

Modified Ranked Ordering Set Samples for Estimating the Population Mean

Hyun Gee Kim¹⁾ and Dong Hee Kim²⁾

Abstract

We propose the new sampling method, called modified ranked ordering set sampling (MROSS). Kim and Kim (2003) suggested the sign test using the ranked ordering set sampling (ROSS), and showed that the asymptotic relative efficiency (ARE) of ROSS against RSS for sign test increases as sample size does. We propose the estimator for the population mean using MROSS. The relative precision (RP) of estimator of the population mean using MROSS method with respect to the usual estimator using modified RSS is higher, and when the underlying distribution is skewed, the bias of the proposed estimator is smaller than that of several ranked set sampling estimators.

Keywords: Biased estimator; modified ranked ordering set sampling (MROSS); relative precision (RP); skewed distribution.

1. Introduction

Over recent decades many authors studied several modification for the ranked set samples (RSS), first proposed by McIntyre (1952) for estimating the population mean. Samawi *et al.* (1996) used the extreme ranked set sampling (ERSS) to estimate the population mean. They showed that the ERSS estimator is more efficient than the simple random sampling (SRS) estimator. Sinha *et al.* (1996) used the median ranked set sampling (MRSS) to modify the RSS estimators of the population mean. Muttalak (2003) studied the usage on the MRSS to estimate the population mean more efficiently than the usual RSS method. Muttalak (2003) suggested the percentiles ranked set sampling (PRSS) to come up with estimators that are more efficient if the underlying distributions are

1) Instructor, Department of Statistics, Pusan National University, Pusan 609-735, Korea.

Correspondence : hgeekim@pusan.ac.kr

2) Professor, Department of Statistics, Statistical Research Institute, Pusan National University, Pusan 609-735, Korea.

E-mail : heekim@pusan.ac.kr

skewed. Muttlak and Abu-Dayyeh (2004) suggested the weighted extreme RSS (WERSS) and Weighted median RSS (WMRSS). The ranked ordering set sampling (ROSS) method is introduced by Kim and Kim (2003), they studied that ROSS is more efficient than RSS for sign test statistics.

In this paper, we propose the modified ranked ordering set sampling method. The population mean estimator using the modified ranked ordering set sampling (MROSS) is more efficient than the estimators of modified RSS[ERSS, MRSS, PRSS, WERSS and WPRSS].

2. Sampling Method

2.1. Brief review of ranked ordering set sampling

In order to use more informations than RSS, we introduce a Ranked Ordering-Set Sampling (ROSS). On RSS, we use one of k observations and the remaining $k-1$ observations are returned to the population, but on ROSS, $k-1$ observations are not returned to the population. We use them by the following methods. We take a sample of size k from a specified population and an expert ranks the observations using a visual ordering of the observations or using the ordering of a concomitant variable in samples. This process is repeated until we have k different samples each having k ordered units ; $X_{(11)}, \dots, X_{(1k)}, X_{(21)}, \dots, X_{(2k)}, \dots, X_{(k1)}, \dots, X_{(kk)}$, where $X_{(ij)}$ is the j^{th} order statistic in sample i . At this stage we perform another ordering among the units having the same ranking from different sample, i.e. $X_{(1j)}, \dots, X_{(kj)}$, for $j = 1, \dots, k$ and measure the j^{th} ordered units for quantification. The quantified observations will be denoted by $X_{(j)1}^{(j)}, j = 1, \dots, k$. This sampling method is explained in Figure 2.1.

2.2. Modified ranked ordering set sampling

Now we introduce the new sampling method, which we call it Modified Ranked Ordering-Set Sampling (MROSS). The MROSS method is similar to ROSS, but the quantified observations for MROSS are not ordered and independent in each cycle. The sampling method is as follows.

