

The Implications of Parametric and Non-Parametric Statistics in Data Analysis in Marketing Research

Dr. Egboro Felix O.

Department of Business Administration
And Management
Delta State Polytechnic
Otefe-Oghara, Nigeria

Abstract

Statistical needs of science, technology and governments had grown. Reviewers of research reports, frequently criticize the choice of statistical methods. Researchers ought to have a large suite of statistical techniques to be successful in analyzing data. Analysis can be viewed as the categorization, the aggregation into constituent parts, and manipulation of data to obtain answers to the research questions or questions underlying the research project. A special aspect of analysis is interpretation, which involve taking the results of analysis and making inferences relevant to the research relationship studied, and drawing managerially useful conclusions about these relationships. In a typical research work, there might be statistical errors and short comings due to the incorrect use of statistical test thereby leading to incorrect conclusions. These incorrect conclusions may have negative effect on the reliability, validity and verifiability of the research results. The focus of this paper is on the nature of data and the correct statistical techniques in relation to parametric and non-parametric statistics that amount to reliable and valid research results. Recommendations were made as per the best inferential statistics that lead to reliable business/marketing decisions, putting into consideration the implications of robustness

Keywords: Parametric Statistics, Research Results, Inferential Statistics, Data Analysis, Robustness

1.0 Introduction

Statistics is the scientific process of collection, organization, analysis and interpretation of data with a view to providing useful concise information for decision making.

The use of statistical techniques in marketing research has created concern for many years. Statisticians have long expressed concern about the slow adoption of statistical ideas by researchers and the frequent misuse of statistics when statistical methods were used (Garfield, 1995).

According to Egbule (2001), Statistics serves as a means for reducing large volumes of data to manageable and comprehensible form; it is used for the study of population and samples; acts as a suitable means of making decisions on data collected and analyzed and also helps in the process of making reliable inferences from observed data.

Nachmias and Nachmias (2009) stated that there are basically two types of statistics which include descriptive and inferential statistics. Descriptive statistics enable the researcher to summarize and organize data in an effective and meaningful way. It involves the use of tables, charts, graphs, mean, modes, median, standard scores and correlation to treat collected data.

Inferential statistics is concerned with making inferences from a unit of a population. Inferential statistics allow the researcher to make decisions or inferences by interpreting data patterns. Researchers use inferential statistics to determine whether an expected pattern designated by the theory and hypotheses is actually found in the observations. We might hypothesize that quality product lead to increase in demand. To decide whether this hypothesis is true, we might survey consumers of the same products with different quality. We would then use descriptive statistics to make comparison between these groups and we would employ inferential statistics to determine whether the differences between the groups consumer support our expectations.

1.2 Statement of the Problems

In a typical research work, there might be statistical errors and shortcomings due to incorrect use of statistical test thereby leading to wrong conclusions. These incorrect usage of statistics and wrong conclusions may have negative effect on the reliability of the result.

Wrong research result can mislead an individual, organization and even a whole nation. This is due to the fact that research is a systematic investigation into problems in order to find solution to them. It involves collection, analysis and interpretation of data. It also involves formulation of hypotheses and subjecting these hypotheses into statistical test. The subjecting of these hypotheses into statistical test involves the use of inferential statistics which is an embodiment of parametric and non-parametric statistics. Except the right statistical technique is used on a right data, the research result might not be valid and reliable.

Kim (2006) reasoned that as the technology for conducting basic research continues to evolve, further analytical challenges could be expected. In addition the increasing trends on the use of statistical methods would have an impact on how current and future investigators were trained to successfully conduct research. He suggested that the increasing complexity of methods used in research would make it more difficult for investigators to effectively conduct their own research.

Strask, et al (2007) identified some common statistical errors and shortcomings that can be observed in a typical research work, which among others include:

- Incorrect use of statistical test.
- Failure to adjust for multiple comparisons.
- Failure to define and explain the statistical test used.
- Failure to conduct a priori sample size estimation
- Robustness of statistical techniques.

Strasak et al (2007) argued that errors in research can lead to incorrect conclusions. These conclusions had a considerable negative effect on the scientific impact of a research project by diminishing the reliability, validity and the verifiability of the study results.

1.3 Objectives of the Research

- ❖ To examine the implication of parametric and non parametric statistics on reliability of research results.
- ❖ To examine the relevance of robustness on research results.

