

SJMLS - 4 (2) - 001

Choice of Parametric and Nonparametric Statistical Procedures in Clinical and Biomedical Research**Okoroiwu, H. U. *¹ and Akwiwu E. C.¹**

Haematology Unit, Department of Medical Laboratory Science, University of Calabar, Nigeria¹
Author for Correspondence: okoroiwuhenshaw@gmail.com/+234- 803-883-3901*

Summary

Statistical literacy is essential in biomedical research as an author or reader. Lack of basic statistical knowledge as well as inappropriate statistical tool choice hinders understanding of biomedical research as well as its application. This research detailed basics of parametric and non-parametric statistical procedures. The advantages, disadvantages, indications, examples and analysis of parametric and non-parametric procedures are discussed.

Keywords: *Biostatistics; Statistical errors; parametric statistics; nonparametric statistics, choice of statistics; biometry.*

Introduction

Statistical analysis is a universal method with which to assess the validity of a conclusion. It is one of the most important aspects of a medical and biomedical paper. It gives meaning to otherwise meaningless series of numbers and permit researchers to draw conclusions from uncertain facts (Nahm, 2016). Statistical literacy is a critical skill for biomedical researchers, users of medical literatures and clinicians (Qualls *et al.*, 2010). Despite the increased emphasis on statistical literacy, many biomedical researchers cannot demonstrate competence with basic statistical concepts (AAMC, 2018; Qualls *et al.*, 2010; West, 2007). This statistical illiteracy cum inappropriate choice of statistical techniques result in mistakes in statistical analysis and faulty conclusions, hence, undermining the importance of the research (Qualls *et al.*, 2010). A study by Parsons and Colleagues (2012) on 100 representative orthopedic papers showed that in

39% (30 – 49%) of the studies, a different analysis should have been undertaken and in 17% (10 – 26%) a different analysis should have made difference to the overall conclusions. More so, statistical analysis in biomedical fields has been made a lot easier by the advent of user-friendly statistical analysis packages. However, the bad side of this is that the risk of choosing incorrect statistical tool or misinterpretation of result is increased as these software can permit one to input data in the software and simply press “analysis” or “OK” to compute p-value or confidence interval without understanding the basic concept of the statistics (Rosebaum, 2010). One of the prevalent statistical errors found in journals is the application of parametric statistical techniques to non-parametric data (Yim *et al.*, 2010). This is presumed to be due to the fact that biomedical researchers have had relatively few opportunities to use non-parametric statistical techniques as compared to parametric techniques because they have been trained mostly on parametric statistics (Rosebaum, 2010). This paper sets off to improve the understanding of parametric and non-parametric statistical techniques as well as review literatures that will aid biomedical researchers in choosing statistical tools for biomedical research.

Parametric Statistics

Parametric statistics is a branch of statistics that assumes that the data have come from a type of probability distribution and makes inference about the parameters (Geisser and Johnson, 2006). It assumes the data are on a quantitative (numerical) scale, with a normal distribution of

the underlying population; the samples have the same variance (homogeneity of variances); the samples are randomly drawn from the population, and the observations within the group are independent of each other (Ali and Bhaskar, 2016). The three most basic prerequisite for parametric statistic choices are:

- The assumption of normality which specifies that the means of the samples group are normally distributed.
- The assumption of equal variance which specifies that the variances of the samples and their corresponding population are equal.
- The data is measured in interval or ratio scale.

Generally speaking, parametric methods make more assumptions than non-parametric methods. If those extra assumptions are correct, parametric methods can produce more accurate and precise estimates. They are said to have more statistical power. However, if the assumptions are incorrect, parametric methods can be very misleading, for that reason, they are often not considered robust (Corder and Foreman, 2009).

Advantages of parametric tests

- They have more statistical power (when their assumptions are not violated), hence are more likely to detect a significant difference when it really exists.
- Inference drawn from parametric tests is easy to interpret and more meaningful.
- Information in the data is fully utilized.

Disadvantages of parametric tests

- They may lead to invalid conclusions when the assumptions are violated.
- They are prone to greater effect by outliers if present.
- Its validity is restricted to interval or ratio scale.

Indications for parametric statistics

It is a common practice that as long as there is no reason to believe that one or more of the assumptions of parametric tests have been violated, when the level of measurement for a set of data is interval or ratio, the data should be evaluated with appropriate parametric tests.

Nonparametric statistics

Non parametric tests are usually referred to as distribution free or assumption free tests (Sheskin, 2011). In truth they are not assumption free, and, in view of this, Marascuilo and McSweeny (1977) suggests that it might be more appropriate to employ the term “assumption freer”. A non-parametric statistic does not necessitate any condition to be fulfilled about the parametric of the population from which the sample was drawn. Non-parametric tests can be used on nominal or ordinal data (Walsh, 1962). Non-parametric tests are also applied to interval or ratio scale which do not follow the normal distribution. Non-parametric statistical analysis differs from parametric statistical analysis in that it only uses + or – sign or the rank of data sizes instead of the original values of the data per se (Nahm, 2016).

