Communications for Statistical Applications and Methods
2013, Vol. 20, No. 4, 241–246

DOI: http://dx.doi.org/10.5351/CSAM.2013.20.4.241

On the Study for the Simultaneous Test

Hyo-Il Park[1,a]

[a]Department of Statistics, Chongju University

Abstract

In this study, we propose a nonparametric simultaneous test procedure for the location translation and scale parameters. We consider the Wilcoxon rank sum test for the location translation parameter and the Mood test for the scale parameter with the quadratic and maximal types of combining functions. Then we derive the limiting null distributions of the combining functions. We illustrate our procedure with an example and compare efficiency by obtaining the empirical powers through a simulation study. Finally, we discuss some interesting features related to the nonparametric simultaneous tests.

Keywords: Combining function, location translation parameter, nonparametric test, permutation principle, scale parameter.

1. Introduction

One can use a non-parametric test such as the famous Wilcoxon rank sum test under the location translation assumption to compare two types of treatments or a treatment with control. In addition, one may consider applying the Ansari-Bradley test for the equality of variances or scales under some suitable conditions. However, one may sometimes consider to test the location translation and scale parameters simultaneously with the nonparametric approach. For this matter, Lepage (1971) proposed a nonparametric test procedure that compared the Wilcoxon and Ansari-Bradley statistics (Randles and Wolfe, 1979) with some discussions for the distributional aspects of the proposed statistic. Also, Lepage (1973) tabulated the exact critical values and significance levels for some selected sample sizes. Since then various modifications and new procedures have been reported and proposed by several authors (Murakami, 2007; Rublik, 2009; Neuhäuser *et al.*, 2011). Almost all the results have been based on the quadratic form for the combining function (Pesarin, 2001) to combine two kinds of test statistics to test the location translation and scale parameters.

To obtain the critical values or more generally p-values, one may consider the permutation principle (Good, 2000) for the exact null distribution of the chosen combining function for the small or reasonable sample sizes (Lepage, 1973). However, for a large sample case, it would be necessary to derive the limiting null distribution for the given test statistic through the large sample approximation theorem. In this study, we will obtain the limiting null distributions of the combining functions.

In this research, we consider to propose nonparametric simultaneous tests for the location translation and scale parameters using the Wilcoxon rank sum and Mood tests. The rest of this paper will be organized with the following order and content. In Section 2, we propose nonparametric simultaneous tests that combine two nonparametric statistics and derive the limiting null distributions with the large sample approximation theorem. Then we illustrate our procedure with a numerical example and compare the efficiency between the proposed tests by obtaining empirical powers through computer

This work was supported by the research grant of Chongju University in 2012.

[1] Professor, Department of Statistics, Chongju University, Chongju 360-764, Korea. E-mail: hipark@cju.ac.kr

simulation in Section 3. Finally, we discuss some interesting features related to the simultaneous tests in Section 4.

2. Simultaneous Test

Let $X_{11}, \ldots, X_{1n_1}$ and $X_{21}, \ldots, X_{2n_2}$ be two independent random samples from populations with continuous but unknown distribution functions F_1 and F_2, respectively. In the sequel, $n = n_1 + n_2$. In this study, we assume that we are interested in the following location-scale model such as

$$F_2(x) = F_1\left(\frac{x-\delta}{\eta}\right),$$

where $\delta \in (-\infty, \infty)$ and $\eta \in (0, \infty)$ are the location translation and scale parameters, respectively. Then our main interest would be to test

$$H_0 : \{\delta = 0\} \cap \{\eta = 1\} \quad \text{against} \quad H_1 : \{\delta \neq 0\} \cup \{\eta \neq 1\}. \tag{2.1}$$

To test the sub-null hypothesis $H_{10} : \delta = 0$, various nonparametric tests have been proposed that may be optimal according to the underlying distributions. To test the sub-null hypothesis $H_{20} : \eta = 1$, several nonparametric tests have also been proposed. In this study, we consider using the Wilcoxon rank sum test for $H_{10} : \delta = 0$ and the Mood test for $H_{20} : \eta = 1$ (Randles and Wolfe, 1979). Let W_n and M_n be the Wilcoxon rank sum statistic and Mood statistic, respectively. Then W_n and M_n can be defined as

$$W_n = \sum_{j=1}^{n_2} R_{2j} \mathrm{M}_n = \sum_{j=1}^{n_2} \left\{R_{2j} - \frac{n+1}{2}\right\}^2,$$

where R_{2j} is the rank of X_{2j} from the combined sample. Also let E_0 and V_0 be the mean and variance under H_0. Then we may propose a nonparametric test statistic T_n as follows:

$$T_n = \frac{(W_n - E_0(W_n))^2}{V_0(W_n)} + \frac{(M_n - E_0(M_n))^2}{V_0(M_n)}.$$

