International Research Journal of Marketing & Economics
ISSN (2348-9766) Impact Factor: 5.779
Vol. 7, Issue 6, June 2020

Website-www.aarf.asia,Email id : editoraarf@gmail.com

# ROLE AND IMPORTANCE OF STATISTICS IN PSYCHOLOGICAL RESEARCH

**AUTHOR**

**Ms. PRIYANKA CHAUDHARY**

HOD & Assistant Professor, Department of Psychology,
Stani Memorial P.G. College, Mansarovar, Jaipur, Rajasthan

## ABSTRACT

Statistics is a wide subject usefull in almost all disciplines especially including Psychology in Research studies. Each and every researcher should have some knowledge in Statistics and must use statistical tools in his or her research, one should know about the importance of statistical tools and how to use them in their research or survey. The quality assurance of the work must be dealt with: the statistical operations necessary to control and verify the analytical procedures as well as the resulting data making mistakes in analytical work is unavoidable. This is the reason why a multitude of different statistical tools is, required some of them simple, some complicated, and often very specific for certain purposes. In analytical work, the most important common operation is the comparison of data, or sets of data, to quantify accuracy (bias) and precision. Fortunately, with a few simple convenient statistical tools most of the information needed in regular psychological analysis work can be obtained: the "t-test, the "F-test", and regression analysis. Clearly, statistics are a tool, not an aim. Simple inspection of data, without statistical treatment, by an experienced and dedicated psychological analyst may be just as useful as statistical figures on the desk of the disinterested. The value of statistics lies with organizing and simplifying data, to permit some objective estimate showing that an analysis is under control or that a change has occurred. Equally important is that the results of these statistical procedures are recorded and can be retrieved. The key is to sift through the overwhelming volume of data available to organizations and businesses and correctly interpret its implications. But to sort through all this information, you need the right statistical data analysis tools. Hence in this paper, author have made an attempt to give a brief report or study on role and importance of Statistical tools used in psychological research studies.

## INTRODUCTON

The subject Statistics is widely used in almost all fields but predominatey in this research paper psychological importance of statstcis is been highlighted. .. While doing research in the psychology fields, the researchers should have some awareness in using the statistical tools which helps them in drawing rigorous and good conclusions. The most well known Statistical tools are the mean, the arithmetical average of numbers, median and mode, Range, dispersion , standard deviation, inter quartile range, coefficient of variation, etc. There are also software packages like SAS and SPSS which are useful in interpreting the results for large sample size.

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 is 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. You can also use software packages like SPSS, EPInfo, STATA, Minitab, Open Epi, Graph pad and many others depending on your usage and familiarity with the software. You should also start looking at the distributions of age, gender, race and any measures of socio-economic status that you have ( income, education level, access to medical care). These distributions will help to inform your analysis in terms of possible age- adjustment, weighting and another analytical tools available to address issues of bias and non representative samples.
Survey analysis is one of the most commonly used research methods, scholars, market researchers and organization of all sizes use surveys to measure public opinion. Researchers use a wide range of statistical methods to analyze survey data. They do this using statistical software packages that are designed for research professionals. Popular programs include SAS, SPSS and STATA. However, many forms of survey data analysis can be done with a spread sheet program such as EXCEL, which is part of Microsoft‗s popular office package. EXCEL and other spreadsheet programs are user-friendly and excellent for entering, coding and storing survey data.

# RESEARCH METHOD

There are many forms of empirical studies in psychology, including case reports, controlled experiments, quasiexperiments, statistical simulations, surveys, observational studies, and studies of studies (meta-analyses). Some are hypothesis generating: They explore data to form or sharpen hypotheses about a population for assessing future hypotheses. Some are hypothesis testing: They assess specific a priori hypotheses or estimate parameters by random sampling from that population. Some are meta-analytic: They assess specific a priori hypotheses or estimate parameters (or both) by synthesizing the results of available studies.

Some researchers have the impression or have been taught to believe that some of these forms yield information that is more valuable or credible than others (see Cronbach, 1975, for a discussion). Occasionally proponents of some research methods disparage others. In fact, each form of research has its own strengths, weaknesses, and standards of practice.

