The Korean Communications
in Statistics Vol. 10 No. 2, 2003
pp. 535-543

Ranked-Set Sample Wilcoxon Signed Rank Test For Quantiles Under Equal Allocation

Dong Hee Kim[1] and Hyun Gee Kim[2]

Abstract

A ranked set sample version of the sign test is proposed for testing hypotheses concerning the quantiles of a population characteristic by Kaur, et. al(2002). In this paper, we proposed the ranked set sample Wilcoxon signed rank test for quantiles under equal allocation. We obtain the asymptotic property and the asymptotic relative efficiencies of the proposed test statistic with respect to Wilcoxon signed rank test of simple random sample for quantiles under equal allocation. We calculate the ARE of test statistics, the proposed test statistic is more efficient than simple random sampling for all quantiles. The relative advantage of ranked set sampling is greatest at the median and tapers off in the tails.

Keywords : Wilcoxon signed rank test, ranked-set sampling, Asymptotic relative efficiency, equal allocation.

1. Introduction

Many authors discuss the sampling method, which is called the Ranked-Set Sampling(RSS) method. This sampling method is introduced by McIntyre(1952). The RSS method is useful when measurements are destructive or expensive or invasive, while ranking of the observation is relatively easy. McIntyre(1952) suggested the RSS method for assessing yields of pasture plots without actually mowing and weighing the hay for a large number of plots. Takahasi and Wakimoto(1968) proposed the same procedure and developed an extensive body of theory that is, however, based upon the assumption of perfect ranking, Dell and Clutter(1972) considered a useful technique for improving estimates of the mean in the situation that measurements of observations are not perfectly judged. Stokes(1980) investigated the estimation of a population variance and asymptotic relative efficiency on the RSS. In one

1) Professor, Department of Statistics, Pusan National University, Pusan 609-735, Korea.
E-mail : heekim@pusan.ac.kr
2) Instructor, Department of Statistics, Pusan National University, Pusan 609-735, Korea.
E-mail : hgeekim@pusan.ac.kr

sample location problem, Hettmansperger(1995) studied the sign test on RSS. Koti and Babu(1996) calculated the exact null distribution of the ranked-set sample sign test statistic. Kim and Kim(1996) studied the Wilcoxon signed rank test on RSS. In two sample location problem, Bohn and Wolfe(1992, 1994) proposed the Mann-Whitney-Wilcoxon statistic and investigated the properties of the test procedures based on RSS for perfect and imperfect judgements. Öztürk(1999a,b) studied the one- and two-sample sign tests for RSS and extended one sample sign test to two sample sign test on RSS. Kaur et al.(2002) studied ranked set sample sign test for quantiles under equal and unequal allocation, search the optimal allocation under unequal allocation. Their research is more efficient than simple random sample sign test for quantiles.

In this paper, we discuss the ranked-set sampling method Wilcoxon signed rank test for quantiles under equal allocation. In the quantile problem, we will propose test statistic using RSS Wilcoxon signed rank test to have more efficient than simple random sampling(SRS) for the Wilcoxon signed rank test. In section 2, we introduce the test statistic and study the asymptotic property of the test statistic. Section 3 shows the asymptotic relative efficiency(ARE) of the proposed test statistic with respect to Wilcoxon signed rank test on SRS. In section 4, we have conclusions and further work.

2. Test Statistic

2.1 Review of Ranked-Set Sampling

Now we introduce the Ranked-Set Sampling method. A set of k items is from the population, the items of the set are ranked by judgment and only the item ranked the smallest is quantified, but $k-1$ remaining observations are returned to the population. The another set of size k is drawn and ranked and only the item ranked the second smallest is quantified. This process is repeated until we measure k ordered observations ; $X_{(1)1}, \cdots, X_{(k)1}$, where $X_{(j)1}$ is the measurement taken on the item judged to be rank j among k observations and the second subscript indicates the first cycle. This sampling method is called the RSS and is explained in Figure 1.

$$\begin{array}{cccccccc}
X_{(11)} & X_{(12)} & \cdots & X_{(1(k-1))} & X_{(1k)} & \Rightarrow & X_{(1)1} \\
X_{(21)} & X_{(22)} & \cdots & X_{(2(k-1))} & X_{(2k)} & \Rightarrow & X_{(2)1} \\
\vdots & \vdots & & \vdots & \vdots & \Rightarrow & \vdots \\
X_{(k1)} & X_{(k2)} & \cdots & X_{(k(k-1))} & X_{(kk)} & \Rightarrow & X_{(k)1}
\end{array}$$

[Figure 1] Display of k measurement observations of the first cycle(RSS)

2.2 The Proposed Test Statistic

We consider one sample location problem on RSS. Let $X_{(1)1}, \cdots, X_{(1)n}, \cdots, X_{(k)1}, \cdots, X_{(k)n}$ be a RSS of size nk as shown in Figure 1. $X_{(j)i}$ are independent and identically distributed for fixed j.

