As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 7(11): 98-103, 2021

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0711017
Available online at: http://www.ijmtst.com/vol7issue11.html

Use of Statistics in Research

Dr. Seema Amit Agarwal

Assistant Professor, Dept. of Mathematics/ Statistics and Computer System, Ghanshyamdas Saraf College of Arts and Commerce, Mumbai, Maharashtra

To Cite this Article
Dr. Seema Amit Agarwal. Use of Statistics in Research. *International Journal for Modern Trends in Science and Technology* 2021, *7*, pp. 98-103. https://doi.org/10.46501/IJMTST0711017.

Article Info
Received: 21 October 2021; Accepted: 03 November 2021; Published: 09 November 2021

ABSTRACT

The function of statistics in research is to purpose as a tool in conniving research, analyzing its data and portrayal of conclusions there from. Most research studies result in a extensive quantity of raw data which must be properly concentrated so that the same can be examined easily and can be used for further investigation.

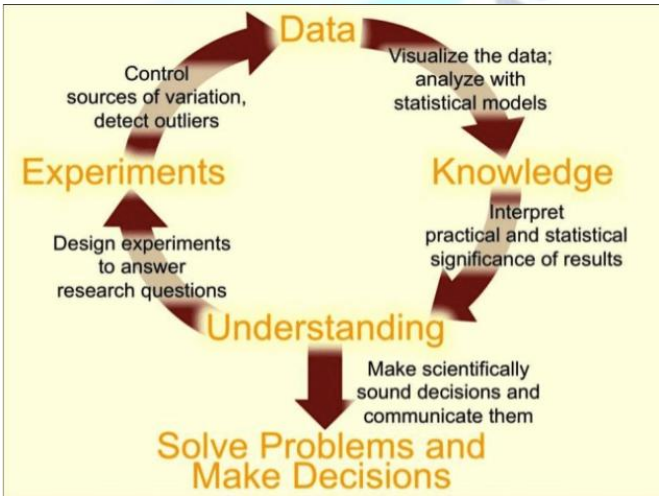
Undoubtedly the science of statistics cannot be overlooked by any research worker, even though he may not have juncture to use statistical methods in all their information and ramifications. Classification and tabulation, however, achieve this purpose to some extent, but we have to go a step auxiliary and develop certain indices or measures to summarize the collected/classified data. Only after this we can assume the process of generalization from small groups (i.e., samples) to populace. If fact, there are two major areas of statistics viz., descriptive statistics and inferential statistics. Descriptive statistics is based on the development of certain indices from the basic initial raw data, whereas inferential statistics concern with the process of generalization.

Keywords: statistics, use, research, investigation, data

INTRODUCTION

The important statistical actions that are used to summarize the survey/research data are:

1. measures of central tendency or statistical averages;
2. measures of dispersion;
3. measures of asymmetry (skewness);
4. measures of relationship; and
5. other measures.

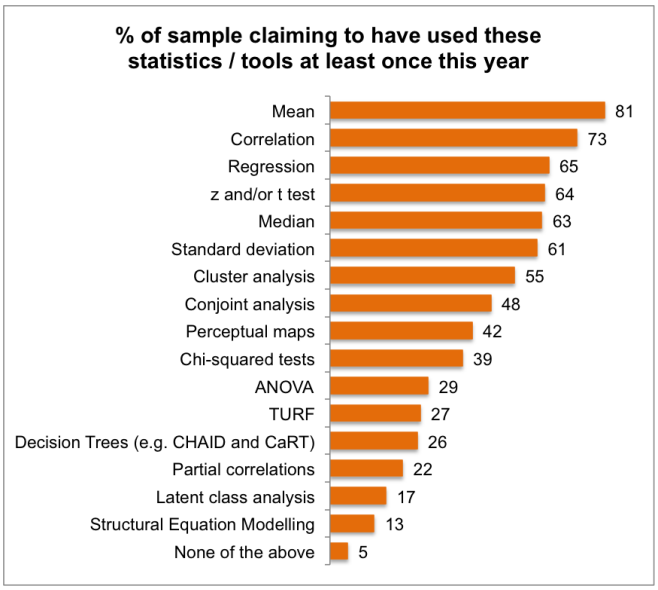


Importance of statistics

Amongst the measures of central tendency, the three most significant ones are the arithmetic average or mean, median and mode. Geometric mean and

harmonic mean are also frequently ustilised. [1] From among the measures of dispersion, variance, and its square root—the standard deviation are the usual used measures. Other measures such as mean deviation, range, etc. can also be used. For comparison purpose, we use a lot, the coefficient of standard deviation or the coefficient of variation.

