Journal of the Korean Statistical Society (2001), 30: 4, pp 667–678

On Quantiles Estimation Using Ranked Samples with Some Applications

HANI M. SAMAWI[1]

ABSTRACT

The asymptotic behavior and distribution for quantiles estimators using ranked samples are introduced. Applications of quantiles estimation on finding the normal ranges (2.5% and 97.5% percentiles) and the median of some medical characteristics and on finding the Hodges-Lehmann estimate are discussed. The conclusion of this study is, whenever perfect ranking is possible, the relative efficiency of quantiles estimation using ranked samples relative to SRS is high. This may translates to large savings in cost and time. Also, this conclusion holds even if the ranking is not perfect. Computer simulation results are given and real data from Iowa 65+ study is used to illustrate the method.

Keywords: Simple random sample, ranked set sample, normal ranges, quantiles, extreme ranked set samples, Hodges-Lehmann, order statistics.

1. INTRODUCTION

Before the diagnostic value of the medical measures can be evaluated, we need the normal ranges for these measures. By definition, the normal range of a continuous variable contains 95% of all disease-free individuals in a population and therefore, 5% of the disease-free people can have abnormal values. In case of simple random sample (SRS), estimation of those normal ranges (2.5% and 97.5% percentiles) using sample quantiles require a sample of size 400 to 2000 quantified population units.

Ranked set sample (RSS) was introduced by McIntyre (1952) for estimating the pasture yield. RSS procedure involves randomly drawing n sets of n units each from the population for which some characteristics of the distribution is to be estimated. It is assumed that the units in each set can be ranked visually or at little cost. From the first set of n units the unit ranked lowest is measured

[1]Department of Statistics, Yarmouk University, Irbid-Jordan, E-mail: hsamawi@yu.edu.jo

accurately, i.e., it is quantified. From the second set of n units the unit ranked second lowest is quantified. The process is continued until the cycle is completed by quantifying the n-th ranked unit from the n-th set.

Takahasi and Wakimoto (1968) warned that in practice the number of units which are easily ranked cannot be more than four. If a large sample size is required, then the sampling process can be repeated independently and identically m times producing a total sample size of mn. Also, Samawi et al. (1996) investigated the use of extreme ranked set samples (ERSS) for estimating the population mean. They indicated that ERSS reduces the ranking error for larger set size since it did not need complete ranking.

Stokes and Sager (1988), showed that the empirical distribution function using RSS is unbiased and has greater precision than the one obtained by using SRS. Also, they indicated that RSS based procedures can also be applied to human populations.

One application of RSS empirical distribution is to find the normal range for some medical measure (e.g., level of the hemoglobin in the blood.) This task is not feasible when using SRS, because of the high cost of the required medical tests to establish those medical measures. However, if we can rank the individuals at little cost, then the RSS can be used. This can be done in many ways. For example, one can distribute a questionnaire contains some leading questions associated with the medical measure of interest to the units in the SRS that has been selected and then ranking the units according to their answers. Then the RSS can be obtained and be quantified by using some costly medical examination. Another method is, by using available medical records for finding some medical characteristics in those records which are highly related to the medical measure under consideration or by using some an inexpensive medical screening test. Then RSS can be selected by ranking the selected random sample according to the information in those medical records or according to the result of that screening test by quantifying the units in RSS using some costly medical examination. On the other hand, some medical conditions can be can be screened visually and then the patients can be ranked according to their conditions. For example, to establish the normal ranges for the level of bilirubin in the blood of the jaundice premature babies, ranking on the level of bilirubin in the blood can be done visually by an expert physician by observing the following:

i- Color of the face.

ii- Color of the chest.

iii- Color of the lower parts of the body.

iv- Finally, the color of the terminal parts of the whole body.
Then as the yellowish goes from i to iv the level of bilirubin in the blood goes higher.

