

# Estimating the Population Mean by Using Stratified Double Extreme Ranked Set Sample

Mahmoud I. Syam, Kamarulzaman Ibrahim, Amer I. Al-Omari

**Abstract**—Stratified double extreme ranked set sampling (SDERSS) method is introduced and considered for estimating the population mean. The SDERSS is compared with the simple random sampling (SRS), stratified ranked set sampling (SRSS) and stratified simple set sampling (SSRS). It is shown that the SDERSS estimator is an unbiased of the population mean and more efficient than the estimators using SRS, SRSS and SSRS when the underlying distribution of the variable of interest is symmetric or asymmetric.

**Keywords**—Double extreme ranked set sampling, Extreme ranked set sampling, Ranked set sampling, Stratified double extreme ranked set sampling.

## I. INTRODUCTION

NOONE denies the importance and the benefit of ranked set sampling method which was first proposed by McIntyre [11] for estimating the mean pasture and forage yields, and without proving he claimed that  $\bar{X}_{RSS}$  is an unbiased estimator for the population mean  $\mu$ , his method was developed and modified by many authors. Takahasi and Wakimoto [16] have established an accurate mathematical theory of ranked set sampling and they get the same results. Dell and Clutter [9] showed that the mean of the RSS is an unbiased estimator of the population mean, whatever or not there are errors in ranking. Samawi et al. [14] investigated the variety of extreme ranked set sample (ERSS) for estimating the population mean. Samawi [13] introduced the stratified ranked set sample method for estimating the population mean. Al-Saleh and Al-Kadiri [5] introduced double ranked set sampling for estimating the population mean. Samawi [15] suggested double extreme ranked set sample with application to regression estimator. Jemain et al. [10] suggested multistage extreme ranked set samples for estimating the population mean and they showed that the efficiency of the mean estimator using MERSS can be increased for specific value of the sample size  $n$  by increase the number of stages. For more about RSS see [4], [7], [12], [1] and [2], [3], and [8].

In this paper, new estimator using stratified double extreme ranked set sampling is suggested to estimate the population mean of symmetric and asymmetric distributions. The organization of this paper is as follows: In Section II, we present the stratified double extreme ranked set sampling. In

Section III some notations and basic results are introduced. A simulation study is considered in Section IV. Finally, conclusions are introduced in Section V.

## II. STRATIFIED DOUBLE EXTREME RANKED SET SAMPLING

In stratified sampling the population of  $N$  units is first divided into  $L$  subpopulations of  $N_1, N_2, \dots, N_L$  units, respectively. These subpopulations are no overlapping and together they comprise the whole population, so that  $N_1 + N_2 + \dots + N_L = N$ . The subpopulations are called strata. To obtain the full benefit from stratification, the values of the  $N_h$  ( $h = 1, 2, \dots, L$ ) must be known. When the strata have been determined, a sample is drawn from each, the drawings being made in different strata. The sample sizes within the strata are denoted by  $n_1, n_2, \dots, n_L$ , respectively.

The appropriate allocation of samples to different strata is very important in stratified sampling, in this article, the type of allocation method is proportional to stratum size. If a simple random sample is taken in each stratum, the whole procedure is described as stratified simple random sampling (SSRS).

The stratified double extreme ranked set sampling (DERSS) is described as:

Step 1. Identify  $n^2$  elements from the target population and divide these elements randomly into  $n^2$  sets each of size  $n$  elements.

Step 2. For each set in step 2, if the sample size is even, select from the first  $\frac{n^2}{2}$  sets the smallest ranked unit, and from the second  $\frac{n^2}{2}$  sets the largest ranked unit. If the sample size is odd, select from the first  $\frac{n(n-1)}{2}$  sets

choose the smallest ranked unit, and from the next  $n$  sets choose the median of each set, and from the other  $\frac{n(n-1)}{2}$  sets choose the largest ranked unit. This step yields  $n$  sets each of size  $n$ .

Step 3. Apply the ERSS procedure again on the sets obtained from Step (2) to obtain a DERSS of size  $n$ .

The cycle can be repeated  $m$  times if needed to get a sample of size  $nm$  units.

Step 4. If the double extreme ranked set sample is used in each stratum, the whole procedure is described as stratified double extreme ranked set sampling (SDERSS).

Mahmoud. I. Syam is with Qatar University, Foundation Program, Mathematics Department, P.O.Box 2713, Doha, Qatar (Phone: 00974-55468238; e-mail: M.syam@qu.edu.qa).

K. Ibrahim, Jr., is with School of Mathematical Sciences, Universiti Kebangsaan Malaysia, 43600 UKM Bangi Selangor, Malaysia.

A. I. Al-Omari is with Al al-Bayt University, Faculty of Science, Department of Mathematics P.O. Box 130095, Mafraq 25113, Jordan.

To illustrate the method, let us consider the following example, which combines an even and an odd sample size in two strata.

