USE OF STATISTICS AND MISCONCEPTIONS ABOUT ITS USES AMONGST USERS OF STATISTICS

Dr. Mrs. J.V. Bhalerao.
Assistant Professor,
MGV's Institute of Management and Research,
Nashik, India.

ABSTRACT

The word 'Statistics' conveys meaning in singular as well as plural form. It has its origin in languages like Latin, German, and Italian. It was used to show political situations from ancient times.

The development in the areas like business, science and Government remarkably increased the use of Statistics. Various functions of statistics are to compare, forecast, simplify, analyze and interpret data to convey meaning which is helpful to researcher to large extent. The basic question in use of statistics is that **'Whether it is a science or art?'** If both the definitions are studied carefully one can conclude that statistics is both Art as well as Science.

The purpose of this article is to list out some uses of statistics, to elaborate upon the misconcepts of the users of statistics and the role of a statistician. It can be said that misuse of statistics is only, because of lack of technical knowledge, use of unsuitable test of significance or even unsuitable data, biased opinions and perception of investigators and not remembering limitations while analyzing and interpreting any data and problem.

Keywords: - Statistics, uses of statistics, misconcepts of users.

Introduction:-

The word 'Statistics' can be used in plural or singular form. When it is used in singular form; it refers to the branch of knowledge which deals with principles, devices or methods of collecting, analyzing and interpreting numerical data. If the statistics is used in plural form; it means the numerical data collected in systematic manner with definite aim or object.

The word statistics has its origin from the Latin word 'Status' or Italian word 'Statista' or German word 'Statistik' each of which means political state, the word statistics was used to show the facts and figures in various areas like economics Demographics and Political situations of a country which is usually brought out by the local government over the decades together.

It was the 19th century from when statistics has grown immeasurably as a science and has tried to touch every sphere of human activity. The basic question in use of statistics is that **'Whether it is a science or art?'** The dictionary meaning of the word science is an accumulated and accepted knowledge that has been systematized and formulated with ref. to the discovery of the general truth or operation of general law. The dictionary meaning of the word art is a skill, dexterity, or the power of performing certain actions acquired by experience, study or observation. If both the definitions are studied carefully one can conclude that statistics is both Art as well as Science. It is a science because its methods are basically systematic and have general applications. It is an art as its successful application depends to a considerable degree on the skill and the special experience of a statistician.

During last 100 years the use of statistical methods is so widened that today there is hardly any branch of knowledge which does not use these methods. In other words it can be said that all the sciences are handicapped without statistics. It alone may not be able to convey the meaning but the interdisciplinary use of statistics imparts meaning to many other sciences. It is been said that **"Science without statistics bears no fruits; statistics without science has no roots"**. The concepts in every branch of science or research or management are formulated, tested or modified with the help of statistics.

Use Of Statistics:-

It is been observed that application and use of statistics has grown remarkably because of three major areas business science and government. Government was using statistics from the time of emperor Akbar, but application of statistics are found in the areas of science and business latter on.

Various invention in science when got transformed to technology resulted in to industrial revolution. That was the time from when businesses started becoming complicated because of numerous government regulations, increasing taxes, advancement in technology, labor problems, and management of recourses; one of which is manpower and bottom line that is cutting of cost and maximization of profits. To understand the status of all these and to keep itself healthy businesses needed data, its analysis and interpretation which has increased demand of statistics. To evaluate the strengths and weaknesses of its economy, government collects and analyses dada. All branches of economics like consumption, production, distribution, exchange etc. are described, correlated, compared with the help of statistics. Thus it has also increased the demand for statistics.

In general it can be observed that growth of statistic, use of statistical methods in various fields have triggered development of science of statistics. The statistical techniques have improved to

such an extent that the cost and time of collecting and analyzing data has gone down drastically. Due to use of technology like computers and digitization; it has further reduced and made use of statistics more structured.

(1) **Statistics simplifies the marks of data: -** It organizes and presents figure in such a way that they start communicating meaning. At a glance representation, pictorial, diagrammatic or graphical presentation of data reduces complexity, increases attractiveness and facilitates comparison. It also helps understanding irrelativeness.