2.0 Review of Literature

2.1 List and Description of Statistical Techniques

Ndou (2011) emphasized on different types of statistical techniques of which one has no statistical component. The purpose of this explanation is to acquaint researchers the different types of statistical techniques and the rate at which they are been used, as indicated in table 1.

Table I

| Column Number | Category | Brief Description |
|---------------|--|--|
| 1. | No statistics | No tabulated data, does not include any of the methods defined in the other categories |
| 2. | Basic descriptive statistics | Descriptive statistics only e.g. percentage estimates, central tendency estimates, dispersion estimates, bar graphs, pie charts, etc. |
| 3. | Basic parametric inferential statistics. | Confidence intervals, t-test (one, two independent samples and paired samples). |
| 4. | Correlation | Classical product moment correlation to measure the strength and significance of relationships (parametric and non parametric). |
| 5. | Contingency tables | Two – way tables (counts and / or percentages). Chi-square tests, Fischer’s exact test, McNemar’s test. Multi-way summaries (counts and or percentages), Mantel-Haenszel procedure, etc. |
| 6. | General non-parametric tests. | Sign test, Wilcoxon, Kruskal-Wallis and Mann- Whitney test. |
| 7. | Dimension reducing techniques | Principal component analysis, factor analysis, discriminant analysis, multidimensional scaling, structural equation modeling correspondence analysis, etc. |
| 8. | Regression methods | Simple, Multiple, Partial, Logistic and Stepwise regression. |
| 9. | ANOVA | One - way, two - way, and higher order models. |
| 10. | Multiple comparisons | All statistical procedures for handling multiple inferences on same data sets e.g. Benferroni, Scheffe, Duncan, etc. |
| 11. | Time series techniques. | Autoregressive models, Moving average models (MA) and Autoregressive and Integrated Moving Average (ARIMA) model, (Trend and seasonality analysis, Box & Jenkins and autocorrelations. |
| 12. | Simulation | Re-sampling techniques, bootstrapping, Jack-knife, all Monte Carlo simulations, etc. |
| 13. | Optimization | Linear and non linear programming, game theory, network analysis. e.t.c |
| 14. | Advanced newer techniques | Designed experiments, repeated measures, missing data methods, power analysis, e.t.c. |

Source: Aifheli, Amos Ndou (2011) assessing the statistical methodologies of Business Research in South African.

The first category does not use any statistical component. The research papers included in this category generally is of a theoretical nature, a literature review or a qualitative assessment.

The second category is a basic descriptive statistics used for describing sets of measurements. They are used to summarize and describe the important characteristics of a set of measurements. E.g. pie charts, bar graphs, frequency histograms.

The inferential statistical method which is the third category is used to make inferences about population characteristics, from information contained in a sample drawn from a population.

The coefficient of correlation is used to measure the degree of the association between bivariate variables. When a researcher is interested in measuring two variables on a single experimental unit, the resulting data becomes bivariate data. The two most popular correlation coefficients considered and classified in this category are Spearman’s correlation coefficient, and Pearson’s product moment correlation coefficients. The coefficient of correlation is a single number that indicates the strength and direction of the relationship between two variables. The value lies in the range - $1 \leq r \leq 1$.

The fifth category labeled contingency table classified all statistical techniques which analyzes table or counts. Contingency tables generally included measurements that would be considered qualitative or categorical. For example, in a two-dimensional contingency table, two variables may be identified; one variable would determine the row categories and the other would define the column categories.

The Sixth category is non-parametric statistical procedures. Non-parametric statistics are used to analyze if the assumptions of parametric statistics under the equality of variances and / or normality are not met. Sign test, Mann – Whitney U test and Kruskal – Wallis test are examples of non-parametric statistics.

The dimension reducing techniques are based on the statistical principle of multivariate statistics which involves observation and analysis of more than one statistical variable at a time. Dimension reduction techniques are used to study multiple dimension, while taking into account the effects of all variable “y” and a set of predictor variables (x_1, \dots, x_p). The type of reduction techniques classified under dimension reduction techniques include discriminant analysis, cluster analysis and factor analysis.

The eighth category is the regression. The regression function is a linear combination of one or more model parameters, called regression coefficients. Multiple regression is a regression equation with more than one predictor variables affecting the response variable. The data can be used to estimate the beta coefficients of the regression model. The model could be used for data explanation, inference and prediction.

