Parametric and Non- Parametric Analysis: A Review

Dr Manpreet Kaur*
Dr Arshveer Kaur**
Dr Savneet Kaur***

Abstract

This paper explores and compares parametric and non-parametric tests, including commonly used tests such as the Student t-test, Z-test, ANOVA, Correlation, and Regression for parametric analysis, and Spearman's Rank Correlation, Mann-Whitney U test, Wilcoxon Rank Sum test, Logistic Regression and Chi-square for non-parametric analysis. It concludes that when the assumptions for parametric tests are not met, non-parametric tests should be employed, as the results from parametric tests can be misleading. However, when the assumptions are satisfied, parametric tests tend to be more efficient.

Keywords: Parametric test, Non-Parametric test, Student t-test, Z-test, ANOVA, Correlation, Regression, Mann- Whitney test, Wilcoxon Rank Sum test, Chi-square, Logistic Regression

Introduction

Researchers are always concerned about finding results using statistical tests based on sample data that are applicable to whole population, be it very large or very narrow. However, it is very difficult to collect data from each and every unit in the population due to time and money constraint. Besides, sometimes it is not even feasible to collect data from entire population, for example, pathologists draw blood sample to conduct blood test. As a result, data is collected from a sample which is subset of population. Statistical tests are typically conducted on a sample, and

* Assistant Professor, PG Department of Commerce, Sri Guru Gobind Singh College, Sector 26, Chandigarh
** Assistant Professor, PG Department of Commerce, Guru Gobind Singh College for Women, Sector 26, Chandigarh
*** Assistant Professor, PG Department of Commerce, Guru Gobind Singh College for Women, Sector 26, Chandigarh

inferences are made to the population as a whole. The validity of these inferences, however, depends on the choice of statistical methods, particularly the assumptions underlying these methods.

Objectives

In this paper, an attempt has been made to understand the basic assumptions on which parametric tests are based. Further, reviews of both parametric and non-parametric tests have been done to better understand their applicability. It will ultimately enable the researchers to suitably apply them to attain dependable outcome about the population.

Parametric Tests

While performing statistical analysis on samples, researchers tend to perform various parametric tests (like regression, ANOVA, t- test, Z- test) on data set. Parametric tests are based on assumptions about distribution of data. Specifically, they require data to meet the following criteria (Field, 2005):

Normality: The data should come from normally distributed population. Most researchers use histograms to check normality of their data, which is subjective method. The Kolmogorov-Smirnov test, the Shapiro-Wilk test and Q-Q plots are some objective tests to decide whether or not a distribution is normal.

Homogeneity of variance: It specifies that the variances across the data should be equal. Levene's test is one of the widely used tests to check this assumption. A non-significant result from Levene’s test indicates that the assumption of homogeneity of variances is likely satisfied.

Interval or Ratio data: The data should be measured on an interval or ratio scale, where the distance between points is consistent. For instance, time is interval data as difference between 3 pm and 5 pm is the same as 6 pm and 8 pm.

Independence: It means that data collected from sample units should be independent of each other.

If the above listed assumptions about the data are met, researchers widely use the following parametric tests:

T-test: Developed by William Sealy Gosset (1908), also popularly known as Student's t-test. This test is applicable if size of the sample is small (usually 30 or less) and variance of the population is not known (Gupta and Gupta, 2010). The t-test is used to determine if there is significant difference between sample mean and population mean (one sample t-test). To examine whether means of two populations are equal or not, two sample t-test is applied (Snedecor and Cochran, 1989). Further two sample t-test is of two types, namely, independent t-test, which is applied when the two samples are randomly and independently drawn from two normally distributed populations and the paired t-test is used when two samples are dependent (Kim, 2015).

Z-test: It is very similar to the Student's t-test as both the tests facilitate in determining the significance of data. It is used to verify whether means of two populations are equal or not, provided the populations are normally distributed. Main difference between Z-test and t-test is that the former is applicable when sample is of large size, and also the variance of the population is known. Besides, t-test is relevant when the size of the sample is small and also variance of the population is not known. The z-test is infrequently used due to difficulty in determining variance of the population.

