

VERSATILE TWO-SAMPLE RANK TESTS BASED ON WILCOXON TEST

Fueda, Kaoru
Graduate School of Mathematics, Kyushu University

Ohori, Katsumasa
Graduate School of Mathematics, Kyushu University

<https://doi.org/10.5109/13449>

出版情報 : Bulletin of informatics and cybernetics. 27 (1), pp.159-164, 1995-03. Research
Association of Statistical Sciences

バージョン :

権利関係 :



VERSATILE TWO-SAMPLE RANK TESTS BASED ON WILCOXON TEST

By

Kaoru FUEDA* and Katsumasa ÔHORI*

Abstract

This paper deals with a class of nonparametric two-sample tests for location, scale and more higher orders. The proposed tests are the extension of the Lepage [5] test for location and scale. It is shown that our tests are asymptotically normal and keep high powers in the situation where it is difficult to specify the alternatives.

Key words and phrases: Wilcoxon test, location and scale test, asymptotic normality, power of test.

1. Introduction

Two samples of m and n independent observations from populations with continuous distribution functions, $F(x)$ and $G(y)$ respectively, are considered. The problem is to test the hypothesis $F(x) \equiv G(x)$ versus alternative $F(x) \not\equiv G(x)$.

For testing for location or for scale, we have many two-sample tests (for example, see Hájek and Šidák [3]). Also Lepage [5][6][7][8], Puri and Sen [9], Hirakawa [4], Duran, Tsai and Lewis [1], Goria [2], Shiraishi [10][11] and others have proposed two-sample tests for location and scale. In this paper we consider a class of two-sample tests which can detect location, scale and higher order difference of the two distributions.

In section 2, we construct the test statistics. In section 3, we show the asymptotic normality of our statistics under null hypothesis and also under contiguous location and scale alternatives. In section 4, we perform simulations to compare the powers of our test with those of Van der Waerden's location test, Wilcoxon's location test, Klotz's scale test, Ansari-Bradley's scale test, and Lepage's location and scale test. Although our test is inferior to Van der Waerden, Wilcoxon and Lepage tests if the type of alternative is location shift only, and inferior to Klotz and Ansari-Bradley tests if the type of alternative is scale difference only, it is shown that our test is superior to those tests when it is difficult to specify the alternatives. Also it is indicated that our test is superior to Lepage's location and scale test if the difference is higher order than the location and scale such as kurtosis.

* Graduate School of Mathematics, Kyushu University 33, Fukuoka 812-81, Japan.
e-mail: fueda@math.kyushu-u.ac.jp

2. The test statistics

2.1. Definitions and Assumptions

Let X_1, \dots, X_m and Y_1, \dots, Y_n be observations from populations with continuous distribution functions $F(x)$ and $G(y)$, respectively. Setting $X_{m+j} = Y_j, j = 1, \dots, n$, and $N = m + n$, we denote a pooled sample by X_1, \dots, X_N . Let $R_{Ni} (i = 1, \dots, N)$ be the rank of the observation X_i in the ordered sequence $X^{(1)} < X^{(2)} < \dots < X^{(N)}$. We consider a simple linear rank statistics

$$T_N = \sum_{i=m+1}^N a_N(R_i),$$

and call vectors $\mathbf{a}_N = (a_N(1), \dots, a_N(N))$ the score vector.

DEFINITION 2.1. For $k \in \mathbf{N}$, we define

$\mathfrak{F}_k = \{\varphi | \varphi : [0, 1] \rightarrow \mathbf{R}, \int_0^1 \varphi(x)^{2k} dx < \infty, \int_0^1 (\varphi(x) - \bar{\varphi})^2 dx > 0\}$, where $\bar{\varphi} = \int_0^1 \varphi(x) dx$. Furthermore for $\varphi_1, \varphi_2 \in \mathfrak{F}_k$, we define inner product of φ_1 and φ_2 as $\langle \varphi_1, \varphi_2 \rangle = \int_0^1 \varphi_1(x) \varphi_2(x) dx$, and the norm of φ_1 as $\|\varphi_1\|^2 = \langle \varphi_1, \varphi_1 \rangle$.