Advantages of Nonparametric test

- Non-parametric methods require no or very limited assumption to be made about the format of the data and may be preferred when assumptions required for parametric methods are not valid.
- There is less possibility to reach incorrect conclusions because assumptions about the population are unnecessary.
- They can be useful in dealing with unexpected outliers that might pose problems to the parametric method as it deals with signs and ranks.
- This method can be useful for small samples.
- It is more intuitive and does not require much statistical knowledge.

Disadvantages of nonparametric statistics

- Non-parametric statistics may lack power when compared with the more traditional parametric method.
- It is geared towards hypothesis testing rather than estimation of effects.
- Information acquired from non-parametric statistics is limited compared to that from parametric method and it is more difficult to interpret.

- The information in the data is not fully utilized.
- Tied values can be problematic when these are common and adjustment to the test statistics may be necessary.

Indications for nonparametric statistics

- When sample size is small and the researcher is unsure about the normality of the data.
- When data is ordinal, nominal or some outliers that cannot be removed.
- When the assumptions of the alternative parametric method are violated.
- If the data is better represented by median rather than mean, then, non-parametric method is preferred. For instance, when dealing on data on income, median better represent the data than mean as a few outliers will skew the data significantly.

Parametric versus nonparametric analysis

Introductory statistics textbooks typically advice against the use of parametric methods unless data approximate to normal distribution (Vickers, 2005; Altman, 1991). In heeding this caution, biomedical authors and readers often assume that inappropriate use of inferential statistic tests can result in low p-value or erroneously narrow 95% confidence interval. This translates reaching a conclusion that a difference between groups exist in the source population when in reality such difference does not exist. In other words, it is widely and incorrectly assumed that inappropriate use of inferential statistics potentiates making a type 1 error (Qualls *et al.*, 2010). In actual fact, the major problem with parametric tests is that they lack power when their assumptions are violated as occurs when they are applied

to non-parametric data (Goodacre, 2008). When analyzing non-normally distributed data with parametric tests, the analyst is more likely to fail to detect any difference in the source population when one truly exists; which defines type II error (Goodacre, 2008). Qualls and Colleagues (2010) substantiated this assertion using the 2006 National Ambulatory Medical Care Survey (NHAMCS) emergency department length of stay that was rightward skewed. They reported that using student t test to analyze covariates led to the conclusion that care by a mid-level provider and patient gender were not associated with emergency department length of stay; while the Wilcoxon rank sum test found these variables to be significantly associated with emergency department length of stay. More so, empirical statistical research has clearly demonstrated that t- test does not inflate type I error (false positive) (Vickers, 2005). In one of such studies, Heeren and Colleagues (1987) using non-normal data found that where there was truly no difference between groups, t-test rejected the null hypothesis close to 5% of the time. Consequently, the concern over the relative advantage of parametric and nonparametric methods have focused on type II error (Vickers *et al.*, 2005; Sawilowski, 1993). However, consistent data has shown that where data are sampled from a normal distribution, t- test has very slightly higher power than Mann-Whitney, the non-parametric alternative, but when data is sampled from non-normal distribution, Mann-Whitney is superior, often by a large amount (Vickers *et al.*, 2005; Bridge and Sawilowski, 1999). Parametric methods and their corresponding non-parametric

analysis are listed in table 1 for easy comparison and choice making.

Some common parametric and non-parametric statistical tests employed in medical and biomedical research

Inferential statistical tests employed with single sample

- **The single-sample z test (Parametric):**
It is used in hypothesis testing situation involving a single sample in order to determine whether or not a sample with mean of \bar{X} is derived from a population with a mean of μ (Sheskin, 2011).
- **The single sample t- test (Parametric):**
The single sample t- test is employed in hypothesis testing situation involving a single sample in order to determine whether or not a sample with a mean of \bar{X} is derived from a population with a mean of μ (Sheskin, 2011).
- **The single sample chi-square test for population variance (Parametric):**
It is employed in hypothesis testing situations involving a single sample in order to determine whether or not a sample with an estimated population variance of \bar{S}^2 is derived from a population with a variance of σ^2 (Sheskin, 2011).
- **The single sample test for evaluation population skewness (Parametric):**
It is used to measure the degree of asymmetry of a given distribution around its mean (Cisar, 2010).
- **The single sample tests for evaluating population kurtosis (Parametric):**
It is used to access the relative peakedness or flatness of a distribution compared with the normal distribution (Cisar, 2010).
- **Wilcoxon signed-rank test (Nonparametric):**
It is employed in hypothesis testing situations involving a single sample in order

to determine whether or not a sample is derives from a population in which the median (Θ) is equal to a specific value (Sheskin, 2011).