One way ANOVA: The variation in the means of two independent groups satisfying both the assumptions of normality and homogeneity of variance is significant or not is determined by Student's t-test. But, it is not suitable to compare the means of three or more independent groups. In such circumstances, one-way analysis of variance (ANOVA) is best suited method. In case of comparing the means of three or more groups, the null hypothesis would be "the population means of three groups are all the same," and alternative hypothesis would be "at least one of the population means of three groups is different." This method generates F-statistic or F-ratio, therefore it is also known as F test. Unlike the t-test, ANOVA avoids the augment in Type I error that would otherwise arise from performing multiple t-tests. According to Kim (2017) ANOVA is regarded as omnibus test as it does not specify the groups that are affected rather it provides for total experimental effect.

Repeated measures ANOVA: 'Repeated measure' means when same subjects are exposed to multiple conditions or are measured multiple times in an experiment (Field, 2005). Specifically, it tests whether there are statistically significant differences between the means of three or more groups when those groups are related (same subjects participate in each group). Basically, this method is an extension of the paired t-test that compares the means of two dependent groups. Where, Repeated measures ANOVA compares the means of three or more dependent groups. This method is also known as within-subjects ANOVA or ANOVA for paired samples (Mishra *et al*, 2019).

Pearson's correlation coefficient: A statistical measure represented by symbol "r" that quantifies the strength and direction of linear relationship amid two variables, with values ranging between +1 and -1. A perfect positive linear relationship is indicated by +1, where -1 signifies a perfect negative linear relationship. Besides, 0 indicates no linear relationship amid two variables. However, it does not explain cause and effect relationship between variables and does not measure non linear relationship amid variables. The probable error of r helps in testing the significance of the correlation. If r is less than the probable error, it suggests that the correlation is insignificant. Where, if r is more than six times the probable error, it indicates that the correlation is significant and correlation exists between the variables (Gupta and Gupta, 2010).

Regression: It was first proposed by Sir Francis Galton in 1877. It is a statistical method to predict the values of one dependent variable based on the values of one or more independent variable. Predicting the value of a dependent variable from one independent variable is known as simple linear regression. When value of one dependent variable is predicted from several independent variables, it is known as multiple regression (Kumari and Yadav, 2018). Besides, regression assists in quantifying the degree of relationship between dependent and independent variables by computing the coefficient of determination (R^2). It calculates the strength of association between variables and determines how much variance in dependent variable is explained by independent variable (Gupta and Gupta, 2010).

Non-Parametric Tests

Researchers apply parametric tests without checking data's distributional properties which makes the results highly suspicious. According to Yim *et al* (2010), performing parametric analysis to non-parametric (distribution free) data is common statistical mistake found in research papers. Parametric tests are performed on the basis of various assumptions and when these are not met, the results of the analysis can be false or completely invalid. In that case, special kind of statistical tests should be used, which are identified as non-parametric tests (Jeffrey and Carol, 2008). They are also recognized as distribution-free tests as no assumption is made about distribution of data. Non-parametric analysis can be applied on nominal or ordinal data (Walsh, 1962). Besides, they can also be used on interval or ratio data that do not meet normality assumption. Non-parametric tests are mostly based on signs and ranks (Francis, 2016). Further, according to Field (2005) majority of non-parametric tests first rank the data and then perform analysis on the ranked data instead on actual data

Non-parametric tests widely used by researchers include the following:

Spearman's rank correlation coefficient: A non-parametric method used when data does not meet the assumption of normality as required by Pearson's correlation. This method is useful when quantitative measure of certain factors like judgement of female beauty cannot be measured, but they can be arranged in order on basis of her rank in the group. Spearman's correlation measures the monotonic relationship between two variables. It means that as one variable increases, the other variable tends to increase (or decrease), but not necessarily in a straight line (Field, 2005) where Pearson's correlation measures linear relationship amid variables. Like Pearson's correlation, Spearman's correlation also ranges from -1 to +1. Value of +1 indicates a perfect positive monotonic relationship (as one variable increases, the other increases). On the other hand, -1 indicates a perfect negative monotonic relationship (as one variable increases, the other decreases). Besides, 0 indicates no monotonic relationship between variables.

Kendall's Tau: It is a non-parametric statistical method proposed by Maurice Kendall in 1938. This method is an alternate to parametric correlation test and calculates the strength of the association between two ordinal variables on the basis of their ranks. It is suitable when the sample size is small having many tied ranks. Kendall's Tau statistic is regarded as superior estimate of correlation as compared to Spearman's rank correlation (Howell, 2007). Therefore, it provides more reliable generalizations about the population.

