Research Methodology in Digital Age Using Statistical Software.

Name of Research scholar: - RAMAN MHATRE
Registration no: - 9201
Name of Research Guide : - Dr. S.L. Sananse
Professor, Department of statistics,
Dr. B.A. Marathwada University,
Aurangabad.

ABSTRACT

Research methodology in digital age using statistical softwares is the need of era. With the advent of powerful personal computers and user friendly softwares, it has become possible for researcher to work on computers sitting right on their desks. The opportunities for computer application in the entire field have increased tremendously. Statistical packages now provide the most popular way of handling statistical analyses. A statistical package has its own command language for specifying the form of data input, the manipulations required in the data and the analysis or analyses required. The digital revolution and the widespread use of the Internet throughout the globe are also responsible for the development of new methods and research tools. The geographic information systems (GPS) are a modern technique of data collection is available in most of the research institutes in the country. Choosing of the appropriate software for any researcher is critical. Software requirement must be assessed on the basis of the nature of applications, objectives of the study statistical, computational and ease of operation to common researcher. New and revised packages, for data analysis are being released continuously by research institutes and commercial organizations therefore researcher may select on the basis of individual requirement. Almost all the statistical techniques of research can be analyzed through the statistical softwares. SPSS, SAS, SYSSTAT, MATLAB, MINITAB and number of other useful softwares are available in the country. Some of them are free and others have to be purchased. The commonly used research methods are Descriptive Statistics, ANOVA, Test Of Significance, Multivariate analysis, Parametric and non parametric methods etc can easily be analyzed using statistical softwares. In the era of digital age statistical software is increasing the quality of data analysis

Key Words: Digital Age, Statistical Softwares, Research Methods.

1.0 INTRODUCTION

In this era computers the research methods are being supported by increasingly sophisticated Statistical Softwares, Internet technologies, also the new methods, techniques and digital tools. With the application of

statistical software,the application of research methodology has become easy , increased accuracy and saved much time of researcher. The traditional, routine or conventional research methods earlier used by the researcher have become outdated as more and more sophisticated digital technology is available for its analysis. With the innovations in the statistical software for analysis of quantitative and qualitative research data the complexity in the analysis has gone. Digital technologies are also available for research, in or via online environments, the data collection, classification, tabulation and data analysis methods are available to the researcher. Therefore, the simple and conventional research methodology has become out dated. It is interesting to note that the past two decades have witnessed an increasing sophistication in most of digital technologies, a shift responsible for new opportunities for online research in studies. Now the mostoftraditional research methods and tools have been transformed into digital versions and is available on the internet. The digital revolution and the widespread use of the Internet throughout the globe were also responsible for the development of new methods and research tools. The geographic information systems (GPS) is a modern technique of data collection is available in most of the research institutes in the country

1.1 Research Methods in Digital Age.

Research methodology in digital age is not essentially different from conventional research designs; however with the availability of highly sophisticated statistical softwares, the way it was analyzed was different. The multivariate analysis is possible with the help of statistical softwares. The research methods like estimation of measures of central tendency, measures of dispersion, test of significance, fitting probability distributions, ANOVA techniques and all types of research methodology for qualitative and quantitative data is available on the web or statistical softwares. This has improved the quality of analysis and helped to ease the calculations.

In this paper an attempt has been made to study research methods in digital age and various statistical softwares and their applications for the data classification, tabulation and analysis. Further, in view to study Research methodology in digital age using highly sophisticated statistical softwares the present study was undertaken with following objectives.

1.2 Objectives of the study

The following are some important objectives of the present study

- To study research methodology in digital age
- To study important statistical softwares available on the web.
- Analytical Facilities available in MS EXCEL
- To study the analytical facilities available in SPSS Softwares and their features for Research Data.
- To study the MINITAB Softwares and their features of analysis of Research Data.
- Analytical Facilities available in SAS.

1.3 Research Hypothesis.

- There are advances in the research methodology in the era of digital age.
- The statistical softwares and other available online softwares helped to improve the quality of analysis of research data and ultimately helped for making valid conclusions.