Analysis of variance (ANOVA) techniques are used to compare the means of several populations. When an experiment is performed, the responses that would be generated in an experimental situation would always exhibit a certain amount of variability. In ANOVA, the total variation in the response measurement is divided into portions that are attributable to various factors of interest to the experimenter. These portions are then used to answer the effects of the various factors on the response of interest.

Multiple comparisons occur when one considers a set or a family of statistical inferences simultaneously. These methods are used in the same context as ANOVA to check whether there exists a difference in population means between more than two populations. In contrast to ANOVA, which simply tests the null hypothesis that all means were equal, multiple comparison procedures help the researcher to determine where the difference among the means occur. (Bacchetti, 2002).

Analysis of time series is based on the assumption that successive values in the data file represent consecutive measurements taken at equally spaced time intervals. Some series are aggregated over a period e.g. industrial production, while others are defined only at discrete points in time e.g. month, quarterly, etc. The main goal of time series analysis is to identify the nature of the phenomenon represented by the sequence of observation and to forecast the future values of the series. e.g. forecast of production, sales volume, financial requirement etc.

The methods for extrapolation in this category include

- ❖ Autoregressive (AR) Models;
- ❖ Moving Average (MA) Models;
- ❖ Autoregressive and Integrated Moving Average (ARIMA) Models;
- ❖ Method for Testing and Removing Trend and Seasonality
- ❖ Methods for Time series forecasting.

Simulation techniques refer to methods where computer programmes are used to imitate real data. The starting point in simulation is to construct a probability model and make probabilistic assumptions. Simulation techniques would include bootstrapping and Monte Carlo. Bootstrapping is a resampling technique used for estimating bootstrapping intervals. Monte Carlo methods are described as a class of computational algorithms, that rely on repeated model fitting to compute results. Monte Carlo methods are often useful for modeling phenomena with uncertainty inputs such as the calculation of risking marketing.

Optimization is a method for finding the best available solution. The method is used in marketing in relation to profit maximization and cost minimization. In the simplest case, this means solving problems in which one seeks to minimize or maximize a real function by systematically choosing the values of real or integer variables from within an allowed set.

Higher new techniques include method like designed experiments, repeated measures, missing data methods and power analysis, experimental error while at the same time accommodating the treatment design in an experiment (Kuehl, 2000). Experimental design involves designing a set of experiments, in which all relevant factors are varied randomly. The results from the experiments would be analyzed so that they can:

- Identify optimal conditions.
- Identify the factors that most influence the results and those that do not.
- Give details such as the existence of interactions and synergies between factors.

Methods for handling missing data include imputation, deletion, substitution etc. Imputation is the substitution of some value for the missing data point. Once all missing values have been imputed, the dataset can then be analyzed using standard techniques.

The main goal of statistical power analysis in a sample size estimating is to allow the researcher to decide on the size of a sample. Power analysis can also be used to test how likely the researchers statistical test would be able to detect the effects of a given size in a particular situation.

Performing power analysis and sample size estimation is regarded as an important aspect of experimental design. Without these calculations, the sample size may be very high or very low. If sample size is very low, the experiment would lack precision to provide reliable answer to the questions it is investigating. If the sample is very large, time and resources would be wasted, often for minimal gain.

2.2 Empirical Evidence of using different Statistical Methods

Ndou (2011) x-rayed the extent to which each of these statistical methods are been used in business research, assessing them with their p-values. For the purposes of interpretation, p-values less than 0.05 ($p < 0.05$) is considered strongly significant. It also shows the rate at which more or less researchers are using each statistical method.

| Statistical method | X ² values | p-values | Significance | Direction of change |
|-----------------------------------|-----------------------|----------|--------------|------------------------|
| No statistical method | 13.69 | 0.003 | Strong | Decreasing |
| Basic descriptive statistics | 6.36 | 0.095 | Weak | Increasing |
| Elementary inferential statistics | 7.11 | 0.068 | Weak | Increasing |
| Contingency table | 7.83 | 0.049 | Strong | Increasing |
| Regression | 8.29 | 0.039 | Strong | Increasing |
| Coefficient of correlation | 5.34 | 0.144 | - | Increasing |
| Reduction techniques | 6.63 | 0.085 | Weak | Increasing |
| ANOVA | 0.90 | 0.82 | - | Insignificant increase |
| Multiple comparison | 2.05 | 0.56 | - | Constant |
| New advance | 10.35 | 0.016 | Strong | Increasing |
| Time series methods | 9.91 | 0.021 | Strong | Increasing |

