AISM (2006) 58: 93–100
DOI 10.1007/s10463-005-0019-3

**Nariaki Sugiura · Hidetoshi Murakami · Seong Keon Lee
Yasutomo Maeda**

Biased and unbiased two-sided Wilcoxon tests for equal sample sizes

Received: 4 October 2004 / Published online: 15 March 2006

Abstract This paper answers the long-standing question of whether the two-sided Wilcoxon rank test for equal sample sizes is unbiased against a location parameter family of distributions by giving a counterexample. It is also shown that the non-randomized two-sided Wilcoxon test for equal sample sizes with the least positive significance level is unbiased.

Keywords Rank test · Unbiasedness · Two-sample problem · Power curve · Location parameter · Beta distribution

1 Introduction

Lehmann (1959, p 240; 1986, p 322) has shown that the one-sided Wilcoxon rank test is unbiased against one-sided location parameter family of distributions under a more general setting. He then raised the question of whether the two-sided Wilcoxon test is unbiased against two-sided location parameter family of distributions. Sugiura (1965) gave a counterexample for the case of unequal sample sizes. It is given by nonrandomized Wilcoxon test with the least positive significance level against exponential distributions. However, when the sample sizes are equal, the test is shown to be unbiased. Then the same question for equal sample sizes was

N. Sugiura · H. Murakami · S.K. Lee · Y. Maeda
Graduate School of Science and Engineering, Chuo University,
1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

N. Sugiura (✉)
12-26 Onogawa, Tsukuba, Ibaraki 305-0053, Japan
E-mail: sugiura@math.tsukuba.ac.jp

Y. Maeda
Research Associate, Institute of Statistics, Korea University,
Anam-dong, Sungbuk-ku, Seoul 136–701, Korea

raised by Professor E.L. Lehmann to the first author in his personal communication just after Sugiura (1965). Amrhein (1995) also gave a counterexample showing that the two-sided Wilcoxon signed rank test for one-sample problem is biased against a two-sided location parameter family of symmetric distributions. It is provided for the discrete distributions with continuous component.

In this paper, we shall show that two-sided Wilcoxon test is not unbiased against two-sided parameter family of distributions even when two sample sizes are equal. The counterexample is provided by the nonrandomized two-sided Wilcoxon test with the second smallest positive significance level when the distributions are generated by beta distribution with singularity at the origin.

2 Unbiased and biased two-sided Wilcoxon tests

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be random samples of equal size n from absolutely continuous distribution functions $F(x)$ and $G(x)$ with respect to Lebesgue measure and their density functions be $f(x)$ and $g(x)$, respectively. We consider a location parameter family of distributions defined by $G(x) = F(x - \Delta)$, where Δ is an unknown location parameter. The hypotheses to be tested are $H\colon\ \Delta = 0$ against alternatives $K\colon\ \Delta \neq 0$, where the functional form of $F(x)$ is unknown. Let $X_{(1)} < \cdots < X_{(n)}$ be order statistics obtained from Xs and similarly $Y_{(1)} < \cdots < Y_{(n)}$ from Ys. Note that the event $X_1, \ldots, X_n < Y_1, \ldots, Y_n$ is equivalent to $X_{(n)} < Y_{(1)}$ and that $Y_1, \ldots, Y_{\mathrm{n}} < X_1, \ldots, X_n$ to $Y_{(n)} < X_{(1)}$. We shall consider the following two-sided Wilcoxon tests, ϕ_1 and ϕ_2:

$$\phi_1(X_1, \ldots, X_n; Y_1, \ldots, Y_n) = \begin{cases} 1 & \text{if } X_{(n)} < Y_{(1)} \text{ or } Y_{(n)} < X_{(1)} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