**Example 1:** Suppose we have two strata,  $L=2$  ( $h=1,2$ ), and in the first stratum we have 64 elements divided into 16 sets, 4 elements in each set, and in the second stratum we have 27 elements divided into 9 sets, 3 elements in each set, as the following:

**Stratum (1):** Assume the 64 elements are

$$X_{11}^{(1)}, X_{12}^{(1)}, \dots, X_{44}^{(1)}, X_{11}^{(2)}, X_{12}^{(2)}, \dots, X_{44}^{(2)}, X_{11}^{(3)}, X_{12}^{(3)}, \dots, X_{44}^{(3)}, X_{11}^{(4)}, X_{12}^{(4)}, \dots, X_{44}^{(4)}$$

After ranking the elements in each set we obtain

$$\begin{aligned} & \{X_{11}^{(1)}, \dots, X_{14}^{(1)}, X_{21}^{(1)}, \dots, X_{24}^{(1)}, X_{31}^{(1)}, \dots, X_{34}^{(1)}, X_{41}^{(1)}, \dots, X_{44}^{(1)}\} \\ & \{X_{11}^{(2)}, \dots, X_{14}^{(2)}, X_{21}^{(2)}, \dots, X_{24}^{(2)}, X_{31}^{(2)}, \dots, X_{34}^{(2)}, X_{41}^{(2)}, \dots, X_{44}^{(2)}\} \\ & \{X_{11}^{(3)}, \dots, X_{14}^{(3)}, X_{21}^{(3)}, \dots, X_{24}^{(3)}, X_{31}^{(3)}, \dots, X_{34}^{(3)}, X_{41}^{(3)}, \dots, X_{44}^{(3)}\} \\ & \{X_{11}^{(4)}, \dots, X_{14}^{(4)}, X_{21}^{(4)}, \dots, X_{24}^{(4)}, X_{31}^{(4)}, \dots, X_{34}^{(4)}, X_{41}^{(4)}, \dots, X_{44}^{(4)}\}. \end{aligned}$$

We will apply the ERSS on each of the 16 elements to get four sets as:

- **Set (1):**  $S_1 = \{X_{(11)}^{(1)}, X_{(21)}^{(1)}, X_{(31)}^{(1)}, X_{(41)}^{(1)}\}$ ,
- **Set (2):**  $S_2 = \{X_{(11)}^{(2)}, X_{(21)}^{(2)}, X_{(31)}^{(2)}, X_{(41)}^{(2)}\}$
- **Set (3):**  $S_3 = \{X_{(14)}^{(3)}, X_{(24)}^{(3)}, X_{(34)}^{(3)}, X_{(44)}^{(3)}\}$ ,
- **Set (4):**  $S_4 = \{X_{(14)}^{(4)}, X_{(24)}^{(4)}, X_{(34)}^{(4)}, X_{(44)}^{(4)}\}$

The notation for the sets  $S_1, S_2, S_3, S_4$  after ordering as:

$$\begin{aligned} S_1^* &= \{X_{(11)}^{(1)*}, X_{(21)}^{(1)*}, X_{(31)}^{(1)*}, X_{(41)}^{(1)*}\}, \\ S_2^* &= \{X_{(11)}^{(2)*}, X_{(21)}^{(2)*}, X_{(31)}^{(2)*}, X_{(41)}^{(2)*}\} \\ S_3^* &= \{X_{(14)}^{(3)*}, X_{(24)}^{(3)*}, X_{(34)}^{(3)*}, X_{(44)}^{(3)*}\}, \\ S_4^* &= \{X_{(14)}^{(4)*}, X_{(24)}^{(4)*}, X_{(34)}^{(4)*}, X_{(44)}^{(4)*}\}. \end{aligned}$$

We apply ERSS again on the 16 elements in ERSS to get the elements of the double extreme ranked set sample in the first stratum ( $h=1$ ),  $X_{(11)}^{(1)*}, X_{(11)}^{(2)*}, X_{(14)}^{(3)*}, X_{(14)}^{(4)*}$ .

**Stratum (2):** Assume the 27 elements are

$$Y_{11}^{(1)}, Y_{12}^{(1)}, \dots, Y_{33}^{(1)}, Y_{11}^{(2)}, Y_{12}^{(2)}, \dots, Y_{33}^{(2)}, Y_{11}^{(3)}, Y_{12}^{(3)}, \dots, Y_{33}^{(3)}.$$

After ranking the elements in each set we obtain

$$\begin{aligned} & \{Y_{(11)}^{(1)}, \dots, Y_{(13)}^{(1)}, Y_{(21)}^{(1)}, \dots, Y_{(23)}^{(1)}, Y_{(31)}^{(1)}, \dots, Y_{(33)}^{(1)}\}, \\ & \{Y_{(11)}^{(2)}, \dots, Y_{(13)}^{(2)}, Y_{(21)}^{(2)}, \dots, Y_{(23)}^{(2)}, Y_{(31)}^{(2)}, \dots, Y_{(33)}^{(2)}\} \end{aligned}$$

$$\{Y_{(11)}^{(3)}, \dots, Y_{(13)}^{(3)}, Y_{(21)}^{(3)}, \dots, Y_{(23)}^{(3)}, Y_{(31)}^{(3)}, \dots, Y_{(33)}^{(3)}\}.$$

