

Review Article

Parametric and Nonparametric Tests: A Brief Review

Banda Gerald*, Tailoka Frank Patson

Itimpi Campus, Mukuba University, Kitwe, Zambia

Email addresses:

gerald1banda@gmail.com (B. Gerald), tailokaf@cbru.ac.zm (T. F. Patson)

*Corresponding author

To cite this article:

Banda Gerald, Tailoka Frank Patson. Parametric and Nonparametric Tests: A Brief Review. *International Journal of Statistical Distributions and Applications*. Vol. 7, No. 3, 2021, pp. 78-82. doi: 10.11648/j.ijstd.20210703.12

Received: April 2, 2021; **Accepted:** April 15, 2021; **Published:** October 28, 2021

Abstract: Context: Statistics is the cornerstone of markets, businesses, policy makers and other sectors that require analysis and interpretation of data. From generation-to-generation, statistics has proved useful in everyday life, not only that it helps improving the quality of life through counting and record keeping, but it also allows people to predict the future events and to make their own analysis. Before making a conclusion, data should be collected, analysed and interpreted. Evidence Acquisition: In this study, the paper reviewed parametric and nonparametric tests. Researchers sampled some articles where parametric and nonparametric tests were used without considering assumptions. Results: In this study, researchers provided a review of parametric tests; namely, independent sample t-test and dependent sample t-test, and nonparametric tests; namely, Mann-Whitney U test and Wilcoxon signed-rank test. The formulae for calculating parametric and nonparametric tests have been provided in the study. Procedures on how to conduct Mann-Whitney U test and Wilcoxon signed-rank test in SPSS have been written in this article. Test of normality has been discussed in brief as a key component in analysing parametric and nonparametric tests. Conclusions: Most of the studies that have been carried out have not been considering assumptions when analysing data using either parametric tests or nonparametric tests. This study looked at parametric and nonparametric tests. In parametric tests, the paper looked at independent and dependent sample t-test, while in nonparametric test, the paper looked at Mann-Whitney U test and Wilcoxon signed-rank test.

Keywords: Parametric, Nonparametric, Test of Normality, Statistics

1. Introduction

Markets, businesses, policy makers and other sectors require Mathematics and statistics, in particular. From generation-to-generation, statistics has proved useful in everyday life, not only that it helps improving the quality of life through counting and record keeping, but it also allows people to predict the future events and to make their own analysis. Before making a conclusion, data should be collected, analysed and interpreted. These procedures should be followed in order to make a concrete conclusion. In 2019, early December, China was hit by Corona Virus. Using statistics and looking at the number of people from different countries world over doing business with China, British Broadcast Corporation (BBC) predicted that Corona Virus would cause problems worldwide. Within three months, the whole world was affected by Corona Virus, and thousands of

people were dying every day. This is just a simple example of the importance of statistics. When we talk about statistics, we are looking at the collection of data, which can either be primary or secondary data, analysis and interpretation of data. Different methods of collecting and analyzing data should be known before any statistical approach is used. According to [1], there are two sets of statistical tests for comparing means of the population and these are: Parametric tests and non-parametric tests. Parametric and non-parametric tests focus on coming up with the conclusion about the population. According [2], there are a lot of techniques used to test the significance differences between groups. Here, we can look at the significance differences between two groups that are either coming from the normal population or not. It can also be significance differences between population mean and the sample mean; that is, the performance of the population based on the sample mean. [2] Identified one important difference between parametric and non-parametric test. According to [3],

parametric test focuses on the sample coming from the normal population, while in non-parametric test, the sample does not come from the normal population. [3] indicated that the term 'parametric' comes from the characteristic of population. [3], further explained that non-parametric test can only be used when the assumptions of parametric test have been violated. According to [4], non-parametric test has two assumptions: The first assumption is that sample from the population should be picked at random and the second assumption is observations should be independent. Nonparametric test is mainly about signs and ranks [5]