We take a sample of size k from a specified population and an expert ranks the observations using a visual ordering of the observations or using the ordering of a concomitant variable in the samples. This process is repeated until we have k different samples each having k ordered units; $X_{(11)}, \dots, X_{(1k)}, X_{(21)}, \dots, X_{(2k)}, \dots, X_{(k1)}, \dots, X_{(kk)}$, where $X_{(ij)}$ is the j^{th} order statistic in the i^{th} sample. At this stage, we perform another ordering among the units having the same ranking from different sample, i.e. $X_{(1j)}, \dots, X_{(kj)}$, for $j = 1, \dots, k$ and measure the $(k-j+1)^{th}$ ordered units for quantification. The quantified observations will be denoted by $X_{(j)1}^{(k-j+1)}, j = 1, \dots, k$. For example, if j is 1, we judge the largest observation among $X_{(11)}, \dots, X_{(k1)}$, which will be denoted by $X_{(1)1}^{(k)}$. Figure 2.2 explains this sampling method well.

$X_{(11)}$	$X_{(12)}$	\cdots	$X_{(1(k-1))}$	$X_{(1k)}$
$X_{(21)}$	$X_{(22)}$	\cdots	$X_{(2(k-1))}$	$X_{(2k)}$
\vdots	\vdots		\vdots	\vdots
$X_{(k1)}$	$X_{(k2)}$	\cdots	$X_{(k(k-1))}$	$X_{(kk)}$
\Downarrow	\Downarrow	\Downarrow	\Downarrow	\Downarrow
$\mathbf{X}_{(1)1}^{(1)}$	$\mathbf{X}_{(2)1}^{(2)}$	\cdots	$\mathbf{X}_{(k-1)1}^{(k-1)}$	$\mathbf{X}_{(k)1}^{(k)}$

 Figure 2.1: Display of k measurement observations of the first cycle (ROSS)

$X_{(11)}$	$X_{(12)}$	\cdots	$X_{(1(k-1))}$	$X_{(1k)}$
$X_{(21)}$	$X_{(22)}$	\cdots	$X_{(2(k-1))}$	$X_{(2k)}$
\vdots	\vdots		\vdots	\vdots
$X_{(k1)}$	$X_{(k2)}$	\cdots	$X_{(k(k-1))}$	$X_{(kk)}$
\Downarrow	\Downarrow	\Downarrow	\Downarrow	\Downarrow
$\mathbf{X}_{(1)1}^{(k)}$	$\mathbf{X}_{(2)1}^{(k-1)}$	\cdots	$\mathbf{X}_{(k-1)1}^{(2)}$	$\mathbf{X}_{(k)1}^{(1)}$

 Figure 2.2: Display of k measurement observations of the first cycle (MROSS)

2.3. The population mean estimators for the several RSS sampling

First, the unbiased estimator of the population mean using RSS is defined as

$$\bar{X}_{RSS} = \frac{1}{n} \sum_{i=1}^n X_{(i:n)}.$$

The variance of \bar{X}_{RSS} is given by

$$\text{Var}(\bar{X}_{RSS}) = \frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:n)}^2,$$

where $\sigma_{(i:n)}^2 = E[X_{(i:n)} - E(X_{(i:n)})]^2$. For more details, see Dell and Clutter (1972).

Second, the estimator of the population mean using the extreme ranked set sampling (ERSS) is defined as

$$\bar{X}_{ERSS} = \frac{1}{n} \sum_{i=1}^n X_{(i:e)},$$

where $X_{(i:e)}$ denote the smallest of the i^{th} sample ($i = 1, 2, \dots, L$) and the largest of the i^{th} sample ($i = L + 1, L + 2, \dots, n$) if the sample size n is even. Also denote the smallest of the i^{th} sample ($i = 1, 2, \dots, L_1$), the median of the i^{th} sample ($i = (n + 1)/2$) and the largest of the i^{th} sample ($i = L_1 + 2, L_1 + 3, \dots, n$) if the sample size n is odd. The variance of \bar{X}_{ERSS} can be written as

$$\text{Var}(\bar{X}_{ERSS}) = \frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:e)}^2,$$

where $\sigma_{(i:e)}^2 = E[X_{(i:e)} - E(X_{(i:e)})]^2$. For more details, see Samawi *et al.* (1996).