We will apply ERSS on each of the nine elements to get three sets as:

$$S_1 = \{Y_{(11)}^{(1)}, Y_{(21)}^{(1)}, Y_{(31)}^{(1)}\}, S_2 = \{Y_{(12)}^{(2)}, Y_{(22)}^{(2)}, Y_{(32)}^{(2)}\}, S_3 = \{Y_{(13)}^{(3)}, Y_{(23)}^{(3)}, Y_{(33)}^{(3)}\}.$$

These sets after ordering will be denoted as:

$$\begin{aligned} S_1^* &= \{Y_{(11)}^{(1)*}, Y_{(21)}^{(1)*}, Y_{(31)}^{(1)*}\}, S_2^* = \{Y_{(12)}^{(2)*}, Y_{(22)}^{(2)*}, Y_{(32)}^{(2)*}\}, \\ S_3^* &= \{Y_{(13)}^{(3)*}, Y_{(23)}^{(3)*}, Y_{(33)}^{(3)*}\}. \end{aligned}$$

If we apply ERSS again on the 9 elements in the sets  $S_1^*, S_2^*, S_3^*$ , we will get the elements of the double extreme ranked set sample in the second stratum ( $h=2$ )  $Y_{(11)}^{(1)*}, Y_{(22)}^{(2)*}, Y_{(33)}^{(3)*}$ .

So the elements of SDERSS are

$$X_{(11)}^{(1)*}, X_{(11)}^{(2)*}, X_{(44)}^{(3)*}, X_{(44)}^{(4)*}, Y_{(11)}^{(1)*}, Y_{(22)}^{(2)*}, Y_{(33)}^{(3)*}.$$

### III. NOTATIONS AND SOME BASIC RESULTS

Al-Saleh and Al-Kadiri [2] studied the double ranked set sampling, and they showed that if  $X_{(i)}^*$ ,  $i=1, 2, \dots, n$  are elements of DRSS with mean  $\mu_{(i)}^*$  and variance  $\sigma_{(i)}^{*2}$  then

$$\mu = \frac{1}{n} \sum_{i=1}^n \mu_{(i)}^* \quad \text{and} \quad \sigma^2 = \frac{1}{n} \left[ \sum_{i=1}^n \sigma_{(i)}^{*2} + \sum_{i=1}^n (\mu_{(i)}^* - \mu)^2 \right].$$

Assume that the variable of interest  $X$  has density  $f(x)$  and cumulative distribution function  $F(x)$ , with mean  $\mu$  and variance  $\sigma^2$ . Let  $X_{11}, X_{12}, \dots, X_{1n}; X_{21}, X_{22}, \dots, X_{2n}; \dots; X_{n1}, X_{n2}, \dots, X_{nn}$  be  $n$  independent simple random samples each of size  $n$ . Let  $X_{(i1)}, X_{(i2)}, \dots, X_{(in)}$  be the order statistics of the  $i$ th sample  $X_{i1}, X_{i2}, \dots, X_{in}$ ,  $i=1, 2, \dots, n$ .

Now, based on DERSS, if the sample size  $n$  is even, we replace the ordered statistics  $X_{(i1)}, X_{(i2)}, \dots, X_{(in)}$  ( $i=1, 2, \dots, n$ ) obtained from an iid sample  $X_{i1}, X_{i2}, \dots, X_{in}$ , ( $i=1, 2, \dots, n$ ), by  $X_{(11)}^*, X_{(21)}^*, \dots, X_{(\frac{n}{2}(1))}^*, X_{(\frac{n}{2}+1(n))}^*, \dots, X_{(n(n))}^*$ , and if the sample size  $n$  is odd, we replace the ordered statistics  $X_{(i1)}, X_{(i2)}, \dots, X_{(in)}$  ( $i=1, 2, \dots, n$ ) by  $X_{(1(1))}^*, X_{(2(1))}^*, \dots, X_{(\frac{n-1}{2}(1))}^*, X_{(\frac{n+1}{2}(\frac{n+1}{2}))}^*, X_{(\frac{n+3}{2}(n))}^*, \dots, X_{(n(n))}^*$  which are independent but not identically distributed random variables, where  $X_{(i1)}^*$  is the minimum of the  $i$ th sample,  $X_{(i(\frac{n+1}{2}))}^*$  is the median of the  $i$ th sample and  $X_{(i(n))}^*$  is the maximum of the  $i$ th sample in DERSS.

Now, using SDERSS if the sample size  $n$  is even, we replace the ordered statistics  $X_{hi(1)}, X_{hi(2)}, \dots, X_{hi(n_h)}$  ( $i = 1, 2, \dots, n_h$ ) and stratum  $h$  ( $h = 1, 2, \dots, L$ ) by  $X_{h1(1)}^*, \dots, X_{h\frac{n_h}{2}(1)}^*, X_{h\frac{n_h}{2}+1}^*, \dots, X_{hn_h(n_h)}^*$ , and if the sample size  $n_h$  is odd, we replace the ordered statistics  $X_{hi(1)}, X_{hi(2)}, \dots, X_{hi(n_h)}$  ( $i = 1, 2, \dots, n_h$ ) and stratum  $h$  ( $h = 1, 2, \dots, L$ ) by  $X_{h1(1)}^*, \dots, X_{h\frac{n_h-1}{2}(1)}^*, X_{h\frac{n_h-1}{2}+1}^*, X_{h\frac{n_h+1}{2}}^*, \dots, X_{hn_h(n_h)}^*$  which are independent but not identically distributed random variables, where  $X_{hi(1)}^*$  is the minimum of the  $i$ th sample in  $h$ th stratum,  $X_{hi(\frac{n_h+1}{2})}^*$  is the median of the  $i$ th sample in  $h$ th stratum and  $X_{hi(n_h)}^*$  is the maximum of the  $i$ th sample in  $h$ th stratum. Let  $F_{hi(n)}(x)$  be the distribution function of the  $i$ th order statistic  $X_{hi(n_h)}$  based on a random sample of size  $n_h$  from a distribution  $F_h(x)$  in stratum  $h$ . Then