***Population***

*The interpretation of the results of any study depends on the characteristics of the population intended for analysis. Define the population (participants, stimuli, or studies) clearly. If control or comparison groups are part of the design, present how they are defined.*

Psychology students sometimes think that a statistical population is the human race or, at least, college sophomores. They also have some difficulty distinguishing a class of objects versus a statistical population-that sometimes we make inferences about a population through statistical methods, and other times we make inferences about a class through logical or other nonstatistical methods. Populations may be sets of potential observations on people, adjectives, or even research articles. How a population is defined in an article affects almost every conclusion in that article.

***Sample***

*Describe the sampling procedures and emphasize any inclusion or exclusion criteria. If the sample is stratified (e.g., by site or gender) describe fully the method and rationale. Note the proposed sample size for each subgroup.*

Interval estimates for clustered and stratified random samples differ from those for simple random samples. Statistical software is now becoming available for these purposes. If you are using a convenience sample (whose members are not selected at random), be sure to make that procedure clear to your readers. Using a convenience sample does not automatically disqualify a study from publication, but it harms your objectivity to try to conceal this by implying that you used a random sample. Sometimes the case for the representativeness of a convenience sample can be strengthened by explicit comparison of sample characteristics with those of a defined population across a wide range of variables.

***Assignment***

***Random assignment.*** *For research involving causal inferences, the assignment of units to levels of the causal variable is critical. Random assignment (not to be confused with random selection) allows for the strongest possible causal inferences free of extraneous assumptions. If random assignment is planned, provide enough information to show that the process for making the actual assignments is random.*

There is a strong research tradition and many exemplars for random assignment in various fields of psychology. Even those who have elucidated quasi-experimental designs in psychological research (e.g., Cook & Campbell, 1979) have repeatedly emphasized the superiority of random assignment as a method for controlling bias and lurking variables. "Random" does not mean "haphazard." Randomization is a fragile condition, easily corrupted deliberately, as we see when a skilled magician flips a fair coin repeatedly to heads, or innocently, as we saw when the drum was not turned sufficiently to randomize the picks in the Vietnam draft lottery. As psychologists, we also know that human participants are incapable of producing a random process (digits, spatial arrangements, etc.) or of recognizing one. It is best not to trust the random behavior of a physical device unless you are an expert in these matters. It is safer to use the pseudorandom sequence from a welldesigned computer generator or from published tables of random numbers. The added benefit of such a procedure is that you can supply a random number seed or starting number in a table that other researchers can use to check your methods later.

***Nonrandom assignment.*** *For some research questions, random assignment is not feasible. In such cases, we need to minimize effects of variables that affect the observed relationship between a causal variable and an outcome. Such variables are commonly called confounds or covariates. The researcher needs to attempt to determine the relevant covariates, measure them adequately, and adjust for their effects either by design or by analysis. If the effects of covariates are adjusted by analysis, the strong assumptions that are made must be explicitly stated and, to the extent possible, tested and justified. Describe methods used to attenuate sources of bias, including plans for minimizing dropouts, noncompliance, and missing data.*

Authors have used the term "control group" to describe, among other things, (a) a comparison group, (b) members of pairs matched or blocked on one or more nuisance variables, (c) a group not receiving a particular treatment, (d) a statistical sample whose values are adjusted post hoc by the use of one or more covariates, or (e) a group for which the experimenter acknowledges bias exists and perhaps hopes that this admission will allow the reader to make appropriate discounts or other mental adjustments. None of these is an instance of a fully adequate control group.

If we can neither implement randomization nor approach total control of variables that modify effects (outcomes), then we should use the term "control group" cautiously. In most of these cases, it would be better to forgo the term and use "contrast group" instead. In any case, we should describe exactly which confounding variables have been explicitly controlled and speculate about which unmeasured ones could lead to incorrect inferences. In the absence of randomization, we should do our best to investigate sensitivity to various untestable assumptions.

