on study of sample data are part of inferential statistics. It must be noted that no guess work is being made. The size of the sample, criteria for inclusion of a person/ activity/thing into the sample, etc are all clearly defined. Based on the results of the study on the sample, inferences about the total population can be drawn.
- (v) **Inductive Statistics** : Statistical methods that help in arriving at general consensus based on a study of random observations are part of inductive statistics. Unlike inferential statistics, the sample chosen under inductive statistics may not 'representative'. Hence, the conclusions arrived at need to be double checked.

Nature of Statistics

Statistics as Science : 'Science' is a systematic study of knowledge. It studies cause and effect relationships between different variables and comes up with principles that are

universally applicable. The law of gravity, Archimedes principle etc are all examples of the various studies in science. (From this stand point, statistics can also be considered as 'Science' Statistics is also involved in study of cause and effect.) It has come up with its own set of 'principles' that are widely used across various fields. However, the principles of Science have a very high degree of precision. This is on account of our ability to isolate the factors being studied in a laboratory environment. This is not possible in case of Statistics. Hence, statistical laws or principles cannot be expected to have the same degree of precision. Thus Statistics can be termed as a 'Social Science'. It is more of a 'Scientific Method' Than a science in itself.

Statistics as Art

(Art is an expression of skill or creativity. (It is not concerned with cause and effect analysis. It is concerned with doing things that cannot be done by everybody. It demands rigorous practice to continuously hone such skills.) Painting, Music, story telling etc are examples of Art. It is very evident that a person with many years of rigorous practice would be a better artist than a novice, even if the novices is 'gifted'. Statistics can very well be considered an Art in as much as it is concerned with the skill of handling facts and figures. Statistics are being very innovatively and creatively used in analysis and presentation of data. However, Art is an end in itself. Pursuit of art gives tremendous satisfaction. Statistics is only a means to an end. It is used to arrive at conclusions.

Thus, Statistics is both a Science and an Art. However, it is neither pure science nor pure Art. It is a scientific method that can be artistically used for the benefit of mankind.

Laws of Statistics

1. **The Law of Statistical Regularity :** The Law of Statistical Regularity states that "a moderately large number of items chosen at random from a large group are almost sure on the average to possess the characteristics of the large group. "According to this law, if a large 'population' has to be studied, a statistician will be able to get the same results by studying a sample, provided the sample is moderately large and is chosen at random. To illustrate, if a company, say M/s Pantaloons, wished to find out the preferences of today's youth in fashion apparel, it may not be worthwhile to try and ask each person about their preferences. Pantaloons may decide to interview about 10 to 15 persons every day at every outlet for a period of 7 to 15 days and try to arrive at some conclusions. These conclusions would not be materially different from what a study of the entire target youth in that city would result in.
2. **The Law of Inertia of Large Numbers :** The Law of Inertia of large numbers is a corollary to the Law of Statistical Regularity. According to this law, "Other things being equal, as the sample size increases, the result tends to be more reliable and accurate." To illustrate, let us consider a person attempting to calculate the average height of boys who are 15 years of age. If this person records the height of 10 boys and then calculates the average he will arrive at a figure. If he records the height of the 11th boy and then again computes the average, there will be significant change in the average. However, if this person has already recorded the heights of 1000 boys, there will be no major change in the average after the 1001th record is captured. As the number of observations tend to become very large, the results from such observations tend to remain constant, as extremes on one side get compensated by extremes on the other side.

Functions of Statistics : Following are the important functions of Statistics

1. **Systematic Collection and Presentation of Facts :** Facts backed by objective numbers carry lot more conviction. Statistics ensures that data is collected in a systematic

manner and presented without any subjective bias. Statistics adds a bit of precision and definiteness to general statements, resulting in greater conviction.

2. **Simplification of Mass Figures** : Statistics cuts through the clutter of data normally available and presents the facts effectively. The large mass of data collected is transferred into a few critical figures, which help in overall analysis and interpretation.

3. **Facilitates Comparison** : Statistics enables impartial and fair comparison of data across a wide range of competing alternatives. For example, Inflation rate compares the current prices with prices of previous year and indicates the trend of rising prices. Thus, Statistics performs the function of providing a definite meaning to the state of affairs by quantification of data and comparison.

4. **Helps in Formulation and Testing of Hypothesis** : Statistics has developed into a separate discipline, with its own theories, principles and methodology, which can be used to answer specific queries. For example, if a contractor has to decide between two alternative choices of cement, he has to check whether the quality of cement is satisfactory. He can take the help of statistics and use sampling techniques and principles to arrive at a decision. Similarly, Statistics can be used in determining the result of a medicine on a patient, the impact of newly introduced Government policy on the general public, and whether the purchase of a computer has resulted in efficiency gains etc.

