Wits Journal of Clinical Medicine, 2025, 7(1) 41–44 http://dx.doi.org/10.18772/26180197.2025.v7n1a9

Statistics @ WJCM

Navigating parametric and non-parametric statistical analyses: A practical guide in health sciences research

Sfiso Mabizela*

Centre for Health Science Education, Faculty of Health Sciences, University of Witwatersrand, York Road, Parktown, South Africa

Corresponding Author: *sfiso.mabizela@wits.ac.za*

INTRODUCTION

The selection of an appropriate statistical method primarily hinges on the research data's characteristics and the foundational assumptions. Broadly, statistical techniques are classified into parametric and non-parametric methods.(1,2) Parametric tests are preferred for their higher statistical power, as they can better detect effects when present. These tests require data to meet specific assumptions, including a normal distribution and measurement on interval or ratio scales.(1) Conversely, non-parametric tests offer flexibility when these assumptions are unmet, making them applicable to a broader range of data types. Non-parametric statistical tests do not require the data to follow a specific distribution, making them suitable for ordinal, nominal, or non-normally distributed continuous data.(1)

PARAMETRIC STATISTICS: COMPARING MEANS

One sample *t*-test

The One-Sample t-test is employed when comparing the mean of a single sample to a known value or population mean.(3,4) For example, a researcher might examine whether the average score of a group significantly deviates from a benchmark value, such as an average test score of 50. This test is ideal when the data is normally distributed, continuous and the sample is randomly drawn from the population.(4,5)

Independent Samples *t*-test

The Independent Samples t-test compares the means of two groups, such as males and females, to assess whether a statistically significant difference exists between them.(2,4,5) This test commonly applies in research involving continuous variables such as height, weight, or age. The key assumptions here include normally distributed data, homogeneity of variances, and the independence of the groups being compared.(4,5)

Paired Samples *t*-test

The Paired Samples t-test compares means from the same group at different times or under two different conditions, such as pre-and post-intervention scores.(4,5) This test is used when the same participants are measured more than once, assuming normality and that the data is continuous.(4) A statistically significant result indicates a significant change or difference between the two measurements (e.g., pre- and post-intervention).

One-Way Between-Groups ANOVA

The One-Way Between-Groups ANOVA extends this comparison to three or more independent groups, making it suitable for studies that compare the effects of different conditions or groups, such as comparing performance across different teaching methods.(4) This test assumes normally distributed data, homogeneity of variances, and independent groups, much like the t-test, but allows for more complex comparisons. If the p-value is statistically significant, it suggests that at least one group's mean differs significantly from the others, and, as a result, post-hoc tests should be interpreted to explore the mean differences among the groups compared.

PARAMETRIC STATISTICS: RELATIONSHIPS AND PREDICTIONS

Bivariate correlation

Bivariate correlation measures the strength and direction of the linear relationship between two continuous variables, such as the relationship between height and weight.(4) It assumes that the data is normally distributed, the relationship between variables is linear, and the data exhibits homoscedasticity, where the variance of one variable is similar across all values of the other. When interpreting the results of bivariate correlation, the strength and direction of the relationship between two continuous variables are assessed using the correlation coefficient (r). The values of r range from -1 to 1, where a value of 0 indicates no relationship, values approaching 1 indicate a strong positive relationship and values nearing -1 indicate a strong negative relationship. A significant p-value ($p < 0.05$) suggests that the observed relationship is statistically significant.

c09.indd 41 14-02-2025 10:54:51

Multiple regression

Multiple regression extends this analysis by examining the relationship between one dependent variable and two or more independent variables. For example, it can assess how income and education level predict spending habits. This method requires continuous or categorical independent variables and continuous dependent variables. Multiple regression assumes linearity, independence of errors, homoscedasticity, no multicollinearity among the independent variables, and that the data follows a normal distribution. The Mahalanobis distance is also checked to identify any multivariate outliers.(4,6) The primary result is the R^2 value, which reflects the extent to which the independent variables account for the variation in the dependent variable. Additionally, individual beta coefficients (β) are analysed to understand the specific contribution of each independent variable.(4)