In this paper, we are interested in studying the quantiles estimation and using them for estimating the normal ranges of some possible medical measures and for finding the Hodges-Lehmann estimate. We show that using RSS for estimating the quantiles and for estimating a location parameter using Hodges-Lehmann estimate are more efficient than using the SRS. Asymptotic results for quantiles estimators as well as some properties of the Hodges-Lehmann estimator using RSS and ERSS are introduced. This study showed that the efficiency obtained by using RSS rather than SRS for estimating the population quantiles have the same pattern as that given by Stokes and Sager (1988) for estimating the distribution function. In fact, this work is partially an extension of their work. In Section 2, we describe some sampling plans, discuss estimation of the population quantiles using these plans, and give some definitions and general results. Hodges-Lehmann estimate using RSS and ERSS and its properties are discussed in Section 3 for perfect ranking. Simulation results from some known distributions are given in Section 4. In Section 5, we illustrate the method using real data from the RHS blood study (Brock et al., 1986). The findings are discussed in Section 6.

2. SAMPLES AND ESTIMATING POPULATION QUANTILES

Let $\{X^*_{11k}, X^*_{21k}, \cdots, X^*_{n1k}; X^*_{12k}, X^*_{22k}, \cdots, X^*_{n2k}; \cdots; X^*_{1nk}, X^*_{2nk}, \cdots, X^*_{nnk}; k = 1, 2, \cdots, m\}$ be nm independent random samples of size n each which are taken from a population. Assume that each unit X^*_{ijk} in the sample has the same distribution function $F(x)$ with finite mean μ, variance σ^2 and p-th quantiles ζ_p. For the simplicity of notation we will assume that X_{ijk} denote the measure of the X^*_{ijk} unit. Then according to our description the SRS may be chosen as $\{X_{11k}, X_{12k}, \cdots, X_{1nk}; k = 1, 2, \cdots, m\}$ and denoted by $X_1, X_2, \cdots, X_{nm}$. Let $X^*_{(1)ik}, X^*_{(2)ik}, \cdots, X^*_{(n)ik}$ be the ordered statistics of the i-th sample of the k-th cycle $X^*_{1ik}, X^*_{2ik}, \cdots, X^*_{nik} (i = 1, 2, \cdots, n; k = 1, 2, \cdots, m)$. Then $\{X_{(1)k}, X_{(2)k}, \cdots, X_{(n)k}; k = 1, 2, \cdots, m\}$ denotes the RSS, where $X_{(i)k} = X_{(i)ik}$. Also, $\{X_{(1)1k}, X_{(n)2k}, \cdots, X_{(1)\{n-1\}k}, X_{(n)nk}; k = 1, 2, \cdots, m\}$ denotes ERSS when n is even, which is the only case that will be considered in this paper.

In the r-th sample of the k-th cycle, $X_{(r)k}$ denotes the measure of r-th smallest judged observation. The density and the cdf of $X_{(r)k}$ will be denoted by $f_{(r)}$ and

$F_{(r)}$, respectively. The population density and cdf will be denoted by f and F, respectively.

For $0 < p < 1$, the p-th quantile is define as $\zeta_p = inf\{x : F(x) \geq p\}$ and is alternately denoted by $F^{-1}(p)$. Suppose we have selected and measured the SRS $X_1, X_2, \cdots, X_{nm}$ of size nm from the population with respect to some variable X. Then the order statistics of the SRS is denoted by $X_{(1)}, X_{(2)}, \cdots, X_{(mn)}$, thus the sample p-th quantile may be expressed as

$$\hat{\zeta}_p = \begin{cases} X_{(mnp)} & \text{if } mnp \text{ is an integer,} \\ X_{([nmp]+1)} & \text{if } mnp \text{ is not an integer,} \end{cases} \tag{1}$$

where $[q]$ is the greatest integer less than or equal to q.

