**Journal of Computer and Mathematical Sciences,** *Vol.6(6),336-342, June 2015* *ISSN 0976-5727 (Print)*
(An International Research Journal), *www.compmath-journal.org* *ISSN 2319-8133 (Online)*

# Comparative Analysis of Parametric and Non-Parametric Tests

**Amandeep Kaur and Robin Kumar**

Department of Mathematics,
GNDU, Amritsar, Punjab, INDIA.
Assistant Professor,
Department of Mathematics, A.S. College, Khanna, Punjab, INDIA.

(Received on: June 23, 2015)

**ABSTRACT**

This Paper is dedicated to the comparative analysis of Parametric and Non-Parametric hypothesis testing. Here we discuss some parametric tests such as Student t-test, Z-test, chi-square, ANOVA (Analysis of variance) and Non-Parametric tests such as Sign test, Wilcoxon Sign-Rank Test and Mann-Whitney Test. And after analysis we can say that if the assumptions do not meet correctly then the parametric tests do not gives the right conclusion. Whereas, in the same case the Non-parametric tests gives the right conclusions. However, if normality assumptions meet then the parametric tests are more efficient than the non-parametric tests.

**Keywords:** Parametric, Non-Parametric, Student t-test, Z-test, chi-square, ANOVA, Sign test, Wilcoxon Sign-Rank Test and Mann-Whitney Test.

## INTRODUCTION

A statistical test provides a mechanism for making quantitative decisions about a process or processes. The intent is to determine whether there is enough evidence to "reject" a conjecture or hypothesis about the process. The conjecture is called the null hypothesis. Not rejecting may be a good result if we want to continue to act as if we "believe" the null hypothesis is true. Or it may be a disappointing result, possibly indicating we may not yet have enough data to "prove" something by rejecting the null hypothesis.

Hypothesis tests also address the uncertainty of the sample estimate. However, instead of providing an interval, a hypothesis test attempts to refute a specific claim about a population parameter based on the sample data. Basically we have Two types of Tests based parameters i.e. Parametric and Non-Parametric. The hypothesis tests lie in the category of parametric tests when they assume the population follows some specific distribution such as normal distribution with a set of parameters. Nonparametric tests, on the other hand, are applied when certain assumptions cannot be made about the population. Rank or ordinal data usually require nonparametric analysis. Nonparametric tests are also referred as distribution-free methods. Since nonparametric tests make fewer assumptions, they are more robust than their corresponding parametric ones. Non-parametric models differ from parametric models in that the model structure is not specified a priori but is instead determined from data. The term non-parametric is not meant to imply that such models completely lack parameters but that the number and nature of the parameters are flexible and not fixed in advance.

## PARAMETRIC TEST

Parametric tests are more robust and for the most part require less data to make a stronger conclusion than nonparametric tests. However, to use a parametric test, 3 parameters of the data must be true or are assumed. First, the data need to be normally distributed, which means all data points must follow a bell-shaped curve without any data skewed above or below the mean. Ca-125 levels are an example of non-normally distributed data. In the general population, normal Ca-125 values range from 0 to 40. The median is 15, which leads to a skewed rather than a normal distribution. The data also need to have equal variance and have the same standard deviation. Finally, the data need to be continuous. Commonly used parametric tests are described below.

### A. Student t-Test

The Student t-test is probably the most widely used parametric test. It was developed by a statistician working at the Guinness brewery and is called the Student t-test because of proprietary rights. A single sample t-test is used to determine whether the mean of a sample is different from a known average. A 2-sample t-test is used to establish whether a difference occurs between the means of 2 similar data sets. The t-test uses the mean, standard deviation, and number of samples to calculate the test statistic. In a data set with a large number of samples, the critical value for the Student t-test is 1.96 for an alpha of 0.05, obtained from a t-test table. The calculation to determine the t-value is relatively simple, but it can be found easily on-line or in any elementary statistics book. As an example, given 1000 men measured for height in China and Japan, are the mean heights different? China's mean is 169.1 cm with a standard deviation of 6.21 cm, and Japan's mean height is 168.6 cm with a standard deviation of 5.7 6cm. The t-value is 1.88; therefore, the mean heights are not statistically different.

### B. The z-Test

The next test, which is very similar to the Student t-test, is the z-test. However, with the z-test, the variance of the standard population, rather than the standard deviation of the study groups, is used to obtain the z-test statistic. Using the z-chart, like the t-table, we see what percentage of the standard population is outside the mean of the sample population. If, like the t-test, greater than 95% of the standard population is on one side of the mean, the p-value is less than 0.05 and statistical significance is achieved. As some assumption of sample size exists in the calculation of the z-test, it should not be used if sample size is less than 30. If both the n and the standard deviation of both groups are known, a two sample t-test is best.

**C. Chi-square**

A chi-square test can be used to test if the variance of a population is equal to a specified value. This test can be either a two-sided test or a one-sided test. The two-sided version tests against the alternative that the true variance is either less than or greater than the specified value. The one-sided version only tests in one direction. The choice of a two-sided or one-sided test is determined by the problem. For example, if we are testing a new process, we may only be concerned if its variability is greater than the variability of the current process.

