

**Research article**

# **A Survey Report on Non-Parametric Hypothesis Testing Including Kruskal-Wallis ANOVA and Kolmogorov– Smirnov Goodness-Fit-Test**

**Prof. (Dr.) Vishwa Nath Maurya**

Professor & Ex Principal,  
Shekhawati Engineering College, Rajasthan Technical University, India  
E-mail: prof.drwnmaurya@gmail.com, prof\_vnmaurya@yahoo.in

**Diwinder Kaur Arora**

Inspector of Police,  
Group Centre, Central Reserve Police Force, Lucknow-226002, U.P.  
Ministry of Home Affairs, Govt. of India  
E-mail: hkdkarora@rediffmail.com, diwi.kaur1992@gmail.com

**Er. Avadhesh Kumar Maurya**

Assistant Professor, Department of Electronics & Communication Engineering  
Lucknow Institute of Technology, U.P. Technical University, Lucknow-226002, India  
E-mail: avadheshmaurya09@gmail.com

---

## **Abstract**

This article deals a survey report of nonparametric hypothesis testing. In the present article, we have discussed five different nonparametric hypotheses testing including sign test, signed-rank test, rank-sum test, Kruskal-Wallis ANOVA and goodness-fit-test. Sign test section gives an overview of nonparametric testing, which begins with the test on sample median without assumption of normal distribution. Signed-rank test section and rank-sum test section concern improvements of sign test. However, the prominence of signed-rank test is to be able to test sample mean based on the assumption about symmetric distribution. Here, it is demonstrated that the rank-sum test has two advantages in comparison of signed-rank test along with special feature that the rank-sum test discards the task of assigning and counting plus signs and so it is the most effective method among ranking test methods. Kruskal-Wallis ANOVA section discusses application of analysis of variance (ANOVA) in nonparametric model. Kruskal-Wallis ANOVA is useful to compare and evaluate various data samples at the same time. Kolmogorov–Smirnov goodness-fit-test section focuses on different hypothesis, which measure the distribution similarity between two samples. It is further explored that the goodness-fit-test determines whether two samples have the same distribution without concerning how the form of distribution is.

**Keywords:** Nonparametric hypothesis testing, sign test, Wilcoxon sign-rank test, rank-sum test, Kruskal-Wallis ANOVA, Kolmogorov–Smirnov goodness-fit-test, binomial distribution, sample distribution, degree of confidence etc.

---

## 1. Introduction

Hypothesis tests can be classified in two categories—first one is parametric test and the second one is called non-parametric tests. 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. Nonparametric 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, nonparametric 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, nonparametric tests play a central role for statisticians and decision makers. Among various noteworthy researchers, Stuart et al [13] proposed that statistical hypotheses concern the behavior 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.

It will have been noticed that in the examples (a) and (b) the distribution underlying the observations was taken to be of a certain form (the normal) and the hypothesis was concerned entirely with the value of one or both of its parameters. Such a hypothesis, for obvious reasons, is called parametric. However, Hypothesis (c) was of a different nature, as no parameter values are specified in the statement of the hypothesis; we might reasonably call such a hypothesis non-parametric. Hypothesis (d) is also non-parametric but, in addition, it does not even specify the underlying form of the distribution and may now be reasonably termed distribution-free. Notwithstanding these distinctions, the statistical literature now commonly applies the label "non-parametric" to test procedures that we have just termed "distribution-free", thereby losing a useful classification.

Nonparametric tests find their wide applications for studying populations that take on a ranked order (such as movie reviews receiving one to four stars). The application of non-parametric approaches requires when data have a ranking but no clear numerical interpretation, such as when assessing preferences. In terms of levels of measurement, non-parametric methods result in "ordinal" data. As nonparametric methods need fewer assumptions, their applicability is much wider than the corresponding parametric methods. In particular, they may be applied in situations where less is known about the application in question. Also, due to the reliance on fewer assumptions, nonparametric methods are more robust. Another justification for the use of non-parametric methods is simplicity. In certain cases, even when the use of parametric methods is justified, non-parametric methods may be easier to use. Due both to this simplicity and to their greater robustness, non-parametric methods are seen by some statisticians as leaving less room for improper use and misunderstanding. However, the wider applicability and increased robustness of non-parametric tests comes at a cost in cases where a parametric test would be appropriate, non-parametric tests have less power. In other words, a larger sample size can be required to draw conclusions with the same degree of confidence. In the light of wide applications of nonparametric methods in hypothesis testing, several noteworthy researchers [1, 2, ..., 5, 8, 10, ..., 12 & 15] focused their attention in this connection.