Suppose that F is twice differentiable at ζ_p , with $F'(\zeta_p) = f(\zeta_p) > 0$. Then Bahadur (1966) showed that under SRS,

$$\hat{\zeta}_p = \zeta_p + \frac{p - \hat{F}(\zeta_p)}{f(\zeta_p)} + \mathbf{R}_{mn}, \tag{2}$$

where $\hat{F}$ is the sample empirical distribution function and with probability 1 (wp. 1)

$$R_{nm} = O((nm)^{-3/4}(\log mn)^{1/2}(\log\log mn)^{1/4}), m \to \infty. \tag{3}$$

Therefore, $\sqrt{mn}(\hat{\zeta}_p - \zeta_p)$ converge in distribution to $N(0, \frac{p(1-p)}{f^2(\zeta_p)})$. Using this result, the approximate $100(1-\alpha)\%$ confidence interval is given by

$$\hat{\zeta}_p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{nm\hat{f}^2(\hat{\zeta}_p)}}, \tag{4}$$

where $\hat{f}$ is any consistent estimate of f using SRS and $Z_{\alpha/2}$ is the upper $(1-\alpha/2)$ quantile of the standard normal distribution.

Now suppose we have selected and measured the RSS $\{X_{(1)k}, X_{(2)k}, \cdots, X_{(n)k};\ k = 1, 2, \cdots, m\}$ of size nm from the population with respect to the same variable X. Let the order statistics of the RSS be denoted by $Y_1, \cdots, Y_{nm}$. Then the sample p-th quantile using RSS may be expressed as

$$\hat{\zeta}_p^* = \begin{cases} Y_{nmp} & \text{if } nmp \text{ is an integer,} \\ Y_{[nmp]+1} & \text{if } nmp \text{ is not an integer.} \end{cases} \tag{5}$$

Stokes and Sager (1988) defined the empirical cdf of RSS by

$$F^*(t) = \frac{1}{nm}\sum_{r=1}^{n}\sum_{k=1}^{m}\chi[X_{(r)k} \leq t], \tag{6}$$

where $\chi[\cdot]$ is the indicator function. Also, the following theorem summarized their results:

Theorem1. *Stokes and Sager (1988):*

1 - F^* is an unbiased estimator for F.

2 - $var(F^*) = \frac{1}{n^2 m}\sum_{r=1}^{n} F_{(r)}(t)[1-F_{(r)}(t)]$

$= \frac{1}{nm}\left\{F(t) - \sum_{r=1}^{n}[I_{F(t)}(r, n-r+1)]^2/n\right\},$

where $I_{F(t)}(r, n-r+1)$ (see Arnold, 1992) is the incomplete beta ratio function.

3 - $[F^*(t) - E(F^*)]/\sqrt{var(F^*)}$, converges in distribution to a standard normal variable as $m \to \infty$, when n and t are held fixed.

In this paper we extend Stokes and Sager (1988) theorem further as follows: Define $Z_k = \frac{1}{n}\sum_{r=1}^{n}\chi[X_{(r)k} \le t], \{k = 1, 2, \cdots, m\}$, then $Z_1, Z_2, \cdots, Z_m$ are i.i.d with mean $E(Z_k) = F(t)$ and variance

$$var(Z_k) = \frac{1}{n^2}\sum_{r=1}^{n} F_{(r)}(t)[1-F_{(r)}(t)] = \frac{1}{n}\left\{F(t) - \sum_{r=1}^{n}[I_{F(t)}(r, n-r+1)]^2/n\right\}. \tag{7}$$

Clearly, for fixed n and t, the mean and the variance of Z_k are finite. Therefore, by the strong law of large numbers, $F^*(t) \to F(t)$ wp. 1. Now let

$$F(\zeta_p) = \int_{-\infty}^{\zeta_p} dF(x), \tag{8}$$

assuming the existence of the integral, define the functional $T_p(F) = F^{-1}(p) = \zeta_p$, then $\hat{\zeta}_p^* = T_p(F^*)$. Since $F^*(t) \to F(t)$ wp. 1 and $F(t)$ is assume to be absolutely continuous, it follows that $\hat{\zeta}_p^* \to \zeta_p$ wp. 1.