Hierarchical multiple regression

Hierarchical multiple regression involves entering independent variables into the equation in steps based on theoretical justification.(4) Hierarchical regression, like multiple regression, requires that the same assumptions be inspected and requires careful sequencing of variable entry to evaluate incremental predictive power.(4,6) Interpretation involves comparing the R^2 values between models to determine if the added variables significantly improve the prediction. The change in R^2 is important to determine if additional variables significantly contribute beyond those already entered in previous steps and should include a discussion on the theoretical or practical rationale for adding variables in a particular sequence.(4)

NON-PARAMETRIC STATISTICAL ANALYSIS

Factorial analysis

Exploratory Factorial Analysis (EFA) is designed to uncover the underlying structure of a set of variables by grouping them into factors based on their correlations.(4,7) The main goal is to identify latent variables that explain the observed relationships among measured variables.(7) By doing so, it reduces the dimensionality of the data, making it easier to interpret complex relationships. The overall sample should be large enough for each variable to provide sufficient observations to detect meaningful patterns and relationships among the variables. Yong and Pearce (7) provide further insights on determining if data is suitable for exploratory factorial analysis.(4,7)

Inter-item reliability

Inter-item reliability measures the consistency of items within a test, survey, or scale. It assesses whether multiple items that measure the same construct produce consistent results.(4,8) For instance, if a survey contains several items to gauge job satisfaction, inter-item reliability would indicate whether those items consistently assess job satisfaction. A Cronbach alpha of 0.7 or higher is acceptable.(4,9) If Cronbach's alpha is below the acceptable threshold, it may be necessary to remove certain items to raise the alpha to a satisfactory level.

NON-PARAMETRIC STATISTICS: COMPARING MEANS

Mann-Whitney *U* Test

The Mann-Whitney *U* Test compares two independent groups when the data is ordinal or not normally distributed.(4,10) It is compatible with independent samples and is a non-parametric alternative to the independent t-test.(4) It assesses whether there is a significant difference between the groups by comparing the mean ranks of the two groups rather than the actual data points.

Kruskal-Wallis ANOVA Test

The Kruskal-Wallis ANOVA Test compares three or more independent groups on an ordinal or continuous non-normally distributed variable.(4,11) Like the Mann-Whitney test, it uses mean ranks to assess group differences. The Kruskal-Wallis ANOVA Test is compatible with independent samples and is a non-parametric alternative to one-way ANOVA.(12,13) A significant result suggests that at least one group is different from the others, and *post-hoc* analysis is necessary to determine where these differences are.

NON-PARAMETRIC STATISTICS: ASSOCIATIONS AND RELATIONSHIPS

Chi-square tests

The Chi-Square Test is a non-parametric test used to determine whether an association exists between two categorical variables.(14,15) For instance, it can be used to test if there is a significant association between gender and voting preference (e.g., yes/no responses). This test suits categorical data, where both variables are measured at the nominal or ordinal level.(4) It is crucial to ensure that each observation is independent, and there should be an expected frequency of at least 5% in each contingency table cell.(5)

Spearman Rho

The Spearman Rho is a non-parametric test that measures the strength and direction of the association between two ranked or ordinal variables.(4,14) Unlike Pearson correlation, Spearman's correlation does not assume a linear relationship but instead looks for a monotonic relationship between the variables.(16) The strength of the relationship is represented by Spearman's rho, which ranges from −1 to 1.

Logistic regression

Logistic regression is used to predict the probability of a binary outcome based on one or more predictor variables.