Third, the estimator of the population mean using the percentile ranked set sampling (PRSS) is defined as

$$\bar{X}_{PRSS} = \frac{1}{n} \sum_{i=1}^n X_{(i:p)},$$

where $X_{(i:p)}$ denote the $(p(n + 1))^{th}$ order statistic of the i^{th} sample ($i = 1, 2, \dots, L_1$) and the $(q(n + 1))^{th}$ order statistic of the i^{th} sample ($i = L_1 + 1, L_1 + 2, \dots, n$) if the sample size n is even, where $q = 1 - p$ and $0 \leq p \leq 1$. Also denote the $(p(n + 1))^{th}$ order statistic of the i^{th} sample ($i = 1, 2, \dots, L_2$), the median of the i^{th} sample ($i = (n + 1)/2$) and the $(q(n + 1))^{th}$ order statistic of the i^{th} sample ($i = L_2 + 2, L_2 + 3, \dots, n$) if the sample size is n odd. The variance of \bar{X}_{PRSS} can be written as

$$\text{Var}(\bar{X}_{PRSS}) = \frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:p)}^2,$$

where $\sigma_{(i:p)}^2 = E[X_{(i:p)} - E(X_{(i:p)})]^2$. In this paper, we will consider the case of $p = 3$ only. For more details, see Muttlak and Abu-Dayyeh (2004).

Fourth, the estimator of the population mean using the weight extreme ranked set sampling (WERSS) is defined as

$$\bar{X}_{WERSS} = \frac{1}{n} \left(\sum_{i=1}^{n_1} X_{(i:1)} + \sum_{i=1}^{n_2} X_{(i:n)} \right),$$

where $n_1 = nw$, $n_2 = n(1 - w) = n - n_1$, and $0 \leq w \leq 1$, is the weight to be chosen to make \bar{X}_{WERSS} unbiased estimator of the population mean. The variance of \bar{X}_{WERSS} is

$$\text{Var}(\bar{X}_{WERSS}) = \frac{1}{n} [w\sigma_{(1)}^2 + (1 - w)\sigma_{(n)}^2],$$

where $\sigma_{(1)}^2$ and $\sigma_{(n)}^2$ are the variance of the smallest and the largest order statistic, respectively. For more details, see Muttlak and Abu-Dayyeh (2004).

Last, the estimator of the population mean using the weight percentile ranked set sampling (WPRSS) is defined as

$$\bar{X}_{WPRSS} = \frac{1}{n} \left(\sum_{i=1}^{n_1} X_{(i:k)} + \sum_{i=1}^{n_2} X_{(i:n-k+1)} \right),$$

where $k = p(n + 1)$ and $n - k + 1 = q(n + 1)$, where $q = 1 - p$ and $0 \leq p \leq 1$, and $n_1 = nw, n_2 = n(1 - w) = n - n_1$, and $0 \leq w \leq 1$, is the weight to be chosen to make \bar{X}_{WPRSS} unbiased estimator of the population mean. The variance of \bar{X}_{WPRSS} is

$$\text{Var}(\bar{X}_{WPRSS}) = \frac{1}{n} [w\sigma_{(k)}^2 + (1 - w)\sigma_{(n-k+1)}^2],$$

where $\sigma_{(k)}^2$ and $\sigma_{(n-k+1)}^2$ are the variance of the k^{th} and $(n - k + 1)^{th}$ order statistic, respectively. For more details, see Muttalak and Abu-Dayyeh (2004).

3. The RP of Estimating the Population Mean using Proposed Estimators

We now introduce the estimator of the population mean using modified ranked ordering set sampling (MROSS). The mean estimator is defined as

$$\bar{X}_{MROSS} = \frac{1}{n} \sum_{i=1}^n X_{(k)}^{(n-k+1)},$$

where $X_{(k)}^{(n-k+1)}$ denotes the $(n - k + 1)^{th}$ order statistic in the k^{th} ordered sample.