- **The single sample chi-square test (Nonparametric):**

It is employed in hypothesis testing involving single samples testing if there is a difference between the observed frequencies of the k cells and their expected frequencies (Cochran, 1952).

Inferential statistical tests employed with two independent samples (and related measures of association and correlation)

- **The t-test for two independent samples (Parametric):**
It is used in hypothesis testing involving comparison of means of two groups that are independent of each other (Kim, 2015).
- **The Mann-Whitney U test (Nonparametric):**
The test examines whether the distribution of ranked responses between the two samples being compared are significantly different (Gaddis and Gaddis, 1990).
- **The Kolmogorov-Smirnov test for two independent samples (Nonparametric):**
The test is used to see whether two independent samples of data have been drawn from the same or different population (Gaddis and Gaddis, 1990).
- **The chi-square test for $r \times c$ tables (Nonparametric):**
It is an extension of the chi-square goodness of fit test to two dimensional tables. It is applied to categorical data in the form of $r \times c$ tables (contingency table) (Sheskin, 2011).
- **Fisher exact test (Non-parametric):**
It is a hypergeometric distribution-based test used mainly for 2×2 contingency tables involving small sample sizes when one or neither of the marginal sums is predetermined by the researcher (Daniel, 1990).

Inferential statistical tests employed with two dependent samples (and related measure of association/correlation)

- **The t-test for two dependent samples (Parametric):**
The test is used in hypothesis testing involving comparison of means of two groups that are dependent (paired) on each other (Kim, 2015).
- **The Wilcoxon matched-pairs signed-rank test (Nonparametric):**
It is employed in hypothesis testing situation involving design with two dependent samples. It is an extension of the Wilcoxon signed-ranks test (Sheskin, 2011). It is the non-parametric analog of paired (dependent) t-tests (Gaddis and Gaddis, 1990).
- **The McNemar test (Nonparametric):**
It is a non-parametric procedure for categorical data employed in a hypothesis testing situation involving a design with two dependent samples (Sheskin, 2011).

Inferential statistical tests employed with two or more independent samples (and related measures of association).

- **The single-factor between subject's analysis of variance (one-way ANOVA) (Parametric):**
It is used to evaluate whether or not there is a difference between at least two means in a set of data for which two or more means can be compared (Sheskin, 2011). It is employed in a hypothesis testing situation involving independent samples (Anderson, 2001).
- **The Kruskal-Wallis one-way analysis of variance by ranks (Nonparametric):**
It is employed with (rank-order) data in hypothesis testing involving a design with two or more independent samples. It is an extension of the Mann-Whitney U test (Sheskin, 2011). It is the non-parametric analog of one-way ANOVA (Gaddis and Gaddis, 1990).

Inferential statistical tests employed with two or more dependent samples (and related measures of association/correlation)

- **The single-factor within-subjects analysis of variance (Parametric):**
It is also called single factor repeated measure ANOVA (Anderson, 2001). It is employed in hypothesis involving k dependent samples (Sheskin, 2011). It is an extension of dependent t- test. It compares means across one or more variables that are based on repeated observation.
- **Friedman Two-way analysis of variance by ranks (Nonparametric):**
It is a rank test used in hypothesis testing involving two or more dependent samples (Gaddis and Gaddis, 1990).
- **The Cochran Q test (Non-parametric):**
It is a nonparametric procedure for categorical data employed in a hypothesis testing situation involving a design with k = 2 or more dependent samples. It is employed to evaluate an experiment in which a sample of n subjects is evaluated on a dichotomous dependent variable (Sheskin, 2011).

Inferential statistical tests employed with factorial design (and related measures of association/correlation)

- **The between-subjects, factorial analysis of variable (X-way ANOVA) (Parametric):**
It is employed to simultaneously evaluate the effect of two or more independent variables on a dependent variable (Sheskin, 2011). The commonest type is the two-way ANOVA.

Inferential statistical tests employed in measure of homogeneity of variance

- **The Levene' test for homogeneity of variance (Parametric):**
It is used in determining within each group an absolute deviation score from the group mean for each of the n scores in a group (Sheskin, 2011). It is used in assessing assumption of homogeneity of variance assumed by parametric tests.