Mann- Whitney U test and Wilcoxon Rank Sum test: According to Field (2005) both the non-parametric tests are equal to and applied as an alternative to independent t-test where data does not fulfill the assumptions of normality and homogeneity of variance. They are used to test significance of difference between two groups where different subjects participated in each group and they compare ranks of both the groups than their actual data. The Mann-Whitney U test checks whether the ranks of one group are higher or lower than the ranks of another group where, Wilcoxon Rank-Sum test compares the sum of ranks for each group. Both tests are mathematically equivalent, but the Mann- Whitney U test is more widely used.

Wilcoxon Signed-rank test: This method is equal to dependent (paired) t-test which assumes that differences between the paired scores are normally distributed. However, when this assumption is not met, Wilcoxon Signed-rank test is applied. This non-parametric test is used to compare two related (paired) groups that provide two sets of data scores for comparison from same subjects. Dalgaard (2008) also stated that this statistical method is most suitable to check whether difference between before and after treatment is significant or not provided normality assumption is not satisfied.

Kruskal-Wallis test: Non-parametric method developed by Kruskal and Wallis in 1952, is based on ranked data and is equivalent to one way independent ANOVA. Boslaugh & Watters (2008) stated that ANOVA is a parametric statistical method used to check whether there exist any significant differences between three or more independent groups meeting the assumption of normality and equal variance. But, if these assumptions are violated Kruskal-Wallis test is an alternate. According to Kanji (2006) Kruskal-Wallis test identifies if there exists any significant difference between more than two independent groups. This method is an extension of Mann-Whitney U test as there are more than two groups (Sheskin, 2011). Similar to other non-parametric tests, this method also uses ranks of data to analyze it and draw conclusions

Friedman's ANOVA: It is a non-parametric test proposed by Milton Friedman and is an alternate to repeated measures ANOVA test. Hoffman (2019) stated that when the assumption of normality is not met or data is of ordinal nature, distribution free Friedman's ANOVA should be employed. According to Field (2005) and Gaddis and Gaddis (1990), this method is used to check whether there exists significant difference between three or more related groups based on ranked data rather than on actual figures.

Non-parametric logistic regression: This method is counterpart of parametric regression method and is used to determine the relationship between one or more independent variables (categorical or continuous) and one dichotomous (0, 1) dependent variable (Peng *et al*, 2002). This statistical method can be applied to anticipate the probability of various binary events like forecasting whether team will win match or not; candidate will win elections or not; debtor will repay loan or not.

Chi-square test: The test was proposed by Karl Pearson in 1990 and is based on chi-square distribution. It is one of the simplest tests and makes no assumption about the population being sampled (Gupta and Gupta, 2010). Chi-square test verifies whether there exists any significant relationship between two attributes or not which are measured at nominal or ordinal level. According to Kothari (2010), chi-square does not compute strength of relationship between attributes; rather it is a tool to evaluate statistical significance of relationship between them.

Conclusion

Selection of tests should primarily be based on distribution of data i.e. whether it is normally distributed or not and other assumptions of parametric tests as discussed in this paper. However, while conducting statistical analysis, many researchers neglect these assumptions and presume that sample data is normally distributed (based on central limit theorem) without checking whether population distribution is normally distributed or not. In such circumstances, researchers most likely commit type II error also known as false negatives which can have severe consequences in many areas such as medicine, criminal justice, product testing. Type II error implies that statistical test failed to identify relationship between variables or effect or difference in population. Thus, if assumptions of parametric analysis are violated, researchers should employ their alternative non-parametric test for comparing two or more groups as summarized in Table 1.

Table 1: Parametric and alternate Non-parametric tests for comparing 2 or more groups

	Two independent groups	Two dependent groups	More than two independent groups	More than two dependent groups

Parametric Test	1. Independent t-test (if sample size is small) 2. Z-test (if sample size is large)	Paired t-test	One-way analysis of variance (ANOVA)	Repeated measures ANOVA
Non-parametric test	Mann-Whitney U test and Wilcoxon Rank Sum test	Wilcoxon Signed-rank test	Kruskal-Wallis test	Friedman's ANOVA

Further, Table 2 indicates summarized results of parametric and their counterpart non-parametric tests used for computing correlation between variables and predicting value of one variable from another.