2.0 Research Methodology.

The present study was conducted to study research methods in digital age and various statistical softwares and their applications for the data classification, tabulation and analysis.

For this purpose the information were collected from INTERNET and also from the instruction manuals of the various softwares.

There are many statistical softwares are available in the market for analysis of statistical data. Some are ONLINE and some are available in the market. All are equally good and their aim is same. However we have considered here the statistical software

i) MS-EXCEL

ii) Statistical Packages for Social Sciences (SPSS)

iii) MINITAB

iv) Statistical Analysis Software (SAS).

3.0 Results and Discussion

The statistical methods help the researcher to classify, tabulate, analyse data and draw conclusions. The statistical softwares are user friendly and in this digital era they help the researcher. Any scientific research and investigation often contain much more data or information than the researcher needs. This data or information is mostly in the form of raw data. Depending on the research, the researcher use statistical methods suitable to his research. Descriptively or for exploratory research.

The statistical softwares commonly used in data analysis along with their features are described below. In this digital era the computations of complex statistical methods is not possible.

3.1 Statistical Softwares in Digital Age

The computer has immense power of information and data processing and dissemination, storage and retrieval. Now days, through computer networking with use of modem, computer has become extremely useful for communication. The tremendous speed of computer is another advantage. Initially mainframe computers were used in data analysis which required specially trained manpower and one has to depend on such trained personnel for the processing of data or information. During the last ten years computer technology has

undergone a revolution. With the advent of powerful personal computers and user friendly softwares, it has become possible for everyone to work on computers sitting right on their desks. Thus, the opportunities for computer application in the entire field have increased tremendously.

Statistical packages now provide the most popular way of handling statistical analyses. A package is distinguished from a library of subroutines because it is complete in itself. It has its own command language for specifying the form of data input, the manipulations required in the data and the analysis or analyses required.

Packages vary widely both in quality and in what they will do, and the choice between them may not be easy. Some have been written by expert and others have not. Some print out plenty of warning, error and help messages, but others do not and may indeed go 'happily' producing meaningless results. Some packages are written for one small area of methodology while others are more general. Some are written for expert users, while others are intended for statistical novices. Some packages allow completely automatic analysis of data, where the analyst abdicates all responsibility to the computer, but interactive analyses are usually preferable.

Software needs to be appraised on various criteria which include statistical, computational and commercial considerations. Keeping these considerations, new and revised packages are being released continuously and it is very difficult to provide an up-to-date review of them all here. In this regard it is necessary that one should keep up-to-date information by referring to reviews of computer software in various journals.

The recent advances in computer and communication technology have made computer hardware and software more user friendly, affordable and have resulted in faster movement of information and its utilization. Several renowned and multinational companies and educational Institutes are engaged in development of statistical software packages. Software needs to be appraised on various criteria, which include statistical, computational and commercial considerations. Keeping these considerations, new and revised packages, for data analysis is being released continuously and it is difficult to provide and up to date review of them all here. However, some of the commonly used packages are described for the researcher.

3.2 Software Packages Used for Data Analysis

Choosing of the appropriate software for any researcher is critical. Software requirement must be assessed on the basis of the nature of applications. Some of the important packages which are in common use are given below. The information has been collected from the Internet/ companies manual which are described below.

3.2.1MINITAB :

It is an interactive, command- driven package which covers such topics as exploratory data analysis, Significance tests, regression and time-series analysis. It is very easy to use and is widely employed by commercial, academic and teaching institutions. The functionality of MINITAB is accessible through interactive windows and menus, or through a command language called session commands. There are three windows viz. Data window, Session window and Project Manager.Datawindow is a worksheet in spread sheet format, with rows and columns that intersect to form individual cells. A worksheet can contain up to 4000

columns, 1000 constants, and up to 10,000,000 rows depending on memory of the computer. The text output generated by the analyses is displayed in Session window. The Project Manager contains folders that allow one to navigate, view, and manipulate various parts of the project. Minitab has the advanced Design of Experiments (DOE) capabilities. One can screen the factors to determine which are important for explaining process variation. It also allows one to perform one-way analysis of variance, two-way analysis of variance for balanced data, test for equality of variances, and generate various plots. Balanced ANOVA models with crossed or nested and fixed or random factors can also be analyzed. The option General MANOVA analyzes balanced or unbalanced MANOVA models with crossed or nested and fixed or random factors. The analysis of covariance is also possible with option General MANOVA. The more advanced versions are also available now days.