This article presents a survey report on nonparametric hypothesis testing procedures which covers sign test, Wilcoxon sign-rank test, rank-sum test, Kruskal-Wallis ANOVA and Kolmogorov–Smirnov goodness-fit-test. The present article contains 7 sections. First section itself presents introduction. In section 2, sign test gives an overview of nonparametric testing, which begins with the test on sample median without assumption of normal distribution. In sections 3 & 4, signed-rank test and rank-sum test concern improvements of sign test. The prominence of signed-rank test is to be able to test sample mean based on the assumption about symmetric distribution. Rank-sum test discards the task of assigning and counting plus signs and so it is the most effective method among ranking test methods. In section 5, nonparametric ANOVA discusses application of analysis of variance (ANOVA) in nonparametric model along with providing special focus on its application aspect in order to compare and evaluate various data samples at the same time. In section 6, nonparametric goodness-fit-test section focuses on different hypothesis, which measure the distribution

similarity between two samples. It determines whether two samples have the same distribution without concerning how the form of distribution is. Finally, we have drawn valuable observations as conclusions based on survey report in the he last section. Note that in this report terms sample and data sample have the same meaning. A sample contains many data points. Each data point is also called an observation.

## 2. 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  $\tilde{\mu}$  is the data point at which the left side data and the right side data are of equal accumulate probability.

$$P(D < \tilde{\mu}) = P(D > \tilde{\mu}) = 0.5$$

If data is not large and there is no assumption about normality, the median is approximate to population mean. Given null hypothesis  $H_0: \tilde{\mu} = \tilde{\mu}_0$  and alternative hypothesis  $H_1: \tilde{\mu} \neq \tilde{\mu}_0$ , the test so-called sign test [Walpole, Myers, Myers, Ye 2012] is performed as below steps:

- **Step 1.** Assigning plus signs to sample data points whose values are greater than  $\tilde{\mu}_0$  and minus signs to ones whose values are less than  $\tilde{\mu}_0$ . Note that values which equal  $\tilde{\mu}_0$  are not considered. Plus signs and minus signs represent the right side and left side of  $\tilde{\mu}_0$ , 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  $H_0$  is false. In other words, that the proportion of plus signs is significantly different from 0.5 cause to rejecting  $H_0$  in favor of  $H_1$ .

The reason of  $H_0$  acceptance is that the probability that data points (or observations) fall in both left side and right side of  $\tilde{\mu}_0$  are of equal value 0.5 and of course, it is asserted that  $\tilde{\mu}_0$  is a real median. Note that terms *data point*, *sample point*, *sample value* and *observation* are identical.

In the case that alternative hypothesis  $H_1: \tilde{\mu} < \tilde{\mu}_0$ , if the proportion of plus signs is less than 0.5 then rejecting  $H_0$  in flavour of  $H_1$ . In the case that alternative hypothesis  $H_1: \tilde{\mu} > \tilde{\mu}_0$ , if the proportion of plus signs is greater than 0.5 then rejecting  $H_0$  in flavour of  $H_1$ . Now let  $X$  be the discrete random variable representing the number of plus signs and suppose that  $X$  confirms binomial distribution  $B(X; n; p)$  where  $n$  and  $p$  are the total number of sample data points and the probability that plus sign is assigned to a data point, respectively. Because the proportion of plus signs gets 0.5 when  $H_0: \tilde{\mu} = \tilde{\mu}_0$  is true, the parameter  $p$  is set to be 0.5. Given the distribution of plus signs is  $B(X; n; 0.5)$  and significant level  $\alpha$  and let  $x$  be the instance of  $X$  where  $x = \frac{\text{The number of plus signs}}{n}$ , there are three following tests:

- (i)  $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} \neq \tilde{\mu}_0$ : In case of  $x < n/2$ , if  $2P(X \leq x) < \alpha$  then rejecting  $H_0$ . In case of  $x > n/2$ , if  $2P(X \geq x) < \alpha$  then rejecting  $H_0$ . This test belongs to two-sided test family.
- (ii)  $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} < \tilde{\mu}_0$ : if  $P(X \leq x) < \alpha$  then rejecting  $H_0$ . This test belongs to one-sided test family.
- (iii)  $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} > \tilde{\mu}_0$ : if  $P(X \geq x) < \alpha$  then rejecting  $H_0$ . This test belongs to one-sided test family.

Note that  $P(\dots)$  is accumulated probability of binomial distribution  $B(X; n; 0.5)$ , for example,  $P(X \leq x) = \sum_{k=0}^x \binom{n}{k} 0.5^k 0.5^{n-k}$ . In case that  $n$  is large enough, for instance  $n > 10$ ,  $B(X; n; 0.5)$  is approximate to standard

normal distribution  $N(Z; 0; 1)$  where  $Z = \frac{x-0.5n}{\sqrt{0.25n}}$ . Let  $z$  be the instance of  $Z$  where  $z = \frac{x-0.5n}{\sqrt{0.25n}}$ , there are three following tests:

- $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} \neq \tilde{\mu}_0$ : if  $|z| > z_{\alpha/2}$  then rejecting  $H_0$  where  $z_{\alpha/2}$  is 100 $\alpha/2$  percentage point of standard normal distribution.
- $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} < \tilde{\mu}_0$ : if  $z < -z_{\alpha/2}$  then rejecting  $H_0$ .
- $H_0: \tilde{\mu} = \tilde{\mu}_0$  and  $H_1: \tilde{\mu} > \tilde{\mu}_0$ : if  $z > z_{\alpha/2}$  then rejecting  $H_0$ .