We repeat the process for n cycles to get nk measurements.

$$
\begin{array}{lcl}
X_{(1)1}, \cdots, X_{(1)n} & \sim & iid \;\; h_{(1)}(t) \\
X_{(2)1}, \cdots, X_{(2)n} & \sim & iid \;\; h_{(2)}(t) \\
\vdots & & \vdots \\
X_{(k)1}, \cdots, X_{(k)n} & \sim & iid \;\; h_{(k)}(t)
\end{array}
$$

The density function $h_{(j)}(t)$ represents the pdf of the j-th order statistic of size k from a distribution with pdf $h(t)$ and cdf $H(t)$ and $H_{(j)}(t)$ is cdf and its density function is given by;

$$
\begin{aligned}
h_{(j)}(t) &= \frac{k!}{(j-1)!(k-j)!} H^{j-1}(t)[1-H(t)]^{k-j} h(t), \\
H_{(j)}(t) &= \int_{-\infty}^{t} h_{(j)}(x)\,dx.
\end{aligned}
$$

We suppose the population is distributed as $H(x) = F(x-\xi_p)$ and wish to test for $H_0: \xi_p = 0$ against $H_A: \xi_p \neq 0$. For $p=0.5$, this hypothesis reduces to a test of the median. Wilcoxon signed rank test is a very useful test for one sample location problem. We propose the test statistic using Wilcoxon signed rank test on RSS for quantiles.

The proposed statistic is given by

$$
W_{RSS}^{+} = \sum_{j=1}^{k} \sum_{i=1}^{n} i \cdot T_{(j)i}.
$$

where $T_{(j)i} = \sum_{l=i}^{n} I(X_{(j)i} + X_{(j)l} > 0)$.

Let $V_{(j)}^{+} = \sum_{i=1}^{n} i \cdot T_{(j)i} / \binom{n}{2}$. We consider $V_{(j)}^{+}$ for convenience calculations instead of W_{RSS}^{+}.

The $V^+_{(j)}$ can be written as

$$\begin{aligned} V^+_{(j)} &= \frac{1}{\binom{n}{2}} \sum_{i=1}^{n} \sum_{l=i}^{n} I(X_{(j)i} + X_{(j)l} > 0) \\ &= \frac{1}{\binom{n}{2}} \sum_{i=1}^{n} I(X_{(j)i} > 0) + \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{l=i+1}^{n} I(X_{(j)i} + X_{(j)l} > 0) \\ &= \frac{2}{n-1} U_1 + U_2, \end{aligned}$$

where $U_1 = \frac{1}{n} \sum_{i=1}^{n} I(X_{(j)i} > 0)$ and $U_2 = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{l=i+1}^{n} I(X_{(j)i} + X_{(j)l} > 0)$.

First of all, we evaluate the expectation of $V^+_{(j)}$,

$$\begin{aligned} E(V^+_{(j)}) &= E\left[\frac{2}{n-1} U_1 + U_2\right] \\ &= \frac{2}{n-1} P(X_{(j)i} > 0) + P(X_{(j)1} + X_{(j)2} > 0), \end{aligned}$$

where $P(X_{(j)1} > 0) = 1 - H_{(j)}(0)$ and

$$P(X_{(j)1} + X_{(j)2} > 0) = \int_{-\infty}^{\infty} [1 - H_{(j)}(-x - 2\xi_p)] dH_{(j)}(x).$$

The expectation of proposed test statistic, W^+_{RSS},

$$\begin{aligned} E(W^+_{RSS}) &= \binom{n}{2} \sum_{j=1}^{k} E(V^+_{(j)}) \\ &= n \sum_{j=1}^{k} (1 - H_{(j)}(0)) \\ &\quad + \binom{n}{2} \sum_{j=1}^{k} \int_{-\infty}^{\infty} (1 - H_{(j)}(-x - 2\xi_p)) dH_{(j)}(x). \end{aligned}$$