***Measurement***

***Variables.*** *Explicitly define the variables in the study, show how they are related to the goals of the study, and explain how they are measured. The units of measurement of all* *variables, causal and outcome, should fit the language you use in the introduction and discussion sections of your report.*

A variable is a method for assigning to a set of observations a value from a set of possible outcomes. For example, a variable called "gender" might assign each of 50 observations to one of the values male or female. When we define a variable, we are declaring what we are prepared to represent as a valid observation and what we must consider as invalid. If we define the range of a particular variable (the set of possible outcomes) to be from 1 to 7 on a Likert scale, for example, then a value of 9 is not an outlier (an unusually extreme value). It is an illegal value. If we declare the range of a variable to be positive real numbers and the domain to be observations of reaction time (in milliseconds) to an administration of electric shock, then a value of 3,000 is not illegal; it is an outlier.

Naming a variable is almost as important as measuring it. We do well to select a name that reflects how a variable is measured. On this basis, the name "IQ test score" is preferable to "intelligence" and "retrospective self-report of childhood sexual abuse" is preferable to "childhood sexual abuse." Without such precision, ambiguity in defining variables can give a theory an unfortunate resistance to empirical falsification. Being precise does not make us operationalists. It simply means that we try to avoid excessive generalization.

Editors and reviewers should be suspicious when they notice authors changing definitions or names of variables, failing to make clear what would be contrary evidence, or using measures with no history and thus no known properties. Researchers should be suspicious when code books and scoring systems are inscrutable or more voluminous than the research articles on which they are based. Everyone should worry when a system offers to code a specific observation in two or more ways for the same variable.

***Instruments.*** *If a questionnaire is used to collect data, summarize the psychometric properties of its scores with specific regard to the way the instrument is used in a population. Psychometric properties include measures of validity, reliability, and any other qualities affecting conclusions. If a physical apparatus is used, provide enough information (brand, model, design specifications) to allow another experimenter to replicate your measurement process.*

There are many methods for constructing instruments and psychometrically validating scores from such measures. Traditional true-score theory and item-response test theory provide appropriate frameworks for assessing reliability and internal validity. Signal detection theory and various coefficients of association can be used to assess external validity. Messick (1989) provides a comprehensive guide to validity.

It is important to remember that a test is not reliable or unreliable. Reliability is a property of the scores on a test for a particular population of examinees (Feldt & Brennan, 1989). Thus, authors should provide reliability coefficients of the scores for the data being analyzed even when the focus of their research is not psychometric. Interpreting the size of observed effects requires an assessment of the reliability of the scores.

Besides showing that an instrument is reliable, we need to show that it does not correlate strongly with other key constructs. It is just as important to establish that a measure does *not* measure what it should not measure as it is to show that it *does* measure what it should.

Researchers occasionally encounter a measurement problem that has no obvious solution. This happens when they decide to explore a new and rapidly growing research area that is based on a previous researcher's well-defined construct implemented with a poorly developed psychometric instrument. Innovators, in the excitement of their discovery, sometimes give insufficient attention to the quality of their instruments. Once a defective measure enters the literature, subsequent researchers are reluctant to change it. In these cases, editors and reviewers should pay special attention to the psychometric properties of the instruments used, and they might want to encourage revisions (even if not by the scale's author) to prevent the accumulation of results based on relatively invalid or unreliable measures.

***Procedure.*** *Describe any anticipated sources of attrition due to noncompliance, dropout, death, or other factors. Indicate how such attrition may affect the generalizability of the results. Clearly describe the conditions under which measurements are taken (e.g., format, time, place, personnel who collected data). Describe the specific methods used to deal with experimenter bias, especially if you collected the data yourself*

Despite the long-established findings of the effects of experimenter bias (Rosenthal, 1966), many published studies appear to ignore or discount these problems. For example, some authors or their assistants with knowledge of hypotheses or study goals screen participants (through personal interviews or telephone conversations) for inclusion in their studies. Some authors administer questionnaires. Some authors give instructions to participants. Some authors perform experimental manipulations. Some tally or code responses. Some rate videotapes.

