Journal of the Nigerian Association of Mathematical Physics
Volume 52 (*July & Sept., 2019 Issue*), **pp9 – 14**

A NEW METHOD OF NONPARAMETRIC TEST AGAINST MANN-WHITNEY U TEST BETWEEN TWO SAMPLE GROUPS

Babarinsa O. and Shobanke D.A.

Department of Mathematical Sciences,
Federal University Lokoja, P.M.B 1154, Lokoja, Kogi State, Nigeria

Abstract

This article presents a new method comparable to Mann-Whitney U test. The method makes use of combine ranking where ties between the groups, but not within the group, are considered. In the method, if no ties occur between the groups, then zero is considered otherwise average of the rank of ties common between groups are considered. The result obtained using the new method is about 7% more accurate than the traditional Mann-Whitney U test.

Keywords**:** Nonparametric, Mann-Whitney test, samples, ranks

1. Introduction

Nonparametric tests or distribution-free methods are used when the original data consists of ranks instead of value [1]. Although nonparametric tests are less sensitive and thus more liable to be influenced by Type II error, they make no assumptions about the form of the underlying distribution [2]. They do not utilize all the information provided by the sample, but the computations involved in nonparametric methods are appealing and intuitive. One of the most effective methods of nonparametric test is the Mann-Whitney U test. Over half of a century, Mann-Whitney U test has been used for nonparametric test of null hypothesis for two samples [3]. Mann-Whitney U test or Mann-Whitney-Wilcoxons is statistically equivalent to the Wilcoxons rank sum test. Thus, the Mann-Whitney U is the equivalent of the independent group t-test for determining whether two sampled groups are from a single population and used to test whether two sample means are equal or not. When data do not meet the parametric assumptions of the t-test, the Mann-Whitney U test tends to be more appropriate [4].

Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed [5]. The assumptions to use Mann-Whitney U test are that the sample drawn from the population is random, ordinal measurement scale is assumed and Independence within the samples is presumed. However, Mann-Whitney U test has some limitations when Monte Carlo methods are used [6].

The rest of the paper is organised as follows. In Section 2, we gave an in-depth description of the Mann-Whitney U test; we highlighted its strength and pointed out some of the limitations of the technique. In section 3 we develop a new technique for nonparametric test, gave the algorithm for its implementation and compared the new procedure with the Mann-Whitney U test using two independent examples. Then, we examine the advantage of the new method over the traditional Mann-Whitney U test. Section 4 consists of results and discussion of results obtained from applying the two methods, while section 5 provides concluding remarks.

2. Mann-Whitney U test

The concept of testing nonparametric statistics where two independent groups have the same distribution and are homogeneous was proposed in [7]. The Mann-Whitney U test is established on the principle of comparing each variable from the first group with a corresponding variable from the second group, where the data need to be orderly then the data in each group are individually compared together. The test can be calculated with the following statistical formula

$$U_a = n_a n_b + \frac{n_a(n_a+1)}{2} - \sum R_a \quad (1)$$

and

$$U_b = n_b n_a + \frac{n_b(n_b+1)}{2} - \sum R_b \quad (2)$$

Corresponding Author**:** Babarinsa O., Email: olaiwola.babarinsa@fulokoja.edu.ng, Tel: +2348060032554, +2348036074881 (SDA)

where R_a and R_b are ranked scores and, n_a and n_b are the number in group A and B respectively. Let H_0 be the null hypothesis and H_1 the alternative hypothesis. Now, a case where two-tailed test is required, the alternative hypothesis against the null hypothesis to be tested specifies that the distribution of the first group data differs from the distribution of the second group data. Then, the null hypothesis is rejected if the values of the test statistic fall outside either of the tails of its sampling distribution. Contrarily, if one-tailed test is required, the alternative hypothesis (according to the test direction) indicates that the variable of one group is stochastically larger than the other group. Thus, the null hypothesis is rejected only for values of the test statistic falling into one tail of its sampling distribution [5].

Mann-Whitney U test has strengths and weaknesses. For the strengths, the Mann-Whitney U needs not depend on assumptions on the data distribution of the target population. When data do not meet the conditions of normality of the parametric assumptions of the t-test, the Mann-Whitney U tends to be more appropriate [4]. The test has higher-ranking than the t-test for non-normal populations. Furthermore, Mann-Whitney U test can be used when the sample is small, and the data are at least ordinal. Mann-Whitney U test is one of the most powerful nonparametric tests where the probability of rejecting a false null hypothesis has a close similarity to the statistical power. For average size samples between 10 and 20 observations, Mann-Whitney U provides significant results where it has approximately 95% of the Student's t-test statistical power [8]. The Mann-Whitney U test has less risk of giving a wrong significant result even when there is presence of few extreme values in the investigated sample [9].