In respect of the measures of skewness and kurtosis, we however use the first measure of skewness based on mean and mode or on mean and median. Other measures of skewness, based on quartiles or on the methods of moments, are also used rarely. Kurtosis is also used to measure the peakedness of the curve of the frequency distribution in graphical format.[2]

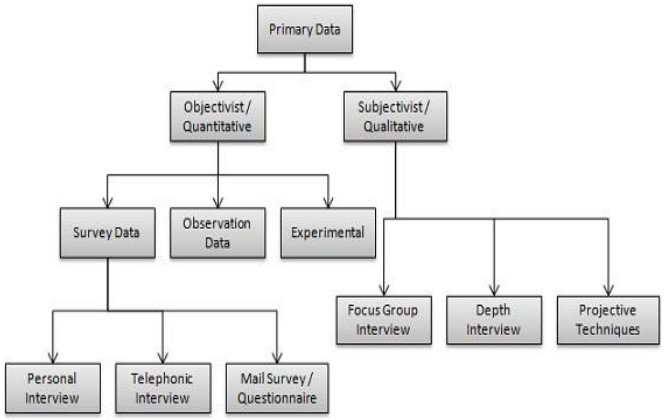


Various statistical tools

Amongst the measures of relationship, Karl Pearson's coefficient of correlation is the much more used measure in case of statistics of variables, whereas Yule's coefficient of association is used in the study of statistics of attributes. [3] Multiple correlation coefficient, partial correlation coefficient, regression analysis, etc., are other significant measures often used by a researcher. Index numbers, analysis of time series, coefficient of contingency, etc., are other measures that may as well be utilised by a researcher, depending upon the type of the problem under study.[4]

OBSERVATIONS

It is obvious that society can't be run smoothly on the basis of hunches or trial and error, and that in business and economics much depends on the right and accurate analysis of numerical information. Decisions based on data will provide better results than those based on intuition or gut feelings. What applies to this complete world applies to undertaking research into the wider world. And learning to use statistics in the investigation or survey will have a wider benefit than helping towards a qualification. [5]

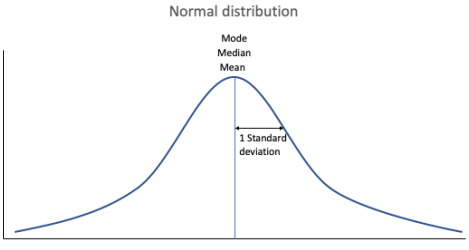


Data collection

Once mastered the language and some of the techniques in order to make sense of the complete research or investigation, supply of a knowledge and understanding will enable for coping with the information encountered in everyday life. Statistical thoughts permeate all social interaction.

Much of everyday life depends on preparing forecasts, and business can't progress without being able to audit revolution in research or plan action.[6]

In research, the areas such as purchasing, production, capital investment, long-term development, quality control, human resource development, recruitment and selection, marketing, credit risk assessment or financial forecasts or others are calculated. And that is why the informed use of statistics is of direct significance during collecting data and analysing them.

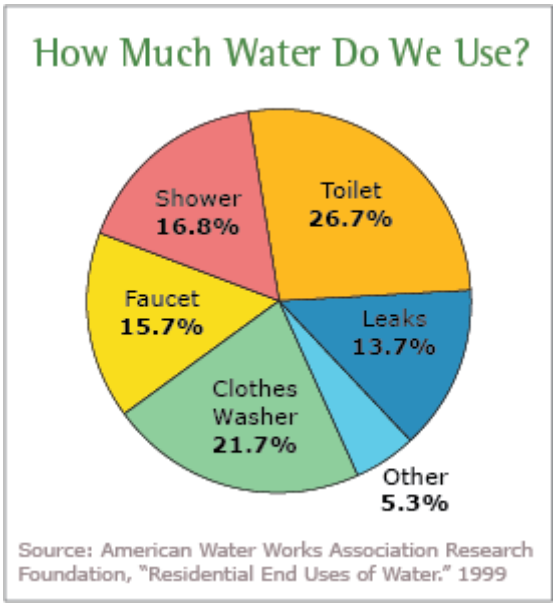


Beginners guide (mean, median and mode)