An author's self-awareness, experience, or resolve does not eliminate experimenter bias. In short, there are no valid excuses, financial or otherwise, for avoiding an opportunity to double-blind. Researchers looking for guidance on this matter should consult the classic book of Webb, Campbell, Schwartz, and Sechrest (1966) and an exemplary dissertation (performed on a modest budget) by Baker (1969).

***Power and sample size.*** *Provide information on sample size and the process that led to sample size decisions. Document the effect sizes, sampling and measurement assumptions, as well as analytic procedures used in power calculations. Because power computations are most meaningful when done before data are collected and examined, it is important to show how effect-size estimates have been derived from previous research and theory in order to dispel suspicions that they might have been taken from data used in the study or, even worse, constructed to justify a particulars ample size. Once the study is analyzed, confidence intervals replace calculated power in describing results.*

Largely because of the work of Cohen (1969, 1988), psychologists have become aware of the need to consider power in the design of their studies, before they collect data. The intellectual exercise required to do this stimulates authors to take seriously prior research and theory in their field, and it gives an opportunity, with incumbent risk, for a few to offer the challenge that there is no applicable research behind a given study. If exploration were not disguised in hypothetico-deductive language, then it might have the opportunity to influence subsequent research constructively.

Computer programs that calculate power for various designs and distributions are now available. One can use them to conduct power analyses for a range of reasonable alpha values and effect sizes. Doing so reveals how power changes across this range and overcomes a tendency to regard a single power estimate as being absolutely definitive.

Many of us encounter power issues when applying for grants. Even when not asking for money, think about power. Statistical power does not corrupt.

***Complications***

*Before presenting results, report complications, protocol violations, and other unanticipated events in data collection. These include missing data, attrition, and nonresponse. Discuss analytic techniques devised to ameliorate these problems. Describe nonrepresentativeness statistically by reporting patterns and distributions of missing data and contaminations. Document how the actual analysis differs from the analysis planned before complications arose. The use of techniques to ensure that the reported results are not produced by anomalies in the data (e.g., outliers, points of high influence, nonrandom missing data, selection bias, attrition problems) should be a standard component of all analyses.*

As soon as you have collected your data, before you compute *any* statistics, *look at your data.* Data screening is not data snooping. It is not an opportunity to discard data or change values to favor your hypotheses. However, if you assess hypotheses without examining your data, you risk publishing nonsense.

Computer malfunctions tend to be catastrophic: A system crashes; a file fails to import; data are lost. Less well-known are more subtle bugs that can be more catastrophic in the long run. For example, a single value in a file may be corrupted in reading or writing (often in the first or last record). This circumstance usually produces a major value error, the kind of singleton that can make large correlations change sign and small correlations become large.

Graphical inspection of data offers an excellent possibility for detecting serious compromises to data integrity. The reason is simple: Graphics broadcast; statistics narrowcast. Indeed, some international corporations that must defend themselves against rapidly evolving fraudulent schemes use real-time graphic displays as their first line of defense and statistical analyses as a distant second. The following example shows why.
stacked like a histogram) and scales used for each variable. The three variables shown are questionnaire measures of respondent's age *(AGE)*, gender *(SEX)*, and number of years together in current relationship *(TOGETHER)*. The graphic in Figure 1 is not intended for final presentation of results; we use it instead to locate coding errors and other anomalies before we analyze our data. Figure 1 is a selected portion of a computer screen display that offers tools for zooming in and out, examining points, and linking to information in other graphical displays and data editors. SPLOM displays can be used to recognize unusual patterns in 20 or more variables simultaneously. We focus on these three only.

Modem statistical packages offer graphical diagnostics for helping to determine whether a model appears to fit data appropriately. Most users are familiar with residual plots for linear regression modeling. Fewer are aware that John Tukey's paradigmatic equation, *data = fit + residual*, applies to a more general class of models and has broad implications for graphical analysis of assumptions. Stemand- leaf plots, box plots, histograms, dot plots, spread/level plots, probability plots, spectral plots, autocorrelation and cross-correlation plots, co-plots, and trellises (Chambers, Cleveland, Kleiner, & Tukey, 1983; Cleveland, 1995; Tukey, 1977) all serve at various times for displaying residuals, whether they arise from analysis of variance (ANOVA), nonlinear modeling, factor analysis, latent vari- able modeling, multidimensional scaling, hierarchical linear modeling, or other procedures.