and

$$\phi_2(X_1, \ldots, X_n; Y_1, \ldots, Y_n) = \begin{cases} 1 & \begin{array}{l} \text{if } X_{(n)} < Y_{(1)} \text{or } Y_{(n)} < X_{(1)} \\ \text{or } X_{(n-1)} < Y_{(1)} < X_{(n)} < Y_{(2)} \\ \text{or } Y_{(n-1)} < X_{(1)} < Y_{(n)} < X_{(2)} \end{array} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The test ϕ_1 has the smallest positive significance level $\alpha = 2(n!)^2/(2n)!$ among the nonrandomized two-sided Wilcoxon tests. The level of the test ϕ_2 is $\alpha = 4(n!)^2/(2n)!$ which is the second smallest. The power function of the test ϕ_1 is written by

$$\begin{aligned} \beta_1(\Delta) &= P(X_{(n)} < Y_1, \ldots, Y_n) + P(Y_1, \ldots, Y_n < X_{(1)}) \\ &= n \int_{-\infty}^{\infty} (1 - G(x))^n F(x)^{n-1} f(x)\, \mathrm{d}x \\ &\quad + n \int_{-\infty}^{\infty} G(x)^n (1 - F(x))^{n-1} f(x)\, \mathrm{d}x \end{aligned} \tag{3}$$

and that of ϕ_2 by

$$\begin{aligned}\beta_2(\Delta) &= \beta_1(\Delta) + nP(X_{(n-1)} < Y_1 < X_{(n)} < Y_2, \dots, Y_n) \\ &\quad + nP(Y_1, \dots, Y_{n-1} < X_{(1)} < Y_n < X_{(2)}) \\ &= \beta_1(\Delta) + n^2(n-1) \iint\limits_{-\infty<x<y<\infty} (G(y) - G(x))(1 - G(y))^{n-1} \\ &\quad \times F(x)^{n-2} f(x) f(y)\, \mathrm{d}x\mathrm{d}y + n^2(n-1) \\ &\quad \times \iint\limits_{-\infty<x<y<\infty} G(x)^{n-1}(G(y) - G(x))(1 - F(y))^{n-2} f(x) f(y)\, \mathrm{d}x\mathrm{d}y.\end{aligned} \tag{4}$$

First, we shall note the symmetry of the power function of these tests in the following theorem.

Theorem 1 *The power functions are symmetric with respect to* $\Delta = 0$*, namely, we have* $\beta_1(\Delta) = \beta_1(-\Delta)$ *and* $\beta_2(\Delta) = \beta_2(-\Delta)$.

Proof Put $Y_i - \Delta = Z_i$ for $i = 1, \dots, n$. Then $X_1, \dots, X_n$ and $Z_1, \dots, Z_n$ are all independent and $Z_1, \dots, Z_n$ have the same joint distribution as that of $X_1, \dots, X_n$. Noting that the order statistics $Y_{(i)} - \Delta = Z_{(i)}$ and that $(X_{(1)}, Z_{(n)})$ and $(Z_{(1)}, X_{(n)})$ are exchangeable, we can write

$$\begin{aligned}\beta_1(\Delta) &= P(X_{(n)} < Z_{(1)} + \Delta) + P(Z_{(n)} + \Delta < X_{(1)}) \\ &= P(X_{(n)} < Z_{(1)} + \Delta) + P(X_{(n)} + \Delta < Z_{(1)}),\end{aligned} \tag{5}$$

from which we can see that $\beta_1(\Delta) = \beta_1(-\Delta)$. For the test ϕ_2, we can rewrite

$$\begin{aligned}\beta_2(\Delta) &= \beta_1(\Delta) + P(X_{(n-1)} < Z_{(1)} + \Delta < X_{(n)} < Z_{(2)} + \Delta) \\ &\quad + P(Z_{(n-1)} + \Delta < X_{(1)} < Z_{(n)} + \Delta < X_{(2)}).\end{aligned}$$

Note that $(X_{(1)}, X_{(2)}, Z_{(n-1)}, Z_{(n)})$ and $(Z_{(1)}, Z_{(2)}, X_{(n-1)}, X_{(n)})$ are exchangeable. We can rewrite the third term in R.H.S., giving

$$\begin{aligned}&= \beta_1(\Delta) + P(X_{(n-1)} < Z_{(1)} + \Delta < X_{(n)} < Z_{(2)} + \Delta) \\ &\quad + P(X_{(n-1)} < Z_{(1)} - \Delta < X_{(n)} < Z_{(2)} - \Delta),\end{aligned} \tag{6}$$

which implies that $\beta_2(\Delta) = \beta_2(-\Delta)$.