$$F_{hi(n)}(x) = \frac{1}{B(i, n_h - i + 1)} \int_0^{F_h(x)} u^{i-1} (1-u)^{n_h-i} du = B_{i, n_h} [F_h(x)], \quad -\infty < x < \infty$$

where  $B_{i, n_h}$  is the Beta distribution with parameters  $(i, n_h - i + 1)$ , and the beta function is  $B(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)}$ , where  $\Gamma(\alpha) = (\alpha - 1)!$ . The mean and the variance of  $X_{hi(n)}$ , respectively are given by

$$\mu_{hi(n_h)} = \int_{-\infty}^{\infty} x f_{hi(n)}(x) dx \text{ and variance } \sigma_{hi(n_h)}^2 = \int_{-\infty}^{\infty} (x - \mu_{hi(n_h)})^2 f_{hi(n)}(x) dx.$$

The suggested estimator of the population mean using SDERSS is defined as

$$\bar{X}_{SDERSS}^* = \begin{cases} \bar{X}_{SDERSS1}^* = \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} X_{hi(1)}^* + \sum_{i=\frac{n_h}{2}+1}^{n_h} X_{hi(n_h)}^* \right), & \text{if } n_h \text{ is even} \\ \bar{X}_{SDERSS2}^* = \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h-1}{2}} X_{hi(1)}^* + X_{hi(\frac{n_h+1}{2})}^* + \sum_{i=\frac{n_h+3}{2}}^{n_h} X_{hi(n_h)}^* \right), & \text{if } n_h \text{ is odd} \end{cases}$$

where  $W_h = \frac{N_h}{N}$ ,  $N_h$  is the stratum size and  $N$  is the total population size.

The variance of  $\bar{X}_{SDERSS}^*$  for even and odd sample size is given, respectively by

$$Var(\bar{X}_{SDERSS}^*) = \begin{cases} Var(\bar{X}_{SDERSS1}^*) = \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( \sum_{i=1}^{\frac{n_h}{2}} \sigma_{hi(1)}^{2*} + \sum_{i=\frac{n_h}{2}+1}^{n_h} \sigma_{hi(n_h)}^{2*} \right), & \text{if } n_h \text{ is even} \\ Var(\bar{X}_{SDERSS2}^*) = \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( \sum_{i=1}^{\frac{n_h-1}{2}} \sigma_{hi(1)}^{2*} + \sum_{i=\frac{n_h+3}{2}}^{n_h} \sigma_{hi(n_h)}^{2*} + \sigma_{hi(\frac{n_h+1}{2})}^{2*} \right), & \text{if } n_h \text{ is odd} \end{cases}$$

**Lemma 1.** Let  $X_{h1(1)}^*, \dots, X_{h\frac{n_h}{2}(1)}^*, X_{h\frac{n_h}{2}+1}^*, \dots, X_{hn_h(n_h)}^*$ ,

where  $i = 1, 2, \dots, n_h$  and  $h = 1, 2, \dots, L$  be SDERSS with even sample size. Let  $X_h$  be a continuous random variable in stratum  $h$ , with pdf  $f_h(x)$ , cdf  $F_h(x)$ , mean  $\mu_h^*$  and variance  $\sigma_h^{2*}$ . Let  $X_{h(1)}^*$  be the minimum of  $X_{h1(1)}, X_{h2(1)}, \dots, X_{h\frac{n_h}{2}(1)}$

with cdf  $F_{h(1)}^*(x)$ , and  $X_{h(n_h)}^*$  be the maximum of  $X_{h\frac{n_h}{2}+1}, X_{h\frac{n_h}{2}+2}, \dots, X_{hn_h(n_h)}$  with cdf  $F_{h(n_h)}^*(x)$ , then:

$$\begin{aligned} (1) \quad F_{h(1)}^*(x) &= 1 - [1 - F_h(x)]^{n_h^2}, \text{ and } F_{h(n_h)}^*(x) = [F_h(x)]^{n_h^2}. \\ (2) \quad f_{h(1)}^*(x) &= n_h^2 [1 - F_h(x)]^{n_h^2-1} f_h(x), \text{ and } f_{h(n_h)}^*(x) = n_h^2 [F_h(x)]^{n_h^2-1} f_h(x). \end{aligned}$$