The component of T_n tends to have a large value when the corresponding H_{0i}, $i = 1, 2$, is not true; therefore, we may reject H_0 in favor of H_1 for large values of T_n. Then in order to obtain the critical value for any given significance level (or more generally, p-value) we have to derive the null distribution of T_n. For the small or reasonable sample sizes, one may obtain the null distribution with applying the permutation principle. However, for the large sample case, we have to obtain the asymptotic normality through the large sample approximation theory. First, we need the following simple results.

Lemma 1. *Under H_0, we have*

$$E_0(W_n) = \frac{n_2(n+1)}{2}, \qquad V_0(W_n) = \frac{n_1 n_2(n+1)}{12},$$
$$E_0(M_n) = \frac{n_2(n^2-1)}{12}, \qquad V_0(M_n) = \frac{n_1 n_2(2n+1)(8n+11)}{180}.$$

Then it is easy to show that with the continuity theorem (Bickel and Doksum, 1977),

$$\frac{(W_n - E_0(W_n))^2}{V_0(W_n)} \quad \text{and} \quad \frac{(M_n - E_0(M_n))^2}{V_0(M_n)}$$

converge in distribution to chi-square random variables with 1 degree of freedom.

Theorem 1. *Under $H_0 : \{\delta = 0\} \cap \{\eta = 1\}$, we have that $COV_0(W_n, M_n)$, the covariance between W_n and M_n is that*

$$COV_0(W_n, M_n) = 0.$$

Proof: First of all, we note that $\mathrm{COV}_0(W_n, M_n) = E_0(W_n M_n) - E_0(W_n)E_0(M_n)$. Then we have

$$\begin{aligned} W_n M_n &= \left(\sum_{j=1}^{n_2} R_{2j}\right)\left[\sum_{j=1}^{n_2}\left\{R_{2j} - \frac{n+1}{2}\right\}^2\right] \\ &= \left(\sum_{j=1}^{n_2} R_{2j}\right)\left(\sum_{j=1}^{n_2} R_{2j}^2\right) - (n+1)\left(\sum_{j=1}^{n_2} R_{2j}\right)^2 + \frac{n_2(n+1)^2}{4}\left(\sum_{j=1}^{n_2} R_{2j}\right). \end{aligned}$$

In addition, we note that

$$E_0(R_{21}) = \frac{1}{n}\sum_{i=1}^{n} i = \frac{n+1}{2}, \qquad E_0\left(R_{21}^2\right) = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{(n+1)(2n+1)}{6},$$

$$E_0\left(R_{21}^3\right) = \frac{1}{n}\sum_{i=1}^{n} i^3 = \frac{n(n+1)^2}{4}, \qquad E_0\left(R_{21}R_{22}^2\right) = \frac{1}{n(n-1)}\sum\sum_{i\neq j} i j^2 = \frac{n(n+1)^2}{6}.$$

Since $(\sum_{j=1}^{n_2} R_{2j})(\sum_{j=1}^{n_2} R_{2j}^2) = \sum_{j=1}^{n_2} R_{2j}^3 + \sum\sum_{i\neq j} R_{2i}R_{2j}^2$, we have

$$E_0\left\{\left(\sum_{j=1}^{n_2} R_{2j}\right)\left(\sum_{j=1}^{n_2} R_{2j}^2\right)\right\} = \frac{n_2(2n_2+1)n(n+1)^2}{12}.$$

Thus, we have

$$E_0(W_n M_n) = \frac{n_2^2(n+1)^2(n-1)}{24}.$$

From Lemma, we see that

$$E_0(W_n)E_0(M_n) = \frac{n_2^2(n+1)^2(n-1)}{24},$$

which completes the proof of Theorem 1. □

Theorem 2. *Under H_0, the limiting distribution of T_n is a chi-square distribution with 2 degrees of freedom.*

Then using Theorem 2, we may complete a simultaneous test procedure for any given significance level. However, since T_n is of the quadratic form, it would be difficult to apply the one-sided type of alternatives such as $H_{11} : \delta > 0$ or $H_{12} : \eta < 1$. In order to accommodate these partially one-sided alternatives, one may consider the following maximal type of statistic.