**D. ANOVA**

Analysis of Variance (ANOVA) is a statistical method used to test differences between two or more means. It may seem odd that the technique is called “Analysis of Variance” rather than “Analysis of Means.” As you will see, the name is appropriate because inferences about means are made by analyzing variance. Non-specific null hypothesis is sometimes called the omnibus null hypothesis. When the omnibus null hypothesis is rejected, the conclusion is that at least one population mean is different from at least one other mean. However, since the ANOVA does not reveal which means are different from which, it offers less specific information than the Tukey HSD test. The Tukey HSD is therefore preferable to ANOVA in this situation. Some textbooks introduce the Tukey test only as a follow-up to an ANOVA. However, there is no logical or statistical reason why you should not use the Tukey test even if you do not compute an ANOVA. You might be wondering why you should learn about ANOVA when the Tukey test is better. One reason is that there are complex types of analyses that can be done with ANOVA and not with the Tukey test. A second is that ANOVA is by far the most commonly-used technique for comparing means, and it is important to understand ANOVA in order to understand research reports.

Analysis of variance (ANOVA) is a test that incorporates means and variances to determine the test statistic. The test statistic is then used to determine whether groups of data are the same or different. When hypothesis testing is being performed with ANOVA, the null hypothesis is stated such that all groups are the same. The test statistic for ANOVA is called the F-ratio.

**NON-PARAMETRIC TEST**

Non-parametric covers techniques that do not rely on data belonging to any particular distribution. These include, among others: distribution free methods, which do not rely on assumptions that the data are drawn from a given probability distribution. As such it is the opposite of parametric statistics. It includes non-parametric descriptive statistics, statistical models, inference and statistical tests. In other words, non-parametric tests can be referred to be a function on a sample that has no dependency on a parameter, whose interpretation does not depend on the population fitting any parameterized distributions. In hypothesis testing, non-parametric tests play a central role for statisticians and decision makers. Among various noteworthy researchers, statistical hypotheses concern the behaviour of observable random variables. For example, the hypothesis (a) that a normal distribution has a specified mean and variance is statistical; so is the hypothesis (b) that it has a given mean but unspecified variance; so is the hypothesis (c) that a distribution is of normal form with both mean and variance unspecified; finally, so is the hypothesis (d) that two unspecified continuous distributions are identical. Following are some non-identical methods.

### A. Sign Test

Nonparametric testing is used in case of without knowledge about sample distribution; concretely, there is no assumption of normality. The sign test can be used to test the hypothesis that there is "no difference in medians" between the continuous distributions of two random variables X and Y, in the situation when we can draw paired samples from X and Y. It is a non-parametric test which makes very few assumptions about the nature of the distributions under test this means that it has very general applicability but may lack the statistical power of other tests such as the paired samples t-test or the Wilcoxon signed-rank test. The nonparametric testing begins with the test on sample median. If distribution is symmetric, median is identical to mean. Given the median is the data point at which the left side data and the right side data are of equal accumulate probability.

$$P(D < ) = P(D > ) = 0.5$$

If data is not large and there is no assumption about normality, the median is approximate to population mean. Given null hypothesis $H_0$: = and alternative hypothesis $H_1$:, the test so-called sign test is performed as below steps:

**Step 1:** Assigning plus signs to sample data points whose values are greater than and minus signs to ones whose values are less than . Note that values which equal are not considered. Plus signs and minus signs represent the right side and left side of , respectively.

**Step 2:** If the number of plus signs is nearly equal to the number of minus signs, then null hypothesis $H_0$ is true; otherwise $H_1$ is false. In other words, that the proportion of plus signs is significantly different from 0.5 cause to rejecting $H_0$ in flavour of $H_1$.

### B. Wilcoxon Sign-Rank Test

As we have noticed in section previous section-2 that sign test focuses on whether or not the observations are different from null hypothesis but it does not consider the magnitude of such difference. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used when comparing two related samples, matched samples, or repeated measurements on a single sample to assess whether their test, t-test for matched pairs, or the t-test for dependent samples when the population cannot be assumed to be normally distributed, we refer Lowry for more details. The test is named for Frank Wilcoxon (1892–1965) who, in a single paper, proposed both it and the rank-sum test for two independent samples. For more details we refer Wilcoxon. The test was popularized by Siegel in his influential text book on nonparametric statistics. Siegel used the symbol T for the value defined below as W. In consequence, the test is sometimes referred to as the Wilcoxon T test, and the test statistic is reported as a value of T. Other names may include the "t-test for matched pairs" or the "t-test for dependent samples". Walpole et al examined that Wilcoxon signed-rank test based on assumption of symmetric and continuous distribution considers both difference and how much difference is. The median is identical to the mean μ according to symmetric assumption. It includes four following steps:

**Step 1.** Calculating all deviations between data points and $\mu_0$, we have $D = \{d_1, d_2, \ldots, d_n\}$ where $d_i = x_i - \mu_0$ and $di \neq 0$. Note that data point $x_i$ is instance of random variable $X$.

