IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ROLE OF STATISTICS IN RESEARCH

Dr Pratiksha Pandey
Awadhesh Pratap Singh University, Rewa
[Assistant Professor]

Dr Ajay Kumar Govind Rao
Th Har Narayan Singh Degree College, Prayagraj
[Principal]

ABSTRACT

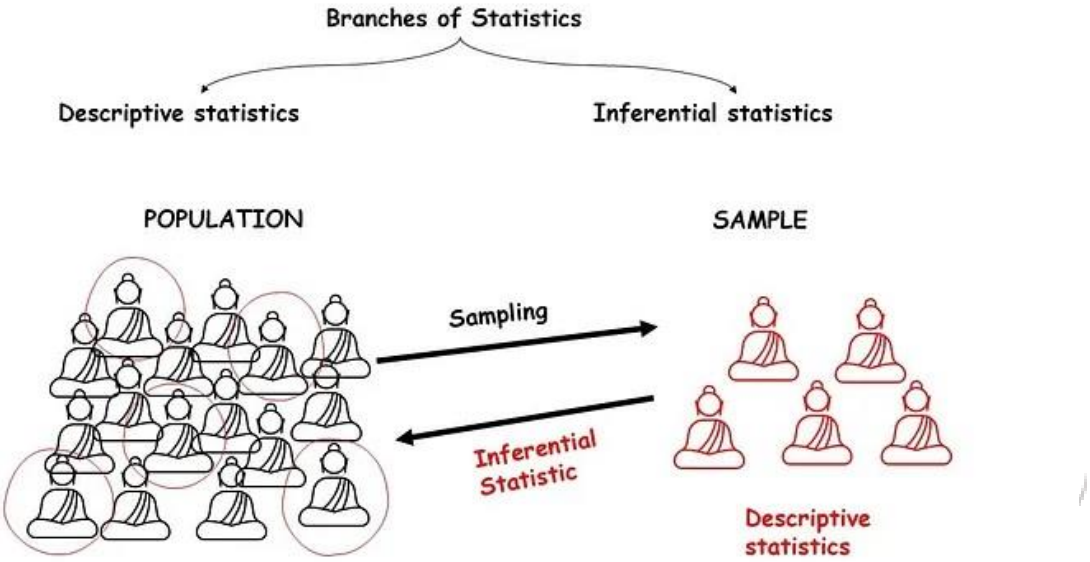
Most of the researcher's time went through in amassing and examining information: information around temperatures, events, concentrations, lengths, etc. The field of insights bargains with the investigation of data. Using fitting insights, you may be able to create sense of the large amount of information you have collected so that you can simply tell your investigated story coherently and with defense. There is a run of factual devices in research which can offer assistance analysts to oversee their investigated information and make strides the result of them inquire about by way better elucidation of information. You may utilize measurements in inquire about by understanding the inquire about address, information of measurements and your individual involvement in coding.

INTRODUCTION

Statistics is a branch of science that deals with the collection, organization and analysis of data based on a sample of the entire population. In addition, it helps to plan the research work more carefully and also provides a logical justification for concluding the hypothesis. However, statistics is a more complex field of study that defines and explains research models based on the sample size used. More specifically, the statistics provide the trend of the studies conducted. In research, research must be properly planned and the research sample must be selected in such a way that it has a positive effect on the results. In addition, the statistical method must be appropriately chosen. The effectiveness of clinical or epidemiological studies requires sufficient statistical data. When statistical methods are flawed, erroneous conclusions can be drawn, leading to unethical practices. Statistics can be very broadly divided into two -descriptive and inferential statistics.

1. **Descriptive statistics**: Descriptive statistics presents facts and proven results about a population, while inferential statistics analyzes samples to make predictions about larger populations. Descriptive statistics focus on the characteristics of the data set, the confidence level is very high. Outliers and other factors can be omitted from general observations to ensure greater accuracy, but calculations are often much simpler and

can lead to firm conclusions. Descriptive statistics are used to present or describe the characteristics or characteristics of data. They summarize specific numerical data or multiple sets and provide a quantitative view of the data through numerical or graphical representation. Descriptive statistics reflect only the data to which they are applied. Descriptive statistics can be: A measure of central tendency, such as the mean, median or mode: These are used to identify the mean or center of a data set. A measure of spread or variation, such as variance, standard deviation, skewness, or range: These reflect the spread of data points. Measure of distribution, such as the number or percentage of a given result: these express the frequency of that result in a data set example of a clinical drug trial, the percentage distribution of side effects and average age represent statistical measures of central tendency and normal distribution of that data set.