3.2.2 SPSS (Statistical Package for the Social Sciences) :.

SPSS stands for "Statistical Package for the Social Sciences," this software is now widely used for data analysis and report generation across disciplines. SPSS for Windows or Mac OS features a spreadsheet- like data editor, integrated graphics, GUI interface, data management facilities (the number of variables and cases it can handle is limited only by the amount of RAM and disk space available), and comprehensive statistical procedures.

3.2.3 SAS (Statistical Analysis System) :

This programming is widely used in almost all the countries in the world. It is suitable for the more advanced statistician. It is composed of basically three sets of statements which can create data sets, run statistical analysis and print data. The first type of statement is called DATA statements while the other types are called PROC's (short for procedures). The interesting feature of SAS is that it has its own IML (Interactive Matrix Language) which includes many functions and subroutines. From traditional statistical analysis of variance and predictive modeling to exact methods and statistical visualization techniques, SAS/STAT software is designed for both specialized and enterprise wide analytical needs. SAS/STAT software provides a complete, comprehensive set of tools that can meet the data analysis needs of the entire organization.

3.2.4Other Packages :

There are numerous other general and more specialized packages are available and all of them cannot be covered here. To name a few are R- SOFTWARE, SYS STAT, MSTATC, STATGRAPHICS, SX, MICROSTAT, STRATA, PC-CARP, EXACT, STATLAB, SIGMAPLOT and FREELANCE are some useful specialist graphics packages. There are number of packages which are mathematically oriented .they include MATLAB, MAPLE and MATHEMATICA. As well as some data-analysis procedures, they enable the user to work with a variety of mathematical functions, carry out numerical analysis and matrix algebra, plot 2-D and 3-D functions and graphs, etc.

One class of packages not mentioned so far are spreadsheet packages. They allow easy manipulation of data in the form of two-or three-way table of numbers and also allow some simple statistical analysis to be carried out on them. They are widely used, and the best known packages are LOTUS 1-2-3, EXCEL, FOXRPO, REFLEX and QUATTRO PRO.

There are many more packages which are available and developed by the softwares companies; one can choose the package as per the suitability and requirement.

3.2.5 Statistical Methods Generally Needed to Researcher:

The most commonly used statistical methods which can be generally needed by the researcher. The statistical softwares described here provide analysis of these techniques. Which is also available in the manual of the software development companies. Some of the commonly used statistical methods are described here.

Descriptive Statistics: Measures of Central Tendency, dispersion and their Relative measures.

Analysis of variance (ANOVA) : Balanced and unbalanced designs; multivariate analysis of variance and repeated measurements; linear and nonlinear mixed models. Mixed models: Linear mixed models, Nonlinear mixed models., Generalize linear mixed models.

Correlation and Regression : Correlation, Least squares regression with nine model selection techniques, including stepwise regression, Diagnostic measures, Robust regression; Loess regression, Nonlinear regression and quadratic response surface models, Partial least squares.

Categorical data analysis: Contingency tables and measures of association, Logistic regression and log linear models; generalized linear models, Bioassay analysis, generalized estimating equations. Weighted least squares regression. Exact methods.

Bayesian analysis: Bayesian modeling and inference for generalized linear models, accelerated life failure models, Cox regression models and piecewise exponential models, General procedure fits Bayesian models with arbitrary priors and likelihood functions.