**Step 2.** Assigning a rank $r_i$ to each deviation $d_i$ without regard to sign, for instance, rank value *1* and rank value *n* to be assigned to smallest and largest **absolute** deviation (without sign), respectively. If two or more absolute deviations have the same value, these deviations are assigned by average rank. For example, if 3rd, 4th and 5th deviations get the same value, they receive the same rank $(3+4+5) / 3 = 4$. We have a set of ranks $R = \{r_1, r_2, \ldots, r_n\}$ where $r_i$ is the rank of $d_i$.

**Step 3.** Let $w^+$ and $w^-$ be the sum of ranks whose corresponding deviations are positive and negative, respectively. We have $w^+ = \sum_{d_i>0} r_i$ and $w^- = \sum_{d_i<0} r_i$ and $w = min(w^+, w^-)$. Note that $w$ is the minimum value between $w^+$ and $w^-$

**Step 4.** In flavor of $H_1$: $\mu < \mu_0$, $H_0$ is rejected if $w^+$ is sufficiently small. In flavour of $H_1$: $\mu > \mu_0$, $H_0$ is rejected if $w^-$ is sufficiently small. In case of two-sided test $H_1$: $\mu \neq \mu_0$, $H_0$ is rejected if $w$ is sufficiently small. The concept sufficiently small is defined via thresholds or pre computed critical values. The value $w^+$, $w^-$ or $w$ is sufficiently small if it is smaller than a certain critical value with respect to significant level α.

### C. Mann-Whitney Test

Use this when two different groups of participants perform both conditions of your study: i.e., it is appropriate for analysing the data from an independent-measures design with two conditions. Use it when the data do not meet the requirements for a parametric test (i.e. if the data are not normally distributed; if the variances for the two conditions are markedly different; or if the data are measurements on an ordinal scale). Otherwise, if the data meet the requirements for a parametric test, it is better to use an independent-measures t-test (also known as a "two-sample" t-test). The logic behind the Mann-Whitney test is to rank the data for each condition, and then see how different the two rank totals are. If there is a systematic difference between the two conditions, then most of the high ranks will belong to one condition and most of the low ranks will belong to the other one. As a result, the rank totals will be quite different. On the other hand, if the two conditions are similar, then high and low ranks will be distributed fairly evenly between the two conditions and the rank totals will be fairly similar. The Mann-Whitney test statistic "U" reflects the difference between the two rank totals. The SMALLER it is (taking into account how many participants you have in each group) then the less likely it is to have occurred by chance. A table of critical values of U shows you how likely it is to obtain your particular value of U purely by chance. Note that the Mann-Whitney test is unusual in this respect: normally, the BIGGER the test statistic, the less likely it is to have occurred by chance).

## CONCLUSION

Parametric and nonparametric are two broad classifications of statistical procedures. Parametric tests are based on assumptions about the distribution of the underlying population from which the sample was taken. The most common parametric assumption is that data are approximately normally distributed. Non-parametric tests do not rely on assumptions about the shape or parameters of the underlying population distribution. If the data deviate strongly from the assumptions of a parametric procedure, using the parametric procedure could lead to incorrect conclusions. You should be aware of the assumptions associated with a parametric procedure and should learn methods to evaluate the validity of those assumptions. If you determine that the assumptions of the parametric procedure are not valid, use an analogous nonparametric procedure instead. The parametric assumption of normality is particularly worrisome for small sample sizes ($n < 30$). Nonparametric tests are often a good option for these data. It can be difficult to decide whether to use a parametric or non-parametric procedure in some cases. Nonparametric procedures generally have less power for the same sample size than the corresponding parametric procedure if the data truly are normal. Interpretation of nonparametric procedures can also be more difficult than for parametric procedures. Non-parametric model is less efficient than parametric model because it lacks valuable information under sample when it has no knowledge about the distribution. All properties of distribution such as mean, variance, standard deviation, median, mode, skewness, kurtosis, etc are essential information of which nonparametric model does not take advantages.

## REFERENCES

1. Walsh, J.E. Handbook of Nonparametric Statistics, New York: D.V. Nostrand (1962).
2. Conover, W.J. Practical Nonparametric Statistics, New York: Wiley & Sons (1980).
3. Rosner, B. Fundamentals of Biostatistics, California: Duxbury Press (2000).
4. Motulsky, H. Intuitive Biostatistics, New York: Oxford University Press (1995).
5. Bagdonavicius V., Kruopis J. and Nikulin M.S., "Non-parametric tests for complete data", ISTE & WILEY: London & Hoboken, ISBN 978-1-84821-269-5, (2011).

6. Corder G.W. and Foreman D.I., Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach, Wiley, ISBN 978-0-470-45461-9, (2009).
7. Gibbons, Jean Dickinson and Chakraborti, Subhabrata,Nonparametric Statistical Inference, 4th Edition, CRC, ISBN 0-8247-4052-1, (2003).
8. Wilcoxon Frank, "Individual comparisons by ranking methods". Biometrics Bulletin 1 (6), pp. 80–83, (1945).
9. Todd Neideen, MD, and Karen Brasel, MD, MPH, “Understanding Statistical Tests”, Division of Trauma and Critical Care, Department of Surgery, Medical College of Wisconsin, Milwaukee, Wisconsin.