Journal of Data Science **10**(2012), 129-141

Estimation of a Scale Parameter of Morgenstern Type Bivariate Uniform Distribution by Ranked Set Sampling

Saeid Tahmasebi[1]* and Ali Akbar Jafari[2]
[1] *Persian Gulf University and* [2] *Yazd University*

Abstract: In this paper, we obtain several estimators of a scale parameter of Morgenstern type bivariate uniform distribution (MTBUD) based on the observations made on the units of the ranked set sampling regarding the study variable Y which is correlated with the auxiliary variable X, when (X, Y) follows a MTBUD. Efficiency comparisons among these estimators are also made in this work. Finally, we illustrate the methods developed by using a real data set.

Key words: Best linear unbiased estimator, concomitants of order statistics, extreme ranked set sampling, Morgenstern family, ranked set sampling, relative efficiency.

1. Introduction

Ranked set sampling (RSS) was first proposed by McIntyre (1952) for estimating the mean pasture yields. McIntyre indicates that RSS is a more efficient sampling method than simple random sampling (SRS) method for estimating the population mean. In the RSS technique, the sample selection procedure is composed of two stages. At the first stage of sample selection, n simple random samples of size n are drawn from an infinite population and each sample is called a set . Then, each of observations are ranked from the smallest to the largest according to variable of interest, say X, in each set. Ranking of the units is done with a low-level measurement such as using previous experiences, visual measurement or using a concomitant variable. At the second stage, the first observation unit from the first set, the second observation unit from the second set and going on like this nth-observation unit from the nth-set are taken and measured according to the variable X with a high level of measurement satisfying the desired sensitivity. The obtained sample is called an RSS. Stokes (1980) introduced a modified ranked set sampling procedure in which only the largest or the smallest

*Corresponding author.

judgment ranked unit is chosen for quantification. Samawi *et al.* (1996) investigated the use of a variety of extreme ranked set samples (ERSS) for estimating the population mean. Another scheme of ranked set sampling was investigated by Al-Odat and Al-Saleh (2001) which is the moving extreme ranked set sampling (MERSS). It is a modification of the RSS that only the lowest or largest unit of sets of varied sizes is measured. Stokes (1977) applied RSS for bivariate random variable (X, Y), where X is the variable of interest and Y is a concomitant variable that is not of direct interest but is relatively easy to measure. Let $X_{(r)r}$ be the observation measured on the variable X in the rth unit of the RSS and let $Y_{[r]r}$ be the corresponding measurement made on the study variable Y of the same unit, $r = 1, 2, \cdots, n$. Then clearly $Y_{[r]r}$ is the concomitant of rth order statistic arising from the rth sample. The procedure of RSS described by Stokes (1977) is as follows:

Step 1. Randomly select n independent bivariate samples, each of size n.

Step 2. Rank the units within each sample with respect to a variable of interest X together with the Y variate associated.

Step 3. In the rth sample of size n, select the unit $(X_{(r)r}, Y_{[r]r})$, $r = 1, 2, \cdots, n$.

A general family of bivariate distributions is proposed by Morgenstern (1956) with specified marginal distributions $F_X(x)$ and $F_Y(y)$ as

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)\left[1 + \alpha\,(1 - F_X(x))(1 - F_Y(y))\right], \quad -1 \le \alpha \le 1, \quad (1.1)$$

where α is the association parameter between X and Y. A member of this family is Morgenstern type bivariate uniform distribution (MTBUD) with the probability density function (pdf),

$$f_{X,Y}(x, y) = \frac{1}{\theta_1\theta_2}\left[1 + \alpha\left(1 - \frac{2x}{\theta_1}\right)\left(1 - \frac{2y}{\theta_2}\right)\right], \quad 0 < x < \theta_1;\ 0 < y < \theta_2. \quad (1.2)$$

From Scaria and Nair (1999), the pdf of $Y_{[r]r}$ for $1 \le r \le n$ is given by

$$\begin{aligned} g_{[r]r}(y) &= \int f_{Y|X}(y|x)f_r(x)dx \\ &= \frac{1}{\theta_2}\left[1 + \alpha\left(\frac{n - 2r + 1}{n + 1}\right)\left(1 - \frac{2y}{\theta_2}\right)\right], \quad 0 < y < \theta_2, \end{aligned} \quad (1.3)$$

where $f_r(x)$ is density function of $X_{(r)r}$, i.e.