If nothing else, the research results and findings will be more accurate, more believable and, consequently, more useful. Some of the reasons to analyse the research data

are the same reasons of investigating or doing research. Ignoring the possibility of researching because the project or dissertation element of your qualification is compulsory, rather than because of the following:

measure things;
examine relationships;
making predictions;
test hypotheses;
construct concepts and develop theories;
exploring issues;
explain activities or attitudes;
describing what is happening;
presenting information;[7]
make comparisons to find similarities and differences;
drawing conclusions about populations based only on sample results. If you didn't want to do at least one of these things, there would be no point to doing your research at all.



Descriptive statistics

DISCUSSION

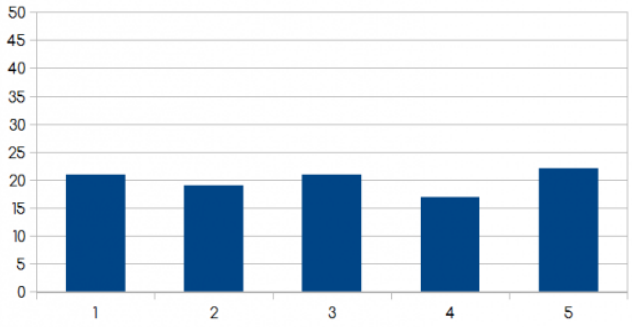
Parametric And Non-Parametric Tests

Numerical data (quantitative variables) that are usually distributed are analysed with parametric tests.

Two most basic requirements for parametric statistical analysis are:

- The assumption of normality which specifies that the means of the sample group have normal distribution[8]
- The assumption of equal variance which specifies that the variances of the samples and of their corresponding population have equal approach

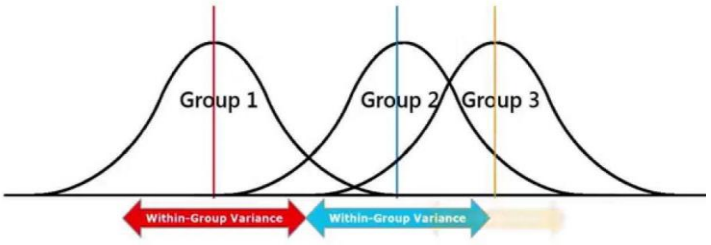
However, if the distribution of the sample is skewed towards one side or the distribution is unknown because of the small sample size, non-parametric, statistical methods are used. Non-parametric tests are utilised to analyze ordinal and categorical data.[9]



Beginners guide to statistics (bar diagrams)

Parametric tests

The parametric tests have an assumption that the data are on a quantitative (numerical) scale, with a normal distribution of the underlying population. The samples have similar variance (homogeneity of variances). The samples are randomly drawn from the population, and the observations within a group are not dependent on each other. The commonly used parametric tests are the Student's *t*-test, analysis of variance (ANOVA) and repeated measures ANOVA.[10]



ANOVA-lean manufacturing

Student's *t*-test

Student's *t*-test is used to determine the null hypothesis that there is no difference between the means of the two groups. It is used in three circumstances:

1. To test if a sample mean (as an estimate of a population mean) differs very much from a given population mean (this is a one-sample *t*-test)

The formula for one sample *t*-test is $t = \frac{X - u}{SE}$ [11]

where X = sample mean, u = population mean and SE = standard error of mean

2. To test if the population means estimated *by two independent samples* differ significantly (the unpaired *t*-test). The formula for unpaired *t*-test is:

$$t = \frac{X_1 - X_2}{SE_{x_1 - x_2}}$$

where $X_1 - X_2$ is the difference between the means of the two groups and SE is the standard error of the difference.[12]

3. To test if the population means estimated *by two dependent samples* differ significantly (the paired *t*-test). A usual setting for paired *t*-test is when measurements are done on the same subjects before and after a treatment.