Now fix $k \in \mathbf{N}$ and assume that for score vectors \mathbf{a}_N there exists $\varphi \in \mathfrak{F}_k$ such that $\lim_{N \rightarrow \infty} \int_0^1 (a_N(1 + [xN]) - \varphi(x))^2 dx = 0$, where $[x]$ denotes the largest integer not exceeding x . This assumption is not extraordinary. For location testing problem such that hypothesis $G(x) \equiv F(x)$ versus alternative $G(x) \equiv F(x + \theta)$, where $\theta \neq 0$, the score vectors \mathbf{a}_N of the locally most powerful rank test satisfy $\lim_{N \rightarrow \infty} \int_0^1 (a_N(1 + [xN]) - \varphi(x))^2 dx = 0$ for $\varphi(x) = -f'(F^{-1}(x))/f(F^{-1}(x))$, where f is the density function, and for many well known underlying distributions, $-f'(F^{-1}(x))/f(F^{-1}(x)) \in \mathfrak{F}_1$.

2.2. Construction of the test statistics

For $\varphi \in \mathfrak{F}_k$, we construct the sequence of functions $\psi_0(x), \psi_1(x), \psi_2(x), \dots, \psi_k(x)$ from the sequence of functions $\varphi(x)^0 \equiv 1, \varphi(x), \varphi(x)^2, \dots, \varphi(x)^k$ by Gram-Schmidt orthonormalization. That is

$$\begin{aligned} \psi_0(x) &= \varphi(x)^0 \equiv 1 \\ \tilde{\psi}_1(x) &= \varphi(x) - \langle \varphi, \psi_0 \rangle \psi_0(x) \\ \psi_1(x) &= \frac{\tilde{\psi}_1(x)}{\|\tilde{\psi}_1\|} \\ &\vdots \\ \tilde{\psi}_k(x) &= \varphi(x)^k - \sum_{i=0}^{k-1} \langle \varphi^k, \psi_i \rangle \psi_i(x) \\ \psi_k(x) &= \frac{\tilde{\psi}_k(x)}{\|\tilde{\psi}_k\|} \end{aligned}$$

Note that $\langle \psi_0, \psi_i \rangle = 0$ means $\int_0^1 \psi_i(x) dx = 0$ for $i = 1, \dots, k$.

Next we take score vectors \mathbf{b}_{Nj} such that

$$\lim_{N \rightarrow \infty} \int_0^1 (b_{Nj}(1 + [xN]) - \psi_j(x))^2 dx = 0 \quad (j = 1, \dots, k).$$

For example, $b_{Nj}(i) = \int_{(i-1)/N}^{i/N} \psi_j(x) dx$, $b_{Nj}(i) = \psi_j(i/(N+1))$, or $b_{Nj}(i) = E[\psi_j(U_N^{(i)})]$, where $U_N^{(1)} < U_N^{(2)} < \dots < U_N^{(N)}$ are an ordered sample from the uniform distribution on $[0, 1]$. Now we propose $T_{Nj} = \sum_{i=m+1}^N b_{Nj}(R_{Ni})$ as the test statistics for our testing problem.

3. Asymptotic normality of the statistics

THEOREM 3.1. Put $\bar{b}_{Nj} = \frac{1}{N} \sum_{i=1}^N b_{Nj}(i)$. Then under the null hypothesis, the statistic

$$\begin{pmatrix} (T_{N1} - n\bar{b}_{N1})/\sqrt{\frac{mn}{N}} \\ \vdots \\ (T_{Nk} - n\bar{b}_{Nk})/\sqrt{\frac{mn}{N}} \end{pmatrix} \text{ are asymptotically normal } N \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \right)$$

for $\min(m, n) \rightarrow \infty$.

PROOF. By exact calculation, we have

$$\begin{aligned} E(T_{Nj}) &= n\bar{b}_{Nj} \quad (j = 1, \dots, k), \\ \text{Var}(T_{Nj}) &= \frac{mn}{N} \left(\frac{1}{N} \sum_{i=1}^N b_{Nj}(i)^2 - \bar{b}_{Nj}^2 \right) \quad (j = 1, \dots, k), \\ \text{Cov}(T_{Nj_1}, T_{Nj_2}) &= \frac{mn}{N} \left(\frac{1}{N} \sum_{i=1}^N b_{Nj_1}(i)b_{Nj_2}(i) - \bar{b}_{Nj_1}\bar{b}_{Nj_2} \right) \quad (j_1 \neq j_2). \end{aligned}$$