2.3 The Seven Deadly Sins of Statistical Analysis

Kuzon et al, (1996) wrote on seven sins of statistical analysis. These 'sins' can make research result not to be valid and reliable. The seven deadly sins or errors according to them include:

- 1) The use of parametric analysis for ordinal data.
- 2) The inappropriate use of parametric analysis in general.
- 3) The failure to consider the possibility of committing type II statistical error.
- 4) The use of unmodified t-tests for multiple comparisons
- 5) The failure to employ analysis of covariance, multivariate regression and nonlinear regression, and logistical regression when indicated.
- 6) The habit of reporting standard error instead of standard deviation
- 7) The under use or over use of statistical consultation.

Sin 1: Using Parametric Analysis for Ordinal Data

The basis of both Sin 1 and Sin 2 is disregarding specific conditions about the parameters of the population being studied. Sin 1 is the use of a parametric statistical test for ordinal data analysis. Expressing ordinal data using integers does not justify the use of parametric statistics. To avoid committing sin 1, for nominal or ordinal scaled data, non-parametric analysis should be used.

Sin 2: Inappropriate Use of Parametric Analysis

Before parametric statistical analysis is appropriate, certain sampling criteria must be met.

- (1) The study sample must be randomly drawn from a normally distributed population
- (2) The sample size must be large enough to be "representative" of the study population.

Although several common parametric tests (the t-test in particular) are “tolerant” of relaxation of these two criteria, in strict terms, parametric analysis should only be employed if they can be fulfilled.

Sin 3: Failure to Consider Type II Statistical Error

When using the calculation of a probability to decide whether two means are “different” or “the same,” a widely accepted significance level of 0.05 or 5% is used. If we compute that the likelihood of two samples being drawn from a single population is less than 5%, we conclude that the two means are “statistically significantly different.” The interpretation of this conclusion is clear. Our null hypothesis was that there is no difference between two means. Rejection of this null hypothesis signifies that there is less than a 5% chance that our conclusion (that the means are different) is erroneous. We accept a 5% chance of incorrectly rejecting the null hypothesis. This wrongful rejection of the null hypothesis when it is true is referred to as a type I error. Alpha is the probability associated with committing a type I error. By preselecting a (usually 5%), rejection of the null hypothesis is associated with a known and controllable chance of error.

A further consideration arises when we accept the null hypothesis, concluding that we fail to find a real difference between the two sample means. A type II error occurs when the null hypothesis is false and we fail to reject it. The probability of committing a type II error is termed beta. Alpha and beta are inversely related and vary inversely with sample size. To decrease the probability of committing both type I and type II errors, the sample size is increased. When comparing two sample means given a level and sample size (N), the estimated depends on the magnitude of the difference between the two means and sample variance.

Sin 4: Using Unmodified t-Tests for Multiple Comparisons

This problem is again related to the calculation of probabilities and, specifically to type I error.

In comparing three groups A, B, and C. in an experiment, we must perform three pair wise comparisons. A vs B, B vs C and A vs C. The cumulative probability of erroneously rejecting the null hypothesis is 5% (for A vs B) + 5% (for B vs C) + 5% (for A vs C) = 15% overall. As more groups are compared, this cumulative chance of type I error is compounded. Thus, multiple unmodified pair wise comparisons are not valid.

A strategy to diminish the chance of reaching invalid conclusions when comparing multiple group means is analysis of variance.

Sin 5: Underutilization of Analysis of covariance (ANCOVA), Multivariate Regression, Nonlinear Regression, and Logistic Regression

While most research is conducted using relatively straightforward experimental designs that are adequately handled with pair wise comparisons, ANOVA, or standard least-squares regression analysis, there are numerous circumstances in which more sophisticated statistical methods should be considered. For example, a research work with many independent variables require multiple regression to establish their relationship on the dependent variable rather than using ANCOVA.

When there is more than one important covariate that could affect a particular outcome, the use of more complex regression analysis should be considered. In a multivariate regression, a least squares computational method is employed for any number of variables in an attempt to account for the variation observed in the dependent variable. The variance due to an individual, independent variable is compared to the total variation in the data set and an F ratio is computed. If the probability of a larger F ratio is less than 5%, that variable is considered to be “significant” in explaining the variation in the outcome measure.