- **The Brown-Forsythe test for homogeneity of variance (Parametric):**
It involves determining within each group an absolute deviation score from the group median for each of the n scores in a group (Sheskin, 2011). It is used in confirming assumption of homogeneity of variance.
- **The Bartlett's test for homogeneity of variance (Parametric):**
It is an inferential procedure used to assess the equality of variance in different populations (Box, 1953).
- **The Siegel-Tukey test for equal variability (Nonparametric):**
It is a non-parametric procedure employed with ordinal (rank-order) data in hypothesis testing situations involving two independent samples (Sheskin, 2011). It is used in contrasting the variances of two independent samples. It is a test of dispersion. When population median is unknown and one is unwilling to assume, they are equal, Siegel-Tukey test for equal variability is not the appropriate non-parametric test of dispersion to use.
- **The spearman's rank-order correlation coefficient (Nonparametric):**
It is a bivariate measure of association/correlation which is employed with rank-order data (Sheskin, 2011).
- **The Kendall's Tau test (Nonparametric):**
It is a bivariate measure of correlation/association that is employed with rank-order data (Sheskin, 2011).
- **The Goodman and Kruskal's Gamma test (Nonparametric):**
It is a bivariate measure of correlation/association that is employed with rank-order data which are summarized within the format of an ordered contingency (Sheskin, 2011).
- **Simple linear regression (Parametric):**
It uses scores on one predictor variable to predict the scores on the criterion variable (Sheskin, 2011).

Multivariate Analysis

Inferential statistical tests employed in measure of normality

- **Shapiro-Wilk test for normality:**
It is used in assessing normality of a data. It is based on the correlation between the data and the corresponding normal scores (Peat and Barton, 2005).
- **Kolmogorov-Smirnov test of normality/goodness of fit:**
It is used in assessing normality of a data set. It is an empirical distribution function (EDF) in which the theoretical cumulative distribution function of the test distribution is contrasted with the EDF of the data (Sheskin, 2011).

- **Multiple Regression (Parametric):** It is used to compute correlation coefficient between criterion variable and k predictor variables (where $k \geq 2$).
- **Multivariate analysis of variance (MANOVA) (Parametric):**
It can be employed to analyze data for an experiment which involves one or more independent variables and multiple dependent variables. It is a multivariate analog of the single-factor between subject's ANOVA (Sheskin, 2011).
- **Multivariate analysis of covariance (MANCOVA) (Parametric):**
It is a procedure which combines the multivariate analysis of variance with an analysis of covariance. It is the multivariate analog of single-factor between-subjects analysis of variance (Sheskin, 2011).
- **Logistic regression analysis (Nonparametric):**
It analyses the relationship between multiple independent variables and a categorical dependent variable and estimate the probability of occurrence of an event by fitting data to a logistic curve (Park, 2013).

Measures of association/correlation

- **The Pearson product moment correlation coefficient (Parametric):**
It is strength measure of linear association between variables (Hauke and Tomasz, 2011). It is rather a descriptive statistic than inferential.

Meta-Analysis (Parametric):

It is a statistical method used to summarize numerical data from the individual studies into one overall estimate. The result of the individual studies and the overall studies from the meta-analysis are usually presented in a graph called forest plot (Callcut, 2009).

A schematic diagram as adapted from Sheskin's table (Sheskin, 2011). is presented in figures 1-4 as a guide in selection of statistical tools whereas table 1 shows parametric tests with their non-parametric analogs.

Conclusion

The appropriate use of parametric and nonparametric methods has been discussed. Table 1 and figures 1-4 summarizes key information to guide the reader in the selection of proper test technique. Nonparametric statistic methods are the correct choice when assumption of normality is clearly violated; however, they are not always the top choice for cases with small sample size as they have less statistical power compared to parametric methods. Parametric methods may lead to significant results in some cases while nonparametric methods may result in

more significant results in other cases. When parametric methods are selected, the researcher should ensure the necessary assumptions are all satisfied. If otherwise, it is more valid to use nonparametric methods as they are “always valid but not always efficient”, while parametric methods are “always efficient but not always valid”. Confirmation for assumption of normality can be accessed using Shapiro-Wilk or Kolmogorov-Smirnov tests, while assumption of homogeneity of variance can be assessed using Bartlett's, Levene's or Hartley's Fmax tests. Conveniently, some of these tests have been incorporated in most statistical software's such as SPSS (Statistical Package for Social Sciences). Figure 1 shows the flow chart for choice of inferential statistics for single sample data. Figure 2 shows the flow chart for choice of inferential statistics for two sample data. Figure 3 shows the flow chart for choice of inferential statistics for two or more sample data. Figure 4 shows the flow chart for choice of statistical tests for measure of association/correlation.

Table 1: Analog of parametric and nonparametric tests

Sample characteristics	Statistical method	
	Parametric tests	Non- parametric tests
One sample	One sample t test	Sign test Wilcoxon’s signed rank test
Two sample	Paired t test	Sign test Wilcoxon’s signed rank test
	Unpaired t test	Mann-Whitney test Kolmogorov-Smirnov test
k-sample	Analysis of variance	Kruskall-Wallis test Jonckheer test
	2- way analysis of variance	Friedman test
Correlation	Pearson’s test	Spearman’s rho

