“Use – Misuse” of Statistics: A Domain of Mathematics

Harpreet Singh Maan
Assistant Professor, Department of Mathematics, SGGS College, Chandigarh

ARTICLE DETAILS

ABSTRACT

A section of mathematics deal with statistics. The foundational tool for assessing data relationships and analyzing data study is related to statistics. Unfortunately, it is observed that these tools are mostly misused, either inadvertently because of ignorance or lack of planning, or conspicuously to attain desired result. Data abuses include the incorrect application of statistical tests, lack of transparency and disclosure about decisions that are made, incomplete or incorrect multivariate model building, or exclusion of outliers. This leads to sum up the information inaccurately. Some researchers follow choose and pick method to select relevant data that can lead them to attain their required result as compared to exact and precise results.

Introduction

Statistics is a word coined to analyse the information related to states of a country only. It efficiently portrays the usage of mathematics for data calculations. Even in this modern era, statistics has not lost its charm and grace. While collecting information from clients or users, one keeps a track of various statistical tools that can be deployed to analyse data for a meaningful interpretation. Records of Scientific experimental values, accountancy, weather forecast data and many more are treated effectively with tools of statistics to observe pattern and inference new predictions on their basis.

In earlier or developmental stage, Statistics was only treated as a mechanism to organize both demographic as well as economic data of states for the purpose of human and material analysis. This observation was used to levy taxes on the masses or promote services for military purpose.

Earlier the collection of data was manual but nowadays, data is collected through electronic devices. To analyze this data, a dedicated data analyst is required who can treat large volumes of data. This processing is very tedious and while treating data accuracy vanished. Thus, certain researchers attempted to develop a computer intensive methods to minimize permutations and quantify accurate answers which could be not be attained otherwise.

In the 19th century, statistics increasingly used “probability theory” whose initial results were found in the 17th and 18th centuries, particularly in the analysis of games of chance (gambling). By 1800, astronomy used probability models and statistical theories, particularly the method of least squares. Johnson and Kotz (1970) produces a four volume compendium on statistical distribution (first edition 1969-1972), which is still valuable currently. [1]

The Statistical analysis depends on the objective of the study. The objective of a survey is to obtain information about the situation of the population study. The first Statistical task is therefore to do a descriptive analysis of variables. In this analysis it is necessary to present results obtained for each type of variable. For qualitative and dichotomous variables, results must be presented as frequencies and percentages. For quantitative variables, the presentation is as means and deviations. After this analysis, you can access the association between variables and predictive analysis based on multiple regression models.[2]

Use of Statistics

Statistical methods, theory, techniques, and models play an important role in several stages of the scientific method, but focus on just two stages. First, statistics is essential to good experimental design as in randomized clinical trials, for example. In order to obtain a rigorous test of a hypothesis, it important to obtain data that can provide evidence for or against the hypothesis. Second, statistics is important in an analyzing and interpreting data. There are many different statistical tools that one may use to analyze data, ranging from simple procedures, such as t-tests and linear regression, to more complex ones, such as analysis of covariance and statistical modeling. To apply any statistical method correctly, one must have information about the variables used (continuous or discrete, gaussian or bimodal, and so on), information about the sampling process used (sample size, independence, randomness, representativeness and so on), and a sound understanding of the theory and assumptions underlying that method. If a researcher does not use a method correctly, then conclusions may overestimate or underestimate an important relationship or effect. If we think of statistics as a tool for distinguishing between random “noise” in the data and the real signal, then someone who incorrectly uses statistics may produce a result that is distorted or even artificial. A person who correctly uses statistics will amplify and clarify the signal without distorting it. [3]
The use of statistics is mainly to handle various concept related with the quality of the data. To ensure the quality of data, one must ensure following parameters:

- Accuracy
- Bias
- Error
- Precision

Some of the basic statistics are:

- Mean: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ (1)
- Median
- Mode
- Variance
- Standard Deviation: $s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$ (2)
- Confidence Limits

Misuse of Statistics

Computers and statistical software packages have increased the complexity with which data can be analyzed and, consequently, the use of statistics in medical research has also increased. Unfortunately, though the types of errors may have changed, the frequency of statistical misuse has not[4]. There are multiple factors that contribute to misuse of statistics. It is observed that in research area, careless or deceptive approach towards use of statistics lead to misuse of statistical tools. Inadequate supervision, improper training or little knowledge act as fuel in this fire. Pressure to attain desired results lead to neglecting the accuracy of statistics.

One of the reason is that computers are fed with the data in order to analyze it effectively but in order to attain researcher’s demands, multiple iterations of data is done. It is also seen that these iterations, alter the values attained and bring the result to the margin of researcher’s demand.

Another reason is, only those values are selected out of the data obtained which can provide positive result. This choose and pick method often leads to faulty analysis of the data.

Even when data is collected with pre-mindset of attaining values that match with the study and neglect the data which can lead to null or negative hypothesis, also leads to misuse of statistics.

The pattern that emerges from this brief history is that initial efforts to approach the issue tend to be optimistically simple. Corrections over time add to the scope and complexity of the documents. The most recent document breaks out areas of ethical responsibility for all people using statistical methods professionally[5].

Conclusions

This paper attempted to throw a light on the conduct and misconduct of statistics in various field. There has been evidence in multiple researches about misuse of tools of statistics which lead to question out genuine usage of mathematics. This regime has been followed and will continue further too. To stop such practices, one need to understand the motive behind churning data into meaningful as well as inferential information. This can be done through proper planning, application and disclosure, combined with guidance and tools, so that hopefully researchers will continue to design, execute and interpret cutting edge.

References