The variance of the mean estimator using MROSS is also defined as

$$\text{Var}(\bar{X}_{MROSS}) = \frac{1}{n^2} \sum_{i=1}^n [\sigma_{(k)}^{(n-k+1)}]^2,$$

where $[\sigma_{(k)}^{(n-k+1)}]^2 = E[X_{(k)}^{(n-k+1)} - E(X_{(k)}^{(n-k+1)})]^2$. The density function $h_{(k)}^{(n-k+1)}(t)$ of the random variable $X_{(k)}^{(n-k+1)}$, represents the pdf of the $(n - k + 1)^{th}$ order statistic of a sample distribution, $H_{(k)}(t)$, where $H_{(k)}(t)$ is the cdf of the k^{th} order statistic from a set of size n , i.e. the density is of the form ;

$$\begin{aligned} h_{(k)}^{(n-k+1)}(t) &= \frac{n!}{(k-1)!(n-k+1)!} H_{(k)}^{n-k+1}(t) [1 - H_{(k)}(t)]^{k-1} h_{(k)}(t) \\ &= \left[\frac{n!}{(k-1)!(n-k+1)!} \right]^2 H_{(k)}^{n-k+1}(t) \\ &\quad \times [1 - H_{(k)}(t)]^{k-1} H(t)^{k-1} [1 - H(t)]^{n-k+1} h(t) \end{aligned}$$

where $H_{(k)}(t) = \int_{-\infty}^t h_{(k)}(x) dx$, $h_{(k)}(t)$ is the k^{th} order statistic of pdf $h(t)$.

To compare MROSS estimators with WERSS, WPRSS, PRSS, ERSS, RSS and SRS estimators of μ , four skewed probability distribution functions are considered for the populations: exponential, gamma, Weibull and half normal. The variance or the mean square error of the sample means for the MROSS, WERSS, WPRSS, PRSS, ERSS and RSS are calculated for the above distributions using the moments of the order statistics.

The relative precision (RP) of MROSS, WERSS, WPRSS, PRSS, ERSS and RSS

with respect to SRS can be obtained as follows:

$$\begin{aligned}
 RP_1(\bar{X}_{SRS}, \bar{X}_{RSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{Var}(\bar{X}_{RSS})}, \\
 RP_2(\bar{X}_{SRS}, \bar{X}_{PRSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{MSE}(\bar{X}_{PRSS})}, \\
 RP_3(\bar{X}_{SRS}, \bar{X}_{ERSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{MSE}(\bar{X}_{ERSS})}, \\
 RP_4(\bar{X}_{SRS}, \bar{X}_{WPRSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{Var}(\bar{X}_{WPRSS})}, \\
 RP_5(\bar{X}_{SRS}, \bar{X}_{WERSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{Var}(\bar{X}_{WERSS})}, \\
 RP_6(\bar{X}_{SRS}, \bar{X}_{MROSS}) &= \frac{\text{Var}(\bar{X}_{SRS})}{\text{MSE}(\bar{X}_{MROSS})},
 \end{aligned}$$

respectively.

4. Result and Discussion

Results are summarized as the RP and biases in Table 4.1-4.6 for RSS, PRSS, ERSS, WPRSS, WERSS and MROSS with respect to SRS.

For each population calculations were done with sample size $n = 4, 5, 6, 7$. We obtain RP for each distribution using software MATHCAD 12.0. From Table 4.1 to 4.6, a gain in efficiency is obtained by using MROSS for different values of for all the distributions except exponential distribution in this study. From tables, we know that the proposed estimator is more efficient than the other estimators for all distributions and sample sizes in study.