***Hypothesis tests.*** *It is hard to imagine a situation in which a dichotomous accept-reject decision is better than reporting an actual p value or, better still, a confidence interval. Never use the unfortunate expression "accept the null hypothesis. " Always provide some effectsize estimate when reporting a p value.* Cohen (1994) has written on this subject in this journal. All psychologists would benefit from reading his insightful article.

***Effect sizes.*** *Always present effect sizes for primary outcomes. If the units of measurement are meaningful on a practical level (e.g., number of cigarettes smoked per day), then we usually prefer an unstandardized measure (regression coefficient or mean difference) to a standardized measure (r or d). It helps to add brief comments that place these effect sizes in a practical and theoretical context.*

APA's (1994) publication manual included an important new "encouragement" (p. 18) to report effect sizes. Unfortunately, empirical studies of various journals indicate that the effect size of this encouragement has been negligible (Keselman et al., 1998; Kirk, 1996; Thompson & Snyder, 1998). We must stress again that reporting and interpreting effect sizes in the context of previously reported effects is essential to good research. It enables readers to evaluate the stability of results across samples, designs, and analyses. Reporting effect sizes also informs power analyses and meta-analyses needed in future research.

Fleiss (1994), Kirk (1996), Rosenthal (1994), and Snyder and Lawson (1993) have summarized various measures of effect sizes used in psychological research. Consult these articles for information on computing them. For a simple, general purpose display of the practical meaning of an effect size, see Rosenthal and Rubin (1982). Consult Rosenthal and Rubin (1994) for information on the use of "counternull intervals" for effect sizes, as alternatives to confidence intervals.

***Interval estimates.*** *Interval estimates should be given for any effect sizes involving principal outcomes. Provide intervals for correlations and other coefficients of association or variation whenever possible.*

Confidence intervals are usually available in statistical software; otherwise, confidence intervals for basic statistics can be computed from typical output. Comparing confidence intervals from a current study to intervals from previous, related studies helps focus attention on stability across studies (Schmidt, 1996). Collecting intervals across studies also helps in constructing plausible regions for population parameters. This practice should help prevent the common mistake of assuming a parameter is contained in a confidence interval.

***Multiplicities***. *Multiple outcomes require special handling. There are many ways to conduct reasonable inference when faced with multiplicity (e.g., Bonferroni correction ofp values, multivariate test statistics, empirical Bayes methods). It is your responsibility to define and justify the methods used.*

Statisticians speak of the curse of dimensionality. To paraphrase, multiplicities are the curse of the social sciences. In many areas of psychology, we cannot do research on important problems without encountering multiplicity. We often encounter many variables and many relationships.

One of the most prevalent strategies psychologists use to handle multiplicity is to follow an ANOVA with pairwise multiple-comparison tests. This approach is usually wrong for several reasons. First, pairwise methods such as Tukey's honestly significant difference procedure were designed to control a familywise error rate based on the sample size and number of comparisons. Preceding them with an omnibus *F* test in a stagewise testing procedure defeats this design, making it unnecessarily conservative. Second, researchers rarely need to compare all possible means to understand their results or assess their theory; by setting their sights large, they sacrifice their power to see small. Third, the lattice of all possible pairs is a straightjacket; forcing themselves to wear it often restricts researchers to uninteresting hypotheses and induces them to ignore more fruitful ones.

As an antidote to the temptation to explore all pairs, imagine yourself restricted to mentioning only pairwise comparisons in the introduction and discussion sections of your article. Higher order concepts such as trends, structures, or clusters of effects would be forbidden. Your theory would be restricted to first-order associations. This scenario brings to mind the illogic of the converse, popular practice of theorizing about higher order concepts in the introduction and discussion sections and then supporting that theorizing in the results section with atomistic pairwise comparisons. If a specific contrast interests you, examine it. If all interest you, ask yourself why. For a detailed treatment of the use of contrasts, see Rosenthal, Rosnow, and Rubin (in press).