We evaluate the variance of statistic W^+_{RSS},

$$\begin{aligned} Var(W^+_{RSS}) &= \left[\binom{n}{2}\right]^2 \sum_{j=1}^{k} Var(V^+_{(j)}) \\ &= n \sum_{j=1}^{k} (1 - H_{(j)}(0)) H_{(j)}(0) \\ &\quad + \binom{n}{2} \sum_{j=1}^{k} P(X_{(j)1} + X_{(j)2} > 0)(1 - P(X_{(j)1} + X_{(j)2} > 0)) \\ &= nk(1 - H(0)) H(0) \delta^2 \\ &\quad + \binom{n}{2} \sum_{j=1}^{k} P(X_{(j)1} + X_{(j)2} > 0)[1 - P(X_{(j)1} + X_{(j)2} > 0)], \end{aligned}$$

where $\delta^2 = 1 - \sum_{j=1}^{k} (H_{(j)}(0) - H(0))^2 / kH(0)(1 - H(0))$

With k fixed and as $n \to \infty$,

$$\sqrt{nk}\left[\, W^{+}_{RSS} - nk(1-H(0)) - \binom{n}{2}\sum_{j=1}^{k}\int_{-\infty}^{\infty}(1-H_{(j)}(-x-2\xi_p))dH_{(j)}(x)\right]$$
$$\sim N\left(0, H(0)(1-H(0))\delta^2 + \frac{n-1}{2k}\sum_{j=1}^{k}P_{(j)12}\cdot[1-P_{(j)12}]\right),$$

where $P_{(j)12} = P(X_{(j)1} + X_{(j)2} > 0)$.

The following theorem provides the mean and variance of the test statistic, W^{+}_{RSS}, on the RSS under the null hypothesis.

Theorem 1 Under $H_0: \xi_p = 0,\ H(0) = F(0) = p$,

$$E_0(W^{+}_{RSS}) = nk(n+1)/4,\ Var_0(W^{+}_{RSS}) = kn(n+1)(2n+1)\,\delta^2_{RSS_p}/24,$$

where $\delta^2_{RSS_p} = 1 - \sum_{j=1}^{k}(H_{(j)}(0)-p)^2/kp(1-p)$

Proof. The expectation and variance formula follow from the fact that $\frac{1}{k}\sum_{j=1}^{k}F_{(j)}(0)$ $= F(0) = p$, and $F_{(j)}(0)$ is given by the change of variable $u = F(t)$.

3. Asymptotic Relative Efficiency

We compare the test under equal allocation using the criterion of the Pitman's asymptotic relative efficiency. For given quantile, we compare the Wilcoxon signed rank statistic on RSS, W^{+}_{RSS}, for quantile under equal allocation will be made with corresponding test under SRS, W^{+}_{SRS}. The efficacy of test statistic is given by;

$$eff(T) = \lim_{n\to\infty}\frac{\partial/\partial\xi_p E(T)}{\sqrt{nkVar(T)}}\,|H_0.$$

Under $H_0: \xi_p = 0$, in order to compute the efficiency of W^{+}_{RSS}, we evaluate

$$E_{\xi_p}\left(\sum_{j=1}^{k}V^{+}_{(j)}\right) = \sum_{j=1}^{k}\left[\frac{2}{n-1}(1-F_{(j)}(-\xi_p)) + \int_{-\infty}^{\infty}(1-F_{(j)}(-x-2\xi_p))dF_{(j)}(x)\right].$$

The derivative of $E_{\xi_p}(\sum_{j=1}^{k}V^{+}_{(j)})$ evaluated at $\xi_p = 0$ is

$$\frac{\partial E_{\xi_p}\left(\sum_{j=1}^{k}\sum_{j=1}^{k}V^{+}_{(j)}\right)}{\partial\xi_p}\,|_{\,\xi_p=0} = \sum_{j=1}^{k}\left(\frac{2}{n-1}f_{(j)}(0) + 2\int_{-\infty}^{\infty}f^2_{(j)}(x)dx\right).$$

Hence.

$$\frac{\partial E_{\xi_p}(W^+_{RSS})}{\partial \xi_p}\Big|_{\xi_p=0} = \binom{n}{2}\sum_{j=1}^{k}\left(\frac{2}{n-1}f_{(j)}(0)+2\int_{-\infty}^{\infty}f^2_{(j)}(x)\,dx\right).$$