(2) **Statistics presents facts in definite form: -** The data collected is presented in numeral forms so that it has preciseness while convening meaning. A statement made like cold is increasing is not a statistics, but when it is said that the temperature has increased by $2^0 C$ than previous day it is statistics, because it precisely shows that cold has increased.

(3) **Statistics facilities comparison: -** The maximum use of statistic is for comparisons and finding out cause an effect relationship if at all they exist. The comparisons help finding causes for change and also helps in finding what changes can occur in future. The green house effect, depletion of ozone layer, global warming are examples where precisely statistics has predicted the changes and has also suggested that what could be the pressed benefits by the remedies those are found out.

(4) **Forecasting: -** It is the important function of statistics which is best on relationships and which tried to predict future events. The most powerful tool for businesses and economics.

(5) **Statistics helps formulate and test hypothesis:-** It not only helps formulating and testing hypothesis, but also infer from the same. It helps making rational decisions. It understands measures and describes the effects of action and reactions. It tries to correlate pieces of isolated events or information to make meaning full picture.

When as a science statistics is providing all these benefits to a mankind the users of it should also be experts. It is a double edged weapon. If it is rightly use it can prove extremely useful, but if misused might be disastrous. It is the statistician who has to take care of presenting facts as they are and not manipulating them to suit the requirements. The sound knowledge of statistical methods, careful investigation, collection and analysis of data should be assured by them. There are certain limitations of statistics due to which many people misuse statistics. Remaining part of the article is the journey through all such misconcepts or misuses of statistics.

Misconcepts Of Statistics: -

Now a days research is gaining more and more importance in every field, so is true further management field. When untrained people employ statistics to draw inferences it may be misleading, because the numbers are innocent the fault lies with the person who is playing with them.

(1) **Research must have testing of hypothesis: -**Every research taken up and is presented in the four of thesis or paper must contain a parametric or non parametric testing of hypothesis is the basic mis concept of researcher. If the data collected under research can not the presented in numbers, it is not possible to apply testing of hypothesis. In fact the biggest limitation of

statistics is that it needs data to be presented in numerical from. It is not possible to analyze qualitative data using statistics directly.

(2)   **Use of percentage method: -** Percentage method can not be used to validate the results because it is very easy method. It is the established fact that degree of change is usually expressed in terms of percentage. Also it converts the data to common base which facilitates comparison. If anybody thus uses percentage method to compare it is valid method for doing so.

(3)   **Sampling must be done: -** Whether the sampling is to be done or not depends upon population size under consideration. Suppose that a quality department of a manufacturing company receives a lot of 11 items from one vendor and a lot of 120 items from another vendor for quality check. It is decided that 10% items are randomly checked and the lot is accepted. How does the quality supervisor take the decision? The obvious answer is one item from 11 and 12 items from 120 should be selected and if they follow quality norms the whole lot should be accepted. How does it sound? Well, for a small population of 11 items why a sampling is required? When one can get exact results with 100% reliability by performing census and that also not involving much of resources why to use sampling? But it happens because of the carelessness and in expertise of a person using statistics.

(4)   **Method of sampling should always be random**:- All the time statistics emphasizes on the unbiasedness in the procedures as well as results. If a researcher is doing a survey related to opinion poll on some current affairs. If he tries to go for random sampling it is quite possible that he may not get the data at all. For the simple reason that random sampling requires up-to-date list of population and then the random selection of respondents. How is it possible to get such a list and then select sample? On the contrary when it is possible to get such a list, people go for convenience sampling and call it as random sampling. For example it a survey of doctors involves say pediatricians and a sample is to be selected, instead of selecting random sample researchers meet known pediatric doctors collect their opinions and declare that they have selected sample randomly.

(5)   **Application of uncertainty make statistics unreliable: -** Suppose it is been found from past data that 20% of certain surgical operations by a particular doctor are failures. You know that last 4 surgeries of the doctor were successful; will you be his 5th patient for the operation? Probability of failure is 20% doesn't mean that in every 5 cases one must fail. Statistical laws are not exact. For the 5 cases considered, all of them can be successful also. By the statement one only means that as sample becomes larger and larger the stated result follows, on an average.