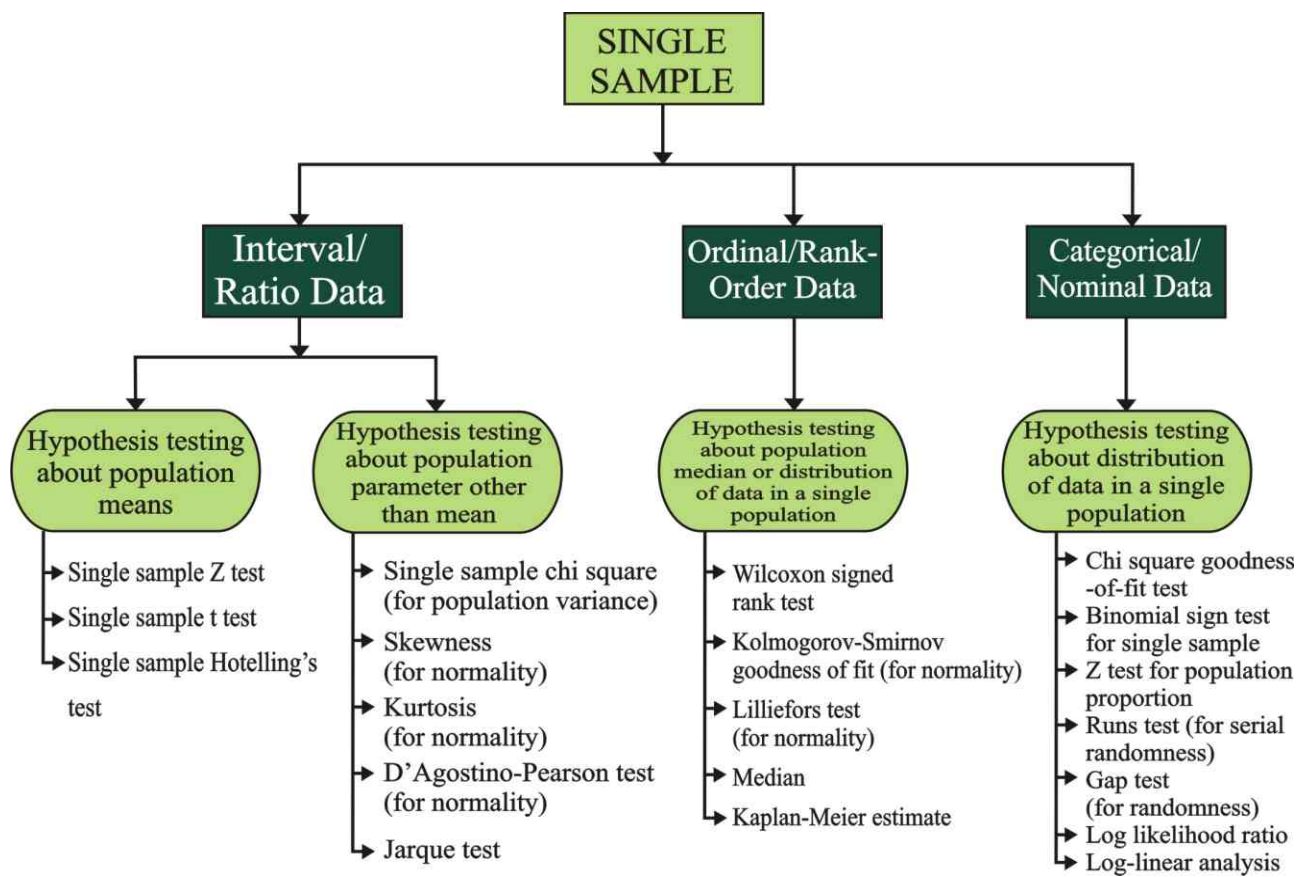


Fig 1: Flow chart for choice of inferential statistics for single sample data.