There is a variant of this preoccupation with all possible pairs that comes with the widespread practice of printing *p* values or asterisks next to every correlation in a correlation matrix. Methodologists frequently point out that these *p* values should be adjusted through Bonferroni or other corrections. One should ask instead why any reader would want this information. The possibilities are as follows:

1. All the correlations are "significant." If so, this can be noted in a single footnote.
2. None of the correlations are "significant." Again, this can be noted once. We need to be reminded that this situation does not rule out the possibility that combinations or subsets of the correlations may be "significant." The definition of the null hypothesis for the global test may not include other potential null hypotheses that might be rejected if they were tested.

3. A subset of the correlations is "significant." If so, our purpose in appending asterisks would seem to be to mark this subset. Using "significance" tests in this way is really a highlighting technique to facilitate pattern recognition. If this is your goal in presenting results, then it is better served by calling attention to the pattern (perhaps by sorting the rows and columns of the correlation matrix) and assessing it directly. This would force you, as well, to provide a plausible explanation.

There is a close relative of all possible pairs called "all possible combinations." We see this occasionally in the publishing of higher way factorial ANOVAs that include all possible main effects and interactions. One should not imagine that placing asterisks next to conventionally significant effects in a five-way ANOVA, for example, skirts the multiplicity problem. A typical five-way fully factorial design applied to a reasonably large sample of random data has about an 80% chance of producing at least one significant effect by conventional $F$ tests at the .05 critical level (Hurlburt & Spiegel, 1976).

Underlying the widespread use of all-possible-pairs methodology is the legitimate fear among editors and reviewers that some researchers would indulge in fishing expeditions without the restraint of simultaneous test procedures. We should indeed fear the well-intentioned, indiscriminate search for structure more than the deliberate falsification of results, if only for the prevalence of wishful thinking over nefariousness. There are Bonferroni and recent related methods (e.g., Benjamini & Hochberg, 1995) for controlling this problem statistically. Nevertheless, there is an alternative institutional restraint. Reviewers should require writers to articulate their expectations well enough to reduce the likelihood of post hoc rationalizations. Fishing expeditions are often recognizable by the promiscuity of their explanations. They mix ideas from scattered sources, rely heavily on common sense, and cite fragments rather than trends.

If, on the other hand, a researcher fools us with an intriguing result caught while indiscriminately fishing, we might want to fear this possibility less than we do now. The enforcing of rules to prevent chance results in our journals may at times distract us from noticing the more harmful possibility of publishing bogus theories and methods (illdefined variables, lack of parsimony, experimenter bias, logical errors, artifacts) that are buttressed by evidently impeccable statistics. There are enough good ideas behind fortuitous results to make us wary of restricting them. This is especially true in those areas of psychology where lives and major budgets are not at stake. Let replications promote reputations.

## RESULTS AND DISCUSSION

***Quantitative and qualitative data :***

*In* advanced studies, a researcher may approach his topics quantitatively, qualitatively or with the use of a mixed methodology. When opting for a qualitative approach, researchers have several options in analyzing the data. The use of matrices, charts, tables and other visual displays are common tools used. With visual displays, the researchers can pare down the often abundant subjective data that has been gathered and determine what will be useful variables in his qualitative data analysis. One way educational researchers work to overcome the challenge of repeatability is to distinguish, in their reports, between repeatable practices and the non repeatable results that emerged from those practices.

Quantitative research can demonstrate rigor by including a wide variety of numerical and statistical data Schroder, K.E., Carey, M.P., Venable, P.A. (2003)7 ., while the rigor of qualitative research is harder to demonstrate because it often involves the qualitative analysis of qualitative data. For example in literary studies, researchers apply interpretive models to texts such as poems or novels. A literary researcher can apply a wide variety of interpretation models and can apply a single interpretive model in multiple ways to a variety of texts. Therefore it is difficult to generate a unifying set of criteria for determining whether that researcher‘s work is truly rigorous. When the researcher is applying qualitative models of analysis to qualitative or numerical data, the research process can be long and tedious because the researcher must carefully pore over the data in detail while crafting the analysis. For example to write a comprehensive historical account, a historian must examine hundreds of primary historical records and secondary historical accounts. Even after spending all his time and energy examining records and accounts, the historian has no guarantee that it covered everything. One way to compensate for the time- consuming problem of qualitative research is to promote qualitative research projects, such as writing historical accounts, as team based or collaborative. After collection of data, the selection of statistical test is more important. To select the right test, two questions arise, What kind of data have you collected ? and what is your goal ? Accordingly you have to select the statistical test.