2. Problem Statement

Test of normality is the key component of determining which test to use when analyzing data using either parametric or nonparametric tests. Test of normality helps researchers to choose either to use parametric or non-parametric test. When a researcher uses, for example, independent sample t-test instead of Man -Whitney U test, the researcher will commit either type I or type II error. Type I error is the rejection of the null hypothesis when in actual sense it is true. Type II error is the acceptance of the null hypothesis when it is false. Researchers have observed that most of the articles published use parametric tests to analyse data without giving reasons. Previous study for example, (6) used parametric tests in his analysis. In this study, parametric test was used without giving reasons. The reason of not using normality test has consequences on determining which one to use between parametric and no-parametric tests more especially on t-tests. Parametric tests are the type of tests that are used when variables have a distribution that fit a certain criterion. For example, if the distribution is symmetric [7]. Non-parametric test is a test that assumes that data under investigation is not coming from the normal population. The study looked at test of normality, parametric and non-parametric tests.

3. Discussion

3.1. Test of Normality

The word normal in statistics is used to describe the curve which is bell shaped or symmetric. According to [2], normal is used to describe bell-shaped curve which has the greatest frequency of scores at the center with smaller frequencies towards the extremes. Checking the normality assumption is key in deciding on what to use between parametric and nonparametric test [8]. Test of normality helps in determining either to use parametric test or non-parametric test when analysing data. The most used tests for determining whether the data is coming from the normal distribution or not is Kolmogorov-Smirnov and Shapiro-Wilk [9]. One important method for normality assumption diagnostics is Shapiro-Wilk test. It is one of the most popular tests for normality, which has good properties of power and it is based on correlation within given observations and associated normal scores [10]. Shapiro-Wilk test works better when the sample size is less than fifty, while Kolmogorov-Smirnov works better when the

sample size is more than fifty. The null hypothesis (H_0), in both the Kolmogorov-Smirnov and Shapiro-Wilk test, states that data come from the normal population or data follow the normal distribution. In SPSS, after coding data in variable view, we go to data view. In data view we go to: analyse, descriptive statistics, explore (scores become dependent list, factor is the independent variable), plots (stem-and-leaf, histogram, normality tests with tests), continue and finally ok. $P - value < \alpha = 1\% \text{ or } 5\% \text{ or } 10\%$ in Kolmogorov-Smirnov test or Shapiro-Wilk test, the null hypothesis is rejected. This implies that data is not normally distributed

3.2. Independent Sample t-test vs Mann-Whitney U Test

An independent sample t-test is the type of the test that compares the mean of two groups that are independent. Researchers are helped to come up with conclusion about the population based on the mean of the two groups. According to [11], an independent sample t-test compares the means of two groups and is only applied when data for each group is normally distributed and the two groups are independent. [11] further explained that we need to have scores in order for independent sample t-test to be used. According to [1], one of the formulae used to calculate an independent sample t-test is given below;

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum_{i=1}^n X_{i1}^2 - \frac{(\sum_{i=1}^n X_{i1})^2}{n_1} + \sum_{i=1}^n X_{i2}^2 - \frac{(\sum_{i=1}^n X_{i2})^2}{n_2}}{n_1 n_2 (1 - \frac{2}{n_1 + n_2})}}$$

[1] Wrote the procedures on how to analyse independent sample t-test in SPSS. Just like an independent sample t-test, Mann-Whitney U test is used to test the differences between two groups that are independent [2]. Mann-Whitney test is sometimes used for comparing the efficacy of two treatments in clinical trials. In most cases, it is presented as an alternative to an independent sample t-test when the data are not normally distributed [12]. The key assumption for using Mann-Whitney U test is the normality test. When the two groups are not coming from the normal population, then Mann-Whitney U test can be used to compare the two groups. In independent sample t-test, the null hypothesis states that the mean of the two groups is the same or is not significantly different ($H_0: \mu_1 = \mu_2$), while the null hypothesis for Mann-Whitney U test states that the two populations are equal. In most cases, Mann-Whitney U test is used for two-sided test. Decision rule in independent sample t-test states that reject the null hypothesis if the test statistic is greater than the critical value, while decision rule in Mann-Whitney U test states that reject the null hypothesis if the test statistic is less than the critical value. Critical value is the value we get from the table using degrees of freedom and level of confidence. Two-sided means that the direction for the alternative hypothesis (H_1) is not specified. The formula used in Mann Whitney U test is