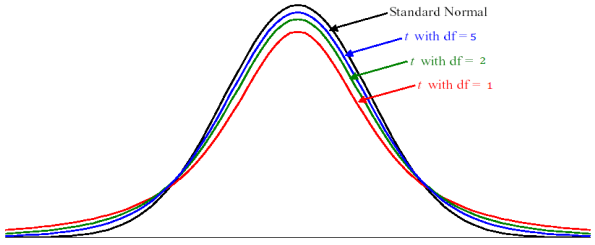
The formula for paired *t*-test is:

$$t = \frac{d}{SE_d}$$

where *d* is the mean difference and SE is the standard error of this difference.

The group variances can be compared using the *F*-test. The *F*-test is the ratio of variances (var l/var 2). If F differs significantly from 1.0, then it is final that the group variances differ significantly.[13]

Student's *t*-distribution



Analysis of variance

The Student's *t*-test cannot be used for comparing three or more groups. The purpose of ANOVA is to test if there is any significant difference between the means of two or more groups.

In ANOVA, we study two variances – (a) between-group variability and (b) within-group variability. The within-group variability (error variance) is the variation that cannot be mathematically resolved for in the study design. It is based on random differences present in our samples.

However, the between-group (or effect variance) is the result of our treatment. These two estimates of variances are comparatively determined using the F-test.[14]

A simplified formula for the *F* statistic is:

$$F = \frac{MS_b}{MS_w}$$

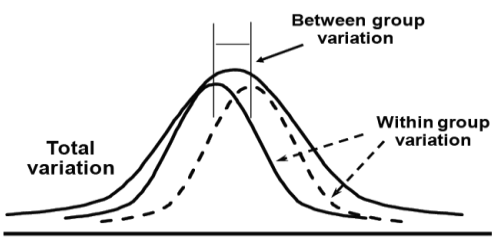
where MS_b is the mean squares among the groups and MS_w is the mean squares within groups.

Repeated measures analysis of variance

As with ANOVA, repeated measures ANOVA analyses the similarity of means of three or more groups. However, a repeated measure ANOVA is used when all variables of a sample are measured under variable conditions or at different points in time.

As the variables are measured from a sample at different points of time, the measurement of the dependent variable is repeated. Using a standard ANOVA in this case is not correct because it fails to model the correlation between the repeated measures: The data violate the ANOVA assumption of independence. Hence, in the measurement of repeated dependent variables, repeated measures ANOVA should be utilized.

ANOVA



Analysis of variance

1) Non-parametric tests

When the assumptions of normality are not undertaken, and the sample means are not normally, distributed parametric tests can lead to erroneous results. Non-parametric tests (distribution-free test) are used in such situation as they do not require the normality assumption. Non-parametric tests may not make the grade to detect a significant difference when compared with a parametric test. That is, they usually have less supremacy[15]

As is done for the parametric tests, the test statistic is compared with recognized values for the sampling distribution of that statistic and the null hypothesis is accepted or discarded.

RESULTS

Numerous statistical software systems are available at this time. The commonly used software systems are Statistical Package for the Social Sciences (SPSS – contrived by IBM corporation), Statistical Analysis System ((SAS – developed by SAS Institute North Carolina, United States of America), R (designed by Ross Ihaka and Robert Gentleman from R core team), Minitab (developed by Minitab Inc), Stata (developed by StataCorp) and the MS Excel (developed by Microsoft).[1]

There are a quantity of web resources which are related to statistical power analyses. A few are:

- StatPages.net – provides links to a number of online power calculators
- G-Power – provides a downloadable power analysis program that runs under DOS
- Power analysis for ANOVA designs an interactive site that calculates power or sample size needed to accomplish a given power for one effect in a factorial ANOVA design
- SPSS makes a program called Sample Power. It gives an output of a complete information on the computer screen which can be cut and paste into another document.[2]

It is important that a researcher knows the concepts of the basic statistical methods used for manner of a research study. This will help to conduct an appropriately well-designed study leading to convincing and reliable results. Inappropriate use of statistical techniques may lead to defective conclusions, inducing errors and dejecting the significance of the article. Bad statistics may lead to bad research, and bad research may lead to immoral practice. Hence, an adequate knowledge of statistics and the suitable use of statistical tests are important. An appropriate knowledge about the basic statistical methods will go a long way in improving the research designs and producing quality medical research which can be utilized for formulating the evidence-based strategies.[3]