$$S_n = \max\left\{\frac{W_n - E_0(W_n)}{\sqrt{V_0(W_n)}}, \frac{M_n - E_0(M_n)}{\sqrt{V_0(M_n)}}\right\}.$$

Table 1: Kerosene heater data

Brand	Time in second
A	69.3 56.0 22.1 47.6 53.2 48.1 23.2 13.8 52.6 34.4 60.2 43.8
B	28.6 25.1 26.4 34.9 29.8 28.4 38.5 30.2 30.6 31.8 41.6 21.1 36.0 37.9 13.9

Theorem 3. *Under H_0, the limiting distribution of S_n is the product of two independent standard normal distributions with the same end value.*

Proof: For any real number s, first of all, we note that

$$\Pr\{S_n \le s\} = \Pr\left\{\frac{W_n - E_0(W_n)}{\sqrt{V_0(W_n)}} \le s, \frac{M_n - E_0(M_n)}{\sqrt{V_0(M_n)}} \le s\right\}.$$

We have the result since each component converges in distribution to a standard normal random variable and the two components are uncorrelated. □

3. A Numerical Example and Simulation Study

In this section, we first illustrate our procedure with a numerical example with the data in Milton and Arnold (2003) tabulated in Table 1. In this example, two brands of kerosene heaters are tested. The observations are the times in seconds required to raise the room temperature $10°F$. Moser and Stevens (1992) raised the question whether only the two-sample t-test would be appropriate for this data set by obtaining 0.0449 as p-value to test the equality of two variances. With this result, they concluded that rather than applying the usual two-sample t-test for this data set in Table 1, one should consider using the Satterthwaite test, which is a testing procedure for the mean difference when the two variances seem to unequal. Furthermore it seems that the symmetry of the data would be questionable, which is a serious drawback for applying the usual two-sample t-test. Therefore, it would be interesting to check the location translation and scale problems simultaneously with a nonparametric test procedure. Then to test (2.1), we have obtained 0.0024 and 0.0003 as p-values for simultaneous tests with T_n and S_n, respectively.

We now compare the performance between T_n and S_n by obtaining empirical powers through a simulation study. For this we consider four different distributions such as normal, Cauchy, double exponential and exponential with $(10, 10)$, $(10, 15)$ $(15, 10)$ pairs of the sample sizes. The values of (δ, η) varies from $(0.0, 1.0)$ to $(2.0, 3.0)$ with 0.5 increment for each component. The simulation has been repeated 10,000 times with SAS/IML PC version with the nominal significance level 0.05. The results are summarized in Table 2 through 5. We note that the nominal significance levels are over-estimated for T_n while under-estimated for S_n for all cases. With this in mind, it appears that S_n achieves more efficiency than T_n does for the normal and double-exponential cases. However the other two cases- Cauchy and exponential distributions-seem to be the same in their performances.

4. Some Concluding Remarks

We have obtained the p-values in the previous section for the example and empirical powers in the simulation study using the limiting distribution approach. In addition, one may obtain all the necessary probabilities through applying the permutation principle which is a re-sampling method. The permutation principle has been initiated by Fisher (1925) but the application to the statistics has begun relatively lately with the development of adequate computer facilities and corresponding software since the re-sampling method requires exceptionally burdensome computational tasks. However the permutation principle has been known to producing an exact procedure.

Table 2: Normal distribution

Test	(n_1, n_2)	(η, δ) (0.0, 1.0)	(0.5, 1.5)	(1.0, 2.0)	(1.5, 2.5)	(2.0, 3.0)
T_n	(10, 10)	0.0791	0.1647	0.2896	0.3943	0.4705
	(10, 15)	0.0816	0.1924	0.3332	0.4390	0.5133
	(15, 10)	0.0806	0.1723	0.3244	0.4390	0.5204
S_n	(10, 10)	0.0302	0.1379	0.2746	0.3919	0.4769
	(10, 15)	0.0253	0.1623	0.3159	0.5383	0.5243
	(15, 10)	0.0294	0.1371	0.2915	0.4209	0.5138

Table 3: Cauchy distribution

Test	(n_1, n_2)	(η, δ) (0.0, 1.0)	(0.5, 1.5)	(1.0, 2.0)	(1.5, 2.5)	(2.0, 3.0)
T_n	(10, 10)	0.0787	0.1092	0.1545	0.1988	0.2353
	(10, 15)	0.0813	0.1172	0.1727	0.2233	0.2676
	(15, 10)	0.0793	0.1037	0.1539	0.2040	0.2466
S_n	(10, 10)	0.0299	0.0775	0.1290	0.1794	0.2182
	(10, 15)	0.0266	0.0814	0.1407	0.1960	0.2448
	(15, 10)	0.0266	0.0717	0.1244	0.1733	0.2185