$$U_1 = \frac{n_1(n_1 + 2n_2 + 1) - 2R_1}{2}$$

$$U_2 = \frac{n_2(2n_1 + n_2 + 1) - 2R_2}{2}$$

The formulae above are used to compute test statistic. A test statistic is the value that is computed from the sample. It is used to capture information in the sample data and decide between the null hypothesis (H_0) and alternative hypothesis (H_1). Null hypothesis is a method of statistical inference by which an experimental factor is tested against a hypothesis of no effect or no relationship based on a given observation [13]. The smallest value of U_1 and U_2 defines U where n_1 is the sample size 1 and n_2 is the sample size 2. R_1 and R_2 are the sum of the ranks for group 1 and group 2. The two groups should be ranked as one. Example from <https://spweb.bumc.bu.edu> and the example is on new approach to prenatal care. Table 1 shows data which was collected on prenatal care.

Table 1. Scores on Prenatal Care.

Usual Care	8	7	6	2	5	8	7	3
New Program	9	8	7	8	10	9	6	

Table 2. Ranked Scores on Prenatal Care.

Usual		8	7	6	2	5	8	7	3
Care	Rank 1	10.5	7	4.5	1	3	10.5	7	2
New		9	8	7	8	10	9	6	
Program	Rank 2	13.5	10.5	7	10.5	15	13.5	4.5	

$$4 \text{ and } 5 = \frac{4+5}{2} = 4.5, 6, 7 \text{ and } 8 = \frac{6+7+8}{3} = 7$$

$$9, 10, 11 \text{ and } 12 = \frac{9+10+11+12}{4} = 10.5, 13 \text{ and } 14 = \frac{13+14}{2} = 13.5$$

$$R_1 = 10.5 + 7 + 4.5 + 1 + 3 + 10.5 + 7 + 2 = 45.5$$

$$R_2 = 13.5 + 10.5 + 7 + 10.5 + 15 + 13.5 + 4.5 = 74.5$$

$$U_1 = \frac{8(8+2 \times 7+1) - 2 \times 45.5}{2} = 46.5$$

$$U_2 = \frac{7(2 \times 8+7+1) - 2 \times 74.5}{2} = 9.5$$

From the calculations above, the value of U is $U_2 = 9.5$ since U_2 is less than $U_1 = 46.5$. In this case our test statistic is $U_2 = 9.5$. $n_1 = 8$ and $n_2 = 7$, using Mann-Whitney U test table for double tail at 5% level of significance, the critical value is 10. Meaning that the null hypothesis is rejected since $U = 9.5 < 10$. The value of U can also be calculated using Statistical Package for the Social Sciences (SPSS) as follows: After coding data in variable view, we go to data view. Within data view now we have: Analyse-Nonparametric-Legacy dialogs-2 independent samples (where test variable is results or scores and grouping variable are the two independent variables, groups are the ones already coded in variable view then tick on Mann-Whitney U box and finally ok). The computed value is compared with the value from Mann-Whitney U table 3. Dependent Sample t-test vs Wilcoxon Signed Rank Test.

Just like an independent sample t-test, dependent sample t-test is also a parametric test and it assumes that data are normally distributed; meaning that test of normality is key when using dependent sample t-test. The term dependent means that one group of the sample is used to determine the

other group. According to [14], the other name of dependent sample t-test is matched-pairs t-test. In research, matched samples t-test is used to compare two groups of scores and their means in which the participants in one group are somehow meaningfully related to the participants in the other group (1). When testing the reliability of the instrument, dependent sample t-test can also be used. Reliability means that the instrument should be consistent in terms of the results. The null hypothesis (H_0) for dependent sample t-test states that the mean of the two groups is the same or is not significantly different ($H_0: \mu_1 = \mu_2$). Procedures on how to analyse dependent sample t-test in SPSS were written by [1]. According to [1], the following formula can be used to calculate dependent sample t-test.

$$t = \frac{1}{\sqrt{\frac{\sum_{i=1}^n \left(\frac{d_i - 1}{d}\right)^2}{n(n-1)}}$$

When you want to test differences between two conditions and different participants have been used in each condition, then you have two choices: You can either use independent sample t-test and dependent sample test or Mann-Whitney Test and Wilcoxon test. Independent sample t-test and dependent sample test are parametric tests, while Mann-Whitney Test and Wilcoxon signed-rank test are nonparametric tests. According to [15] just like dependent t-test, Wilcoxon signed-rank test works in a fairly similar way in that it is based on the differences between scores in the two conditions you are comparing. Differences between the two groups are ranked and the sign of the difference (positive or negative) is assigned to the rank.