**Proof:** To prove (1),

$$F_{hi(n_h)}(x) = \frac{1}{B(i, n_h - i + 1)} \int_0^{F_h(x)} u^{i-1} (1-u)^{n_h-i} du = B_{i, n_h} [F_h(x)].$$

$$\text{For } i = 1, \text{ we get } F_{h(1)}^*(x) = B_{1, n_h} [F_h(x)] = 1 - [1 - F_h(x)]^{n_h} = 1 - [1 - F_h(x)]^{n_h^2}$$

and

$$F_{hi(n_h)}(x) = \frac{1}{B(i, n_h - i + 1)} \int_0^{F_h(x)} u^{i-1} (1-u)^{n_h-i} du = B_{i, n_h} [F_h(x)].$$

$$\text{For } i = n_h, \text{ we get } F_{h(n_h)}^*(x) = B_{n_h, n_h} [F_h(x)] = B_{n_h, n_h} [(F_h(x))^{n_h}] = [F_h(x)]^{n_h^2}.$$

Part (2) of Lemma (1) can be proved by taking the first derivative of part (1), respectively.

**Lemma 2.** Suppose that the population distribution is symmetric, then  $\mu_{h(1)}^* + \mu_{h(n_h)}^* = 2\mu_h^*$  and  $\sigma_{h(1)}^{2*} = \sigma_{h(n_h)}^{2*}$ .

**Proof:** By using the facts that  $f_h(-x) = f_h(x)$  and  $F_h(-x) = 1 - F_h(x)$ , and the results of Lemma 1, it will lead us to the result.

**Lemma 3.**  $\bar{X}_{SDERSS}^*$  is an unbiased estimator of a symmetric population mean.

**Proof.** We have two cases:

First: If the sample sizes in the strata  $n_h, h=1,2,\dots,L$  are even, we have

$$\begin{aligned}\bar{X}_{SDERSS1}^* &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} X_{hi(1)}^* + \sum_{i=\frac{n_h}{2}+1}^n X_{hi(n_h)}^* \right) \\ E(\bar{X}_{SDERSS1}^*) &= E \left( \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} X_{hi(1)}^* + \sum_{i=\frac{n_h}{2}+1}^n X_{hi(n_h)}^* \right) \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} E(X_{hi(1)}^*) + \sum_{i=\frac{n_h}{2}+1}^n E(X_{hi(n_h)}^*) \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} \mu_{h(1)}^* + \sum_{i=\frac{n_h}{2}+1}^n \mu_{h(n_h)}^* \right).\end{aligned}$$

Since the distribution is symmetric about  $\mu$ , then  $\mu_{h(1)}^* + \mu_{h(n_h)}^* = 2\mu_h^*$ . Therefore,

$$E(\bar{X}_{SDERSS1}^*) = \left( \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} \mu_{hi}^* \right) \right)$$

It was shown by [2] that

$$\begin{aligned}\mu_h &= \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h}{2}} \mu_{hi}^* \right), \\ E(\bar{X}_{SDERSS1}^*) &= \left( \sum_{h=1}^L W_h \cdot \mu_h \right) = \mu.\end{aligned}$$

Second: If the sample sizes in the strata  $n_h, h=1,2,\dots,L$  are odd, we have

$$\begin{aligned}\bar{X}_{SDERSS2}^* &= E \left( \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h-1}{2}} X_{hi(1)}^* + X_{hi(\frac{n_h+1}{2})}^* + \sum_{i=\frac{n_h+1}{2}+1}^n X_{hi(n_h)}^* \right) \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h-1}{2}} E(X_{hi(1)}^*) + E(X_{hi(\frac{n_h+1}{2})}^*) + \sum_{i=\frac{n_h+1}{2}+1}^n E(X_{hi(n_h)}^*) \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \sum_{i=1}^{\frac{n_h-1}{2}} \mu_{h(1)}^* + \mu_{h(\frac{n_h+1}{2})}^* + \sum_{i=\frac{n_h+1}{2}+1}^n \mu_{h(n_h)}^* \right)\end{aligned}$$

Since the distribution is symmetric about the  $\mu$ , then we have  $\mu_{h(1)}^* + \mu_{h(n_h)}^* = 2\mu_h^*$ , and the mean = median. Therefore,

$$\begin{aligned}E(\bar{X}_{SDERSS2}^*) &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \left( \frac{n_h-1}{2} \right) \mu_{h(1)}^* + \mu_{h(\frac{n_h+1}{2})}^* + \left( \frac{n_h-1}{2} \right) \mu_{h(n_h)}^* \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \left( \frac{n_h-1}{2} \right) (\mu_{h(1)}^* + \mu_{h(n_h)}^*) + \mu_{h(\frac{n_h+1}{2})}^* \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} \left( \left( \frac{n_h-1}{2} \right) (2\mu_h) + \mu_h \right) \\ &= \sum_{h=1}^L W_h \cdot \frac{1}{n_h} ((n_h-1)\mu_h + \mu_h) = \sum_{h=1}^L W_h \cdot \mu_h = \mu.\end{aligned}$$