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!}\left[\frac{x^{r-1}(\theta_1 - x)^{n-r}}{\theta_1^n}\right], \quad 0 < x < \theta_1, \quad (1.4)$$

and therefore, the mean and variance of $Y_{[r]r}$ for $1 \le r \le n$ are given as

$$E\left[Y_{[r]r}\right] = \theta_2\beta_r, \quad Var\left[Y_{[r]r}\right] = \theta_2^2\lambda_r, \quad (1.5)$$

where

$$\beta_r = \frac{1}{2}\left[1-\alpha\frac{n-2r+1}{3(n+1)}\right], \quad \lambda_r = \frac{1}{12}\left[1-\frac{1}{3}(\frac{\alpha(n-2r+1)}{n+1})^2\right].$$

Stokes (1995) has obtained the estimation of parameters of location-scale family of distribution by RSS. Lam *et al.* (1994) used RSS to estimate the two-parameter exponential distribution. Al-saleh and Ananbeh (2005, 2007) estimated the means of the bivariate normal distribution using moving extremes RSS with concomitant variable. Estimation of a parameter of Morgenstern type bivariate exponential distribution by using RSS was considered by Chacko and Thomas (2008). Al-saleh and Diab (2009) considered estimation of the parameters of Downton's bivariate exponential distribution using RSS scheme. Barnett and Moore (1997) derived the best linear unbiased estimator (BLUE) for the mean of Y, based on a ranked set sample obtained using an auxiliary variable X for ranking the sample units.

The organization of this article is as follows. In Section 2, we present four estimators for the scale parameter, θ_2 in MTBUD: an estimator based on the ranked set sample mean, an estimator using simple random sample, a BLUE using RSS, and a BLUE by using the upper ranked set sample (URSS) based on measurement of Y which is taken on the unit that has the maximum value for the X variable. Also, we consider the efficiency of these four estimators. In Section 3, we obtain different estimators for θ_2 in MTBUD by using ERSS and MERSS methods. Also, we evaluate the efficiency of all estimators considered in this paper. In Section 4, we illustrate the ERSS and MERSS methods using a real data set.

2. Estimators for θ_2 Based on RSS

Suppose that the random variable (X, Y) has a MTBUD as defined in (1.2). Let $Y_{[r]r}$, $r = 1, 2, \cdots, n$, be the RSS observations made on the units of the ranked set sampling regarding the study variable Y which is correlated with the auxiliary variable X. Then an unbiased estimator for θ_2 based on RSS mean in (1.5) is given by

$$\hat{\theta}_{2,\mathrm{RSS}} = \frac{2}{n}\sum_{r=1}^{n} Y_{[r]r}, \tag{2.1}$$

and its variance is

$$Var(\hat{\theta}_{2,\mathrm{RSS}}) = \frac{\theta_2^2}{3n}\left[1-\frac{1}{3n}\sum_{r=1}^{n}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2\right]. \tag{2.2}$$

We know that Y has a uniform distribution, $Y \sim U(0,\theta_2)$. Therefore, an unbiased estimator of θ_2 based on a simple random sample (SRS) of size n from $U(0,\theta_2)$ is $2\overline{Y}$ with variance $\theta_2^2/(3n)$. The relative efficiency of $\hat{\theta}_{2,\mathrm{RSS}}$ to $2\overline{Y}$ is given by

$$e_1 = e\left(\hat{\theta}_{2,\mathrm{RSS}} \mid 2\overline{Y}\right) = \frac{Var\left(2\overline{Y}\right)}{Var\left(\hat{\theta}_{2,\mathrm{RSS}}\right)} = \frac{1}{1-\frac{\alpha^2}{9}\left(\frac{n-1}{n+1}\right)}.$$

Note that $1 \leq e_1 \leq 9/8$. For fixed $n > 1$, the efficiency increases as $|\alpha|$ increases. Thus, we conclude that $\hat{\theta}_{2,\mathrm{RSS}}$ is more efficient than $2\overline{Y}$.