(6)   **Statistical analysis must have minimum one small or large sample test or chi-square test: -** Now days, computerized statistical analysis is available to researchers. Thus to show that analysis is done and statistics is used researchers try to use any of above tests, because it gives weight age to research paper or thesis. Is it really needed? Is it really relevant to the objectives of the research? These are the questions nobody bothers about. Look at this example – The questionnaire had various question related to liking and disliking of various flavours of ice-creams. The researcher also has collected data related to demographic variables one of which is income of a person. Applying chi-square test between income group of a person and liking to specific flavour of an ice-cream the researchers conclusion is high income group people like butterscotch flavour. How is that? It is important to understand that tests can not be

applied to any of the data that is collected, purpose of collection is more important. Secondly even though chi-square reveals such kind of relationships it is spurious correlation so one can not use it as an inference.

(7)   **Sample size can be different for every question in the questionnaire: -** If a questionnaire contains let us say 10 questions. Unless and until there is diversion of data by options to any question. The sample size for analysis of each question must be same. If sample size decided is 200 and suppose that 150 have answered all the questions while remaining 50 have answered few questions in the questionnaire. The relevant sample size will be 150. one can not included the remaining 50 responses on question to question basis. These 50 respondents must be treated as redundant and should be omitted from final analysis. Thus to assure specified sample size researcher must assure to get fully filled responses from the respondents.

(8)   **Collection of data is more important than finalizing objectives: -** Researchers always think that collection of data is more important. Objectives can be formulated afterwards, so as to suit the analysis. Statisticians are not magicians that they can mold or manipulate the data to suit to the researchers expectations. Suppose that researcher collect the information on sales of all kinds of edibles like ready to cook food, readymade food items like grains and oils etc. which are used to cook food and so on. The objective finalized afterwards is the consumption pattern of ready to cook-food in the market. What is the use of all the remaining data that is collected and the resources those are employed for its collection? If the objective is to find which brand is more popular. Now you realize that brand wise information has not been collected at all. Now how will you use the collected data? No statistician can help you in this regard.

(9)   **The research should yield always positive and desired results:-** The frequent question asked to statistician is what should be done to get significant result? Statistics always tries to put forth facts unbiased. When a research is carried out why the researcher should always get positive results? Even it is proved that the data obtained is showing insignificant result it helps future researches in formulating their research problem, and saving time, which is the sole purpose of review of related literature.

(10)   **Convincing people using statistics:-** Have you come across the advertisements which say 80% people believe that our product is best? Well if 8 out of 10 people are rating the product best, will you treat is the best? Never the true sample size is quoted in such advertisements. But at least once every one thinks of buying that particular product. All of us keenly talk about the TRP of programmes on television. How many do really know the way of calculating it? The equipment that is fixed on to the T.V. set counts the number of times you switch over to a particular channel and the time of the same. Suppose you start T.V. with X channel at 9.00 o'clock in the night, it records you as a viewer of that channel for 9.00 pm programme, even if you do not watch the complete programme and switch over to another channel and programme immediately or after some time.

Conclusion: -

According to Prof. Chou "Derogatory statements have been inspired by the misuse of statistics on the part of untrained practitioners of statistical methods, ignorant users of statistical result, dishonest people who employ statistics to mislead and confuse rather than to clarity and illuminate."

It can be said that misuse of statistics is only, because of lack of technical knowledge, use of unsuitable test of significance or even unsuitable data, biased opinions and perception of investigators and not remembering limitations while analyzing and interpreting any data and problem. The biggest distrust about statistics is – **there are three kinds of lies – Lies, Damn lies and statistics, wicked in order of their naming.** It is the researcher and the statistician who has to work hand in hand to disprove the same.

References

Levin Richard and Rubin David (1998) "Statistics for Management" Prentice – Hall, India.

Arora P., Arora S. and Arora S. (2007) "Statistical Methods" S. Chand.

Ghosh R. and Saha S. (2008) "Business Mathematics and Satistics" Central Educational Enterprises, Kolkata.

Cooper D. and Schindler P. (2010) "Busyness Research Methods" Tata Mcgraw Hill.

Krishnaswami O. and Ranganatham M. (2006) "Methodology of Research in Social Sciences".