And $\lim_{N \rightarrow \infty} \int_0^1 (b_{Nj}(1 + [xN]) - \psi_j(x))^2 dx = 0 \quad (j = 1, \dots, k)$ implies

$$\begin{aligned} \lim_{N \rightarrow \infty} \bar{b}_{Nj} &= \lim_{N \rightarrow \infty} \int_0^1 (b_{Nj}(1 + [xN]) dx \\ &= \int_0^1 \psi_j(x) dx = 0 \quad (j = 1, \dots, k), \\ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N b_{Nj}(i)^2 &= \lim_{N \rightarrow \infty} \int_0^1 (b_{Nj}(1 + [xN])^2 dx \\ &= \int_0^1 \psi_j(x)^2 dx = 1 \quad (j = 1, \dots, k), \\ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N b_{Nj_1}(i)b_{Nj_2}(i) &= \lim_{N \rightarrow \infty} \int_0^1 (b_{Nj_1}(1 + [xN])(b_{Nj_2}(1 + [xN]) dx \\ &= \int_0^1 \psi_{j_1}(x)\psi_{j_2}(x) dx = 0 \quad (j_1 \neq j_2). \end{aligned}$$

Thus, $\lim_{N \rightarrow \infty} \text{Var}(T_{Nj}) = 1$ for $j = 1, \dots, k$, and $\lim_{N \rightarrow \infty} \text{Cov}(T_{Nj_1}, T_{Nj_2}) = 0$ for $j_1 \neq j_2$. Asymptotic normality is shown by Theorem V.1.5.b of Hájek and Šidák [3].

THEOREM 3.2. *Let X_1, \dots, X_m and X_{m+1}, \dots, X_{m+n} be observation from populations with continuous distribution functions, $F(x)$ and $G(y)$ respectively, and $G(x) \equiv F(\frac{x - \alpha_N}{e^{\beta_N}})$, where $N = m + n$. Assume that F has a differentiable density f , and*

$$\lim_{N \rightarrow \infty} (\alpha_N, \beta_N) = (0, 0),$$

$$\lim_{N \rightarrow \infty} I(f) \frac{mn}{N} \alpha_N^2 = c_1^2, 0 < c_1 < \infty,$$

$$\lim_{N \rightarrow \infty} I_1(f) \frac{mn}{N} \beta_N^2 = c_2^2, 0 < c_2 < \infty,$$

$$\text{where } I(f) = \int_0^1 \varphi^2(x, f) dx, \quad I_1(f) = \int_0^1 \varphi_1^2(x, f) dx,$$

$$\varphi(x, f) = -\frac{f'(F^{-1}(x))}{f(F^{-1}(x))}, \quad \varphi_1(x, f) = -1 - F^{-1}(x) \frac{f'(F^{-1}(x))}{f(F^{-1}(x))}, \quad (0 < x < 1).$$

Then the statistic

$$\begin{pmatrix} (T_{N1} - n\bar{b}_{N1})/\sqrt{\frac{mn}{N}} \\ \vdots \\ (T_{Nk} - n\bar{b}_{Nk})/\sqrt{\frac{mn}{N}} \end{pmatrix} \text{ are asymptotically normal } N \left(\begin{pmatrix} \mu_{Nj} \\ \vdots \\ \mu_{Nj} \end{pmatrix}, \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \right)$$

for $\min(m, n) \rightarrow \infty$. Where $\mu_{Nj} = \sqrt{\frac{mn}{N}} \int_0^1 \psi_j(x) (\alpha_N \varphi(x, f) + \beta_N \varphi_1(x, f)) dx$.

This theorem is shown by theorem 3.1 of this paper and theorem VI.2.3 of Hájek and Šidák [3].

4. Comparisons of power by simulations

We conducted simulations to compare the powers of our tests with Van der Waerden's location test, Wilcoxon's location test, Klotz's scale test, Ansari-Bradley's scale test, and Lepage's location and scale test. Among the family of our tests, we considered the test $\sum_{i=1}^4 ((T_{Ni} - n\bar{b}_{Ni})/\sqrt{\frac{mn}{N}})^2$ with $\varphi(x) = x$ and $b_{Nj}(i) = E[\psi_j(U_N^{(i)})]$, where $U_N^{(1)} < U_N^{(2)} < \dots < U_N^{(N)}$ are an ordered sample from the uniform distribution on $[0, 1]$. $\varphi(x) = x$ means that T_{N1} is equivalent to Wilcoxon's location test. The following four alternative distributions are considered in the simulations.