## LIMITATIONS TO QUALITATIVE RESEARCH*:*

Qualitative Research is a broad term that refers to research methods most commonly used in fields such as Sociology, anthropology, ethnography and other human and social sciences. The strongest objection to qualitative research is that the quality of the research depends too greatly on the individual researcher (Silverman, S., Manson, M. (2003)8 .. Because the researcher designs the type of questions, he or she can in adherently influence the results due to her own personal beliefs. Because qualitative research is so inextricably entwined with the individual researcher, it is extremely challenging for other researchers to repeat qualitative studies. This makes it hard to confirm or deny the results of the original study. For example, in the field of education, one of the challenges of repeating qualitative study is that different elements of the original study can‘t be repeated, the teachers and students will all be different, as will the school and classroom environment, the methods of teaching and the styles of learning.

**USAGE OF EXCEL:**

Excel, the spread sheet program in Microsoft‘s popular office Software Package is a powerful application used to manage various types of data. Excel‘s capabilities, however are not limited to data management. The program Data Analysis tool enables users to analyze data using an array of statistical procedures that range from descriptive measures to rigorous inferential statistics, such as regression and analysis of variance (Smeeton, N., Goda, D. (2003) .The data analysis tool is included in all versions of Excel but must be installed by the user.

Fortunately, setting up and using the tool is relatively easy. We can use Data Analysis for Random Number Generation, to test a hypothesis in Excel to Analyze data. Excel‘s data analysis capabilities make it possible to conduct some advanced analyses of survey data but not others However a program known as XL Stat expands the analytical capabilities of Excel. Tools such as SAS and SPSS are designed with research professionals in mind and make a full range of analytical methods possible.

Choosing between parametric and non parametric tests is sometimes easy. You should definitely choose a parametric test if you are sure that your data are sampled from a population that follows a Gaussian distribution ( at least approximately). It is not always easy to decide whether a sample comes from a Gaussian population. If you collect many data points (over a hundred or so) you can look at the distribution of data and it will be fairly obvious whether the distribution is approximately bell shaped. (Thompson, B., Noferi, G. 2002)10 A formal statistical test (Kolmogorov Smirnoff test) can be used to test whether the distribution of the data differs significantly from a Gaussian distribution. But the solution depends on sample size. Parametric tests work well with large samples even if the population is non-Gaussian. In other words, Parametric tests are robust to deviate from Gaussian distributions as long as the samples are large. Parametric test is suitable when there are at least two dozen data points in each group.

Non-Parametric tests work well with large samples from Gaussian population. The p Values tend to be a bit too large, but the discrepancy is small. Non parametric tests are only slightly less powerful than parametric tests with large samples. P value is inaccurate for small samples and it tends to be too high.

## CONCLUSIONS

In this paper, different types of Statistical tools were explained for the purpose of Research and dissertations in psychology field .So one should have the skill of selecting a statistical tool for their research which renders good conclusions. Still some more information can be given for the researchers for their future research.

## REFERENCES


- Abelson, R. P. (1995). *Statistics as principled argument.* Hillsdale, NJ: Erlbaum.

- Abelson, R. P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science, 23,* 12-15.

- American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

- Bailar, J. C., & Mosteller, F. (1988). Guidelines for statistical reporting in articles for medical journals: Amplifications and explanations. *Annals of Internal Medicine, 108,* 266-273.

- Baker, B. L. (1969). Symptom treatment and symptom substitution in enuresis. *Journal of Abnormal Psychology, 74,* 42-49.

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society,* 57(Series B), 289-300.

- Chambers, J., Cleveland, W., Kleiner, B., & Tukey, P. (1983). *Graphical methods for data analysis.* Monterey, CA: Wadsworth.