In case of pair-test  $H_0: \tilde{\mu}_1 - \tilde{\mu}_2 = d_0$  which we need to know how much median  $\tilde{\mu}_1$  shifts from other one  $\tilde{\mu}_2$ , sign test is applied in similar way with a little bit of change. If  $d_0 = 0$ ,  $H_0$  indicates whether  $\tilde{\mu}_1$  equals  $\tilde{\mu}_2$ . We compute all deviations between two samples  $X$  and  $Y$  where  $\tilde{\mu}_1$  is sample median of  $X$  and  $\tilde{\mu}_2$  is sample median of  $Y$ . Let  $d_i = x_i - y_i$  be the deviation between  $x \in X$  and  $y \in Y$ . Plus signs (minus signs) are assigned to  $d_i$  (s) which are greater (less) than  $d_0$ . Now signed test is applied into such plus signs and minus signs by discussed method.

### 3. 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 [ 7] 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 [15]. The test was popularized by Siegel [11] in his influential text book on nonparametric statistics. Siegel [11] 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 [14] examined that Wilcoxon signed-rank test based on assumption of symmetric and continuous distribution considers both difference and how much difference is. The median  $\tilde{\mu}_0$  is identical to the mean  $\mu$  according to symmetric assumption. It includes four following steps:

- **Step 1.** Calculating all deviations between data points and  $\mu_0$ , we have  $D = \{d_1, d_2, \dots, d_n\}$  where  $d_i = x_i - \mu_0$  and  $d_i \neq 0$ . 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 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> deviations get the same value, they receive the same rank  $(3+4+5) / 3 = 4$ . We have a set of ranks  $R = \{r_1, r_2, \dots, r_n\}$  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 flavor 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, see pp. 759, Walpole et al [14] for 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  $\alpha$ .

In case of pair test  $H_0: \mu_1 - \mu_2 = d_0$ , the deviation  $d_i$  in step 1 is calculated based  $d_0$  and two samples  $X$  and  $Y$ , so  $d_i = x_i - y_i - d_0$  where  $x \in X$  and  $y \in Y$ . Note that  $\mu_1$  and  $\mu_2$  are taken from  $X$  and  $Y$ , respectively. Steps 2, 3, 4 are performed in similar way.

Let  $W^+$  be random variables of  $w_+$ . If  $n \geq 15$  then  $W^+$  approaches normal distribution with mean  $\mu_{W^+} = \frac{n(n+1)}{4}$  and variance  $\sigma_{W^+}^2 = \frac{n(n+1)(2n+1)}{24}$ . We can normalize  $W^+$  so as to define critical region via percentage point  $z_\alpha$  of normal standard distribution,  $Z_{W^+} = \frac{W^+ - \mu_{W^+}}{\sigma_{W^+}}$

#### 4. Rank-Sum Test

The Wilcoxon Rank Sum test can be used to test the null hypothesis that two populations X and Y have the same continuous distribution. As it is keenly observed in view of Walpole et al [14] that rank-sum test is a variant of signed-rank test. Suppose there are two samples  $X = \{x_1, x_2, \dots, x_{n_1}\}$  and  $Y = \{y_1, y_2, \dots, y_{n_2}\}$  and the null hypothesis is specified as  $H_0: \mu_1 = \mu_2$  where  $\mu_1$  and  $\mu_2$  are taken from X and Y, respectively. We assign ranks to such  $n_1 + n_2$  data points according to their values, for instance, rank value 1 and rank value  $n_1 + n_2$  to be assigned to smallest and largest sample value. If two or more data points have the same value, these points are assigned by average rank. For example, if 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> data points get the same value, they receive the same rank  $(3+4+5) / 3 = 4$ . Let  $R = \{r_1, r_2, \dots, r_{n_1+n_2}\}$  be the set of these ranks. Let  $w_1$  and  $w_2$  be the sum of ranks corresponding to  $n_1$  data points in X and  $n_2$  data points in Y, respectively.

$$w_1 = \sum_{x_i \in X} r_i \text{ and } w_2 = \sum_{y_i \in Y} r_i$$

where  $r_i$  is a rank of a data point in the set  $X \cup Y$  and  $r_i = \overline{1, n_1 + n_2}$

We have  $w_1 + w_2 = \frac{(n_1+n_2)(n_1+n_2+1)}{2}$ . There are three following tests:

- (i) Rejecting  $H_0$  in favor of alternative  $H_1: \mu_1 < \mu_2$  if  $w_1$  is sufficiently small.
- (ii) Rejecting  $H_0$  in favor of alternative  $H_1: \mu_1 > \mu_2$  if  $w_2$  is sufficiently small.
- (iii) In case of two-sided test with  $H_1: \mu_1 \neq \mu_2$  if the minimum of  $w_1$  and  $w_2$  is sufficiently small then rejecting  $H_0$ .