From Theorem 1, we may consider the power function only for $\Delta > 0$. By the same argument as in the proof of Theorem 1, we can see that the symmetry of the power function is extended to $\beta_{m,n}(\Delta) = \beta_{n,m}(-\Delta)$ for the two-sided Wilcoxon test for unequal sample sizes m and n, the special case of which was shown in Sugiura (1965). In particular, the derivative of the power function of the general two-sided Wilcoxon test for equal sample sizes at $\Delta = 0$ vanishes, if it is differentiable. This may be a support for the unbiasedness of the two-sided Wilcoxon test for equal sample sizes against many distributions. This is the case with the test ϕ_1.

Theorem 2 *Assume that the derivative of the power function* $\beta_1(\Delta)$ *is obtainable by differentiation under the integral sign. Then the test* ϕ_1 *is unbiased against the location parameter family of distributions.*

Proof By the assumption, we can differentiate the power function under the integral sign, giving

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}\beta_1(\Delta) = n^2\int_{-\infty}^{\infty} f(x)g(x)\Big\{[(1-G(x))F(x)]^{n-1} - [G(x)(1-F(x))]^{n-1}\Big\}\,\mathrm{d}x$$

$$= n^2\int_{-\infty}^{\infty} f(x)g(x)\,(F(x)-G(x))$$

$$\times\Big\{[(1-G(x))F(x)]^{n-2}+\cdots+[G(x)(1-F(x))]^{n-2}\Big\}\,\mathrm{d}x. \qquad (7)$$

Note that $F(x) - G(x) = F(x) - F(x-\Delta)$ is nonnegative if $\Delta > 0$ and nonpositive if $\Delta < 0$. Combined with expression (7), we can see that $\beta_1(\Delta) \geq \beta_1(0)$ for any Δ, which completes the proof. □

Sugiura (1965) gave an example of biased two-sided Wilcoxon test in a form of ϕ_1 for unequal sample sizes. If the density function $f(x)$ is bounded, the assumption in Theorem 2 is satisfied by the dominated convergence theorem. However, the following example shows that the test ϕ_1 is unbiased against a wider location parameter family of distributions. We now specify a family of unbounded distributions

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^{\rho} & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases} \quad \text{and} \quad f(x) = \begin{cases} \rho x^{\rho-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad (8)$$

for $0 < \rho < 1$. Assume that $0 \leq \Delta \leq 1$, then dividing the domain of integration in Eq. 3 to the disjoint intervals $(0, \Delta)$ and $(\Delta, 1)$, we have

$$\beta_1(\Delta) = \Delta^{\rho n} + n\rho\int_{\Delta}^{1} (1-(x-\Delta)^{\rho})^n x^{\rho n-1}\mathrm{d}x$$

$$+ n\rho\int_{\Delta}^{1} (x-\Delta)^{\rho n}(1-x^{\rho})^{n-1}x^{\rho-1}\mathrm{d}x. \qquad (9)$$

The derivative of the power function is written by

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}\beta_1(\Delta) = n^2\rho^2\int_{\Delta}^{1} \Big\{(1-(x-\Delta)^{\rho})^{n-1}(x-\Delta)^{\rho-1}x^{\rho n-1}$$

$$-(x-\Delta)^{\rho n-1}(1-x^{\rho})^{n-1}x^{\rho-1}\Big\}\mathrm{d}x$$

$$= n^2\rho^2\int_{\Delta}^{1} x^{\rho-1}(x-\Delta)^{\rho-1}\Big\{(1-(x-\Delta)^{\rho})^{n-1}x^{\rho(n-1)}$$

$$-(x-\Delta)^{\rho(n-1)}(1-x^{\rho})^{n-1}\Big\}\mathrm{d}x. \qquad (10)$$

The term in the curly bracket in the integrand in the last expression in Eq. 10 is factorized to