Table 4.1: The RP values for estimating the population mean using RSS, PRSS, ERSS, WPRSS, WERSS and MROSS for sample size and the bias of each estimators for exponential distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS		MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias	
4	1.920	1.455	0.166	1.171	0.333	2.581	1.615	2.907	0.236	
5	2.190	2.620	0.144	1.322	0.150	2.934	1.705	2.629	0.242	
6	2.449	3.030	0.308	0.752	0.097	3.169	1.779	2.372	0.243	
7	2.700	3.309	0.280	2.813	0.076	3.360	1.844	2.163	0.241	

Table 4.2: The RP values for Gamma($r = 2, \lambda = 1$) distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS	MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias
4	2.096	1.313	0.175	1.452	0.492	2.623	1.720	3.777	0.248
5	2.424	2.858	0.158	1.665	0.158	3.051	1.903	3.798	0.255
6	2.742	3.243	0.325	1.089	0.097	3.346	2.006	3.663	0.256
7	3.052	3.577	0.297	1.188	0.081	3.585	2.092	3.485	0.254

Table 4.3: The RP values for Gamma ($r = 4, \lambda = 1$) distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS	MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias
4	2.100	1.222	0.180	1.782	0.713	2.642	1.892	4.615	0.254
5	2.575	3.033	0.162	1.955	0.162	3.126	2.048	5.175	0.260
6	2.939	3.392	0.333	1.476	0.100	3.452	2.176	5.422	0.261
7	3.292	3.764	0.304	1.628	0.083	3.721	2.286	5.475	0.260

Table 4.4: The RP values for Weibull ($a = 1, b = 2$) distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS	MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias
4	2.325	1.168	0.029	1.987	0.166	2.559	2.117	4.893	0.041
5	2.744	2.999	0.026	2.333	0.026	3.042	2.376	6.055	0.042
6	3.155	3.399	0.054	2.038	0.016	3.434	2.603	6.976	0.042
7	3.561	3.816	0.050	2.272	0.013	3.871	2.808	7.689	0.042

Table 4.5: The RP values for Weibull ($a = 1, b = 4$) distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS	MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias
4	2.383	1.034	0.002	2.135	0.092	2.531	2.139	5.772	0.004
5	2.830	3.136	0.002	2.536	0.003	3.106	2.401	8.064	0.004
6	3.243	3.493	0.004	2.615	0.001	3.496	2.634	9.763	0.004
7	3.671	3.927	0.004	2.981	0.001	3.847	2.845	13.46	0.004

Table 4.6: The RP values for Half-normal (1) distribution.

sample size n	RSS	PRSS		ERSS		WPRSS	WERSS	MROSS	
	RP1	RP2	Bias	RP3	Bias	RP4	RP5	RP6	Bias
4	2.239	1.257	0.065	1.768	0.211	2.441	2.096	3.695	0.118
5	2.628	2.721	0.058	2.048	0.058	2.875	2.362	3.969	0.121
6	3.010	3.198	0.120	1.456	0.034	3.268	2.599	4.006	0.122
7	3.386	3.594	0.109	1.599	0.028	3.633	2.813	3.958	0.121

5. Further Work

In this article, we propose the estimator of the population mean using modified ranked ordering set sampling. But, the proposed estimator has slightly small bias. We are working about the unbiased estimator of modified ranked ordering set sampling using weights.

References

- Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, **28**, 545–555.
- Kim, D. H. and Kim, H. G. (2003). Sign test using ranked ordering-set sampling. *Journal of Nonparametric Statistics*, **15**, 303–309.
- McIntyre, G. A. (1952). A method of unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, **3**, 385–390.
- Muttlak, H. A. (2003). Modified ranked set sampling. *Pakistan Journal of Statistics*, **3**, 315–323.
- Muttlak H. A. and Abu-Dayyeh, W. (2004). Weight modified ranked set sampling methods. *Applied Mathematics and computation*, **151**, 645–657.
- Samawi, H., Abu-Dayyeh, W. and Ahmed, S. (1996). Extreme ranked set sampling. *Biometrical Journal*, **30**, 577–586.
- Sinha, B. K., Sinha, B. K. and Purkayastha, S. (1996). On some aspects of ranked set sampling for estimation of normal and exponential parameters. *Statistics Decisions*, **14**, 223–240.

[Received August 2007, Accepted November 2007]