Rank-sum test has two advantages in comparison of signed-rank test:

- There is no need to calculate deviations among samples and to count the number of plus signs and minus signs.
- Samples can have different number of data points, for instance,  $|X| = n_1 \neq n_2 = |Y|$

Setting  $u_1 = w_1 - \frac{n_1(n_1+1)}{2}$  and  $u_2 = w_2 - \frac{n_2(n_2+1)}{2}$  and suppose that  $u_1$  and  $u_2$  are instances of random variables  $U_1$  and  $U_2$ , respectively. If both  $n_1$  and  $n_2$  are greater than 8, variable  $U_1$  (or  $U_2$ ) is approximate to normal distribution with mean  $\mu_{U_1} = \frac{n_1 n_2}{2}$  and variance  $\sigma_{U_1}^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$ . We can normalize  $U_1$  ( $U_2$ ) so as to define critical region via percentage point  $z_\alpha$  of normal standard distribution,  $Z_{U_1} = \frac{U_1 - \mu_{U_1}}{\sigma_{U_1}}$ .

#### 5. Kruskal -Wallis ANOVA Test

In many applications, we process various samples (X, Y, Z, etc.) where each sample is a set of observations (data points) which relate to a concrete method, a way or an approach that creates or produces these observations. Such concrete method is called **treatment**. In other words, we consider a matrix of observations and each row represents a mono-sample attached to a treatment, for instance, X or Y or Z, etc. For convenience, matrix of observations is called multi-sample or sample, in short. Treatments are grouped into categories which are called **factors**. If sample has only one factor, it is single-factor sample; otherwise, it is called several-factor sample. The Kruskal–Wallis one-way analysis of variance by ranks (named after William Kruskal and W. Allen Wallis) is a non-parametric method for testing whether



samples originate from the same distribution. It is used for comparing more than two samples that are independent, or not related. The parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA). When the Kruskal-Wallis test leads to significant results, then at least one of the samples is different from the other samples. The test does not identify where the differences occur or how many differences actually occur. It is an extension of the Mann-Whitney U test to 3 or more groups. The Mann-Whitney would help to analyse the specific sample pairs for significant differences. The Kruskal-Wallis ANOVA is useful as a general nonparametric test for comparing two or more independent samples. It can be used to test whether such samples come from the same distribution. They are powerful alternatives to the one-way analysis of variance.

The Kruskal-Wallis ANOVA uses the sum of difference between mean ranks of these samples as the statistic. The statistic of Mood's median test only relates to the number of larger or smaller than the median value but not their actual distance from the median, so it is not as effective as Kruskal-Wallis ANOVA.

As an example, researchers want to know whether the enhanced eyesight of young patients, who use three different therapies to enhance their eyesight, comes from the same distribution. Thirty students' enhanced eyesight, after adopting these three therapies, was recorded. Following table is an example of single-factor sample.

Treatment 1	$y_{11}$	$y_{12}$	$y_{13}$	$\bar{y}_1 = (y_{11} + y_{12} + y_{13}) / 3$
Treatment 2	$y_{21}$	$y_{22}$	$y_{23}$	$\bar{y}_2 = (y_{21} + y_{22} + y_{23}) / 3$
Treatment 3	$y_{31}$	$y_{32}$	$y_{33}$	$\bar{y}_3 = (y_{31} + y_{32} + y_{33}) / 3$
				$\bar{y} = (\bar{y}_1 + \bar{y}_2 + \bar{y}_3) / 3$

Let  $Y_{ij}$  be the random variable representing  $j^{th}$  data point of  $i^{th}$  treatment.

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

where  $\mu$  so-call overall mean is the mean over whole sample,  $\tau_i$  called treatment effect denotes the parameter of  $i^{th}$  treatment and  $\epsilon_{ij}$  denotes the random error.

There is an assumption that random error  $\epsilon_{ij}$  is independently distributed and confirms normal distribution; moreover, it has mean 0 and variance  $\sigma^2$ . Let  $\mu_i = \mu + \tau_i$  be the treatment mean of  $i^{th}$  treatment. The objective of analysis of variance (ANOVA), refer Montgomery and Runger [9] is to analyse statistics about treatment mean, treatment effect, random error so as to take out conclusions about such statistics. Basically, ANOVA focuses on characteristics relating to deviation, variability, sum of squares, mean square, etc. A typical approach of ANOVA is to test whether  $k$  treatment means  $\mu_1, \mu_2, \dots, \mu_k$  are equal; it means that we test the following hypotheses:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$$

Due to  $\mu_i = \mu + \tau_i$ , this test is re-written:

$$H_0: \tau_1 = \tau_2 = \dots = \tau_k = 0$$

$$H_1: \tau_i \neq 0 \text{ for at least one treatment}$$

If  $H_0$  is true, treatments have no effect on whole sample. Let  $y_{ij}$  be the instance of random variable  $Y_{ij}$ . Let  $y_i, \bar{y}_i, y$  and  $\bar{y}$  be the sum of observations of treatment  $i$ , the average of observations of treatment  $i$ , the sum of whole observations and the average of whole observations.

$$y_i = \sum_{j=1}^{n_i} y_{ij}, \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, y = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}, \bar{y} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}$$

where  $k$  is the number of treatments,  $n_i$  is the number of observations under treatment and  $N = n_1 + n_2 + \dots + n_k$  is the total number of observations.