Under the assumptions of Bahadur theorem (1966) then for fixed p and n

$$\hat{\zeta}_p^* = \zeta_p + \frac{p - F^*(\zeta_p)}{f(\zeta_p)} + \mathbf{R}_m, \tag{9}$$

where F^* is the RSS empirical distribution function and *wp. 1.*

$$\mathbf{R}_m = O((m)^{-3/4}(\log m)^{1/2}(\log\log m)^{1/4}) \tag{10}$$

Therefore, $\sqrt{m}(\hat{\zeta}_p^* - \zeta_p)$ converge in distribution to $N(0, \{p - \sum_{r=1}^{n}[I_p(r, n-r+1)]^2/n\}/nf^2(\zeta_p))$. Using this result, the approximate $100(1-\alpha)\%$ confidence interval is given by

$$\hat{\zeta}_p^* - Z_{\alpha/2}\sqrt{\frac{p - \sum_{r=1}^{n}[I_p(r, n-r+1)]^2/n}{nm\hat{f}^2(\hat{\zeta}_p^*)}},$$

where $I_p(r, n-r+1)$ is the incomplete beta ratio function (see Arnold, 1992), $\hat{f}$ is any consistent estimate of f and $Z_{\alpha/2}$ is the upper $(1-\alpha/2)$ quantile of the standard normal distribution.

2.1. ASYMPTOTIC RELATIVE EFFICIENCY OF $\hat{\zeta}_p^*$

The asymptotic relative efficiency (AREF) of $\hat{\zeta}_p^*$ with respect to $\hat{\zeta}_p$ is given by

$$AREF = \frac{MSE(\hat{\zeta}_p)}{MSE(\hat{\zeta}_p^*)}, \tag{11}$$

where MSE($\cdot$) is the mean square error of the estimator. However, the asymptotic MSE of $\hat{\zeta}_p$ under Bahadur (1966) conditions is

$$MSE(\hat{\zeta}_p) = \frac{p(1-P)}{nmf^2(\zeta_p)}. \tag{12}$$

Similarly, the asymptotic MSE of $\hat{\zeta}_p^*$ is

$$MSE(\hat{\zeta}_p^*) = \frac{\{p - \sum_{r=1}^{n}[I_p(r, n-r+1)]^2/n\}}{nmf^2(\zeta_p)}. \tag{13}$$

Therefore, using (12) and (13),

$$AREF = \frac{p(1-P)}{\{p - \sum_{r=1}^{n}[I_p(r, n-r+1)]^2/n\}}. \tag{14}$$

Stokes and Sager (1988) Table 1, gives for perfect ranking of X the value of AREF for some values of p and $n = 2, \cdots, 5$. They indicated that the AREF is monotone increasing from p=0 to p=0.5, achieve its maximum at p=0.5 and it is symmetric about that point. For example, when n=5 and p=0.01, 0.05, 0.10, 0.30 and 0.50, the values of AREF are 1.04, 1.20, 1.38, 1.88 and 2.03 respectively. Also, they showed that in case of using concomitant random variable Y to X (see Stokers, 1977), and hence imperfect ranking, the loss in precision under bivariate normality, although substantial, is not as great for estimating F, and hence estimating ζ_p , as for estimating the mean of the distribution. In the next section, another application of quantiles estimation, in particular finding the Hodges-Lehmann estimate, is discussed.

3. HODGES-LEHMANN ESTIMATOR

Let $F(x)$ be absolutely continuous symmetric distribution with unique median 0. Let $X_1, X_2, \cdots, X_{nm}$ be a random sample from $F(x-\theta)$. Hence θ is the unique median and mean (when it exists) at the center of the distribution. Then the $nm(nm+1)/2$ Walsh averages are defined by $\frac{X_i+X_j}{2}, i \leq j$, (see for example Hettmansperger, 1984.) As an application for quantiles estimation is the Hodges-Lehmann estimate for θ which is define as the median of these Walsh averages, that is $\hat{\theta} = med_{i \leq j}\left(\frac{X_i+X_j}{2}\right)$. Note that, under the symmetry of the underlying distribution function assumption, the distribution of the $\hat{\theta}$ is symmetric about θ (see Hettmansperger, 1984.)