**Lemma 4.** If the distribution is symmetric about  $\mu$ , then

$$Var(\bar{X}_{SDERSS}) < Var(\bar{X}_{SRS}).$$

**Proof.** If the sample sizes in the strata  $n_h, h=1,\dots,L$  are even, the variance of  $\bar{X}_{SDERSS}$  is

$$Var(\bar{X}_{SDERSS1}^*) = \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( \sum_{i=1}^{\frac{n_h}{2}} \sigma_{hi(1)}^{2*} + \sum_{i=\frac{n_h}{2}+1}^n \sigma_{hi(n_h)}^{2*} \right).$$

Nevertheless,  $\sigma_{hi}^{2*} < \sigma_h^2$  for each stratum  $h=1,2,\dots,L$ , this implies

$$Var(\bar{X}_{SDERSS1}^*) = \sum_{h=1}^L W_h^2 \cdot \frac{1}{n_h^2} \sigma_{hi}^{2*} < \sum_{h=1}^L W_h^2 \cdot \frac{1}{n_h^2} \sigma_h^2 = Var(\bar{X}_{SSRS}) < Var(\bar{X}_{SRS}).$$

The proof in case of odd sample sizes is similar.

**Example 2:** For the uniform distribution with parameters  $\alpha$  and  $\beta$ ;  $U(\alpha, \beta)$ , the probability density function and the cumulative distribution function are defined as

$$f(x) = \begin{cases} 0, & x < \alpha \\ \frac{1}{\beta - \alpha}, & \alpha < x < \beta \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad F(x) = \begin{cases} 0, & x < \alpha \\ \frac{x - \alpha}{\beta - \alpha}, & \alpha \leq x < \beta \\ 1, & x \geq \beta \end{cases}.$$

The mean and the variance are given, respectively by

$$\mu = \frac{\alpha + \beta}{2}, \text{ and } \sigma^2 = \frac{(\beta - \alpha)^2}{12}.$$

Now, for  $U(0, \beta)$ , the mean and the variance, respectively are  $\mu = \frac{\beta}{2}, \sigma^2 = \frac{\beta^2}{12}$ .

From [6], formula for the variance of  $X_{h(1)}^*$  and  $X_{h(n)}^*$  which are equal since the distribution is symmetric about the mean, for stratum  $h$  in case of even sample size, can be defined as

$$Var(X_{h(1)}^*) = Var(X_{h(n_h)}^*) = \frac{n_h^2 \beta^2}{(n_h^2 + 2)(n_h^2 + 1)}.$$

**Case 1.** The efficiency of  $\bar{X}_{SDERSS1}^*$  relative to  $\bar{X}_{SSRS}$ . Now, the variance of SDERSS in case of even sample size can be calculated as

$$\begin{aligned} Var(\bar{X}_{SDERSS1}^*) &= \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( \sum_{i=1}^{\frac{n_h}{2}} Var(X_{h(i)}^*) + \sum_{i=\frac{n_h}{2}+1}^{n_h} Var(X_{h(n_h)}^*) \right) \\ &= \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( \sum_{i=1}^{\frac{n_h}{2}} \frac{n_h^2 \beta^2}{(n_h^2 + 2)(n_h^2 + 1)} + \sum_{i=\frac{n_h}{2}+1}^{n_h} \frac{n_h^2 \beta^2}{(n_h^2 + 2)(n_h^2 + 1)} \right) \\ &= \sum_{h=1}^L \frac{W_h^2}{n_h^2} \left( n_h \cdot \frac{n_h^2 \beta^2}{(n_h^2 + 2)(n_h^2 + 1)} \right) = \sum_{h=1}^L \frac{W_h^2 n_h \beta^2}{(n_h^2 + 2)(n_h^2 + 1)}. \end{aligned}$$

The variance of SSRS for  $U(0, \beta)$  is defined as

$$Var(\bar{X}_{SSRS}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \cdot \sigma_{Xh}^2 = \sum_{h=1}^L \frac{W_h^2}{n_h} \cdot \frac{\beta^2}{12}$$

The efficiency of  $\bar{X}_{SDERSS1}^*$  relative to  $\bar{X}_{SSRS}$  is

$$\begin{aligned} eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) &= \frac{Var(\bar{X}_{SSRS})}{Var(\bar{X}_{SDERSS1}^*)} \\ &= \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \cdot \frac{\beta^2}{12}}{\sum_{h=1}^L \frac{W_h^2 n_h \beta^2}{(n_h^2 + 2)(n_h^2 + 1)}} = \frac{\sum_{h=1}^L \frac{W_h^2}{12 n_h}}{\sum_{h=1}^L \frac{W_h^2 n_h}{(n_h^2 + 2)(n_h^2 + 1)}} > 1. \end{aligned}$$

Algebraically it is not easy to show that the last quantity is greater than 1, so some different cases are considered, two or three strata, the following results are obtained:

1. If  $L=2, h=1, 2$  when  $n_1=4$  and  $n_2=6$  with total  $n=10$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 2.3011$ .
2. If  $L=2, h=1, 2$  when  $n_1=6$  and  $n_2=8$  with total  $n=14$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 4.2737$ .
3. If  $L=2, h=1, 2$  when  $n_1=8$  and  $n_2=10$  with total  $n=18$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 6.9311$ .
4. If  $L=3, h=1, 2, 3$  when  $n_1=4$ ,  $n_2=4$  and  $n_3=6$  with total  $n=14$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 2.0399$ .
5. If  $L=3, h=1, 2, 3$  when  $n_1=4$ ,  $n_2=6$  and  $n_3=8$  with total  $n=18$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 3.1112$ .