Alternative 1: F is normal distribution with mean=0, variance=1, G is normal distribution with mean=0.5 variance=1.

Alternative 2: F is normal distribution with mean=0, variance=1, G is normal distribution with mean=0 variance= $\sqrt{2}$.

Alternative 3: F is logistic distribution with mean=0, variance=1, G is logistic distribution with mean=0.5 variance=1.

Alternative 4: F is normal distribution with mean=0.5, variance=1, G is t distribution with 2 degrees of freedom. This G is different from F in location, scale and kurtosis.

The following table lists the powers of the test at level 0.05 estimated by 1000 trials when $m = n = 50$.

| | Alternative 1 | Alternative 2 | Alternative 3 | Alternative 4 |
|---|---------------|---------------|---------------|---------------|
| Van der Waerden's one sided location test | 0.790 | 0.047 | 0.812 | 0.640 |
| Wilcoxon's one sided location test | 0.767 | 0.054 | 0.834 | 0.672 |
| Klotz's one sided scale test | 0.034 | 0.716 | 0.045 | 0.670 |
| Ansari-Bradley's one sided scale test | 0.039 | 0.582 | 0.043 | 0.329 |
| Lepage's location and scale test | 0.576 | 0.356 | 0.631 | 0.553 |
| Our test | 0.470 | 0.379 | 0.487 | 0.702 |

The table shows that

1. Van der Waerden's test and Wilcoxon's test have high powers for location alternatives, but have poor powers for the other alternatives. Specially for scale alternative, the powers of those test are about 0.05, because these tests are aimed for location test only.
2. Klotz's test and Ansari-Bradley's test have high powers for scale alternative, but have poor powers for the other alternatives. Specially for location alternative, the powers of those test are poor than 0.05, because these tests are again aimed for scale test only.
3. Our test and Lepage's test keep high powers for all alternatives, and our test have highest power for alternative 4 in this simulation. Thus it is indicated that our test is superior to Lepage's test if the difference is in higher order than location and scale such as kurtosis.
4. In the above simulation one-sided test powers are computed for the Van der Waerden, Wilcoxon, Klotz and Ansari-Bradley tests. However, in the situations where it is not easy to specify the alternatives, two-sided tests must be carried out. If these tests were modified to be two-sided tests, the powers of these tests are poorer than those values in the table. Thus our test might be useful in those situation where it is not easy to specify the alternatives.

Acknowledgement

I would like to thank Professor T. Yanagawa for his useful ideas and advices.

References

- [1] Duran, B. S., Tsai, W. S. and Lewis, T. O. : *A class of location-scale nonparametric tests*. *Biometrika*. **63** (1976), 173-76.
- [2] Gorla, M. N. : *A survey of two-sample location-scale problem. Asymptotic relative efficiencies of some rank tests*. *Statistica Neerlandica*, **36** (1982), 3-13.
- [3] Hájek, J. and Šidák, Z. : *Theory of Rank Tests*, New York: Academic Press. (1967).
- [4] Hirakawa, K. : *The comparison of powers of distribution-free two-sample tests*. *TRU Mathematics*, **10** (1974), 65-82.
- [5] Lepage, Y. : *A Combination of Wilcoxon's and Ansari-Bradley's statistics*. *Biometrika*. **58**, (1971) 213-217.
- [6] Lepage, Y. : *A table for a combined Wilcoxon Ansari-Bradley statistics*. *Biometrika*. **60** (1973), 113-6.
- [7] Lepage, Y. : *Asymptotically optimum rank tests for contiguous location and scale alternatives*. *Commun. Statist.* **4** (1975), 671-87.
- [8] Lepage, Y. : *A class of nonparametric tests for location and scale parameters*. *Commun. Statist.* **A.6** (1977), 649-659.
- [9] Puri, M. L. and Sen, P. K. : *Nonparametric Methods in Multivariate Analysis*. Wiley, New York. (1971).
- [10] Shiraishi, T. : *Rank tests for ordered location-scale alternatives*. *J. Japan Statist. Soc.* **18** (1988), NO. 1.
- [11] Shiraishi, T. : *Multivariate multi-sample rank tests for location scale alternatives*. *Commun. Statist.* **18** (1989), 85-105.

Received December 27, 1994

Revised January 20, 1995