Sin 6: Reporting Standard Error Instead of Standard Deviation

Reporting standard error of the mean is perhaps not a sin at all; but, reporting standard deviation in place of standard error is a serious transgression. We all know that standard error is computed as the standard deviation divided by the square root of N, but this equation does not define the meaning of standard error of the mean. The standard error of the mean is the square root of the variance (i.e., the standard deviation) associated with the distribution of sample means that would be derived by repeatedly sampling data elements from the study population. Standard deviation is the square root of the sample variance and is, therefore, a direct measure of the spread of data in a sample. It is well known that two-thirds of the sample data points fall within one standard deviation of the sample mean and that 94% of data points fall within two standard deviations of the mean.

This direct, easily conceptualized meaning of standard deviation makes it preferable when reporting descriptive statistics. The meaning of the standard error of the mean is far more difficult to conceptualize and therefore more difficult to interpret directly (Ndou, 2011). The practice of reporting standard error because it looks better” is a statistical sin. (Kuzon, et al 1996).

Another argument sometimes advance for reporting standard error is that one can easily determine, by looking at the overlap of standard error bars on graph, whether or not two means are significantly different. This belief is incorrect. It is easy to construct scenarios wherein two means will have values within one standard error of each other, yet they are significantly different statistically. It is also easy to construct the alternative scenario (bars don't overlap, means are not significantly different). It is not possible to determine whether two means are significantly statistically different simply by looking at either standard deviation or standard error bars on a graph. Therefore because of its direct and easily understood meaning, we advocate the reporting of standard deviation as the parameter indicating the spread of the sample data.

Sin 7: Failure to rely on a Statistician or Relying too much on a Statistician

This statement is a double-edged sword. In its positive connotation, it indicates that the researcher has sought the expertise of a statistician to assist with the interpretation of data, an obviously desirable maneuver. It may, however, indicate that the researcher has little or no concept of the statistical methods being employed for the analysis of the data, preferring to abdicate all responsibility to a third party. While there are times when statistical analysis may become extraordinarily complex, it is our opinion that it is the responsibility of the primary author or investigator to understand and to agree with the statistical analysis utilized. This may seem unfair, since one cannot become an expert on everything. Nevertheless, if statistical analysis is to be used as a means of evaluating our research results and thereby use it to validate important decisions regarding patient management, we submit that it is a sin simply to “give data to the statistician” and then to get back “results.”

2.4 The Concept of Robustness on Parametric Statistics

Norman, (2010) stated that Reviewers of research reports, frequently criticize the choice of statistical methods. While some of these criticisms are well-founded, frequently the use of various parametric methods such as analysis of variance, regression, correlation are faulted because (a) the sample size is too small, (b) the data may not be normally distributed, (c) the data are from Likert scales, which are ordinal, so parametric statistics cannot be used. Normans (2010), dissected these arguments, and show that parametric statistics are robust with respect to violations of these assumptions. According to him, challenges like those above are unfounded, and parametric methods can be utilized without concern for “getting the wrong answers”.

A model is said to possess robustness if it is difficult for user to obtain incorrect answers. (Schmudt and Svend, 2006). A model is robust if it is stable and insensitive to minor changes in the environment.

Bacchetti (2002), has also pointed out that many of these comments are unfounded or wrong, and appear to result from a review culture that encourages “overvaluation of criticism for its own sake, appropriate statistical dogmatism” and is subject to time pressure, and lack of rewards for good peer reviewing.

The appropriate descriptive and inferential statistics differ for ordinal and interval variable and if the wrong statistical technique is used, the researcher increase the chance of coming to the wrong conclusion. But what is left unsaved is how much it increases the chance of an erroneous conclusion. This is what statisticians call “robustness”. Robustness is the extent to which the test will give the right answer even when assumptions are violated (Norma 2010).

One of the beauties of statistical methods is that, although they often involve heroic assumptions about the data, it seems to matter very little even when these are violated (Gaito 1980). If it doesn't increase the chance very much (or not at all, then we can press on (Fletcher et al, 2010).

In addressing the issue of robustness, we can explore the impact of three characteristics of statistical analysis. (Norman, 2010).

(1) You can't use parametric' tests in a study because the sample size is too small. Norman (2010) argued that, nowhere in the assumptions of parametric statistics is there any restriction on sample size. According to him, it is simply not true, for example, that ANOVA can only be used for large samples, and one should use a t-test for smaller samples.

(2) ANOVA and t-tests are based on the same assumptions, for two groups the F- tests from the ANOVA is the square of the t-test. Nowhere is there any evidence that non-parametric tests are more appropriate than parametric tests when sample sizes are smaller (Jamieson, 2004).