Similarly, let $\{X_{(1)k}, X_{(2)k}, \cdots, X_{(n)k}; k = 1, 2, \cdots, m\}$ be the RSS drown from $F(x-\theta)$, then again θ can be estimated by

$$\hat{\theta}_{RSS} = med_B\left(\frac{X_{(i)k} + X_{(j)l}}{2}\right), \tag{15}$$

where $B = \{(i, j, k, l) : i \leq j \text{ and /or } k \leq l\}$. Also, let $\{X_{(1)1k}, X_{(n)2k}, \cdots, X_{(1)\{n-1\}k}, X_{(n)nk}; k = 1, 2, \cdots, m\}$ be the ERSS, when n is even, drown from $F(x-\theta)$, then θ can be estimated by

$$\hat{\theta}_{ERSS} = med_C\left(\frac{X_{(1)ik} + X_{(1)jl}}{2}, \frac{X_{(n)ik} + X_{(n)jl}}{2}, \frac{X_{(1)ik} + X_{(n)jl}}{2}\right), \tag{16}$$

where $C = \{(i, j, k, l) : i \leq j \text{ and /or } k \leq l; i, j = 1, 2, \cdots, n/2\}$.

Theorem 2. *If F is symmetric about θ, then the distributions of (15) and (16) are symmetric about θ.*

Proof: It is clear that estimators (15) and (16) are translation statistics. Then we have $P_\theta(\hat{\theta}_{RSS} - \theta < x) = P_0(\hat{\theta}_{RSS} < x)$ and $P_\theta(\hat{\theta}_{ERSS} - \theta < x) = P_0(\hat{\theta}_{ERSS} < x)$. Hence without loss of generality take $\theta = 0$. Now since F is symmetric, then it can be shown that $X_{(i)k} \stackrel{d}{=} -X_{(n-i+1)k}$ and $X_{(1)k} \stackrel{d}{=} -X_{(n)k}, k = 1, 2, \cdots, m$ and $i = 1, 2, \cdots, n$, (see for example Arnold, 1992.) Since the number of maxima and the number of minima in ERSS are equal then $\{X_{(1)k}, X_{(2)k}, \cdots, X_{(n)k}; k = 1, 2, \cdots, m\} \stackrel{d}{=} \{-X_{(1)k}, -X_{(2)k}, \cdots, -X_{(n)k}; k = 1, 2, \cdots, m\}$ and $\{X_{(1)1k}, X_{(n)2k}, \cdots, X_{(1)\{n-1\}k}, X_{(n)nk}; k = 1, 2, \cdots, m\} \stackrel{d}{=} \{-X_{(1)1k}, -X_{(n)2k}, \cdots, -X_{(1)\{n-1\}k}, -X_{(n)nk}; k = 1, 2, \cdots, m\}$. From (15), $\hat{\theta}_{RSS}(X)$ and $\hat{\theta}_{RSS}(-X)$ have the same distribution but $\hat{\theta}_{RSS}(-X) = -\hat{\theta}_{RSS}(X)$, therefore, $\hat{\theta}_{RSS}$ and $-\hat{\theta}_{RSS}$ have the same distribution. Similarly, it can be shown that $\hat{\theta}_{ERSS} - \hat{\theta}_{ERSS}$ have the same distribution. Hence the theorem follows.

4. SIMULATION

The normal, exponential and logistic distributions are used in the simulation for perfect ranking. Sample sizes (m=30, n= 3), (m=10, n=10), (m=2, n=40) and (m=1, n=64) and three values of p=0.05, 0.50 and 0.95 are considered.

In case of using concomitant random variable Y to X, the bivariate normal distribution is used. Sample sizes (m=30, n=3) and (m=10, n=10) and three values of p=0.05, 0.50 and 0.95 are considered.

For each of the possible combinations of distributions, sample sizes and different values of p, 5000 data sets were generated. The relative efficiencies for estimating the population ζ_p using RSS with respect to SRS are obtained. The values obtained by simulation for perfect ranking are given in Table 1. In Table 2 we give the simulation results when concomitant random variable Y to X is used.

To study the performance of the Hodges-Lehmann estimate when using RSS and ERSS compare with using SRS, N(2, 1), Logistic(2, 1) and uniform(0, 2) are used in the simulation. Sample sizes n=4, 5, 6, and 8 and m=1 and 4 are considered. The simulation size is based on 5000 generated data sets. The results of the simulation of the performance for the Hodges-Lehmann estimators are given in Table 3.

