



# “Basic Statistical Tools in Research for Data Analysis”

**Ravindra P. Rewaskar**

Department of Mathematics, Shankarlal Khandelwal College, Akola (M.S.), India.

email\_id: ravindrarkpr@gmail.com

## Abstract:

Generally, researchers are in constant demand regarding analysis of statistical data or information and the use of statistical methodologies in statistical treatment of data. Advanced mathematical statistics uses several techniques including statistical approaches, to handle data collected through experiments. The majority of research projects produce a significant amount of raw data, which needs to be appropriately reduced in order to make it easier to understand and utilize for additional analysis. It is evident that a researcher cannot overlook the science of statistics, even if he does not frequently utilize statistical procedures with all of their implications. Classification and well representation help to some extent to achieve this goal, but we still need to establish certain indices or measures to provide a summary of the data that has been gathered and classified. There are multiple types of data analysis techniques that are available. This paper gives idea about some statistical methods that can be used in handling the data in appropriate way.

**Keywords:** Research Methodology, statistical methods, data set.

## Introduction:

One approach to methodically addressing the research challenge is through research technique. It might be viewed as a science that studies scientific research methodology. In it, we examine the many approaches that a researcher often takes to analyse their research challenge and the reasoning behind them. It is imperative that the researcher possess knowledge of both methodology and research methods/techniques. The terms mean, median and mode are indicators of central tendency. The purpose and use of central tendency measures, like the mean, is to characterize the data by highlighting their centrality [1]. In addition to knowing how to create specific indices or tests, compute mean, mode, median, standard deviation, or chi-square and apply specific research techniques, researchers also need to understand which of these strategies or approaches are relevant and those that are not, as well as what their use would entail with reason. Like this, a scientist doing study must submit their decisions for review before carrying them out. For the judgments he makes to be assessed by others as well. They must be quite explicit about what they chooses and its reason.

Research methodologies do form an integral aspect of research methodology, which has several facets. Research techniques are more limited in scope than methodologies for research. In order to ensure that research findings can be assessed by the researcher or by others, when we discuss research methodology, we do more than just discuss the research methods themselves. We also take into account the reasoning associated with the methods we choose for the specific context of our study and provide an explanation for why we choose to use a given method or technique rather than others. When we discuss research methodology related to a research problem or study, common questions are addressed such as why the study was conducted, how the research problem was defined, how and why the hypothesis was formulated, what data was collected, which specific method was used, why a specific approach for data analysis was employed and a host of additional comparable inquiries.

In advanced coursework, a researcher might employ a hybrid technique, a quantitative approach or a qualitative one and when using a qualitative technique, researchers can choose how they want to analyse the data. Tables, charts, matrices and other visual representations are frequently used tools. The researchers can identify what will be important factors in his qualitative data analysis by using visual displays to whittle down the frequently copious subjective data that has been collected [2].

## **DIFFERENT STATISTICAL METHODS ARE USED TO REPRESENT DATA:**

Factual strategies are utilized to investigations and summarize information. Here are a few common and vital factual strategies utilized to represent information:

### **Histograms:**

Histograms are utilized to represent the dispersion of numerical information. They isolate the information into interims and show the recurrence of information focuses inside each interim. Histograms are especially valuable for understanding the shape and spread of information, like whether it takes after a ordinary conveyance or in case there are exceptions.

### **Scatterplots:**

They plot every piece of information as a point on a diagram, with one variable on the X-axis and the other on the Y-axis. Scatterplots are utilized to outwardly survey relationships or examples among factors and recognize any exceptions or groups of pieces of information. Scatterplots address the connection between two mathematical factors.

### **Bar charts:**

Bar outlines address absolute information by showing the recurrence or extent of every class as bars. They are helpful for looking at the dispersion of classifications or gatherings inside a dataset. Bar diagrams can likewise be utilized to imagine examinations between various datasets, across various time spans.

### **Line Graphs:**

By connecting data points with lines, line graphs facilitate the observation of trends and patterns and are useful for displaying changes in numerical data over time or another continuous parameter. When doing a time series analysis, graphs of lines are frequently used to show how variables like stock prices, temperature, or sales have changed during the course of time.