- Chartrand, J. M. (1997). *National sample survey.* Unpublished raw data.

- Cleveland, W. S. (1995). *Visualizing data.* Summit, NJ: Hobart Press.

- Cleveland, W. S., & Devlin, S. (1988). Locally weighted regression analysis by local fitting. *Journal of the American Statistical Association, 83,* 596-640.

- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

- Cohen, J. *(1988). Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

- Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 49,* 997-1003.

- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

- Cronbach, L. J. (1975). Beyond the two disciplines of psychology. *American Psychologist, 30,* 116-127.
- Ehrenberg, A. S. C. (1975). *Data reduction: Analyzing and interpreting statistical data.* New York: Wiley.
- Ehrenberg, A. S. C. (1981). The problem of numeracy. *American Statistician, 35,* 67-71.
- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). Washington, DC: American Council on Education.
- Fisher, R. A. (1935). *The design of experiments.* Edinburgh, Scotland: Oliver & Boyd.
- Fleiss, J. L. (1994). Measures of effect size for categorical data. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 245-260). New York: Sage.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). *What if there were no significance tests?* Hillsdale, NJ: Erlbaum.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81,* 945-960.
- Holland, P. W., & Rubin, D. B. (1983). On Lord's paradox. In H. Wainer & S. Messick (Eds.), *Principals of modern psychological measurement* (pp. 3-25). Hillsdale, NJ: Erlbaum.
- Hotelling, H., Bartky, W., Deming, W. E., Friedman, M., & Hoel, P. (1948). The teaching of statistics. *Annals of Mathematical Statistics, 19,* 95-115.
- Hurlburt, R. T., & Spiegel, D. K. (1976). Dependence of *F* ratios sharing a common denominator mean square. *American Statistician, 20,* 74-78.
- Kahn, J. R., & Udry, J. R. (1986). Marital coital frequency: Unnoticed outliers and unspecified interactions lead to erroneous conclusions. *American Sociological Review, 51,* 734-737.
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R., Donahue, B., Kowalchuk, R. K., Lowman, L. L., Petoskey, M. D., Keselman, J. C., & Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research, 68,* 350-386.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56,* 746-759.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.
- McDonald, R. P. (1997). Haldane's lungs: A case study in path analysis. *Multivariate Behavioral Research,* 32, 1-38.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). Washington, DC: American Council on Education.

- Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231- 244). New York: Sage.
- Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (in press). *Contrasts and effect sizes in behavioral research: A correlational approach.* New York: Cambridge University Press.
- Rosenthal, R., & Rubin, D. B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166-169.
- Rosenthal, R., & Rubin, D. B. (1994). The countemull value of an effect size: A new statistic. *Psychological Science, 5,* 329-334.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology,* 66, 688-701.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63,* 581- 592.
- Rubin, D. B. (1977). Assignment of treatment group on the basis of a covariate. *Journal of Educational Statistics, 2,* 1-26.
- Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics,* 6, 34-58.
- Rubin, D. B. (1990a). Formal modes of statistical inference for causal effects. *Journal of Statistical Planning and Inference, 25,* 279-292.
- Rubin, D. B. (1990b). Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science, 5,* 472-480.
- Scheines, R., Spirites, P., Glymour, C., Meek, C., & Richardson, T. (1998). The TETRAD project: Constraint based aids to causal model specification. *Multivariate Behavioral Research, 33,* 65-117.
- Schmidt, F. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychological Methods, 1,* 115-129.
- Scott, D. W. (1992). *Multivariate density estimation: Theory, practice, and visualization.* New York: Wiley.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* New York: Chapman & Hall.
- Snyder, P., & Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. *Journal of Experimental Education, 61,* 334-349.
- Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher,* 25(2), 26-30.

- Thompson, B., & Snyder, P. A. (1998). Statistical significance and reliability analyses in recent *JCD* research articles. *Journal of Counseling and Development*, 76, 436-441.
- Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison- Wesley.
- Wainer, H. (in press). One cheer for null hypothesis significance testing. *Psychological Methods.*
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.