Let  $SS_T$ ,  $SS_{Treatment}$  and  $SS_E$ , Montgomery, Runger [9] be the total sum of squares, treatment sum of squares and error sum of squares. Please pay attention to  $SS_T$ ,  $SS_{Treatment}$  and  $SS_E$  because they are main research objects in ANOVA.

We have:

$$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

$$SS_{Treatment} = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$$

$$SS_E = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

Following is the sum of squares identity:

$$SS_T = SS_{Treatment} + SS_E$$

Treatment sum of squares  $SS_{Treatment}$  is very important because it reflects treatment effects  $\tau_i$  (s) and treatment means  $\mu_i$  (s). The expected values of treatment sum of squares and error sum of squares are computed as below:

$$E(SS_{Treatment}) = (k - 1)\sigma^2 + \sum_{i=1}^k n_i \tau_i^2$$

$$E(SS_E) = (N - k)\sigma^2$$

$SS_T$  and  $SS_{Treatment}$  and  $SS_E$  have  $N - 1$  and  $k - 1$  degrees of freedom, respective because there are  $N$  observations over whole sample and  $k$  treatments. So  $SS_E$  has  $N - k = (N - 1) - (k - 1)$  due to  $SS_E = SS_T - SS_{Treatment}$ . Based on degrees of freedom, treatment mean square  $MS_{Treatment}$  and error mean square  $MS_E$  is determined as below:

$$MS_{Treatment} = \frac{SS_{Treatment}}{k - 1}$$

$$MS_E = \frac{SS_E}{N - k}$$

If null hypothesis  $H_0: \tau_1 = \tau_2 = \dots = \tau_k = 0$  is true,  $MS_{Treatment}$  is an unbiased estimate of variance  $\sigma^2$  due to  $E(MS_{Treatment}) = \frac{1}{k-1} E(SS_{Treatment}) = \sigma^2 + \frac{1}{k-1} \sum_{i=1}^k n_i \tau_i^2 = \sigma^2 + \frac{1}{k-1} \sum_{i=1}^k n_i 0^2 = \sigma^2$ . Moreover

$MS_E$  is always an unbiased estimate of variance  $\sigma^2$  due to  $E(MS_E) = \frac{1}{N-k} E(SS_E) = \sigma^2$ . So  $MS_{Treatment}$  and  $MS_E$  conform chi-square distribution and the ratio of  $MS_{Treatment}$  to  $MS_E$  conforms  $F$ -distribution with  $k - 1$  and  $n(k - 1)$  degrees of freedom:

$$F_0 = \frac{MS_{Treatment}}{MS_E} \sim F_{k-1, N-k}$$

Hypothesis  $H_0: \tau_1 = \tau_2 = \dots = \tau_k = 0$  is rejected if the ratio  $F_0 > f_{\alpha, k-1, n(k-1)}$  where  $f_{\alpha, k-1, n(k-1)}$  is the  $100\alpha$  percentage point of  $F$ -distribution with  $k-1$  and  $N-k$  degrees of freedom.

We have already discussed about parametric ANOVA with normality assumption, now nonparametric ANOVA is the next topic. Nonparametric ANOVA has no assumption of normality of random error but the independence of random error is required. The noteworthy researchers Montgomery and Runger [ 9] and Walpole et al [14] examined to propose that the Kruskal-Wallis test is a popular nonparametric test. Suppose treatment  $i$  has  $n_i$  observations and there are  $k$  treatment, let  $N = n_1 + n_2 + \dots + n_k$  be the total of observations. Kruskal-Wallis test assigns ranks to such  $N$  observations according to their values, for instance, rank value 1 and rank value  $N$  to be assigned to smallest and largest sample value. If two or more observations have the same value, these observations are assigned by average rank. For example, if 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> observations get the same value, they receive the same rank  $(3+4+5) / 3 = 4$ . Let  $R_{ij}$  be the rank of observation  $Y_{ij}$ . If null hypothesis  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  is true, which means that all treatments have the same mean, then ranks spread over all treatments equally. In other words, the expected value of  $R_{ij}$  (s) is nearly equal to the mid-point of  $N$  ranks, so we have:

$$E(R_{ij}) = (N + 1) / 2$$

Let  $\bar{R}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij}$  be average rank of treatment  $i$ , the expected value of  $\bar{R}_i$  is determined as below:

$$E(\bar{R}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} E(R_{ij}) = \frac{N + 1}{2}$$

