TSUKUBA J. MATH.
Vol. 10 No. 1 (1986). 57–61

OPTIMUM PROPERTIES OF THE WILCOXON SIGNED RANK TEST UNDER A LEHMANN ALTERNATIVE

By

Taka-aki SHIRAISHI

1. Introduction.

Let $X_1, \cdots, X_n$ be a random sample from an absolutely continuous distribution function $F(x)$. The problem is to test the null hypothesis $H: F(x)=G(x)$ where $G'(x)=g(x)$ is assumed to be symmetric about zero. When $G(x)$ is a logistic distribution function, Hájek and Šidák [1] reviewed that the Wilcoxon signed rank test is locally most powerful among all rank tests against the location alternative $A: F(x)=G(x-\theta)$ for $\theta>0$ and showed that the test is asymptotically optimum under the contiguous sequence of alternatives $A_n: F(x)=G(x-\Delta/\sqrt{n})$ for some $\Delta>0$.

In this paper, we consider the alternative of the contaminated distribution

$$K: F(x)=(1-\theta)G(x)+\theta\{G(x)\}^2 \qquad \text{for } 0<\theta<1. \tag{1.1}$$

The alternative K was introduced by Lehmann [2] for a two-sample problem. In order to get an asymptotic optimum property, we consider the sequence of alternatives

$$K_n: F(x)=(1-\Delta/\sqrt{n})G(x)+(\Delta/\sqrt{n})\{G(x)\}^2 \qquad \text{for } \Delta>0, \tag{1.2}$$

which is included in K and approaches the null hypothesis H as $n\to\infty$. In the following Section, we shall show that the Wilcoxon signed rank test is locally most powerful among all rank tests under K and is asymptotically most powerful under K_n. Further in Section 3, we shall compare the Wilcoxon signed rank test with the one-sample t-test by the asymptotic relative efficiency under the contiguous sequence of alternatives of general contaminated distributions

$$K_n': F(x)=(1-\Delta/\sqrt{n})G(x)+(\Delta/\sqrt{n})H(G(x)) \qquad \text{for } \Delta>0. \tag{1.3}$$

2. Optimum properties.

Taking the absolute values of observations, let R_i be the rank of $|X_i|$ among the observations $\{|X_i|; i=1, \cdots, n\}$ and define $\operatorname{sign} X=1$ for $X>0$, 0 for $X=0$ and

Received April 8, 1985. Revised June 27, 1985.

-1 otherwise. Note that $\Pr\{\text{sign } X_i=0\}=0$ since we consider only absolutely continuous distribution. Then we can describe the Wilcoxon signed rank statistic as the following.

$$T=\sum_{i=1}^{n}(\text{sign } X_i)R_i \,. \tag{2.1}$$

At first, we investigate the property of "locally most powerful".

THEOREM 1. *The Wilcoxon signed rank test based on T defined by* (2.1) *is locally most powerful for H versus K defined by* (1.1) *among all rank tests.*

PROOF. Putting $\text{sign } \underline{X}=(\text{sign } X_1, \cdots, \text{sign } X_n)$ and $\underline{R}=(R_1, \cdots, R_n)$, we get for any vector $\underline{v}=(v_1, \cdots, v_n)$ such that $v_i=1$ or -1 and any permutation $\underline{r}=(r_1, \cdots, r_n)$ of $(1, \cdots, n)$, under H, $\Pr\{\text{sign } \underline{X}=\underline{v}\}=1/2^n$ and $\Pr\{\underline{R}=\underline{r}\}=1/n!$. Here since the likelihood function of $(X_1, \cdots, X_n)$ under K is given by

$$p_\theta(x)=\prod_{i=1}^{n}\{(1-\theta)g(x_i)+2\theta G(x_i)g(x_i)\}\,, \tag{2.2}$$

the joint probability of sign vector $\text{sign } \underline{X}$ and rank vector $\underline{R}$ is expressed by

$$\begin{aligned}\beta(\theta)&=P_\theta\{\text{sign } \underline{X}=\underline{v}, \underline{R}=\underline{r}\}\\ &=\int\cdots\int_{\text{sign } \underline{X}=\underline{v}, \underline{R}=\underline{r}} p_\theta(x)d\underline{x}\\ &=1/(2^n\cdot n!)+\sum_{j=1}^{n}\int\cdots\int_{\text{sign } \underline{X}=\underline{v}, \underline{R}=\underline{r}} \prod_{k=1}^{j-1} g(x_k) \prod_{k=j+1}^{n}\{(1-\theta)g(x_k)+2\theta G(x_k)g(x_k)\}\\ &\quad\times[\{(1-\theta)g(x_j)+2\theta G(x_j)g(x_j)\}-g(x_j)]d\underline{x}\,,\end{aligned}$$