### **Boxplots:**

Boxplots show the median, quartiles and any data outliers in a visual representation of the numerical information distribution using quartiles. Boxplots can be used to compare the central tendency and spread of various datasets and to spot possible outliers.

### **Coefficient of Correlation:**

The linear link between the two numerical parameters is quantified using correlation analysis, along with its direction. Pearson's and Spearman's rank correlation coefficients are two examples of common correlation coefficients. Understanding data dependencies may be facilitated by using correlation analysis to determine the existence and strength of a relationship between the two variables.

## Regression Analysis:

The linear link between two numerical variables is quantified using correlation analysis, along with its direction. Pearson's and Spearman's rank correlation coefficients are two examples of common correlation coefficients. Understanding data dependencies may be facilitated by using correlation analysis to determine the existence and strength of a relationship between two variables.

## Measures of Central Tendency:

- a) *Mean*: The average value a set of data. In a collection of data, it has a tendency to depict a typical value.
- b) *Median*: In an organized set of numbers, the middle value is called the median. It is the average of both of the middle values if the total amount of values is even. When there are outliers in the data, it represents the core trend of the data better than the mean and is less susceptible to them.
- c) *Mode*: It is helpful in determining which value in a set of data is the most prevalent or well-liked. The mode is applicable to categorical data as well, in contrast to the mean and median.

## Measures of Dispersion:

- a) *Range*: The lowest value in a dataset is subtracted from the highest value to determine the range, which is the most basic measure of dispersion. It is susceptible to outliers yet provides a sense of the distribution of the data.
- b) *Variance*: The variance quantifies the average squared departure of every data point from the whole set's mean. It offers a gauge for the degree of deviation of the data points from the mean. Interpreting variance is more challenging since it is not on the same scale as the original data. Subtracting one from the number of data points, the variance is equal to the sum of squared deviations between each data point and the mean.
- c) *Standard Deviation*: Square root of variance equals standard deviation. Although it is on the identical scale as the original information, it is a more widely used measure of dispersion. The data units are more widely distributed when the standard deviation is bigger than when it is less, indicating that the data points are nearer to the mean.
- d) *Quartile Range*: The range of roughly half of the total data set is known as the quartile range. It is computed as the difference between the dataset's first quartile and third quartiles. Compared to the range, the quartile range is somewhat resistant to outliers.

It is important for researchers to comprehend the underlying assumptions of diverse approaches and to ascertain the criteria that will enable them to determine which techniques and processes are suitable for a given topic and which ones are not. All of this indicates that, as methodology might vary depending on the topic, the researcher must develop his approach specifically for his issue. An architect, for instance, must deliberately consider the rationale behind his choices while designing a structure. This includes considering the reasons behind his choices about the particular size, quantity and placement of doors, windows and ventilators as well as the materials he chooses to employ rather than others.

## Conclusion:

These statistical techniques offer a variety of tools for data representation and analysis, enabling investigators and analysers to find patterns, anticipate outcomes and assistance in decision-making. The type of data being used and the particular questions being answered determine which approach is best.

**References:**

- [1] Kaur, Parampreet, "Descriptive statistics," *International Journal of Academic Medicine* 4, pp 60 – 63, 2018.
- [2] M. Grabar and K. A. Krasnyanskaya, "Application of mathematical statistics in educational research. Non-parametric methods," Moscow: Pedagogika, 1977.
- [3] K. J. Begum and A. Ahmed, "The Importance of Statistical Tools in Research Work," *International Journal of Scientific and Innovative Mathematical Research*, vol. 3, no. 12, pp. 50-58, 2015.
- [4] Croxton, F.E., Cowden, D.J., Klein, S., *Applied General Statistics*, 3rd ed., New Delhi: Prentice-Hall of India Pvt. Ltd., 1975.
- [5] L. R. Gay, "Educational Research: Competencies for Analysis and Application," Upper Saddle River, New Jersey: Prentice-Hall Inc, A Simon & Schuster Company, , pp. 15-140, 144-305, 1996
- [6] A. Strauss and J. Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Newbury Park, CA: SAGE Publications, 1990.