CONCLUSION

The field of statistics is the discipline of learning from data. Statistical knowledge helps you use the appropriate methods to collect the data, employ the exact analyses, and effectively present the results. Statistics is a critical process behind how we make discoveries in science, make decisions based on data, and make predictions. Statistics allows you to comprehend a subject much more deeply.When analysts use statistical procedures correctly, they have a propensity to produce accurate results. In fact, statistical analyses report for uncertainty and error in the results. Statisticians ensure that all aspects of a study tag along the appropriate methods to produce trustworthy results. [15] These methods include:

- Production of reliable data.
- Analysis of the data appropriately.
- Drawing the reasonable conclusions.

Statistical analyses are used in almost all fields to make sense of the vast sum of data that are available. Even if the field of statistics is not your key field of study, it can help you make an impact in your preferred field. Chances are very high that you'll need working knowledge of statistical methodology both to produce new answers in your field and to understand the job of others.Conversely, as a statistician, there is a high demand for your skills in a wide diversity of areas: universities, research labs, government, industry, etc.[14]

REFERENCES

[1] Winters R, Winters A, Amedee RG. Statistics: A brief overview. *Ochsner J.* 2010;10:213–6. [PMC free article] [PubMed] [Google Scholar]

[2] Sprent P. Statistics in medical research. *Swiss Med Wkly.* 2003;133:522–9. [PubMed] [Google Scholar]

[3] Kaur SP. Variables in research. *Indian J Res Rep Med Sci.* 2013;4:36–8. [Google Scholar]

[4] Satake EB. *Statistical Methods and Reasoning for the Clinical Sciences Evidence-Based Practice.* Ist ed. San Diego: Plural Publishing, Inc; 2015. pp. 1–19. [Google Scholar]

[5] Wilder RT, Flick RP, Sprung J, Katusic SK, Barbaresi WJ, Mickelson C, et al. Early exposure to anesthesia and learning disabilities in a population-based birth cohort. *Anesthesiology.* 2009;110:796–804. [PMC free article] [PubMed] [Google Scholar]

[6] Manikandan S. Measures of central tendency: Median and mode. *J Pharmacol Pharmacother.* 2011;2:214–5. [PMC free article] [PubMed] [Google Scholar]

[7] Myles PS, Gin T. *Statistical Methods for Anaesthesia and Intensive Care.* Ist ed. Oxford: Butterworth Heinemann; 2000. pp. 8–10. [Google Scholar]

[8] Binu VS, Mayya SS, Dhar M. Some basic aspects of statistical methods and sample size determination in health science research. *Ayu.* 2014;35:119–23. [PMC free article] [PubMed] [Google Scholar]

[9] Nickerson RS. Null hypothesis significance testing: A review of an old and continuing controversy. *Psychol Methods.* 2000;5:241–301. [PubMed] [Google Scholar]

[10] Mushtaq K, Ali Z, Shah NF, Syed S, Naqash I, Ramzan AU. A randomized controlled study to compare the effectiveness of intravenous dexmedetomidine with placebo to attenuate the hemodynamic and neuroendocrine responses to fixation of skull pin head holder for craniotomy. *North J ISA.* 2016;1:16–23. [Google Scholar]

[11] Bajwa SJ. Basics, common errors and essentials of statistical tools and techniques in anesthesiology research. *J Anaesthesiol Clin Pharmacol.* 2015;31:547–53. [PMC free article] [PubMed] [Google Scholar]

[12] Das S, Mitra K, Mandal M. Sample size calculation: Basic principles. *Indian J Anaesth.* 2016;60:652–6. [PMC free article] [PubMed] [Google Scholar]

[13] Altman DG, Bland JM. Parametric v non-parametric methods for data analysis. *BMJ.* 2009;338:a3167. [PubMed] [Google Scholar]

[14] Bewick V, Cheek L, Ball J. Statistics review 10: Further nonparametric methods. *Crit Care.* 2004;8:196–9. [PMC free article] [PubMed] [Google Scholar]

[15] Nahm FS. Nonparametric statistical tests for the continuous data: The basic concept and the practical use. *Korean J Anesthesiol.* 2016;69:8–14. [PMC free article] [PubMed] [Google Scholar]