5. **Helps in prediction** : Statistical methods are a very useful tool in attempting to forecast the future. If an investor wishes to purchase the shares of say, Reliance Industries Limited, he can use Statistics to find out whether the share price of Reliance is likely to increase or decrease. He can use correlation and Regression to know the extent of impact of some key factors on the performance of the company and then take his decision. Thus, Statistics is the closest substitute of Crystal Ball.

6. **Helps in Policy Formulation and Decision -Making** : Many a time, Government policy and even business decisions are taken intuitively or on "gut feel" However, Statistics can reveal as to whether such intuitive feeling can be translated into reality. For example, Statistics would have revealed that none of the previous Governments of Andhra Pradesh could successfully implement total prohibition of liquor in the state. Similarly, a businessman would be better off if he conducts a market research before entering into a particular business.

7. **Enlarge Individual Experience** : A Proper function of Statistics is to enlarge human experience and knowledge. Statistics makes it easier for man to understand, describe and measure the impact of any action. The efficient and sound techniques of Statistics that different fields of knowledge are opened up to mankind.

8. **Study Relationships between Various Phenomena** : Statistics enables us to observe and understand the relationships between different phenomena. For example, there is a strong relationship between unemployment related data and data pertaining to GDP.

9. **Measures Uncertainty** : Statistics helps in ascertaining the chance of occurrence of an event. It also helps in finding out the impact of the happening or not happening of an event. It helps mankind to cope with uncertainty. Thus, statistics performs the function of systematically creating knowledge through analysis of facts and help the user of statistics in making informed, intelligent decisions.

Importance of Statistics

Originally, statistics was understood as a tool to collect information. However, today, it is used as not just a device for collecting numerical data but as a means for their handling and analysis, thereby drawing inferences. The use of Statistics has spread to almost all

aspects of our lives. Statistics is an indispensable tool for any science requiring intelligent decision-making. While it is possible that some people may not have heard of or studied statistics, they still use statistics in their daily life and are affected by it. Let us understand some of the areas in which statistics finds its application.

1. **Importance of Statistics to Common Man** : Statistics creates data from the environment and converts it into information with the help of analysis. The common man has always been doing this in a small way to make his life better. The groceries purchased for the month by a housewife, the choice of mode of transport by an officer or a student, the choice of dates for sowing of seeds by a farmer, the modus operandi for searching something that is lost, are all examples of the sub-conscious use of statistical methods by the common man.

2. **Importance of Statistics to State** : The Government has been and continues to be the biggest collector and user of Statistics. Any policy decision taken by the Government is based on detailed statistical analysis. For example, introduction of a securities transaction tax is based on statistics pertaining to extent of trading done in securities. Similarly, a social measure such as ban on smoking in public places is based on data that helps evaluate the 'pros' and 'cons' of such measure. The state would be crippled without the availability of statistics. The use of statistics by the State is so wide that almost all ministries and departments have dedicated statistical units.

A major area for which Statistics is used by the Government is planning. The Annual Budget and the Five year plans are formulated based on wealth of Statistical data relating to production, population, consumption, investments, savings, incomes, spending patterns, prices (inflation), money supply, forex reserves, etc. The National Sample Survey (N.S.S) was set up in 1950 primarily for collection of statistical data for planning in India.

3. **Importance of Statistics in Business and Management** : Statistics has tremendously influenced the way business is managed. Sourcing of raw-material and other resources, selecting the right personnel, infusion of capital, forecasting of sales, control of costs, financing of research, etc are all decisions influenced by Statistics. Some of the key areas of importance of Statistics are

(i) **Insurance** : The Insurance industry almost entirely relies on Statistics to conduct itself. Life expectancy studies, mortality rates and actuarial valuations dictate the premium calculations.

(ii) **Financial Services** : The financial services industry is highly data oriented. For example, 80% of the trading in stock Exchanges is done relying to Technical Analysis, which is nothing but applied statistics. Banks are also prolific users of Statistics. All parameters of performance measurement of Banks are heavily statistics oriented. Credit Analysis also relies on Statistics.

(iii) **Market Research** : Market Research is no longer a qualitative assessment. Huge data is collected and analyzed using multiple tools to provide precise pointers to the prevailing trends. Job placement agencies also use Statistics to improve their ability to match employer, employee expectations.

(iv) **Quality Assurance** : Statistical Principles form the basis of carrying out quality Assurance exercises. For example, a software company specialized in "testing" may decide to adopt the 80-20 rules, which state that 80% of the defects could be found in 20% of the application. Quality standards prescribed for the manufacturing industry for product and Process control are based on Statistical Analysis.

(v) **Accounts and Audit** : Firms specializing in providing Accounting and Audit services have derived their techniques from Statistics. All Accounting information is analyzed through trend analysis and study of correlation amongst different components comprising the accounting information. The increase in marketing expenses and its correlation with increase in sales is an example of use of statistics in Accounting. It is difficult to think of an audit exercise without using the technique of sampling. Even in case of 100% audit (Physical verification of stock), use of principles of classification and tabulation of data are applied.