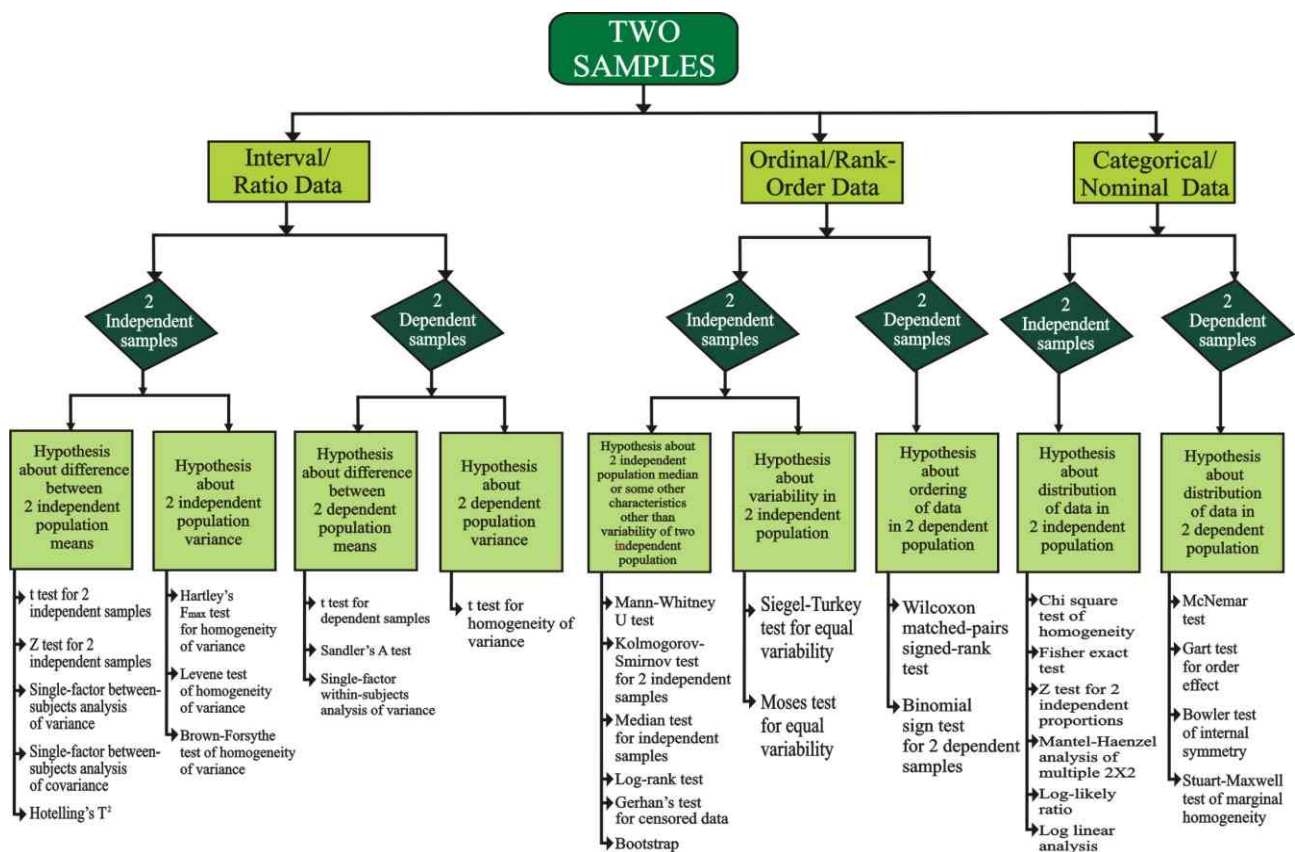


Fig 2: Flow chart for choice of inferential statistics for two sample data.