Multivariate analysis: Factor analysis; principal components; canonical correlation and discriminate analysis; path analysis; structural equations.

Survival analysis: Comparison of survival distributions; accelerated failure time models; proportional hazards models.

Psychometric analysis: Multidimensional scaling; conjoint analysis with variable transformations; correspondence analysis.

Cluster Analysis: Hierarchical clustering of multivariate data or distance data; disjoint clustering of large data sets; nonparametric clustering with hypothesis tests for the number of clusters.

Nonparametric analysis : Nonparametric analysis of variance. Exact probabilities computed for many nonparametric statistics. Kruskal-Wallis, Wilcoxon-Mann-Whitney and Friedman tests.Other rank tests for balanced or unbalanced one-way or two-way designs.

Survey data analysis: Sample selection; descriptive statistics and *t*-tests; linear and logistic regression; frequency table analysis

Multiple imputation for missing values: Regression and propensity scoring for monotone missing patterns. MCMC method for arbitrary missing patterns. Combine results for statistically valid inferences.

Study planning: Power and Sample Size application provides interface for computation of sample sizes and characterization of power for *t*-tests, confidence intervals, linear models, tests of proportions and rank tests for survival analysis.

4.0 Conclusions

- With the advent of powerful personal computers and user friendly softwares, it has become possible for every researcher to use the softwares for their research data.
- Choosing of the appropriate hardware and software for any individual /office/organization is critical. Software requirement must be assessed on the basis of the nature of applications. The researcher can select any one of them for their analysis of data.
- Statistical packages now provide the most popular way of handling statistical analyses. A package is distinguished from a library of subroutines because it is complete in itself. It has its own command language for specifying the form of data input, the manipulations required in the data and the analysis or analyses required.
- In the era of software, there are Several renowned and multinational companies and educational Institutes which are engaged in development of statistical software packages.
- Software needs to be appraised on various criteria, which include statistical, computational and commercial considerations. Keeping these considerations, new and revised packages, for data analysis are being released continuously and it is difficult to provide and up to date review of them all here.
- The statistical software like SPSS, SAS, MINITAB, SYSSTAT, MATLAB etc. are commonly used by the researcher. There are many more packages which are also useful of data analysis.
- There are numerous other general and more specialized packages are available to name a few are R-SOFTWARE ,MSTATC, STATGRAPHICS, SX, MICROSTAT, STRATA, PC-CARP, EXACT, STATLAB, SIGMAPLOT and FREELANCE are some useful specialist graphics packages. There are number of packages which are mathematically oriented they include MATLAB, MAPLE and MATHEMATICA. As well as some data-analysis procedures, they enable the user to work with a variety of mathematical functions, carry out numerical analysis and matrix algebra, plot 2-D and 3-D functions and graphs, etc.
- One class of packages not mentioned so far are spreadsheet packages. They allow easy manipulation of data in the form of two-or three-way table of numbers and also allow some simple statistical analysis to be carried out on them. They are widely used, and the best known packages are LOTUS 1-2-3, EXCEL, FOXRPO, REFLEX and QUATTRO PRO.

- There are many more packages which are available and developed by the softwares companies; one can choose the package as per the suitability and requirement.
- In the digital age the analytical technology is being developed by the scientists. In view of this the use of statistical softwares cannot be ignored.

5.0 References:

Arvind Kumar (2002) : Research Methodology in Social Sciences, Sarup and Sons, New Delhi.

Bhatnagar, O.P. : Research Methods and Measurement in Behavioral Science.Agricole Pub Acadamy. New Delhi.

George W. Snedecor and W.G.Cochran : Statistical Methods, Sixth Edition, Oxford and IBM Publishing Co.

Journal of Statistical Software, http://www.jstatsoft.org/

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL

http ://www.R-project.org/

Sananse S. L ((2004): Scaling Techniques in Agril. And Social Science.Published by Yashwantrao Chavan Mah. Open University, Nashik.

Statistics.com list of commercial software
http://www.statistics.com/resources/software/commercial/fulllist.php3