2. **Inferential Statistics:** Inferential statistics are used to make inferences or draw conclusions based on the data available on a smaller sample set. This is often done by analyzing a random sample of a much larger data set, such as a larger population. Conclusions drawn from this sample apply to the entire population. The importance and quality of the sample population is important to ensure the reliability of the drawn conclusion. This is true whether the population is a group of people, geographical areas, health services etc. A representative sample should be large enough to obtain statistically significant results, but not so large that analysis is impossible. Inferential statistics are designed to test a dependent variable namely, a parameter or outcome of the population being studied and may involve multiple variables. The calculations are more advanced, but the results are less certain. There is also a margin of error. Ultimately, inferential statistics are more educated guesses than statements. Sampling error can affect the results, although some statistical methods can be used to minimize problematic results. Inferential statistical methods include: Hypothesis tests or significance tests: These involve confirming certain results. are significant and not just by chance. Correlation analysis: this helps determine the relationship or correlation between variables Logistic or linear regression analysis: these methods can be used to infer and predict causality and relationships between other variables confidence intervals: these help to identify the probability of

the estimated outcome occurring, for example, inferential statistics can be used in the study of comorbidities. Instead of examining large health data as a whole, researchers can analyze a sample of patients with common characteristics for example, patients with more than two chronic conditions and extrapolate the results to the larger population from which the sample was drawn.

IMPORTANCE

In research, statistics are important for the following reasons:

1. Statistics allow researchers to design studies in such a way that research results can be extrapolated to a larger population. Statistics allow researchers to design studies. Scientists are often interested in answering questions about a population, such as:
   What is the average weight of a particular bird species?
   What is the average height of a certain plant species?
   What percentage of the population of a certain city supports a certain law?

2. Statistics allow scientists to perform hypothesis tests to determine whether a claim about a new drug, a new method, a new manufacturing method, etc., is true. Another way to use statistics in research is in the form of hypothesis tests. These are tests that allow researchers to determine whether different medical procedures or treatments have statistical significance.
   For example, suppose a researcher believes that a new drug can lower blood pressure in obese patients. To test this, he measures the blood pressure of 30 patients before and after a month of using the new drug. He then conducts a paired-sample t-test using the following hypotheses:

   a. H0: μ after = μ before (average blood pressure is the same before and after drug addiction).
   b. HA: μ after < μ before (average blood pressure is lower after drug).
   c. If p If the significance of the test is less than some level of significance (e.g., α = 0.05), he can reject the null hypothesis and conclude that the new drug causes a decrease in blood pressure. This is just one example of a hypothesis test used in research. Other common tests are one-sample t-test, two-sample t-test, one-way ANOVA, and two-way ANOVA.

3. Statistics allow researchers to construct confidence intervals to describe the uncertainty of a population estimate. Another way to use statistics in research is in the form of confidence intervals. A confidence interval is a range of values that are likely to contain a population parameter at a given confidence level. For example, suppose researchers are interested in estimating the average weight of a particular species of turtle. Instead of rounding up and weighing all the turtles in a population, researchers can instead take a simple random sample of turtles with the following information:

a. Sample size n = 25
b. Sample mean weight x = 300
c. Sample standard deviation s = 18.5

Using the confidence interval of the mean formula, researchers can construct the following 95% confidence interval:95% confidence interval: 300 +/- 1.96*(18.5/√25) = [292.75, 307 25] The researchers then state that they are 95% confident that the true average weight of this turtle population is between 292.75 and 307.25 pounds.

STATISTICAL TOOLS IN RESEARCH

For researchers in the field of education, commerce, science, biology etc. statistical analysis in research is the most intimidating aspect of completing research. However, research statistical tools can help researchers understand what to do with the data and how to interpret the results, making this process as easy as possible.

1. **Statistical Package for Social Sciences (SPSS):** It is a widely used software for studying human behavior. SPSS can produce descriptive statistics and graphical representations of results. In addition, it includes the ability to create scripts that automate analysis or perform advanced statistical processing.