If the null hypothesis  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  is true, the average rank  $\bar{R}_i$  does not shift from its expected value  $(N+1) / 2$  much. The difference between  $\bar{R}_i$  and its expected value  $(N + 1) / 2$  is determined by following statistic:

$$K = \frac{12}{N(N + 1)} \sum_{i=1}^k n_i \left( \bar{R}_i - \frac{N + 1}{2} \right)^2$$

This formula is transformed into more practical format as below:

$$K = \frac{12}{N(N + 1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N + 1)$$

where  $R_i = \sum_{j=1}^{n_i} R_{ij}$  is the sum of ranks under treatment  $i$ . It is proved that statistic  $K$  approaches chi-square distribution  $\chi_{\alpha, k-1}$  with  $k - 1$  degrees of freedom where  $k$  is the number of treatments. Null hypothesis  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  is rejected in favour of alternative hypothesis  $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$  if  $K > \chi_{\alpha, k-1}$ .

## 6. Kolmogorov–Smirnov Goodness-Fit-Test

Goodness-fit-test is the test that determines whether a sample confirms specified distribution or whether two samples have the same distribution. Although Kolmogorov–Smirnov goodness-fit-test being a kind of nonparametric testing does not consider the sample distribution, it is based on the definition of Kolmogorov distribution. Kolmogorov distribution is continuous distribution whose accumulative distribution function is defined as below Wikipedia [5]:

$$P(K \leq k) = \frac{\sqrt{2\pi}}{k} \sum_{i=1}^{\infty} e^{-(2i-1)^2 \pi^2 / (8k^2)}$$



The critical value  $K_\alpha$  at significant level  $\alpha$  is  $100\alpha$  percentage point satisfying equation:

$$1 - \alpha = P(K \leq K_\alpha) = \frac{\sqrt{2\pi}}{K_\alpha} \sum_{i=1}^{\infty} e^{-(2i-1)^2 \pi^2 / (8K_\alpha^2)}$$

Kolmogorov–Smirnov goodness-fit-test is to determine whether two samples have the same distribution regardless of the underlying distribution. Given  $X = \{x_1, x_2, \dots, x_n\}$  and  $Y = \{y_1, y_2, \dots, y_n\}$  are two testing samples, the null hypothesis  $H_0$  is that  $X$  and  $Y$  have the same distribution. Let  $F_X$  and  $F_Y$  be the empirical distribution functions of  $X$  and  $Y$  respectively. Note that empirical distribution function is accumulative function which increases gradually according to the order of values.

$$F_X(x_i) = \frac{\text{The number of } x \in X \text{ that } \leq x_i}{n}$$

$$F_Y(y_i) = \frac{\text{The number of } y \in Y \text{ that } \leq y_i}{n}$$

Let  $D$  be the maximum absolute deviation between  $F_X$  and  $F_Y$  over whole samples  $X$  and  $Y$

$$D = \max |F_X(x_i) - F_Y(y_i)| \text{ where } i = \overline{1, n}$$

It is easy to recognize that the process to find out  $D$  is iterative process browsing all pairs of observation  $(x_i, y_i) \in X \times Y$ .

It is proved that  $D\sqrt{n/2}$  confirms  $K$  distribution. Therefore, the null hypothesis  $H_0$  is rejected at significant level  $\alpha$  if  $D\sqrt{n/2} > K_\alpha$ .

## 7. Conclusions

Now we had a general and detailed point of view about nonparametric testing. We can draw two main comments from research over this domain:

- Firstly, nonparametric 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. However, nonparametric testing is very useful and appropriate to cases that knowledge of distribution cannot be extracted or sample does not conform normal distribution. In case that underlying distribution is ignored and nonparametric testing is the best choice. Therefore, we conclude that the most important thing is to choose appropriate model (parametric or nonparametric) which is adaptive to testing situation and testing requirement.
- Secondly, nonparametric model is often based on ranking. Ranking process aims to transform origin sample into simpler sample so-called ranking sample. Ranking sample is the set of ranks; thus, each rank is assigned to respective observation from origin sample. Because nonparametric model does not know valuable information of origin sample such as mean, variance, standard deviation; it will exploit ranking sample to discover such valuable information. Therefore, nonparametric testing, in turn, applies parametric methods into the ranking sample. Concretely, nonparametric testing assumes that statistic (s) on ranking sample conform some pre-defined distributions. For example, sign test assumes that the number of plus signs in ranking data confirms binominal distribution, signed-rank test and sum-rank test apply Wilcoxon distribution into ranking data and nonparametric goodness-fit-test is based on Kolmogorov distribution. We conclude that parametric testing and nonparametric testing have a strongly mutual relationship and so, we should take advantages of both of them.