Table 4: Double exponential distribution

Test	(n_1, n_2)	(η, δ) (0.0, 1.0)	(0.5, 1.5)	(1.0, 2.0)	(1.5, 2.5)	(2.0, 3.0)
T_n	(10, 10)	0.0838	0.1497	0.2553	0.3448	0.4100
	(10, 15)	0.0795	0.1727	0.2961	0.3950	0.4611
	(15, 10)	0.0819	0.1578	0.2793	0.3861	0.4650
S_n	(10, 10)	0.0341	0.1233	0.2333	0.3348	0.4098
	(10, 15)	0.0267	0.1363	0.2737	0.3807	0.4648
	(15, 10)	0.0307	0.1191	0.2437	0.3607	0.4498

Table 5: Exponential distribution

Test	(n_1, n_2)	(η, δ) (0.0, 1.0)	(0.5, 1.5)	(1.0, 2.0)	(1.5, 2.5)	(2.0, 3.0)
T_n	(10, 10)	0.0811	0.1127	0.1392	0.1578	0.1714
	(10, 15)	0.0791	0.1323	0.1713	0.1928	0.2114
	(15, 10)	0.0821	0.1041	0.1318	0.1475	0.1614
S_n	(10, 10)	0.0284	0.0824	0.1197	0.1440	0.1630
	(10, 15)	0.0276	0.0993	0.1490	0.1786	0.2015
	(15, 10)	0.0320	0.0761	0.1121	0.1329	0.1522

In this study, we have used two types of combining functions such as the quadratic and maximal forms. In addition, one may use the summing type of combining function. However we did not consider the summing type of combining function in this study since it would produce the equivalent or asymptotically equivalent results with T_n by the fact that T_n is the sum of two squared statistics. For more discussion of the combining functions, you may refer to Pesarin (2001).

The non-correlatedness between W_n and M_n may resemble the independence between $\bar{X}$ and S^2 under the normality assumption. However, the probabilistic or distributional properties should be noted and discussed in a future study since it could provide insight into the inter-relation between two statistics for the location translation and scale parameters with disregard to the parametric or nonparametric approach.

A referee has recommended an increase of sample sizes to observe if the over- and under-estimation of the nominal significance level may be alleviated. Even for the case that $n_1 = n_2 = 100$, we have seen that the phenomenon does not disappear. This might come from the structural cause of the combinations of two individual tests.

Acknowledgement

The author would like to express his sincere appreciation to two anonymous referees to read carefully and suggest some constructive advices.

References

Bickel, P. J. and Doksum, K. A. (1977). *Mathematical Statistics-Basic Ideas and Selected Topics*, Holden-Day, San Francisco.

Fisher, R. A. (1925). *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh.

Good, P. (2000). *Permutation Tests-A Practical Guide to Resampling Methods for Testing Hypotheses*, Springer, New York.

Lepage, Y. (1971). A combination of Wilcoxon's and Ansari-Bradley's statistics, *Biometrika*, **58**, 213–217.

Lepage, Y. (1973). A table for a combined Wilcoxon Ansari-Bradley statistic, *Biometrika*, **60**, 113–116.

Milton, J. S. and Arnold, J. C. (2003). *Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences*, 4th ed. McGraw-Hill, Boston.

Moser, B. K. and Stevens, G. R. (1992). Homogeneity of variance in the two-sample means test, *The American Statistician*, **46**, 19–21.

Murakami, H. (2007). Lepage type statistic based on the modified Baumgartner statistic, *Computational Statistics and Data Analysis*, **51**, 5061–5067.

Neuhäuser, M., Leuchs, A.-K. and Ball, D. (2011). A new location-scale test based on a combination of the ideas of Levene and Lepage, *Biometrical Journal*, **53**, 525–534.

Pesarin, F. (2001). *Multivariate Permutation Tests with Applications in Biostatistics*, Wiley, New York.

Randles, R. H. and Wolfe, D. A. (1979). *Introduction to the Theory of Nonparametric Statistics*, Wiley, New York.

Rublik, F. (2009). Critical values for testing location-scale hypothesis, *Measurement Science Review*, **9**, 9–15.

Received October 22, 2012; Revised December 17, 2012; Accepted May 29, 2013