2. **Foundation for Statistical Computing:** This software is used in the study of human behavior and other fields. R is a powerful tool. It has a steep learning curve. However, this requires a certain level of coding. It also comes with an active community dedicated to building and improving the software and related plugins.

3. **MATLAB (The Math Works):** It is an analysis platform and programming language. Scientists and engineers use this software and create their own code and help answer their research question. Although MATLAB can be a difficult tool for beginners, it offers the flexibility to meet the needs of a researcher.

4. **Microsoft Excel:** Not the best solution for statistical analysis in research, but MS Excel offers a wide range of tools for data visualization and simple statistics. Creating summaries and customizable charts and graphs is easy. MS Excel is the easiest choice for those who want to start with statistics.

5. **Statistical Analysis Software (SAS):** This is a statistical platform used in business, health and human behavior research. It can perform advanced analysis and produce publication-worthy figures, tables and graphs.

6. **GraphPad Prism:** This is an advanced software primarily used by biological scientists. But it offers a variety of different ones that can be used in many other areas. Like SPSS, GraphPad allows scripts to automate analyzes to perform complex statistical calculations.

7. **Minitab:** This software provides both basic and advanced statistical tools for data analysis. However, like GraphPad and SPSS, mini tab requires coding and can provide automatic analyses.

LIMITATIONS

1. **Qualitative aspect ignored:** Statistical methods do not study the nature of a phenomenon that cannot be quantitatively expressed. Such phenomena cannot be part of statistical studies. These include health, wealth, intelligence, etc. It needs to transform qualitative information into quantitative information. So, experiments are done to measure the man's reactions using the data. Today, statistics are used in all areas of life and universal activities.

2. **It does not deal with individual questions:** From the definition given by Prof. Horace Sacristy, it is clear: "By statistics we mean collections of facts and placed in relation to each other', this statistic deals only with collections of facts or parts and does not identify an individual lot. Thus, individual terms such as the death of 6 people in an accident, 85% of the results of a school lesson in a given year are not statistical because they are not placed in a group of similar units. It does not deal with individual items, although they may be important.

3. **It does not describe the whole history of a phenomenon:** Even if phenomena occur, it is for many reasons, but not all of these reasons can be expressed by data. So, we can't draw any real conclusions. The development of a group depends on many social factors such as the financial status of parents, education, culture, region, government, etc. But not all these factors can be accommodated in the data. Thus, we analyzed the data found only quantitatively, not qualitatively. So, the results or conclusions are not 100% correct because many aspects are omitted.

4. **It can be deformed:** As W.I. King points out, "One of the disadvantages of statistics is that they do not bear the stamp of quality on their face." So, we can say that we can review information and procedures as it approaches conclusions. However, this information may have been collected by inexperienced persons, or they may have been dishonest or biased. Because it is a delicate science and can be easily misused by a careless person. Therefore, the information should be used with caution. Otherwise, the results can be disastrous.

5. **Laws are not precise:** Regarding the two basic laws related to statistics:

(i) the law of inertia of large numbers.

(ii) the law of regularity statistics, are not as good as their laws of science. They are based on probability. So, these results are not always as good as scientific laws. We can only estimate the 2008 net production

based on probabilities or interpolation, but we cannot claim to be 100% accurate. Only estimates are made here.

6. **Results are only true on average:** As mentioned above, here the results are interpolated, for which time series or regression or likelihood can be used. They are not entirely true. In statistics, if two student sections have the same average, it does not mean that all 50 students in A received the same grades as in B. There are many differences between these two. So, we get average results. "Statistics is mostly about averages, and those averages can consist of individual points that are radically different from each other - "W. L. King."

7. **For several methods to investigate problems:** In this topic we use so many methods to find one result. Variations can be found based on interquartile range, standard deviation or standard deviation, and the results will be different in each case. According to Croxten and Cowden "Statistics should not be considered the only method used in research, nor should it be considered the best attack against a problem.

8. **Statistical results are not always irrefutable:** "Statistics only deals with things that can be measured, so it can rarely give a complete solution to a problem. It provides a basis for making a decision, but not the whole judgment - "Prof. L. R. Connor." Although we use many laws and formulas in statistics, the results obtained are not final and decisive. Since they cannot give a complete solution to the problem, the result must be taken and used very wisely.