$$\{x^\rho - (x-\Delta)^\rho\}[\{(1-(x-\Delta)^\rho)x^\rho\}^{n-2} + \cdots + \{(x-\Delta)^\rho(1-x^\rho)\}^{n-2}], \tag{11}$$

which is positive since $\Delta > 0$. Hence $\beta_1(\Delta)$ is a strictly increasing function for $0 \le \Delta \le 1$. We note that $\beta_1(1) = 1$, $\beta_1(0) = \alpha$ and $(\mathrm{d}/\mathrm{d}\Delta)\beta_1(0+0) = 0$. It is clear that $\beta_1(\Delta) = 1$ for $\Delta > 1$. Combined with Theorem 1, we can say that the test ϕ_1 is unbiased against the location parameter family generated by Eq. 8.

We shall now show that the test ϕ_2 is biased. For numerical computation, the expression of the power function (9) is not appropriate because the integral becomes unstable near at $\Delta = 0$. A useful expression is given by putting $x^\rho = z$ in each integral,

$$\beta_1(\Delta) = \Delta^{\rho n} + n\int_{\Delta^\rho}^{1}(1-(z^{1/\rho}-\Delta)^\rho)^n z^{n-1}\mathrm{d}z + n\int_{\Delta^\rho}^{1}(z^{1/\rho}-\Delta)^{\rho n}(1-z)^{n-1}\mathrm{d}z. \tag{12}$$

When $0 < \Delta < 1$, the power function of the test ϕ_2 is written by dividing the domain of integration in Eq. 4 into three disjoint intervals given by $0 < \Delta < x < y < 1, 0 < x < \Delta < y < 1, 0 < x < y < \Delta < 1$ and changing the variables (x,y) to (u, v) by $x^\rho = u$ and $y^\rho = v$,

$$\begin{aligned}\beta_2(\Delta) = {} & \beta_1(\Delta) + n^2(n-1) \\ & \times \iint_{\Delta^\rho<u<v<1} \{(v^{1/\rho}-\Delta)^\rho - (u^{1/\rho}-\Delta)^\rho\}\{1-(v^{1/\rho}-\Delta)^\rho\}^{n-1}u^{n-2}\mathrm{d}u\mathrm{d}v \\ & + n^2\Delta^{\rho(n-1)}\int_{\Delta^\rho}^{1}(v^{1/\rho}-\Delta)^\rho\{1-(v^{1/\rho}-\Delta)^\rho\}^{n-1}\mathrm{d}v \\ & + n^2(n-1)\iint_{\Delta^\rho<u<v<1}(u^{1/\rho}-\Delta)^{\rho(n-1)} \\ & \times\left\{(v^{1/\rho}-\Delta)^\rho - (u^{1/\rho}-\Delta)^\rho\right\}(1-v)^{n-2}\mathrm{d}u\mathrm{d}v.\end{aligned} \tag{13}$$

Note that $\beta_2(1) = 1$ and that clearly $\beta_2(\Delta) = 1$ for $\Delta > 1$. Expressions (13) greatly save the computing time for numerical integrations by Mathematica 4.2.

For $n = 2$, the significance level of Wilcoxon test ϕ_2 is 2/3. Put $\rho = 0.16$ and $\Delta = 0.01$, for example. Then numerical integrations based on Eqs. 12 and 13, yield $\beta_2(0.01) = 0.655287$, showing that the test ϕ_2 is biased. Moreover, we can show the global behavior of the power function in the neighborhood of