Table 2: Parametric and alternate Non-parametric tests for determining association between variables and forecasting value of one variable from another

	Association between variables	Predicting value of one variable (dependent) from another (independent)
Parametric Test	Pearson's correlation coefficient	1. Simple linear regression (one dependent and one independent variable) 2. Multiple regression (one dependent and more than one independent variable)
Non-parametric test	1. Spearman's rank correlation 2. Kendall's Tau	Non-parametric logistic Regression

Many investigators do not prefer to use non-parametric tests as they consider them less powerful. But Bradley (1968); Hunter and May (1993); and Field (2005) have proved that non-parametric tests are as good as parametric tests when the assumptions of later are not satisfied.

Therefore, it can be concluded that parametric tests give reliable results and can be generalized provided its assumptions are satisfied. Therefore, researchers should be vigilant and check the assumptions of statistical procedures before employing them so as to reach consistent results about the population.

References

Boslaugh S and Watters PA (2008). *Statistics in a Nutshell*. O'Reilly Media, Inc.

Bradley JV (1968). *Distribution- Free Statistical Tests.* Prentice Hall, Englewood Cliffs, NJ.

Dalgaard P (2008). *Introductory Statistics with R* (2^{nd} ed). Springer

Field A (2005). *Discovering Statistics Using SPSS*. Sage Publications

Francis SN (2016). Nonparametric Statistical Tests for the Continuous Data. The Basic Concept and the Practical Use. Korean Journal of Anesthesiology. 69 (1): 8-14.

Gaddis GM and Gaddis ML (1990). Introduction to Biostatistics: Part 5, Statistical Inference Techniques for Hypothesis Testing with Non-parametric Data. Annals of Emerging Medicine. 19(9): 1054-59.

Gupta SP and Gupta MP (2010). *Business Statistics*. Sultan Chand and Sons, New Delhi.

Hoffman JIE (2019). *Basic Biostatistics for Medical and Biomedical Practitioners* (2^{nd} ed). Academic Press

Howell DC (2007). *Statistical Methods for Psychology* (7^{th} ed). Cengage Wadsworth, Belmont, CA, USA.

Hunter M and May R (1993). Some Myths Concerning Parametric and Non Parametric Tests. Canadian Psychology. 34 (4): 384- 389.

Jeffrey EH and Carol B (2008). Publishing Nutrion Research: A Review of Nonparamentric Methods, Part 3. Journal of the Academy Nutrtion and Dietetics. 108 (9): 1488-1496.

Kanji G (2006). *100 Statistical Test.* Sage Publications Ltd, New Delhi

Kim TK (2015). T-test as a Parametric Statistic. Korean Journal of Anesthesiology. Dec 68 (6): 540-546.

Kim TK (2017). Understanding One-Way ANOVA Using Conceptual Figures. Korean Journal of Anesthesiology. 70 (1): 22-26.

Kothari CR (2010). *Research Methodology*. New Age International Publishers, Mumbai

Kumari K and Yadav S (2018). Linear Regression Analysis Study. Journal of the Practice of Cardiovascular Sciences. 4(1): 33-36.

Mishra P, Singh U, Pandey CM, Mishra P and Pandey G (2019). Application of Student's t-test, Analysis of Variance, and Covariance. Annals of Cardiac Anaesthesia, 22(4): 407- 411.

Peng J, Lee KL and Ingersoll G (2002). An Introduction to Logistic Regression Analysis and Reporting. Journal of Educational Research, 96(1): 3-14.

Sheskin, DJ (2011). *Handbook of Parametric and Nonparametric Statistical Procedures.* Chapman and Hall Books, New York.

Snedecor GW and Cochran WG (1989). *Statistical Methods* (8^{th}ed). Iowa State University Press, Ames.

Walsh JE (1962). *Handbook of Nonparametric Statistics*. D. Van Nostrand, Princeton, NJ.

Yim KH, Nahm FS, Han KA and Park, SY (2010). Analysis of Statistical Methods and Errors in the Articles Published in the Korean Journal of Pain. Korean Journal of Pain, 23: 35-41.