It follows that

$$\beta'(0)=\sum_{j=1}^{n}\int\cdots\int_{\text{sign } \underline{X}=\underline{v}, \underline{R}=\underline{r}} \{-1+2G(x_j)\}\prod_{k=1}^{n}g(x_k)d\underline{x}\,.$$

Let $|X|^{(i)}$ be the i-th order statistic among the absolute values $\{|X_i|; i=1, \cdots, n\}$. Since $|X|^{(i)}$, $\text{sign } \underline{X}$ and $\underline{R}$ are mutually independent under H from II 1.3 theorem of Hájek and Šidák [1], we can get

$$\begin{aligned}\beta'(0)&=1/(2^n\cdot n!)\cdot\sum_{j=1}^{n}E\{-1+2G(v_j|X|^{(r_j)})\}\\ &=1/(2^n\cdot n!)\cdot\sum_{j=1}^{n}E[v_j\{2G(|X|^{(r_j)})-1\}]\\ &=1/(2^n\cdot n!)\cdot\sum_{j=1}^{n}v_jr_j/(n+1)\,,\end{aligned}$$

which implies the result.

Next we shall show the asymptotic optimum property. Corresponding to (2.2), the joint density of $(X_1, \cdots, X_n)$ under K_n defined by (1.2) is given by

$$q_\Delta(\underline{X})=\prod_{i=1}^{n}[\{(1-\Delta/\sqrt{n})+2\Delta G(X_i)/\sqrt{n}\}g(X_i)]$$

THEOREM 2. *The asymptotic power of the Wilcoxon signed rank test is equal to that of the most powerful test for H versus K_n when Δ and G(u) are known, having critical region* $\{\underline{x}\,;\log\{q_\Delta(\underline{x})/p(\underline{x})\}\geqq t_{n\alpha}\}$.

PROOF. Taylor's series expansion of the logarithm of the likelihood ratio yields

$$\begin{aligned}(2.3)\qquad L_\Delta&=\log\{q_\Delta(X)/p(X)\}\\&=\log\left[\prod_{i=1}^{n}\{(1-\Delta/\sqrt{n})+2\Delta G(X_i)/\sqrt{n}\}\right]\\&=(\Delta/\sqrt{n})\sum_{i=1}^{n}\{2G(X_i)-1\}-\Delta^2/(2n)\cdot\sum_{i=1}^{n}\{2G(X_i)-1\}^2\\&\quad+\Delta^3/(3n\sqrt{n})\sum_{i=1}^{n}[\{2G(X_i)-1\}^3/\{1+\delta_i(\Delta/\sqrt{n})(2G(X_i)-1)\}^3],\end{aligned}$$

where δ_i satisfies $0<\delta_i<1$. Under the null hypothesis H, the first term of the last expression of (2.3), namely $(\Delta/\sqrt{n})\sum_{i=1}^{n}\{2G(X_i)-1\}$, has asymptotically a normal distribution with mean 0 and variance $\Delta^2/3$ by the central limit theorem, the second term converges to $-\Delta^2/6$ in probability by the law of large numbers and the third term tends to zero in probability.

Thus we get

$$(2.4)\qquad L_\Delta \xrightarrow[\mathfrak{L}]{} N(\mu, \sigma^2)\,,$$

where $\xrightarrow[\mathfrak{L}]{}$ denotes convergence in law and

$$(2.5)\qquad \mu=-\Delta^2/6 \quad\text{and}\quad \sigma^2=-2\mu\,.$$

From VI 1.2 corollary of Hájek and Šidák [1], the family of densities $\{q_\Delta(x)\}$ is contiguous to $\{p(x)\}$. So from LeCam's third lemma stated in VI 1.4 of Hájek and Šidák [1], under $\{q_\Delta(x)\}$, $L_\Delta \xrightarrow[\mathfrak{L}]{} N(-\mu, \sigma^2)$, where μ and σ^2 are defined by (2.5).

Therefore the asymptotic power of the test of level α with critical region $L_\Delta>t_{n\alpha}$ under $\{q_\Delta(x)\}$ is