In fact, there is one circumstance where non-parametric tests will give an answer that can be extremely conservative (i.e. wrong). The act of dichotomizing data (for example, using final exam scores to create pass and fail groups and analyzing failure rates, instead of simply analyzing the actual scores), can reduce statistical power enormously.

Sample size is not unimportant. It may be an issue in the use of statistics for a number of reasons unrelated to the choice of test (Jamieson, 2004).

(a) With too small a sample, external validity is a concern.

(b) When the sample size is small, there may be concern about the distributions.

(c) Small samples require larger effects to advance statistical significance. Pedhazur and Schmelkin (1991) describe effect size as “the magnitude of findings” (e.g. a correlation between two variables, the difference between two means). In other words, the effect the independent or predictor variable had on the dependent variable.

(3) You can't use t-test and ANOVA because the data are not normally distributed.

In this case, we must not forget that the assumption of normality is of the distribution of means, not of the data. The central limit theorem shows that, for sample sizes greater than 5 or 10 per group, the means are approximately normally distributed regardless of the original distribution. Empirical studies of robustness of ANOVA date all the way back to Pearson (1981), who found ANOVA was robust for highly skewed non-normal distributions (Norman, 2010).

(4) You can't use parametric test like ANOVA and Pearson Correlations (or regression, which amounts to the same thing) because the data are ordinal, you can't assume normality.

The question, then is how robust are Likert scales to departures from linear, normal distributions. Carifio and Perla (2008) pointed out that those who defend the logical position that parametric methods cannot be used on ordinal data ignore the many studies of robustness.

They argued that while Likert questions or items may well be ordinal, Likert scales, consisting of sum across many items, will be interval. It is completely analogous to the everyday, and perfectly defensible, practice of threatening the sum of correct answers on a multiple choice test, each of which is binary, as an interval scale.

According to Joseph, et al (2009), there are many situations in marketing research where the information problem, defined construct, and data requirements call for the collection of certain type of state of mind, state of behaviour, and state of intention data that are truly ordinal but that have an assumed distance scaling property artificially imposed on them so that the researcher can perform some type of advanced statistical analysis. This is referred to as ordinal interval scale.

Norman (2010), Gaito (1980), Carifio and Perla (2008) summarized their concept of robustness by stressing that parametric statistics can be used with Likert data, with small sample sizes with unequal variances, and with non-normal distributions, with no fear of “coming to the wrong conclusions”. To them, these findings are consistent with empirical literature dating back nearly 80 years.

2.5 Statistical Power in Marketing Research

When planning research, we are usually looking for something, for example, a difference in product ratings between experimental groups in a taste-test, or increased sales after a promotional campaign (Cohen, 1988). In order to find what is being sought, the researcher may measure a number of variables and conduct a statistical analysis relating the measurements to some model of reality. Often in marketing research, the model is compared with the available data, by use of a parametric statistical test of some null hypotheses.

The aim of this is to enable the maximum chance of a statistical test of a statistically significant result, if the null hypothesis is false in the population while staying within budgetary constraints.

Statistical power is the probability that we will find what we are looking for – if it is really there. Power is the probability that a statistical hypothesis test will correctly reveal an effect if, in fact it exists in the population.

If an effect exists in the population, what factors influence the chances that a statistical test on a sample drawn from that population will correctly identify the presence of the effect?.

The ability of a statistical test to reject the null hypothesis when it is false is referred to as the power of the test (Olanye 2006). The more power a statistical test has, the greater the ability of the test, to detect significant differences or relationship if they actually exist.

Summary and Conclusions

The overall process of analyzing and making inferences from sample data can be viewed as a process of refinement that involves a number of separate and sequential steps, that may be identified as part of three broad stages. These stages include tabulation, hypothesis formulation and making inferences. Making inferences has to do with reaching conclusions about variables that are important, their parameters, their differences and the relationships among them.

Errors in reaching conclusions or decisions lead to negative effects in relation to the reliability, validity and verifiability of the results. This emanate from incorrect use of statistical tests, failure to adjust for multiple comparisons, the failure to define and explain the statistical test used, and failure to conduct a priori sample size estimation.

Parametric statistics techniques are commonly used with internal and ratio data, while the non-parametric statistics techniques uses nominal and ordinal data.

However, the concept of robustness emphasizes that there is little or no effect whether interval or ratio data are strictly attached to parametric or non parametric statistics.