Table 1: The relative efficiency of RSS relative to SRS for perfect ranking: m(n).

Distribution function	30(3)	10(10)	2(40)	1(64)
Normal (0, 1)				
p=0.05	1.14	1.52	2.71	3.27
p=0.50	1.64	2.88	6.21	8.33
p=0.95	1.19	1.52	2.30	3.00
Exponential(1)				
p=0.05	1.14	1.66	3.63	4.75
p=0.50	1.65	3.05	5.82	7.88
p=0.95	1.14	1.44	2.16	2.75
Logistic(0, 1)				
p=0.05	1.05	1.50	2.62	3.36
p=0.50	1.77	2.87	6.09	7.51
p=0.95	1.10	1.50	2.30	2.87

Our simulation indicates that using RSS for estimating the quantiles are more efficient than using SRS. Their application to normal ranges may translate to a large saving in time an cost. Also, the efficiency of quantiles estimation using RSS has same pattern of efficiency obtained from the asymptotic results when ranking is assumed to be perfect. For example, in case of n=3, the values of the asymptotic relative efficiency for p=0.05, 0.50, and 0.95 are 1.10, 1.60 and 1.10 respectively which are very close to the values obtained by the simulation, see Table 1. Note that these efficiencies are independent of the cycle size m.

Table 2 indicates that in case of imperfect ranking using RSS is still at least as good as using SRS. Also, Table 2 shows, as expected, that the loss in precision under bivariate normality, is decreasing as ρ (the correlation coefficient between X and Y) increases to 1.

Table 2: The relative efficiency of RSS relative to SRS in case of using concomitant random variable Y to X.

p	m(n)	ρ=0.90	0.80	0.50	0.20
		Normal (0, 0, 1, 1, ρ)			
0.05		1.10	1.03	1.00	1.00
0.50	30(3)	1.49	1.33	1.07	1.00
0.95		1.12	1.09	1.00	1.00
0.05		1.31	1.16	1.04	1.03
0.50	10(10)	1.91	1.52	1.33	1.08
0.95		1.33	1.23	1.10	1.03

Table 3: The relative efficiency of RSS and ERSS relative to SRS for the Hodges-Lehmann estimate

		N(2, 1)		Logistic(2, 1)		Uniform(0, 2)	
m	n	RSS	ERSS	RSS	ERSS	RSS	ERSS
	4	2.45	2.17	2.35	1.82	2.24	3.37
	5	2.67	2.25	2.77	1.96	2.69	4.06
1	6	3.23	2.49	3.19	1.79	3.12	6.60
	7	3.88	2.87	3.84	2.11	3.90	7.07
	8	4.16	2.69	4.11	1.92	4.38	12.75
	4	2.49	2.17	2.47	1.80	2.52	4.04
	5	3.13	2.64	2.98	2.29	3.33	4.84
2	6	3.44	2.51	3.52	1.95	3.53	8.43
	7	4.09	2.83	3.90	2.13	4.08	9.85
	8	4.31	2.67	4.55	2.06	3.04	14.56

Table 3, demonstrate that RSS and ERSS both are superior to SRS for finding the Hodges-Lehmann estimate. It is not surprising that RSS is more efficient than ERSS in case of normal and logistic distribution, however, ERSS is more efficient than RSS in case of uniform distribution (see Samawi et al., 1996.) Note that ERSS is more practical than RSS. Moreover, the efficiency of using RSS and ERSS compared with SRS is increased by increasing the set size. However, in some case the efficiency is slightly increased by increasing the cycle size m.

5. EXAMPLE: Normal Range of Hemoglobin Level in the Blood

This example is just to illustrate the method. The data for this example is drawn from the Iowa 65+ Rural Health Study (RHS). RHS is a longitudinal cohort study of 3,673 individuals (1,420 men and 2,253 women) aged 65 or older living in Washington and Iowa counties of the state of Iowa in 1982. This study is one of the four supported by the National Institute on Aging and collectively referred to as EPESE, (Established Populations for Epidemiologic Studies of the Elderly), National Institute on Aging, 1986.