6. If  $L=3, h=1, 2, 3$  when  $n_1=6$ ,  $n_2=6$  and  $n_3=6$  with total  $n=18$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SSRS}) = 3.2546$ .

**Case 2:** The efficiency of  $\bar{X}_{SDERSS1}^*$  relative to  $\bar{X}_{SRS}$ . Now, the variance of SRS for  $U(0, \beta)$  is given by

$$Var(\bar{X}_{SRS}) = \frac{\beta^2}{12}.$$

The efficiency of  $\bar{X}_{SDERSS1}^*$  relative to  $\bar{X}_{SRS}$  is

$$\begin{aligned} eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) &= \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{SDERSS1}^*)} = \frac{\sum_{h=1}^L \frac{\beta^2}{12}}{\sum_{h=1}^L \frac{W_h^2 n_h \beta^2}{(n_h^2 + 2)(n_h^2 + 1)}} \\ &= \frac{\sum_{h=1}^L \frac{1}{12}}{\sum_{h=1}^L \frac{W_h^2 n_h}{(n_h^2 + 2)(n_h^2 + 1)}} = \frac{\frac{L}{12}}{12 \cdot \sum_{h=1}^L \frac{W_h^2 n_h}{(n_h^2 + 2)(n_h^2 + 1)}} > 1. \end{aligned}$$

1. If  $L=2, h=1, 2$  when  $n_1=4$  and  $n_2=6$  with total  $n=10$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) = 45.9415$ .
2. If  $L=2, h=1, 2$  when  $n_1=6$  and  $n_2=8$  with total  $n=14$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) = 119.6708$ .
3. If  $L=2, h=1, 2$  when  $n_1=8$  and  $n_2=10$  with total  $n=18$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) = 249.5197$ .
4. If  $L=3, h=1, 2, 3$  when  $n_1=4$ ,  $n_2=4$  and  $n_3=6$  with total  $n=14$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) = 85.6810$ .
5. If  $L=3, h=1, 2, 3$  when  $n_1=4$ ,  $n_2=6$  and  $n_3=8$  with total  $n=18$ ,  $eff(\bar{X}_{SDERSS1}^*, \bar{X}_{SRS}) = 167.9995$ .

#### IV. SIMULATION STUDY

A simulation study is conducted to investigate the performance of SDERSS for estimating the population mean with respect to SRS, SSRS and SRSS. Symmetric distributions, namely; uniform and normal and asymmetric distributions namely exponential, gamma and Weibull have been considered for samples of sizes  $n=9, 12, 15, 18$ , assuming that the population is partitioned into two or three strata. Using 100000 replications, estimates of the means, variances and mean square errors were computed.

When the underlying distribution is symmetric, the efficiency of SDERSS relative to SRS, SSRS, SRSS is given by

$$eff(\bar{X}_{SDERSS}, \bar{X}_T) = \frac{Var(\bar{X}_T)}{Var(\bar{X}_{SDERSS})}, T = SSRS, SRSS, SRS,$$

and if the distribution is asymmetric, the efficiency of SDERSS relative to SSRS and SRSS is defined in terms of mean square error (MSE) as

$$eff(\bar{X}_{SDERSS}, \bar{X}_T) = \frac{MSE(\bar{X}_T)}{MSE(\bar{X}_{SDERSS})}, T = SSRS, SRSS,$$

where

$$MSE(\bar{X}_T) = Var(\bar{X}_T) + [Bias(\bar{X}_T)]^2, \\ T = SSRS, SRSS, SDQRSS.$$

The values of the relative efficiency and bias found under different distributional assumptions are provided in Tables I-III.

TABLE I

THE RELATIVE EFFICIENCY FOR ESTIMATING THE POPULATION MEAN USING SDERSS WITH RESPECT TO SRSS, SSRS AND SRS WITH SAMPLE SIZE  $N = 9$  AND  $N = 12$

Distribution	$n = 9:$			$n = 12:$		
	$n_1 = 4, n_2 = 5$			$n_1 = 5, n_2 = 7$		
	SRSS	SSRS	SRS	SRSS	SSRS	SRS
Uniform (0,1)	20.75	26.65	26.25	33.17	34.44	35.67
Normal (0,1)	18.66	21.89	20.64	29.45	31.01	29.85
Exponential (1)	16.95	16.53	16.25	19.35	20.44	19.23
Gamma (1,2)	16.02	15.76	15.54	19.65	19.83	20.21
Weibull (1,2)	14.65	13.95	14.97	19.78	19.82	19.37

TABLE II

THE RELATIVE EFFICIENCY FOR ESTIMATING THE POPULATION MEAN USING SDERSS WITH RESPECT TO SRSS, SSRS AND SRS WITH SAMPLE SIZE  $N = 15$  AND  $N = 18$