CONCLUSION

The field of statistics is a data-based field of study. Statistics knowledge helps you use accurate methods to collect data, employ accurate analysis, and present results effectively. Statistics is an essential process in how we make scientific discoveries, make data-based decisions, and make predictions. Statistics allows you to understand a topic much more deeply. When analysts use statistical procedures correctly, they tend to produce accurate results. In fact, statistical analyzes report uncertainties and errors in the results. Statisticians ensure that all aspects of research follow proper methods to produce reliable results. These methods include:

a. Production of reliable data.
b. Analyze data appropriately.
c. Draw reasonable conclusions.

Statistical analysis is used in almost every field to make sense of the vast amounts of data available. Even if the field of statistics is not your primary field of study, it can help you make an impact in your preferred field. It is very likely that you will need a working knowledge of statistical methods to generate new answers in your field and to understand the work of others. In contrast, as a statistician, your skills are in great demand in many different fields: universities, research labs, government, industry and more.

REFERENCES

1. H. Wold(1948), "On prediction in stationary time series", Ann. Math. Statist., vol. 19, no. 4, pp. 558-567.

2. Shapiro SS, Wilk MB. (1965) An analysis of variance test for normality (complete samples). Biometrika; 52:591–611.

3. J.-F. Chen, W.-M. Wang, and C.-M. Huang, (1995) ‘‘Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting,’’ Electr. Power Syst. Res., vol. 34, no. 3, pp. 187–196.

4. Tabachnick. B.G. & S.L. Fidell. (1996). Using multivariate statistics. (3rd Edition). Harper Collins College Publishers. New York.

5. Unver. O. Gamgam. H. (1999) Applied Statistical Methods. Political Bookstore. Ankara.

6. Armstrong, R. A., Slade, S. V. and Eperjesi, F. (2000) An introduction to analysis of variance (ANOVA) with special reference to data from clinical experiments in optometry. Ophthal Physiol. Opt 20, 235–241.

7. Nickerson RS. (2000), Null hypothesis significance testing: A review of an old and continuing controversy. Psychol Methods.; 5:241–301.

8. Vaughan, L. (2001). Statistical methods for the information is professional: A practical, painless approach to understanding, using, and interpreting statistics (1st ed.).

9 McGraw−TaylorHarris PAGE: 1 SESS: 20 OUTPUT: Thu Oct 4 13:50:42 2007 SUM: 355503E1 /production/mcgraw−hill/booksxml/brownsaunders/ch1.10-K.K. Sharma, A. Kumar, A.Chaudhary,(2009), Statistics in management studies, Krishna Prakashan, Meerut, India, ISBN:81-87224-06-1.

11. Stohl, H., Rider, R., & Tarr, J. (2004). Making connections between empirical and theoretical probability: Students’ generation and analysis of data in a technological environment. Retrieved June 5, 2009, from http://www.probexplorer.com/Articles/LeeRider TarrConnectE&T.pdf

12. Bewick V, Cheek L, Ball J. (2004) Statistics review 10: Further nonparametric methods. Crit Care; 8:196-9.

13. Munro, B. H. (2005). Statistical methods for health care research. Lippincott Williams & Wilkins.

14. T M. Hastie, R. Tibshirani & J.Friedman (2008). The Elements of Statistical Learning. Springer Series in Statistics.

15. K.K. Sharma, A. Kumar, A.Chaudhary,(2009), Statistics in management studies, Krishna Prakashan, Meerut, India, ISBN:81-87224-06-1.

16. Altman DG, Bland JM.(2009), Parametric v non-parametric methods for data analysis. BMJ.;338: a3167.

17. Winters R, Winters A, Amedee RG.(2010) Statistics: A brief overview. Ochsner J.

18. Manikandan S.(2011), Measure of central tendency: Median, Mode, Journal of Pharmacology and Pharmacotherapeutics | July-September | Vol 2 | Issue 3

19. S. K. Næss(2012), Application of the Kolmogorov-Smirnov test to CMB data: Is the universe really weakly random?, Astronomy & Astrophysics, Volume 538

20. Gravetter, F., & Wallnau, L. (2013). Essentials of statistics for the behavioral sciences. Cengage Learning.

21. Bajwa SJ. Basics(2015), common errors, and essentials of statistical tools and techniques in anaesthesiology research. J Anaesthesiol Clin Pharmacol.; 31:547–53.