Then the efficacy of W^+_{RSS} is

$$\begin{aligned} eff(W^+_{RSS}) &= \lim_{n\to\infty}\frac{\frac{n(n-1)}{2}\sum_{j=1}^{k}\left(\frac{2}{n-1}f_{(j)}(0)+2\int_{-\infty}^{\infty}f^2_{(j)}(x)\,dx\right)}{\sqrt{nk\frac{nk(n+1)(2n+1)}{24}\delta^2_{RSS_p}}} \\ &= \frac{2\sqrt{3}\sum_{j=1}^{k}\int_{-\infty}^{\infty}f^2_{(j)}(x)\,dx}{k\delta_{RSS_p}} \end{aligned}$$

Note that the efficacy of W^+_{SRS} is

$$eff(W^+_{SRS}) = 2\sqrt{3}\int_{-\infty}^{\infty}f^2(x)\,dx.$$

The asymptotic relative efficiency of W^+_{RSS} with respect to W^+_{SRS} is

$$\begin{aligned} ARE(W^+_{RSS},\ W^+_{SRS}) &= \frac{eff^2(W^+_{RSS})}{eff^2(W^+_{SRS})} \\ &= \left[\frac{\sum_{j=1}^{k}\int_{-\infty}^{\infty}f^2_{(j)}(x)\,dx}{k\delta_{RSS_p}\int_{-\infty}^{\infty}f^2(x)\,dx}\right]^2, \end{aligned}$$

where

$$\begin{aligned} f^2_{(j)}(x) &= \left[\frac{k!}{(j-1)!(k-j)!}\right]^2 F^{2(j-1)}(x)(1-F(x))^{2(k-j)}f^2(x) \\ &= \left[\frac{k!}{(j-1)!(k-j)!}\right]^2 F^{2(j-1)}\sum_{d=0}^{2(k-j)}(-1)^d\binom{2(k-j)}{d}F^d(x)f^2(x). \end{aligned}$$

Then we consider the asymptotic relative efficiencies for the uniform, double exponential and logistic distributions.

First, for uniform distribution,

$$\begin{aligned} &ARE(W^+_{RSS},\ W^+_{SRS}) \\ &= \left[\frac{\sum_{j=1}^{k}\left[\frac{k!}{(j-1)!(k-j)!}\right]^2\sum_{d=0}^{2(k-j)}(-1)^d\binom{2(k-j)}{d}\frac{1}{2j-1+d}}{k\delta_{RSS_p}}\right]^2. \end{aligned}$$

Second, for double exponetial distribution,

$ARE(W^{+}_{RSS},\ W^{+}_{SRS})$

$$= \left[\frac{4 \sum_{j=1}^{k} \left[\frac{k!}{(j-1)!(k-j)!} \right]^2 \sum_{d=0}^{2(k-j)} (-1)^d \binom{2(k-j)}{d} \frac{1-(\frac{1}{2})^{2j+d-1}}{(2j-1+d)(2j+d)}}{k\delta_{RSS_p}} \right]^2 .$$

Third, for logistic distribution,

$ARE(W^{+}_{RSS},\ W^{+}_{SRS})$

$$= \left[\frac{6 \sum_{j=1}^{k} \left[\frac{k!}{(j-1)!(k-j)!} \right]^2 \sum_{d=0}^{2(k-j)} (-1)^d \binom{2(k-j)}{d} \frac{1}{(2j+1+d)(2j+d)}}{k\delta_{RSS_p}} \right]^2 .$$

In order to calculate the ARE, we tabulated δ_{RSS_p} for different values of p and for $k=2(1)5,10$ in Table 1.

[Table 1] Values of δ_{RSS_p} for $p=0.05,0.1,0.2,0.5,0.2,0.9,0.95$ and $k=2(1)5,10$

k	p						
	0.05	0.1	0.2	0.5	0.8	0.9	0.95
2	0.976	0.954	0.917	0.866	0.917	0.954	0.976
3	0.954	0.914	0.855	0.791	0.855	0.914	0.954
4	0.933	0.880	0.807	0.740	0.807	0.880	0.933
5	0.914	0.851	0.771	0.702	0.771	0.851	0.914
10	0.835	0.746	0.659	0.592	0.659	0.746	0.835

[Table 2] Values of $ARE(W^{+}_{RSS}, W^{+}_{SRS})$ for uniform distribution.