However, Mann-Whitney U test has some limitations. First, the t-test is most of the time more powerful than the Mann-Whitney U test when applying Monte Carlo methods, even if the averages of the populations under investigation do not meet the criteria of normality [10]. This can lead to accepting the alternative hypothesis obtained from two populations with the same average but different variances [11]. In a case of distinct variances, the type I is amplified when Mann-Whitney U test is applied.

3. The new nonparametric method

In this section, we proposed a novelty method for the nonparametric test of two samples. The method makes use of combing ranking, where ties between the groups (not within the group) are considered. Like Mann-Whitney U test, the proposed method also ranks the data and considers the ties. If no ties occur between the groups, then zero is considered ($B = 0$). If more than one ties occurs then $B = \frac{b_1 + b_2 + b_3 + \cdots + b_n}{n}$ where $b_i$is the rank of ties common between groups for $i = 1,2,3 \cdots n$.

The steps for implementing the proposed method are as follows:

1. Arrange the two samples such that only ties from the two samples (between samples not within sample) are matched together
2. After the arrangement, rank the whole data
3. Pick the rank of the ties from the two samples
4. Get the mean of the rank of the ties from step 3 (denote it as B which is approximately normal)
5. Count the number of the data after the combine ranking, without repetition of counting the tie (denote this as n_b)
6. Calculate mean as $\mu = n_b + 1$ (3)
7. Calculate variance as $\sigma^2 = \frac{(n_b+1)(2n_b+1)}{20}$ (4)
8. Evaluate standard deviation as $\sigma = \sqrt{\frac{(n_b+1)(2n_b+1)}{20}}$ (5)
9. Calculate the value of Z as $Z = \frac{B-\mu}{\sigma}$ (6)

Before we proceed to examples, we shall refer to Mann-Whitney U test as Method I and the proposed method as Method II.

Example

1. Instructor A and B both teach a first course in chemistry at XYZ University. On a common final examination, their students received the grades shown in the table below

A	88	75	92	71	63	84	55	64	82	96				
B	72	65	84	53	76	80	51	60	57	85	94	87	73	61

Test at the 0.05 significance level the hypothesis that there is no difference between the two instructors grade [12]

Solution;

Method I: Using Mann-Whitney test

The calculated Z value is -0.9369 and the tabulated z value is -1.96 to 1.96.

Decision: There is no different at the 0.05 level

Method II: Using the proposed nonparametric method

Statement of Hypothesis

$H_0: \mu_1 = \mu_2$ (There is no difference between the two instructors 'grades)
$H_1: \mu_1 \neq \mu_2$ (There is difference between the two instructors 'grades)
$\alpha = 0.05$ (Two-tailed test)

Criterion
Reject the null hypothesis if calculated (6) is less than -1.96 or greater than 1.96

Calculation
Step 1

	Instructors	
S/N	A	B
1	88	
2	75	
3	92	
4	71	
5	63	
6	84	
7	55	
8	64	
9	82	
10	96	
11		72
12		65
13		84
14		53
15		76
16		80
17		51
18		60
19		57
20		85
21		94
22		87
23		73
24		61

Step 2

	Instructors		
S/N	A	B	Combine ranking
1	88		21
2	75		13
3	92		22
4	71		10
5	63		7
6	84	84	17.5
7	55		3
8	64		8
9	82		16
10	96		24
11		72	11
12		65	9
13		53	2
14		76	14
15		80	15
16		51	1
17		60	5
18		57	4
19		85	19
20		94	23
21		87	20
22		73	12
23		61	6

Step 3
The rank ties between the instructors (84 is common between A & B) is 17.5

Step 4
Since there is a rank ties between the samples. Thus, $B = 17.5$
Step 5
The initial number of the two samples was 24 but after the combined ranking, the number is reduced to 23. Thus $n_b = 23$
Step 6
The combined samples mean approximated to population mean $\mu = n_b + 1$
$\mu = 23 + 1 = 24$
Step 7
The variance $\sigma^2 = \frac{(n_b+1)(2n_b+1)}{20}$
Hence $\sigma^2 = \frac{(23+1)(2\times 23+1)}{20}$
$= 56.4$
Step 8
The standard deviation $\sigma = \sqrt{\frac{(n_b+1)(2n_b+1)}{20}}$
$\sigma = \sqrt{56.4}$
$= 7.5099$
Step 9
Thus, (6) $= \frac{17.5-24}{7.5099} = -0.8655$
Decision:
Since $Z = -0.8655$ falls between $z = \pm 1.96$, the null hypothesis must not be rejected. Thus, there is no difference between the instructors' grades at 0.05 level.