The following formula can be used to calculate Wilcoxon Signed Rank Test

$$W_+ = \sum_{i=1}^n R_{i+} \text{ and } W_- = \sum_{i=1}^n R_{i-}$$

In this study,

R_{i+} are the positive ranked scores and R_{i-} are the negative ranked scores.

$R_i = d_i$ and d_i is the result of subtracting first result and the second result of each individual participant in the study.

$$d_i = x_{i1} - x_{i2} \text{ where } i = 1, 2, 3, \dots, n$$

x_{i1} is the first results of each participant in the study and x_{i2} is the second results of each participant in the study.

n is the sample size of the study.

After calculating d_i , scores are arranged in order starting with the smallest and ending with the biggest. Negative numbers become positive when arranging in order and signs are only considered after the scores have been arranged in order. The smallest value of W_1 and W_2 defines W . The value of W can also be calculated using SPSS as follows: Under variable view, write before in first column under the first row and then write after in first column under the second row. We go to data view. Within data view now we have: Analyse-Nonparametric-Legacy dialogs-2 related samples (where after is 1 and 2 is before, then Wilcoxon and finally ok). The computed value is compared with the critical value from

Wilcoxon Signed- Ranks Table. For easily understanding of how to calculate Wilcoxon Signed- Rank Test, let's look at example from <https://spweb.bumc.bu.edu> and the example is on the effectiveness of a new drug designed to reduce

repetitive behaviors in children affected with autism. Table 3 is about data which was collected on the effectiveness of a new drug designed to reduce repetitive behaviors in children affected with autism.

Table 3. Scores on Effectiveness of a New Drug.

Child	1	2	3	4	5	6	7	8
Before Treatment	85	70	40	65	80	75	55	20
After 1 week of Treatment	75	50	50	40	20	65	40	25
d_i	10	20	-10	25	60	10	15	-5

Table 4. Arranged d_{is} .

Child	1	2	3	4	5	6	7	8
Before Treatment	85	70	40	65	80	75	55	20
After 1 week of Treatment	75	50	50	40	20	65	40	25
d_i	10	20	-10	25	60	10	15	-5
Arranged	-5	10	-10	10	15	20	25	60

Table 5. Getting R_i without considering signs.

Child	1	2	3	4	5	6	7	8
Before Treatment	85	70	40	65	80	75	55	20
After 1 week of Treatment	75	50	50	40	20	65	40	25
d_i	10	20	-10	25	60	10	15	-5
Arranged	-5	10	-10	10	15	20	25	60
R_i without considering signs	1	3	3	3	5	6	7	8

$$2,3 \text{ and } 4 = \frac{2+3+4}{3} = 3$$

Table 6. R_i with Signs.

Child	1	2	3	4	5	6	7	8
Before Treatment	85	70	40	65	80	75	55	20
After 1 week of Treatment	75	50	50	40	20	65	40	25
d_i	10	20	-10	25	60	10	15	-5
Arranged	-5	10	-10	10	15	20	25	60
R_i without considering signs	1	3	3	3	5	6	7	8
R_i with signs	-1	3	-3	3	5	6	7	8

This means $R_{i+} = 3,3,5,6,7,8$ and $R_{i-} = 1,3$ $W_+ = \sum_{i=1}^n R_{i+} = \sum_{i=1}^6 R_{i+} = 3 + 3 + 5 + 6 + 7 + 8 = 32$, $W_- = \sum_{i=1}^2 R_{i-} = 1 + 3 = 4$

This implies that the value of W is 4. This is the value which is compared with the critical value.

4. Conclusion

In brief, so many studies have been conducted on parametric and non-parametric tests. However, most of the studies, more especially studies on comparing means, parametric tests have been used without considering assumptions used when analysing data using parametric tests. It is important to test for normality before using either parametric or nonparametric tests to avoid committing either type I or type II error. According to [16], type I error (false-positive) occurred when the researcher decides to reject the correct null hypothesis in the population; a type II error (false-negative) occurred when the researcher fail to reject the false null hypothesis in the population.