In the Iowa 65 RHS there were 550 disease free women aged 70+ reported the Erythrocytes counts (RBC) million/mm^3 blood and Hemoglobin level gm./100ml. blood. The question of interest is to estimate some of the normal ranges (5% and 95% quantiles) of the Hemoglobin level of these women. However, since the RBC is highly related to Hemoglobin level, in this example we ignore the Hemoglobin records and we draw SRS and RSS of size 100 each based on RBC.

The results of this example is given in Table 4. The RSS sample size is (n=10, m=10). However, to find 95% confidence interval for the normal ranges based on the RSS we assume perfect ranking. Based on the asymptotic results we found that $SE(\hat{\zeta}_p) \simeq 2.18$ and $SE(\hat{\zeta}_p^*) \simeq 1.82$ and hence the AREF=1.44.

Table 4: Estimation of the Hemoglobin Normal Range (ζ_p) from the Iowa 65+ RHS

p	Estimate	95% Confidence Limits	
		Lower-limit	Upper limit
		RSS: Size10(10)	
0.05	11.50	7.93	15.07
0.95	15.70	12.52	19.27
		SRS: Size100	
0.05	12.20	7.93	16.47
0.95	15.95	11.68	20.22

6. DISCUSSION

The blood data is a good example where we need to find efficient estimator for the population quantiles (normal ranges) for some medical measures. Whenever RSS procedure can be conducted, it provides us with large saving in quantified sample size and hence money and time. With the use of RSS one can practically establish normal ranges for any medical measure using substantially fewer disease-free subjects. Since one must expect the normal ranges to depend on age, gender and ethnicity, the savings is magnified many times when establishing a comprehensive set of normal ranges.

We recommend using RSS to estimate the normal ranges, whenever, it is possible to conduct RSS. It will give an asymptotically unbiased and more efficient estimate of the normal ranges. Also, in some application where finding Hodges-Lehmann is needed, it is recommended to use either RSS or ERSS; also some further investigation about the estimators asymptotic properties is encouraged.
Acknowledgment: The publication of this paper is supported by Yarmouk University Research council. Also, the author is grateful for the referees for their valuable comments and suggestions which have been important in improving the paper.

REFERENCES

Arnold, B. C. (1992), *A First Course in Order Statistics*, John Wiley & Sons, Inc.

Bahadur, R. R., (1966), "A Note on Quantiles in Large Samples," *Annals of Mathematical Statistics*, **37**, 577-580.

Brock, D. B., Wineland T. Freeman, D. H., Lemke, J. H., Scherr, P. A. Demographic Characteristics. (1986). In: *Established Population for Epidemiologic Studies of the Elderly.* Resource Data Book, Cornoni-Huntley, J. Brock, D. B., Ostfeld, A. M., Taylor, J. O. and Wallace, R. B. (eds). National Institute on Aging, NIH Publication No. 86-2443. U.S. Government Printing Office, Washington, D.C.

Hettmansperger, T. P. (1984), *Statistical Inference Based on Ranks*, John Wiley & Sons, Inc.

McIntyre, G. A. (1952), "A Method of Unbiased Selective Sampling, Using Ranked Sets," *Australian J. Agricultural Research*, **3**, 385-390.

Samawi, H. M., Ahmed, M. S. and Abu-Dayyeh, W. (1996), "Estimating the Population Mean Using Extreme Ranked Set Sampling," *Biom. J.*, **38**, 5, 577-586.

Stokes, S. L. (1977), "Ranked Set Sampling with Concomitant Variables," *Communications in Statistics, Part A-Theory and Methods*, **6**, 1207-1211.

Stokes, S. L. and Sager, T. W. (1988), "Characterization of a Ranked-Set Sample with Application to Estimating Distribution Functions," *Journal of the American Statistical Association*, **83**, 374-381.

Takahasi, K. and Wakimoto, K. (1968), "On Unbiased Estimates of the Population Mean Based on the Stratified Sampling by Means of Ordering," *Annals of the Institute of Statistical Mathematics*, **20**, 1-31.