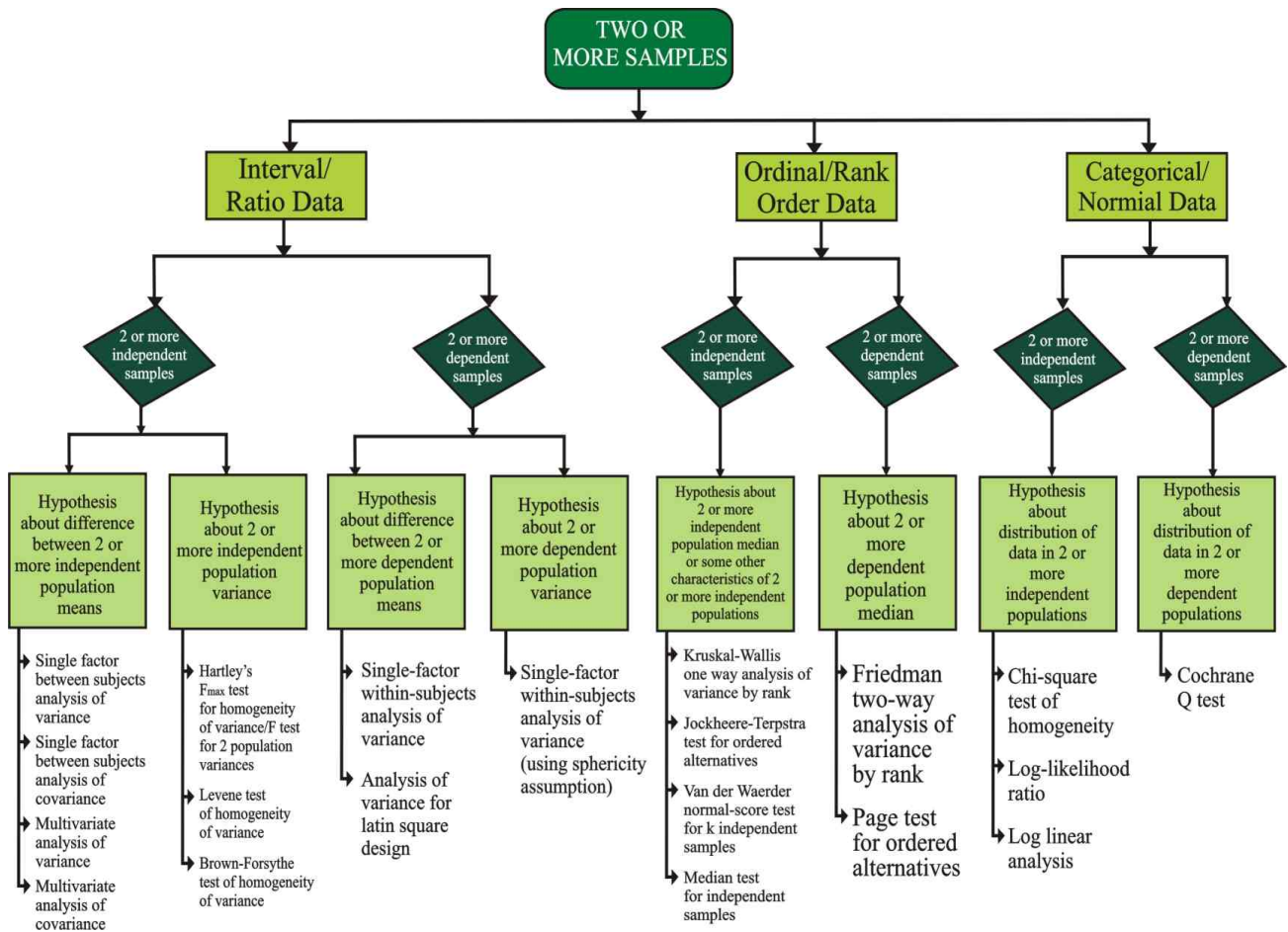


Fig 3: Flow chart for choice of inferential statistics for two or more sample data.

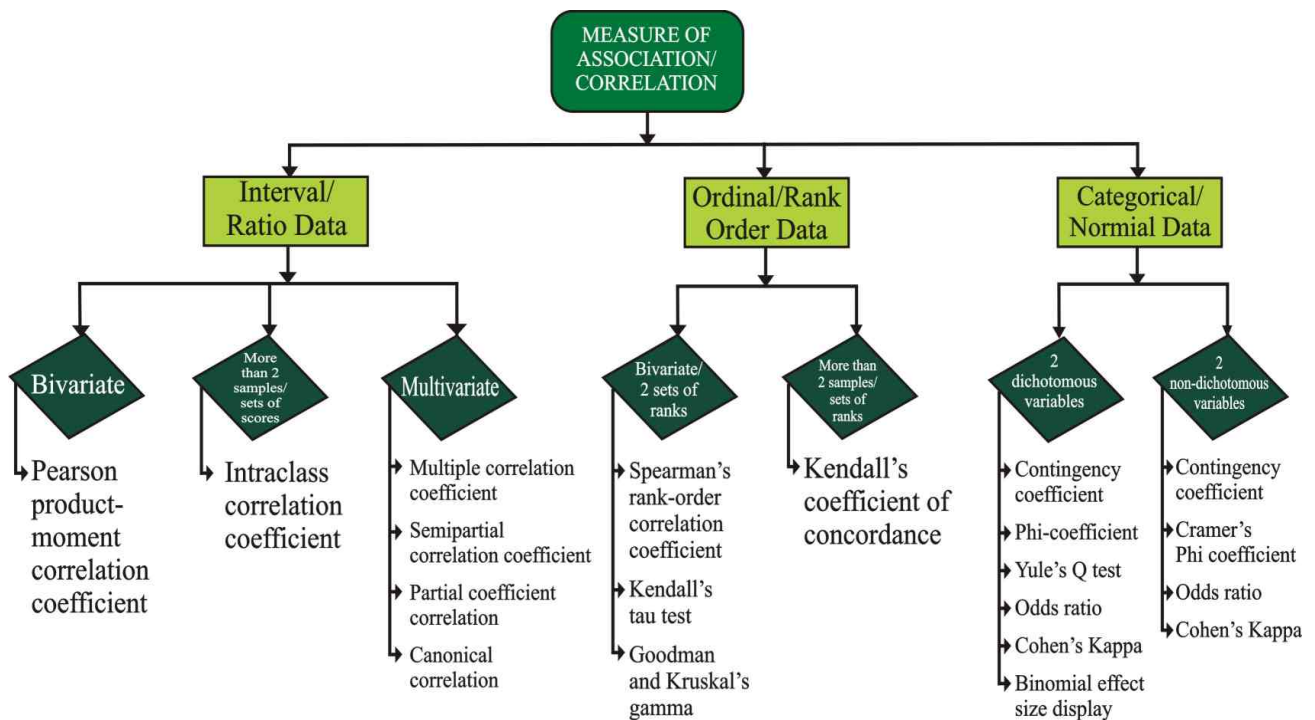


Fig 4: Flow chart for choice of statistical tests for measure of association/correlation.