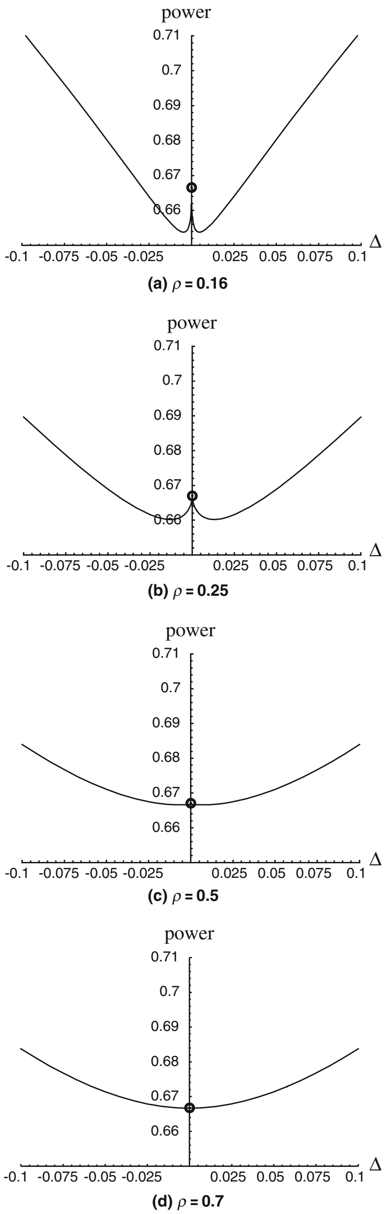

Fig. 1 Power function of two-sided Wilcoxon test ϕ_2 for n=2 with level 2/3

$\Delta = 0$ in Fig. 1a. The parameter $\rho = 0.16$ is so chosen that the value of the power function at $\Delta = 0.01$ is the smallest. The test ϕ_2 is biased also for $\rho = 0.25$ and slightly biased for $\rho = 0.5$, since $\beta_2(0.01) = 0.666605$. However, it is unbiased for $\rho = 0.7$. These graphs of the power function are shown in Fig. 1c, d. The graph of Fig. 1c, d look almost the same.

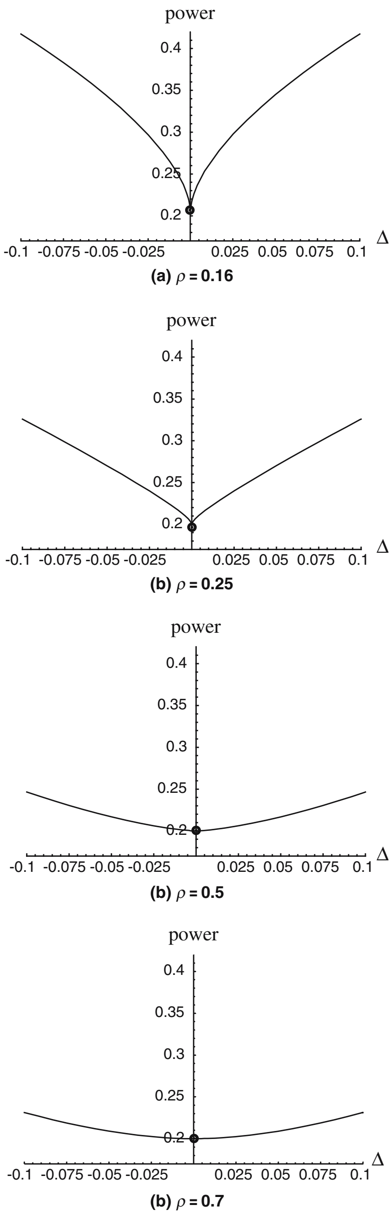

Fig. 2 Power function of two-sided Wilcoxon test ϕ_2 for n=3 with level 1/5

It might be instructive to note that for $n = 3$, these distributions no longer yield a biased Wilcoxon test ϕ_2 as shown in Fig. 2. We failed to find the value of ρ such that the test ϕ_2 for $n = 3$ is biased.

Numerical computations in this paper are performed by Mathematica 4.2, with Power Mac G4/400.

References

Amrhein, P. (1995). An example of a two-sided Wilcoxon signed rank test which is not unbiased. *Annals of the Institute of Statistical Mathematics*, *47*, 167–170.
Lehmann, E.L. (1959). *Testing statistical hypotheses*. New York: Wiley
Lehmann, E.L. (1986). *Testing statistical hypotheses* (2nd ed). New York: Wiley
Sugiura, N. (1965). An example of the two-sided Wilcoxon test which is not unbiased. *Annals of the Institute of Statistical Mathematics*, *17*, 261–263.