(17,18) This analysis is appropriate when the dependent variable is binary (e.g., yes/no) and the independent variables are either continuous or categorical.(4) Logistic regression does not require the assumption of normality, but it assumes linearity between the independent variables and the log odds of the dependent variable. An OR greater than one indicates higher odds of the event occurring, while an OR less than one suggests lower odds. Confidence intervals for the ORs should also be reported to show the precision of the estimates.(4) In logistic regression, assessing model fit and predictive power is crucial to ensure reliable outcomes. Several tests are commonly used to evaluate these aspects, such as the Hosmer-Lemeshow test to examine the goodness-of-fit, the Nagelkerke R^2 to assess the model's explanatory power, and the Omnibus Test of Model Coefficients to evaluate whether the logistic regression model significantly improves predictive accuracy compared to a baseline model without predictors.(19,20,23,24) A classification table further assesses predictive accuracy by comparing predicted and actual outcomes, reporting the percentage of correct classifications.(4)

NON-PARAMETRIC STATISTICS: TIME TO EVENT ANALYSIS

Survival analysis

Survival analysis is a statistical method used to analyse the time until an event of interest occurs, often referred to as time-to-event analysis.(25,26) Survival analysis is unique because it accommodates censored data—instances where the event has not occurred by the end of the study period. It is essential to distinguish between right and left censoring in survival analysis. The former occurs when the event of interest has not happened by the end of the study, and the latter happens when the event has already occurred before the study begins, but the exact timing is unclear.(25,27) Additional types of censoring that should be explored, applying to event timing and study designs, are described by Kartsonaki (25) and Emmert-Streib and Dehmer.(26) In survival analysis, the survival function and hazard function are critically important. The survival function represents the probability that participants will survive beyond a specific point in time.(27) The hazard function describes the rate at which the event of interest occurs over time, providing insights into the likelihood of the event happening at any given moment.(27) To interpret the survival analysis results, survival curves such as Kaplan-Meier plots are helpful to visualise survival probabilities over time, and the statistics used (log-rank, Breslow, and Tarone-Ware tests) compare survival distributions between different groups.(28) The log-rank test compares survival distributions between groups. It gives equal weight to events occurring at all time points.(28) The Breslow test, also known as the generalized Wilcoxon test, compares survival curves between groups but gives more weight to early events.(28) The Tarone-Ware test is like the log-rank and Breslow tests but gives intermediate weight to events occurring early and late in the study.(28)

Cox proportional regression

Other variables may influence survival times, and regression models for survival data can be used to assess the influence of covariates on survival times.(25) The Cox proportional hazard model (CPHM) is a semiparametric regression model to assess the association between multiple covariates simultaneously.(25–27) While there are no set limits on the number of covariates that can be elected, the aims and objectives of the study must be considered. The CPHM is relative in that it assumes that the influence of covariates has an additional effect on the hazard ratio, which implies that, in interpretation, there will be a corresponding increase in the hazard for every one-unit increase in the covariates.(25)

In Cox proportional hazards regression, time represents the duration until an event occurs, serving as the dependent variable. The group is a categorical variable that allows comparison of different categories (e.g., treatment vs. control), estimating hazard ratios to understand relative risks. An event is a binary indicator showing if the event happened (1) or was censored (0). This allows the model to differentiate between completed and censored observations and assess predictors' effects on the time to event. A negative regression coefficient suggests that as the predictor variable increases, there is a corresponding reduction in the hazard, indicating a lower likelihood of the event occurring. Conversely, a positive coefficient indicates reduced survival times until the event.

CONCLUSION

Parametric and non-parametric statistical techniques are versatile tools that can be applied across various quantitative studies in health sciences research. The choice between these techniques depends on the research questions, the nature of the data, and the underlying assumptions of the statistical methods. Selecting the most appropriate technique ensures that the analysis accurately reflects the data and effectively answers the research questions posed.

REFERENCES

1. Orcan F. Parametric or non-parametric: skewness to test normality for mean comparison. Int J Assess Tools Educ. 2020; 7(2):255–265
2. Sedgwick P. A comparison of parametric and non-parametric statistical tests. BMJ. 2015; 350(1):h2053.
3. Ross A, Willson VL. One-sample t-test. In: Ross A, Willson VL, eds. Basic and advanced statistical tests: writing results sections and creating tables and figures. Boston, MA: Sense Publishers; 2017:9–12. doi:10.1007/978-94-6351-086-8_2
4. Allen P, Bennett K, Heritage B. SPSS statistics version 22: a practical guide. 3rd ed. Melbourne: Cengage Learning Australia; 2014.