References

- Ali, Z. and Bhaskar, S.B. (2016). Basic statistical tools in research data analysis. *Indian Journal Anaesthesia*; 60:662-669.
- Altman, D.G. (1991). Practical statistics for medical research. 1991. London: Chapman and Hall.
- Anderson, N.H. (2001). Empirical direction in design and analysis. Mahwah, NJ: Lawrence Erlbaum Associates. 2001.
- Association of American Medical Colleges (2018). Contemporary issues in medicine: Basic science and clinical research. Available at: <http://services.aamc.org/publications>. Accessed October 2, 2018.
- Box GEP (1953). Non-normality and tests on variances. *Biometrika*; 40:318-335.
- Bridge, P.D. and Sawilowsky, S.S. (1999). Increasing physician's awareness of the impact of statistics on research outcomes: comparative power of the t test and Wilcoxon –Rank sum test in small samples applied research. *Journal of Clinical Epidemiology*; 52:229-235.
- Callcut, R.A. and Branson, R.D. (2009). How to read a review paper. *Respiratory Care*; 54(10): 1379-1385.
- Cisar, P. (2010). Skewness and Kurtosis in function of selection of Network Traffic Distribution. *Acta Polytechnica Hungarica*; 7(2): 95-106.
- Cochran, W.G. (1952). The chi-square test of goodness of fit. *Annals of Mathematical statistics*; 23(3): 315-345.
- Corder, G.W. and Foreman, D.I. (2009). Non-parametric statistics for non-statisticians: A step-by-step approach. Hoboken, New Jersey: Wiley and Sons Inc.
- Daniel, W.W. (1990). Applied nonparametric statistics. 2nded, Boston: PWS-Kent Publishing Company: 121-24.
- Gaddis, G.M. and Gaddis, M.L. (1990). Introduction to biostatistics: part 5, statistical inference techniques for hypothesis testing with non-parametric data. *Annals of Emerging Medicine*; 19(9): 1054-1059.
- Geisser, S., Johnson, W. (2006). Modes of parametric statistical inferences. New Jersey: John Wiley and Sons Inc.
- Goodacre, S. (2008). Critical appraisal for emergency medicine 2: statistics. *Emerging Medicine Journal*; 25:362-364.
- Hauke, J. and Tomas Z. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Questiones Geographicae*; 30(2): 87-93.
- Heeren, T., D' Agnostino, R. (1987). Robustness of the two independent sample t test when applied to ordinal scaled data. *Statistics Medicine*; 6:79-90.
- Kim, T.K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*; 68(6): 540-546.
- Marascuilo, L.A. and McSweeny, M. (1997). Non-parametric and distribution-free methods for the social sciences. Monterey, CA: Brooks/Cole Publishing Company.
- Nahm, F.S. (2016). Non-parametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*; 69(1):8-14.
- Park, H.A. (2013). An introduction to logistic Regression: from Basic concept to Interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*; 43(2): 154-164.
- Parsons, N.R., Price, C.L., Hiskens, R., Achten, J., Costa, M.L. (2012). An evaluation of the quality of statistical design and analysis of published medical research; results from a systematic survey of general orthopaedic journals. *BMC Medical Research Methodology*; 12:60.

- Peat, J., Barton, B. (2005). *Medical statistics: A guide to data analysis and critical appraisal*. New Jersey: Blackwell Publishing.
- Qualls, M., Pallin, D.J., Scur, J.D. Parametric versus non-parametric statistical tests: the length of stay example. *Academy of Emerging Medicine*;17(10):1113-1121.
- Rosenbaum, S.H. (2010). Statistical methods in anesthesia. In: Miller RD (ed): *Miller's Anesthesia*. 7th ed. Philadelphia: Elsevier Inc. :3075-3086.
- Sawilowsky, S.S. (1993). Comments on using alternative to normal theory in statistics in social and behavioral sciences. *Canadian Psychology*; 34:432-439.
- Sheskin, D.J. (2011). *Handbook of parametric and nonparametric statistical procedures*. New York: Chapman and Hall Books:109-110.
- Vickers, A.J. (2005). Parametric versus nonparametric statistics in the analysis of randomized trials with non-normal distributed data. *BMC Medical Research Methodology*; 5:35.
- Walsh, J.E. (1962). *Handbook of nonparametric statistics*. New York: Van Nostrand Reinhold.
- West, C.P., Ficalora, R.D. (2007). Clinical attitude towards biostatistics. *Mayo Clinic Proceedings*; 82:939-943.
- Yim, K.H., Nahm, F.S., Han, K.A. and Park, S.Y. (2010). Analysis of statistical methods and errors in the articles published in the Korean Journal of Pain. *Korean Journal of Pain*; 23:35-41.

Citation: Okoroiwu, H. U. and Akwiwu E. C. Choice of Parametric and Non-parametric Statistical Procedures in Clinical and Biomedical Research. *Sokoto Journal of Medical Laboratory Science*;4(2): 5 - 15.

Copyright. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.