k	p						
	0.05	0.1	0.2	0.5	0.8	0.9	0.95
2	1.866	1.954	2.116	2.371	2.116	1.954	1.866
3	2.815	3.062	3.501	4.097	3.501	3.062	2.815
4	3.841	4.315	5.134	6.114	5.134	4.315	3.841
5	4.944	5.705	6.944	8.387	6.944	5.705	4.944
10	11.541	14.465	18.565	22.960	18.565	14.465	11.541

[Table 3] Values of $ARE(W^{+}_{RSS}, W^{+}_{SRS})$ for double exponential distribution

k	p 0.05	0.1	0.2	0.5	0.8	0.9	0.95
2	1.429	1.496	1.620	1.815	1.620	1.496	1.429
3	1.930	2.100	2.401	2.809	2.401	2.100	1.930
4	2.481	2.787	3.316	3.949	3.316	2.787	2.481
5	3.076	3.549	4.320	5.218	4.320	3.549	3.076
10	6.634	8.315	10.672	13.198	10.672	8.315	6.634

[Table 4] Values of $ARE(W^+_{RSS}, W^+_{SRS})$ for logistic distribution

k	p 0.05	0.1	0.2	0.5	0.8	0.9	0.95
2	1.512	1.583	1.714	1.920	1.714	1.583	1.512
3	2.068	2.249	2.572	3.009	2.572	2.249	2.068
4	2.667	2.996	3.565	4.246	3.565	2.996	2.667
5	3.309	3.819	4.648	5.614	4.648	3.819	3.309
10	7.125	8.929	11.461	14.174	11.461	8.929	7.125

When the underlying distribution is uniform, double exponential, logistic distribution, the values of ARE is in Table 2, 3 and 4. It is observed that ranked set sampling is more efficient than simple random sampling for all quantiles. The relative advantage of RSS is greatest at the median and tapers off in the tails. Increasing sample size k further enhances the performance of RSS.

4. Conclusion and Further Work

In one sample problem, we proposed Wilcoxon signed rank test for quantiles on RSS. It is observed that ranked set sampling is more efficient than simple random sampling for all quantiles. The relative advantage of RSS is greatest at the median and tapers off in the tails. Increasing sample size k further enhances the performance of RSS. For further research, we can apply the idea in this thesis to testing under the unequal allocation.

References

[1] Bohn, L. L. and Wolfe, D. A.(1992). Nonparametric two-sample procedures for ranked-set samples data, *Journal of the American Statistical Association*, **87**(418), 552-561.

[2] Bohn, L. L. and Wolfe, D. A.(1994). The effect of imperfect judgement rankings on properties of procedures based on the ranked-set samples analog of Mann-Whitney-Wilcoxon statistic, *Journal of the American Statistical Association*,

89(425), 168-176.
[3] Dell, T. R. and Clutter, J. L.(1972). Ranked-set sampling theory with order statistics background, *Biometrics*, **28**, 545-553.
[4] Hettmansperger, T. P.(1995). The ranked-set sample sign test, *Journal of the Nonparametric Statistics*, **4**, 271-282.
[5] Kaur, A., Patil, G. P., Taillie, C. and Wit, J.(2002) Ranked set sample sign test for quantiles, *Journal of Statistical Planning and Inference*, 100, 337-347.
[6] Kim, D. H. and Kim, Y. C.(1996). Wilcoxon signed rank test using ranked-set sample, *The Korean Journal of Computational and Applied Mathematics*, **3**(2), 235-243.
[7] Koti, K. M. and Babu, G. J.(1996). Sign test for ranked-set sampling, *Communications in Statistics-Theory and Methods*, **25**(7), 1617-1630.
[8] McIntyre, G. A.(1952). A method of unbiased selective sampling using ranked sets, *Australian Journal of Agricultural Research*, **3**, 385-390.
[9] Öztrük, Ö.(1999a). One- and two-sample sign tests for ranked set sample selective designs, *Communications in Statistics-Theory and Methods*, **28**(5), 1231-1245.
[10] Öztürk, Ö.(1999b). Two-sample inference based on one-sample ranked set sample sign statistics, *Journal of Nonparametric Statistics*, **10**, 197-212.
[11] Stokes, S. L.(1980). Estimation of variance using judgement ordered ranked-set samples, *Biometrics*, **36**, 35-42.
[12] Takahasi, K. and Wakimoto, K.(1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering, *Annals of the Institute of Statistical Mathematics*, **20**, 1-13.

[Received Feburary 2003, Accepted July 2003]