Distribution	$n = 15: n_1 = 3, n_2 = 5,$			$n = 18: n_1 = 4, n_2 = 6,$		
	$n_3 = 7$			$n_3 = 8$		
	SRSS	SSRS	SRS	SRSS	SSRS	SRS
Uniform (0,1)	46.74	50.23	48.96	60.42	63.76	61.65
Normal (0,1)	31.78	34.76	33.76	46.54	49.78	48.98
Exponential (1)	20.97	22.64	20.78	22.67	22.96	22.87
Gamma (1,2)	19.87	20.65	19.43	21.73	21.23	19.87
Weibull (1,2)	19.87	20.51	19.98	21.14	21.05	21.79

TABLE III

THE VALUES OF BIAS OF SDERSS FOR DIFFERENT DISTRIBUTIONS AND DIFFERENT NUMBERS OF STRATA

Exponential (1)	Gamma (1,2)	Weibull (1,2)
When $n = 9$ and two strata $n_1 = 4$ and $n_2 = 5$		
0.0713	0.2323	0.8573
When $n = 12$ and two strata $n_1 = 5$ and $n_2 = 7$		
0.0345	0.2427	0.4017
When $n = 15$ and three strata $n_1 = 3, n_2 = 5$ and $n_3 = 7$		
0.0127	0.0834	0.0022
When $n = 18$ and three strata $n_1 = 4, n_2 = 6$ and $n_3 = 8$		
0.0439	0.1119	0.0315

Based on simulation study, we can conclude that:

A gain in efficiency is attained using SDERSS method as opposed to the other contending methods that have been discussed when estimating the population mean of the variable of interest. When the performance of SDERSS are compared to either SRSS, SSRS or SRS, it is found that SDERSS is more efficient, as shown by all the values of relative efficiency which are greater than 1. When the performances of all estimators are compared, the efficiency of SDERSS estimator is found to be more superior when the underlying distributions are symmetric as compared to asymmetric. The relative efficiency of SDERSS estimator to those estimators based on SRS, SSRS and SRSS are increasing as the sample size increases.

## V.CONCLUSIONS

In this paper, we have suggested a new estimator of the population mean using SDERSS. The performance of the estimator based on SDERSS is compared with those found using SRSS, SSRS and SRS for the same number of measured units. It is found that SDERSS produces estimator of the population mean that is unbiased of symmetric distributions and SDERSS is more efficient than SRSS, SSRS and SRS. Thus, SDERSS should be more preferred than SRSS, SSRS and SRS for both symmetric and asymmetric distributions.

## ACKNOWLEDGMENT

This publication was made possible because of the support of Qatar University. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Qatar University.

## REFERENCES

- [1] Al-Omari, A.I. (2011). Estimation of mean based on modified robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 81(8): 1055- 1066.
- [2] Al-Omari, A.I. (2012). Ratio estimation of population mean using auxiliary information in simple random sampling and median ranked set sampling. *Statistics and Probability Letters*, 82(11): 1883-1990.
- [3] Al-Omari, A.I. and Al-Nasser, A.D. (2011). Statistical quality control limits for the sample mean chart using robust extreme ranked set sampling. *Economic Quality Control*, 26: 73-89.
- [4] Al-Saleh, M.F. & Al-Hadhrani, S. A. (2003). Estimation of the mean of exponential distribution using moving extremes ranked set sampling. *Statistical Papers* 44: 367-382.
- [5] Al-Saleh, M.F. and Al-Kadiri, M. (2000). Double ranked set sampling. *Statistics Probability Letters*, 48, 205 - 212.
- [6] Al-Saleh, M.F. and Al-Omari, A.I. (2002). Multistage ranked set sampling. *Journal of Statistical Planning and Inference*, 102, 273-286.
- [7] Bouza, C.N. (2002) Ranked set subsampling the non response strata for estimating the difference of means. *Biometrical Journal*, 44:903915.
- [8] Chen, Z., Bai, Z., and Sinha, B.K. (2004). Ranked set sampling: Theory and Applications. *Lectures Notes in Statistics*, 176. Springer, N. York
- [9] Dell, T.R., and Clutter, J.L. (1972). Ranked set sampling theory with order statistics background. *Biometrika*, 28, 545-555.
- [10] Jemain, A.A., A;-Omari, A.I. & Ibrahim, K. (2007). Multistage extreme ranked set samples for estimating the population mean. *Journal of Statistical Theory and Applications* 6(4): 456-471.
- [11] McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research*, 3, 385-390.
- [12] Ohayama, T., Doi, J., and Yanagawa, T. (2008). Estimating population characteristics by incorporating prior values in stratified random

- sampling/ranked set sampling. *Journal of Statistical Planning and Inference*, 138, 4021-4032.
- [13] Samawi, H. M. (1996). Stratified ranked set sample. *Pakistan Journal of Statistics*, 12(1), 9-16.
- [14] Samawi, H., Abu-Dayyeh, W., and Ahmed, S. (1996). Extreme ranked set sampling. *The Biometrical Journal*, 30, 577-586.
- [15] Samawi, H.M. (2002). On double extreme ranked set sample with application to regression estimator. *Metron. LXn*. 1-2:53-66.
- [16] 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, 1-31.