5. Park HM. Comparing group means: t-tests and one-way ANOVA using STATA, SAS, R, and SPSS. Working Paper. Published online 2009. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University, https://scholarworks.iu.edu/dspace/handle/2022/19735
6. Ghorbani H. Mahalanobis distance and its application for detecting multivariate outliers. Facta Univ Ser Math Inform. 2019; 34(3):583–595.
7. Yong AG, Pearce S. A beginner's guide to factor analysis: focusing on exploratory factor analysis. Tutor Quant Methods Psychol. 2013; 9(2):79–94.
8. Fitzner K. Reliability and validity: a quick review. Diabetes Educ. 2007; 33(5):775–780.
9. Hajjar STE. Statistical analysis: internal-consistency reliability and construct validity. Int J Quant Qual Res Methods. 2018; 6(1):46–57.
10. MacFarland TW, Yates JM. Mann–Whitney U test. In: MacFarland TW, Yates JM, eds. Introduction to nonparametric statistics for the biological sciences using R. Cham: Springer International Publishing; 2016:103–132. doi:10.1007/978-3-319-30634-6_4
11. Guo S, Zhong S, Zhang A. Privacy-preserving Kruskal–Wallis test. Comput Methods Programs Biomed. 2013; 112(1):135–145.
12. Ostertagová E, Ostertag O, Kováč J. Methodology and application of the Kruskal–Wallis test. Appl Mech Mater. 2014; 611:115–120.
13. Sherwani RAK, Shakeel H, Awan WB, Faheem M, Aslam M. Analysis of COVID-19 data using neutrosophic Kruskal–Wallis H test. BMC Med Res Methodol. 2021; 21(1):1–7.
14. Warner P. Quantifying association in ordinal data. J Fam Plann Reprod Health Care. 2010; 36(2):83–85.
15. Franke TM, Ho T, Christie CA. The chi-square test often used and more often misinterpreted. Am J Eval. 2012; 33(3):448–458.
16. Hauke J, Kossowski T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. Quaest Geogr. 2011; 30(2):87–93.
17. Boateng EY, Abaye DA. A review of the logistic regression model with emphasis on medical research. J Data Anal Inf Process. 2019; 7(4):190–207.
18. Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb) 2014; 24(1):12–18.
19. Giovazolias T, Leontopoulou S, Triliva S. Assessment of Greek university students' counselling needs and attitudes: an exploratory study. Int J Adv Couns. 2010; 32(2):101–116.
20. Lai X, Liu L. A simple test procedure in standardizing the power of Hosmer–Lemeshow test in large data sets. J Stat Comput Simul. 2018; 88(13):2463–2472.
21. Nattino G, Pennell ML, Lemeshow S. Assessing the goodness of fit of logistic regression models in large samples: a modification of the Hosmer–Lemeshow test. Biometrics. 2020; 76(2):549–560.
22. Yu W, Xu W, Zhu L. A modified Hosmer–Lemeshow test for large data sets. Commun Stat – Theory Methods. 2017; 46(23):11813–11825.
23. Rachunek-Medved K, Illg C, Einzmann A, et al. Postoperative scaphoid alignment, smoking, and avascular necrosis mostly influence union rate after scaphoid reconstruction: results from a retrospective single-center study involving 370 patients. J Plast Reconstr Aesthet Surg. 2023; 87:430–439.
24. Fernandes AAT, Figueiredo Filho DB, Rocha ECD, Nascimento WDS. Read this paper if you want to learn logistic regression. Rev Sociol Politica. 2020; 28(74):006. doi:10.1590/1678-987320287406en
25. Kartsonaki C. Survival analysis. Diagn Histopathol. 2016; 22(7):263–270.
26. Emmert-Streib F, Dehmer M. Introduction to survival analysis in practice. Mach Learn Knowl Extr. 2019; 1(3):1013–1038.
27. Schober P, Vetter TR. Survival analysis and interpretation of time-to-event data: the tortoise and the hare. Anesth Analg. 2018; 127(3):792.
28. Miot HA. Survival analysis in clinical and experimental studies. J Vasc Bras. 2017; 16:267–269.