Wrong research results may be arrived at if the seven sins of data analysis are committed.

Recommendations

- 1) There is need to identify the nature of data used for hypothesis testing, with associated statistical techniques, quantitative and qualitative data must be identified and the right statistical techniques adopted.
- 2) Where qualitative data are to be tested with parametric statistics, there must be standardization or converting them to interval scale. This result in valid research results. In other words ordinal-interval hybrid scales should be adopted.
- 3) In as much as the exact result cannot be achieved when nominal and ordinal data are tested with parametric techniques by robustness theory, the concept is not encouraged. Quantitative and qualitative data should go with parametric and non-parametric statistics respectively.
- 4) Sample size derived from population of study, should be reasonably large enough to give proper generalization of the result. A good representation of the sample size gives reliable and valid research result.
- 5) The integration of data analysis algorithms into computer software has increased over the last three decades. Statistical software has simplified the use of advance mathematical routines. It is advisable for researchers to adapt this paradigm shift by using different statistical software packages such as SPSS.
- 6) The higher the level of significance, the higher the error made in hypothesis testing, i.e. ($P \leq 0.5$). The level of significance or margin of error should be low in order to increase level of research confidence, viability and reliability.
- 7) Type I and type II errors renders the whole research result invalid and unreliable. Much care must be taken to avoid mistake by using appropriate statistical power test.

References

- Agbonifoh, B. A. and Yomere, G. O. (1999). *Research Methodology in the Social Sciences and Education*, Benin City, Centerpiece Publishers.
- Aweriale, P. E. and Dibua, V. A. (1997). *Research methods and Statistics*, Benin City, Olive Publishers.
- Bacchetti, F. (2002). Peer Review of Statistics in Medical Research: The Other Problem. *British Medical Journal*, 234, 1271 – 1273.
- Cohen, J. (1998). *Statistical Power Analysis for the Behavioural Science*, New Jersey. Hillsdale Publishers.
- Egbule, J. F. and Okobia, D. O. (2001). *Research Methods in Education for Colleges and Universities*, Agbor, Kmensuo Publishers.
- Fletcher, K. E, French, C. T, Corapi, K. M, Irowins, R. S. and Norman, G. R. (2010). Prospective measures provide more accurate assessments than retrospective measures of the minimal important difference in quality of life. *Journal of Clinical Epidemiology*.
- Gaito, J. (1980). Measurement Scale and Statistics: Resurgence of an Old Misconception, *Psychological Bulletin*. 87, 564 - 567.
- Jameison, S. (2004). Likert Scales: How to Abuse Them. *Medical Education*, 38, 1217 – 1218
- Kim, M. (2006). Statistical Methods in Arthritis and Rheumatism, *Arthritis and Rheumatism* Vol. 54, No. 12: 3741 – 3749
- Kress, G. (1982). *Marketing Research*, U.S.A, Prentice Hall Publishers.
- Kuzon, W. M., Urbanchek, M. G. & McCabe, S. (1996). The Seven Deadly Sins of Statistical Analysis. *Annals of Plastic Surgery*. 37, 265 – 272.
- Nachmias, C. F. and Nachmias, D. (2009). *Research Methods in the Social Sciences*, London, Replika Press Ltd.
- Nalou, A. A. (2011). *Assessing the Statistical Methodologies of Business Research in South Africa*.
- Nickerson, R. S. (2000). “Null Hypothesis Significance Testing: A Review of an Old and Continuing Controversy, *Psychological Methods*, 5, 241 – 301.
- Olannye, P. A. (2006). *Research Methods for Business: A Skill Building Approach*. Lagos, Pee Jen Publishers.
- Pearson, E. S. (1981). The Test of Significance for the correction coefficient. *Journal of the American Statistical Association* 27, 128 – 134.
- Rousseau, G. G. (1992). Identifying Criteria of Consumer Awareness in a New South Africa, *Management Dynamics*. Vol. 1, No.: 67 – 86.
- Schmudt, M. J. and Svend, H. (2006). *Marketing Research*. New York, Prentice Hall Publishers.
- Stresak, A. A, Zaman, Q. Marinell, G. Pfeifer. Kp. And Ulmer, H. (2007). The Use of Statistics in Medical Research. A Comparison of the New England. *Journal of Medicine and Nature of Medicine, the American Statistician*. Vol. 61, No. 1: 47 – 55.