(vi) **Other users in Business and Management** : With the growing size and ever-increasing competition, the problems of business enterprises are becoming complex and they are making more and more use of statistics in decision making. According to Wallis Roberts, "Statistics may be regarded as a body of methods for making wise decisions in the face of uncertainty. "Modern statistical tools of collection, classification, tabulation, analysis and interpretation of data are important aids in making wise decisions. Business forecasting techniques are very useful in formulating business policies. Time series analysis gives valuable information regarding trend, seasonal and Cyclical movements of their products. Study of business cycles is of maximum importance to any businessman. The statistical tool of demand analysis enables the businessman to strike a balance between demand and supply. The technique of statistical tool of demand analysis enables the businessman to strike a balance between demand and supply. Random sampling technique can be adopted for the purpose of inventory control. Sample surveys are often made by business houses to know the consumer preferences for different products. On the basis of such survey suitable adjustments can be made in production and marketing of products. Statistics are used to evaluate alternatives and arrive at a decision. For example choice of a location, (Reliance Industries setting up a plant in A.P size of operations, (Tata Motors planning the productions of Indicia on 24-7 basis). Production planning, procurement of resources, product launch, capacity enhancement, outsourcing, Operations Research etc are all decisions that are based on information culled out of statistical Analysis.

4. **Importance of Statistics to Other Sciences** : Statistics is related to various other Sciences, some of which is described below:

(i) **Statistics and Economics** : Statistical data and statistical methods are immensely useful in proper understanding of economic problems and in the formation of economic policies. Statistical methods help in not only formulating appropriate economic policies but also in evaluating their effect. A new field, by the name Econometrics, has emerged which comprises the application of statistical methods to the theoretical economic problems. Economists today are no longer satisfied to theorize in abstract terms. They prefer to use the excellent data that is now available to build a sound factual foundation for their reasoning.

The use of statistical methods in Economics can be examined under the following broad heads.

(a) **Formulation of Economic Laws** : Statistical methods are used for conducting experiments on human behavior. Testing and verification of hypothesis is also done using statistical tools.

(b) **Study of Economic Problems** : Statistical data and methods of statistical analysis render valuable assistance in proper understanding of various economic problems such as Balance of Payments, Unemployment, Inflation, etc.

(c) **Compilation of National Income Accounts** : Data pertaining to various macro economic parameters is systematically collected, classified and analyzed, to understand the state of the economy.

(d) **Economic Planning** : Statistics is not only used to collect information for planning but also to come up with models for optimization of resources. Statistical models help in evaluating performance and triggering mid-course corrections.

(ii) **Statistics and Mathematics** : According to Connor "Statistics is a branch of applied Mathematics which specializes in data. "Statistics is intimately related to and essentially dependent upon Mathematics. The development of statistical techniques and theories are based on fitting different mathematical models of the observed data. This has led to the development of a new branch of Statistics called "Mathematical Statistics".

(iii) **Statistics and Natural Science** : Statistics are invaluable in the study of Natural Sciences. Factual data such as body temperature, blood pressure etc, are used for diagnosis. Statistics can help in estimating the impact of a new medicine for a particular disease. According to Karl Pearson, the whole doctrine of heredity rests on statistical analysis. Most experiments of scientific nature rely heavily on statistical data.

(iv) **Statistics and Social Sciences**- Social sciences involve human behavior and hence, its laws are not universally applicable to all situations. However, some generalization and guiding principles could be evolved only on account of Statistics. The vast applications of statistics in the field of psychology have resulted in a new subject called 'Psychometric'. As Croxton and Cowden have aptly remarked, "without an adequate understanding of statistical methods, the investigator in social sciences may be like the blind man groping in a dark room for a black cat that is not there."

(v) **Statistics and Other Disciplines** : Statistics is widely applied in almost all disciplines of study. Astronomy was one of the first disciplines to make use of statistics to study the movement of heavy bodies. The Normal law of Errors in Astronomy is based on principles of least squares. Meteorologists rely on Statistics to forecast atmospheric and other environmental phenomena such as Monsoons. Statistics is used by traders, agriculturalists, industrialists and service providers such as educational institutions, hospitals and Non Government Organizations (NGO's). There is hardly any Science that does not make use of statistics.

5. **Statistics and Polity** : Politics, or the art of governance, has also been a major user group of Statistics. The general elections of 2004 were fought on the basis of statistical data. The analysis of exit polls was also based on statistical tools. Not only in times of peace, but Statistics can be used effectively even in times of war. The theory of decision functions by A.L Weld can be put to great use in times of war to plan maximum destruction and minimum effort Efforts are made by a country to collect and analyze military data of potential 'enemy' countries, so as to prepare itself better in the event of war.

Conclusion

In the words of Croxton and Cowden," the methods of statistics are useful in an over-widening range of human activities in any field of through in which numerical data may be had. "This sums up the importance and use of Statistics. Statistics can be termed as the science of Universal applicability. However, it must be applied to a particular situation only after understanding the peculiarities of such a situation.