$$(2.6)\qquad 1-\Phi(z_\alpha-\Delta/\sqrt{3})\,,$$

where $t_{n\alpha}=-\Delta^2/6+z_\alpha\Delta/\sqrt{3}+o(1)$, $\Phi(\cdot)$ is a distribution function of the standard normal and z_α is the upper 100α percentage point of the standard normal distribution. On the other hand, let us put $S=\sum_{i=1}^{n}(\text{sign}\, X_i)\{2G(|X_i|)-1\}/\sqrt{n}$, then $T/\{(n+1)\sqrt{n}\}-S$ converges to zero in probability under H from V 1.7 theorem of Hájek and Šidák [1]. Hence $(L_\Delta, T/\{(n+1)\sqrt{n}\})$ and (L_Δ, S) have asymptotically the same normal distribution. Also it follows under H that (L_Δ, S) has asymptotically a bivariate normal distribution with mean $(\mu, 0)$ and singular covariance matrix $\begin{pmatrix}\sigma^2, & \sigma_{12} \\ \sigma_{12}, & \sigma_2^2\end{pmatrix}$, where μ and σ^2 are defined by (2.5), $\sigma_{12}=\Delta/3$ and $\sigma_2^2=1/3$. Hence, from LeCam's third lemma, under $\{q_\Delta(x)\}$, we get that S has asymptotically the normal distribution with mean σ_{12} and variance σ_2^2. Thus the asymptotic power of the test based on T for H versus K_n at level α is given by the expression (2.6). This completes the proof.

3. Comparison with the *t*-test under a contiguous sequence of alternatives of general contaminated distributions.

We extend K_n defined by (1.2) to the contiguous sequence of alternatives of general contaminated distributions K_n' defined by (1.3) and compare the Wilcoxon signed rank test with the t-test based on

$$U=\sqrt{n-1}\sum_{i=1}^{n}X_i\Big/\sqrt{n\sum_{i=1}^{n}(X_i-\bar{X})^2}\,. \tag{3.1}$$

Then we get

THEOREM 3. *Suppose that the derivative of $H(u)$ exists and the derivative $h(u)=H'(u)$ is bounded. Then the asymptotic relative efficiency of the Wilcoxon signed rank test with respect to the t-test based on U under K_n' defined by* (1.3) *is given by*

$$\text{ARE}\,(T, U)=3\sigma^2\left\{\int_0^1(2u-1)h(u)du\right\}^2\Big/\left\{\int_{-\infty}^{\infty}th(G(t))g(t)dt\right\}^2$$

where $\sigma^2=\int_{-\infty}^{\infty}t^2dG(t)$.

PROOF. From the straight similar way to the proof of Theorem 2, $\sqrt{3}\cdot T/\{(n+1)\sqrt{n}\}$ has asymptotically a normal distribution with mean $\sqrt{3}\,\Delta\int_0^1(2u-1)h(u)du$ and variance 1. Further the similar argument as in the proof of Theorem 2 shows that U defined by (3.1) has asymptotically a normal distribution with mean $\Delta\int_{-\infty}^{\infty}th(G(t))dt/\sigma$ and variance 1 under K_n'. The ratio of squares of the two asymptotic means gives the result.

This asymptotic relative efficiency (ARE) equals the ARE of the two-sample

Wilcoxon test with respect to the two-sample t-test under a contiguous sequence of alternatives of contamitated distributions which is given by corollary 2 of Shiraishi [3]. So we find that this ARE is 1 for any bounded function $h(u)$ if $G(x)$ is the distribution function from the uniform random variable on a finite interval. In Table 1 of Shiraishi [3], we showed the values of this ARE for $H(u)=u^k$, $1-(1-u)^k$ with $k=1.1, 1.3, 1.6, 2, 3, 5, 10$ and $G(x)=$uniform, normal, logistic, double exponential distributions. As the numerical results, ARE's are always nearly equal to 1 irrespective of the form of $H(u)$, $G(x)$ and k chosen.

4. Conclusion.

About the exact power, Theorem 1 gives an admissibility of the Wilcoxon signed rank test for the alternative of contaminanted distribution $F(x)=(1-\theta)G(x)+\theta H(G(x))$, which includes K defined by (1.1), as far as we intend to seek a test having higher exact power among all rank tests. Though we found that there does not exist asymptotically a most powerful rank test under a contiguous sequence of alternatives of contaminated distributions for the two-sample problem from corollary 1 of Shiraishi [3], Theorem 2 shows that the Wilcoxon signed rank test is asymptotically most powerful for K_n defined by (1.2) which is included by K'_n. Further we find that the numerical values of ARE of the Wilcoxon signed rank test with respect to the t-test stated by Theorem 3 give no loss of the relative efficiency even against the alternative hypothesis of contaminated distributions discussed in Section 3.

References

[1] Hájek, J. and Šidák, Z., Theory of Rank Tests. Academic Press, New York, 1967.
[2] Lemann, E. L., The power of rank tests. Ann. Math. Statist., **24** (1953), 23–43.
[3] Shiraishi, T., Local powers of two-sample and multi-sample rank tests for Lehmann's contaminated alternative. Ann. Inst. Statist. Math., **37** (1985), 519–527.

Institute of Mathematics
University of Tsukuba
Sakura-mura, Niihari-gun
Ibaraki, 305 Japan