## Acknowledgements



**Dr. V. N. Maurya;** principal author of the present paper has served as founder Director of Vision Institute of Technology, Aligarh (U. P. Technical University, Lucknow (India) and as Principal/Director at Shekhawati Engineering College (Rajasthan Technical University, Kota) after having vast experience in the field of Technical and Management Institutions in the cadre of Professor & Dean Academics such as at Institute of Engineering & Technology, Sitapur, UP, India; Haryana College of Technology & Management (Kuruchhetra University, Kuruchhetra) and Singhania University, Rajasthan. During his tenure as the Director, Vision Institute of Technology, Aligarh (Uttar Pradesh Technical University, Lucknow) and as the Principal, Shekhawati Engineering College (Rajasthan Technical University, Kota); massive expansion of infrastructure, research facilities, laboratories upgradation/augmentation and other relevant facilities and services for B.Tech./M.Tech./MBA academic programmes in different branches had taken place to accommodate and facilitate the campus students. He is the Chief Editor of Editorial Board of American Journal of Modeling and Optimization; Science and Education Publishing, New York, USA and Statistics, Optimization and Information Computing; International Academic Press, Hong Kong and Advisory Editor of World Research Journal of Numerical Analysis and Mathematical Modeling; Bioinfo Publications, Pune, India and Member of Editorial and Reviewer Board of over 50 Indian and Foreign International journals published by leading publishers of USA, Italy, Hong Kong, Austria, U.K., Algeria, Nigeria and other European and African countries. He has been associated with leading Indian Universities-U. P. Technical University, Lucknow during 2005-06 and Chhatrapati Shahu Ji Maharaj University, Kanpur for three terms during 2000-2004 for significant contribution of his supervision as Head Examiner of Central Evaluation for Theory Examinations of UG (B.Tech./B.Pharm.) and PG (MA/M.Sc.) programmes.

Dr. Maurya possesses an outstanding and meritorious academic record. He earned his M.Sc. (1996) with First Division and Ph.D. Degree (2000) in Mathematics & Statistics with specialization in Operations Research from Dr. Ram Manohar Lohia Avadh University, Faizabad, UP, India and thereafter he accomplished another two years Master's Professional Degree-MBA with First Division (B<sup>+</sup> Grade) with specialization in Computer Science from NU, California, USA in 2003. He did Ph.D. on topic titled "A study of use of stochastic processes in some queueing models" under supervision of Prof. (Dr.) S.N. Singh, Ph.D. (BHU). He started his teaching career as Lecturer in 1996 to teach post-graduate courses MBA, MCA and M.Sc. and later he was appointed as Professor & Head, Department of Applied Sciences and Engineering at Singhania University, Rajasthan in the year 2004. Since then, Prof. V. N. Maurya has rendered his services as Professor & Head/Dean as well as keen Researcher for Post-Doctoral research and he has devoted his entire scientific and professional career in teaching at various premier technical institutions of the country such as at Haryana College of Technology & Management, Kaithal (Kuruchhetra University, Kuruchhetra); Institute of Engineering & Technology, Sitapur and United College of Engineering & Research, Allahabad. During the last 17 years of his professional career, Prof. V. N. Maurya has authored three textbooks and published more than 55 scientific and academic research papers including 25 research papers as Principal Author based on his Post-Doctoral work and D.Sc. Thesis in Indian and Foreign leading International Journals in the field of Mathematical and Management Sciences, Industrial Engineering & Technology. He is an approved Supervisor of UGC recognized various Indian Universities for Research Programs leading to M. Phil. & Ph.D. such as Shridhar University, Pilani (Rajasthan), Singhania University, Rajasthan and CMJ University, Sillong, Meghalaya and JJT University Jhunjhunu, Rajasthan and U.P. Technical University Lucknow etc. and since last 7 years, he is actively engaged as Research Supervisor of M. Phil. & Ph.D. Scholars in wide fields of Operations Research, Optimization Techniques, Statistical Inference, Applied Mathematics, Operations Management and Computer Science. He has guided as Principal Supervisor and Co-Supervisor to several Research Scholars of M. Phil. and Ph.D. Most of his published research papers in India, USA, Algeria, Nigeria, Malaysia and other European and African countries are recognized as innovative contributions in the field of Applied Mathematics and Statistics including Operations Research. By virtue of his innovative research contribution to many mathematical, statistical, computer science and industrial engineering related areas basic as well as application oriented, Prof. V. N. Maurya is known as one of the Indian leading experts in Mathematics, Statistics and Operational Research. Apart from this, Prof. Maurya is also on active role of Fellow/Senior/Life Member of various reputed National and International professional bodies of India and abroad including Operations Research Society of India, Kolkata; Indian Society for Technical Education, New Delhi; Indian Association for Productivity, Quality & Reliability, Kolkata; Indian Society for Congress Association, Kolkata; International Indian Statistical Association, Kolkata; All India Management Association, New Delhi; Rajasthan Ganita Parishad, Ajmer and International Association of Computer Science & Information Technology, Singapore etc.