2. To find the best arrangement of instruments on a control panel of an airplane, 2 different arrangements were compared by simulating an emergency condition and measuring the reaction time required to correct the condition. The reaction times (in tenth of a second) of 20 pilots (randomly assigned to the 2 different arrangements) were as follows:

Arrangement I	8	15	10	13	17	10	9	11	12	15
Arrangement II	12	7	13	8	14	6	16	7	10	9

Use the U test at the 0.05 level of significance to check the claim that the second arrangement is better than the first [6]

Solution
Method I: Using Mann-Whitney U test
The calculated Z value is 1.2473 against the tabulated z value.
Decision: The second arrangement is better than the first arrangement.
Method II: Using the proposed nonparametric method
Statement of Hypothesis
$H_0: \mu_1 > \mu_2$ (Second arrangement is better than the first)
$H_1: \mu_1 < \mu_2$ (Second arrangement is not better than the first)
$\alpha = 0.05$ (One-tailed test)
Criterion
Reject the null hypothesis if tabulated value is greater than the (6)
Calculation
Step 1

S/N	Arrangements I	II
1	8	8
2	15	
3	10	
4	13	13
5	17	
6	10	10
7	9	9
8	11	
9	12	12
10	15	
11		7
12		14
13		6
14		16
15		7

Step 2

S/N	Arrangements I	II	Combine ranking	
1	8	8	4	b_1
2	15		12.5	
3	10		6.5	
4	13	13	10	b_2
5	17		15	
6	10	10	6.5	b_3
7	9	9	5	b_4
8	11		8	
9	12	12	9	b_5
10	15		12.5	
11		7	2.5	
12		14	11	
13		6	1	
14		16	14	
15		7	2.5	

Step 3

The ranks are 4, 10, 6.5, 5 and 9 which are represented by b_1, b_2, b_3, b_4 and b_5 respectively

Step 4

The mean of the ties ranked is $B = \frac{b_1 + b_2 + b_3 + b_4 + b_5}{5}$

$$B = \frac{4 + 10 + 6.5 + 5 + 9}{5} = 6.9$$

Step 5

The initial number of the two samples was 20 but after the combined ranking, the number is 15. Thus $n_b = 15$

Step 6

The combined samples means approximated to population mean $\mu = n_b + 1$

$\mu = 15 + 1 = 16$

Step 7

The variance $\sigma^2 = \frac{(n_b+1)(2n_b+1)}{20}$

Hence $\sigma^2 = \frac{(15+1)(2\times15+1)}{20} = 24.8$

Step 8

The standard deviation $\sigma = \sqrt{\frac{(n_b+1)(2n_b+1)}{20}}$

$\sigma = \sqrt{24.8} = 4.98$

Step 9

Thus (6) $= \frac{6.9-16}{4.98} = -1.8273$

Decision:

Since $Z = -1.8273$ is less than $z = 1.645$, the null hypothesis must not be rejected. That is the second arrangement is better than the first arrangement.

4. Results

We shall base our results on the example given above. In the example, the method that is farther away from the negative tail is Method II with 109.45% while Method I is 102.31%. This result shows that Method II is 7.14% better than Method I which tends toward normality. In the case of second example which is one tailed-test hypothesis, the method gives more accurate result as its ranks the samples within group with average of the rank ties being used.

5. Conclusion

The Mann-Whitney U test and the proposed method are similar. However, the versatility of the proposed method over Mann-Whitney U test is that the calculated Z tends toward normality of the data. The main advantage of the method is that its ranks the groups altogether unlike Mann-Whitney U test which ranks each group.

References

[1] Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics: A step-by-step approach*: John Wiley & Sons.

[2] Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric statistical methods* (Vol. 751): John Wiley & Sons.

[3] Lipschutz, S., & Schiller, J. J. (1998). *Schaum's Outline of Introduction to Probability and Statistics*. New York: McGraw Hill Professional.

[4] McKnight, P. E., & Najab, J. (2010). Mann- Whitney U Test. *The Corsini encyclopedia of psychology*, 1-1.

[5] Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in quantitative Methods for Psychology, 4*(1), 13-20.

[6] Johnson, R. A., Miller, I., & Freund, J. E. (2000). *Probability and statistics for engineers* (Vol. 2000): Pearson Education.

[7] Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.

[8] Landers, J. (1981). Quantification in History, Topic 4: Hypothesis Testing II- Differing Central Tendency: Oxford: All Souls College.

[9] Siegel, S. (1956). Nonparametric statistics for the behavioral sciences.
Spiegel, M. R., & Stephens, L. J. (2017). *Schaum's outline of statistics*. New York: McGraw Hill Professional.

[10] Zimmerman, D. W. (1987). Comparative power of Student t test and Mann-Whitney U test for unequal sample sizes and variances. *The Journal of Experimental Education, 55*(3), 171-174.

[11] Robert, C., & Casella, G. (2013). *Monte Carlo statistical methods*: Springer Science & Business Media.