References

- [1] Banda, G (2018). A Brief Review of Independent, Dependent and One Sample t-test. *International Journal of Applied Mathematics and Theoretical Physi.* 4 (2): 50-54. Doi: 10.11648/j.ijamtp.20180402.13.
- [2] Sarah, H and Joseph R. D (2018). Parametric and Nonparametric Tests in Spine Research: Why Do They Matter. *Global Spine Journal.* 8 (6): 652-65. Doi: 10.1177/2192568218782679Kolawole, B. E. (2008). Effects of competitive and cooperative learning strategies on academic performance of Nigerian students in mathematics. *Educational Research and Review*, 34-35.
- [3] Jeffrey, E. H and Carol, B (2008). Publishing Nutrition Research: A Review of Nonparametric Methods, Part 3. *Journal of the Academy Nutrition and Dietetics.* 108 (9): 1488-1496. Doi: 10.1016/j.jada.2008.06.42.
- [4] Verm, J. P and Abdel-Salam, G (2019). *Testing Statistical Assumptions in Research.* Hoboken, New Jersey. 141-174 p. Doi: 10.1002/9781119528388.

- [5] Francis, S. N (2016). Nonparametric Statistical Tests for the Continuous Data. The Basic Concept and the Practical Use. *Korean Journal of Anesthesiology*. 69 (1): 8-14. Doi: 10.4097/kjae.
- [6] Umar, I. (2017). Assessing the Effect of Cooperative Learning on Financial Accounting Achievement among Secondary School Students. *International Journal of Instruction*. 10 (3) 31-46
- [7] Gail, F, Dawson M. D and FAAEP, M, S (2016). *Easy Interpretation of Biostatistics*. Michigan. 63-74. Doi: 10.1016/B978-1-4160-3142-0.50014-6.
- [8] Fatih, O (2020). Parametric or Non-parametric: Skewness to Test Normality for Mean Comparison. *International Journal of Assessment Tools in Education*. 7 (2): 255-265.
- [9] Asghar, G., & Saleh, Z. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology Metabolism*. 10 (2): 486-489. Doi: 10.5812/ijem.3505.
- [10] Keya, D. R., & Rahmatullah, I. A (2016). A Brief Review of Tests for Normality. *American Journal of Theoretical and Applied Statistics*. 2016; 5 (1): *Doing A Quantitative Research in Education with SPSS*. London: Sage Publications 5-12. Doi: 10.11648/j.ajtas.20160501.12.
- [11] Amanda, R and Victor, L. W (2017). *Basic and Advanced Statistical Tests*. Springer. 13-16 p. Doi: 10.1007/978-94-6351-086-8_3.
- [12] Anna. H (2001). Mann-Whitney Test in not Just a Test of Medians: Differences in Spread can be Important. 323 (7309): 391–393. Doi: 10.1136/bmj.323.7309.391. [PMID: 1150943].
- [13] Cyril, P (2016). Null Hypothesis Significance Testing: A Short Tutorial. 621 (3). Doi: 10.12688/fl00research.6963.3. [PMCID: PMC5635437]. [PMID: 29067159].
- [14] Younis, S (2015). The Bread and Butter of Statistical Analysis “t-test”. Uses and Misuses. *Pakistan Journal of Medical Sciences*. 31 (6): 1558-1559. Doi: 10.12669/pjms.316.8984. [PMCID: PMC4744321]. [PMID: 26870136].
- [15] Elise, W and Jonathan B (2002). Statistics Review 6: Nonparametric Methods. *BioMed Central*. 6 (6): 509-513. Doi: 10.1186/cc1820. [PMCID: PMC153434]. [PMID: 12493072].
- [16] Amitav, B, Chitnis, U. B and Chaudhury, S (2009). Hypothesis Testing, Type I and Type II Errors. *Induetria; I Psychiatry Journal*. 18 (2): 127-131. Doi: 10.4103/0972-6748.62274. [PMCID: PMC2996198]. [PMID: 21180491].