**Diwinder Kaur Arora;** co-author of the present paper accomplished MBA Degree with specialization in Human Resources from Pondicherry Central University, Pondicherry and she graduated with B.Sc. (Medical/ZBC Group) Degree in 1987 from Kanpur University, Kanpur, India and did Diploma also from Government Polytechnic College, Amethi, U.P. throughout in First Division. She has vast experience of more than 22 years of general administration and management as Police Officer of Central Reserve Police Force, Ministry of Home Affairs, Govt. of India. She was selected as Assistant Sub-Inspector (Non-Gazetted Officer) in 1991 and after successful completion of her services she was promoted as Sub-Inspector in 2004 and since 2012 she is working in the grade of Inspector of Police at Group Centre, Central Reserve Police Force, Lucknow, U.P. Apart from this, she has published more than 15 research papers in Indian and Foreign International journals of repute in the field of Management, Information Technology and Physical Sciences such as in World of Sciences Journal, Engineers Press Publishing Group, Vienna, Austria; International Journal of Engineering Research and Technology, Engineering Science & Research Support Academy (ESRSA), Vadodara, India; International Journal of Electronics Communication and Electrical Engineering, Algeria; International Journal of Information Technology & Operations Management, Academic and Scientific Publisher, New York, USA.



**Er. Avadhesh Kumar Maurya;** co-author of the paper is having an outstanding academic record and accomplished his M.Tech. Degree with specialization in Digital Communication from Uttarakhand Technical University, Dehradun, UK and he was graduated with B.Tech. Degree in Electronics and Communication Engineering from Rajasthan Technical University, Kota (Rajasthan). He is recipient of four First Divisions in his Student Career with flying colours. Since last one year, Er. A. K. Maurya is serving as Assistant Professor in Department of Electronics and Communication Engineering at Lucknow Institute of Technology, U.P. Technical University, Lucknow. Prior to assuming the post of Assistant Professor at Lucknow Institute of Technology, U.P., he served as a Network Engineer for two years in National Informatics Centre, Department of Information Technology, Govt. of India with collaboration of HCL Co. He has worked on some projects such as Movable Target Shooter using Ultrasonic Radar and Hartley Oscillator. Apart from this, he has got industrial training in Door Darshan Kendra,

Lucknow, U.P. in the field of TV Program Generation and Broadcasting of different channels for partial fulfilment of his Degree and published also over 18 research papers in various Indian and Foreign International journals of repute in the field of Electronics & Communication Engineering, Computer Science & Information Technology and Physical Sciences such as in International Journal of Electronics Communication and Electrical Engineering, Algeria; World of Sciences Journal, Engineers Press Publishing Group, Vienna, Austria; International Journal of Information Technology & Operations Management, Academic and Scientific Publisher, New York, USA; International Journal of Engineering Research and Technology, Engineering Science & Research Support Academy (ESRSA), Vadodara, India; International Journal of Software Engineering & Computing, Serials Publications, New Delhi, India and many more.

## References

- [1] 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
- [2] 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
- [3] Gibbons, Jean Dickinson and Chakraborti, Subhabrata, Nonparametric Statistical Inference, 4<sup>th</sup> Edition, CRC, ISBN 0-8247-4052-1, 2003
- [4] Hettmansperger, T. P. and McKean, J. W., Robust nonparametric statistical methods. Kendall's Library of Statistics, First Edition, London, 1998
- [5] Kolmogorov–Smirnov test, [http://en.wikipedia.org/wiki/Kolmogorov-Smirnov\\_test](http://en.wikipedia.org/wiki/Kolmogorov-Smirnov_test)
- [6] Kruskal and Wallis, "Use of ranks in one-criterion variance analysis", Journal of the American Statistical Association 47 (260), pp. 583–621, 1956
- [7] Lowry Richard. "Concepts & Applications of Inferential Statistics", Retrieved 24 March 2011.

- [8] Maurya V. N., Arora Diwinder Kaur and Maurya Avadhesh Kumar, Performance of approximations to distribution of weighted combination of independent and dependent P-values based on Fisher's inverse chi-square statistic using simulations, Communicated, 2013
- [9] Montgomery Douglas C. and Runger George C., Applied Statistics and Probability for Engineers, 3<sup>rd</sup> edition. Copyright 2003 © John Wiley & Son, Inc., ISBN: 0-471-20454-4, 2003
- [10] Park Hun Myoung , Hypothesis Testing and Statistical Power of Test, University of Information Technology Services Centre for Statistical and Mathematical Computing, Indiana University, 2008
- [11] Siegel Sidney, Non-parametric statistics for the behavioral sciences, McGraw-Hill, New York, pp. 75–83, 1956
- [12] Steven Arnold, Nonparametric Statistics, Statistics-Penn State University
- [13] Stuart A., Ord J.K. and Arnold S, Kendall's Advanced Theory of Statistics: Volume 2A—Classical Inference and the Linear Model, Sixth Edition, 1999
- [14] Walpole Ronald E., Myers Raymond H., Myers Sharon L. and Ye Keying, Probability & Statistics for Engineers & Scientists, 9<sup>th</sup> Edition, Copyright © 2012, Pearson Education, Inc. ISBN: 978-0-321-62911-1
- [15] Wilcoxon Frank, "Individual comparisons by ranking methods". Biometrics Bulletin 1 (6), pp. 80–83, 1945