Now, we study the efficiency of $\hat{\theta}_{2,\mathrm{RSS}}$ relative to the BLUE of θ_2, θ_2^*, based on $Y_{[r]r}$, $r = 1,2,\cdots,n$ of MTBUD, when α is known. Let $\mathbf{Y}_{[n]} = (Y_{[1]1}, Y_{[2]2},\cdots, Y_{[n]n})'$, then by using David and Nagaraja (2003) the BLUE of θ_2 is derived as

$$\theta_2^* = \left(\beta' W^{-1}\beta\right)^{-1}\beta' W^{-1}\mathbf{Y}_{[n]} = \sum_{r=1}^{n} a_r Y_{[r]r}, \tag{2.3}$$

where $\beta = (\beta_1,\beta_2,\cdots,\beta_n)'$, $W = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$, and $a_r = (\beta_r/\lambda_r)\cdot(\sum_{i=1}^{n}\beta_i^2/\lambda_i)^{-1}$, $r = 1,2,\cdots,n$. The variance of θ_2^* is

$$Var\left[\theta_2^*\right] = \left(\beta' W^{-1}\beta\right)^{-1}\theta_2^2 = \frac{\theta_2^2}{\sum_{r=1}^{n}\beta_r^2/\lambda_r}, \tag{2.4}$$

and therefore, the relative efficiency of $\hat{\theta}_{2,\mathrm{RSS}}$ to θ_2^* is given by

$$e_2 = e\left(\theta_2^* \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \sum_{r=1}^{n}\left[\frac{\left(1-\frac{\alpha(n-2r+1)}{3(n+1)}\right)^2}{1-\frac{1}{3}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2}\right]\left[\frac{1}{n}-\frac{\alpha^2(n-1)}{9n(n+1)}\right]. \tag{2.5}$$

The values of e_2 for $n = 2(2)10(5)25$, and $\alpha = \pm.25, \pm.5, \pm.75, \pm1$ have been computed in Table 1. As it can be seen from Table 1, θ_2^* is more efficient than $\hat{\theta}_{2,\mathrm{RSS}}$ and for fixed $n \geq 2$, the efficiency increases with $|\alpha|$.

Remark 1. Our assumption is that α is known, but sometimes α may not be known. We know that the correlation coefficient between X and Y in MTBUD is $\frac{\alpha}{3}$. So by using the sample correlation coefficient q of the RSS observations $(X_{(r)r}, Y_{[r]r})$, $r = 1,2,\cdots,n$ an estimator for α is given

$$\widehat{\alpha} = \begin{cases} -1, & q < -\frac{1}{3}, \\ 3q, & -\frac{1}{3} \leq q \leq \frac{1}{3}, \\ 1, & \frac{1}{3} < q. \end{cases}$$

Table 1: The values of $e_2 = e(\theta_2^* \mid \hat{\theta}_{2,\text{RSS}})$, $e_3 = e(\tilde{\theta}_2 \mid \theta_2^*)$, and $e_4 = e(\tilde{\theta}_2 \mid \hat{\theta}_{2,\text{RSS}})$ in MTBUD

n	α	e_2	e_3	e_4	n	α	e_2	e_3	e_4
2	-1.00	1.0123	0.7805	0.7901	**10**	-1.00	1.0413	0.5944	0.6189
	-0.75	1.0069	0.8345	0.8402		-0.75	1.0201	0.6730	0.6865
	-0.50	1.0031	0.8892	0.8919		-0.50	1.0081	0.7657	0.7719
	-0.25	1.0008	0.9445	0.9452		-0.25	1.0019	0.8739	0.8755
	0.25	1.0008	1.0555	1.0563		0.25	1.0019	1.1484	1.1505
	0.50	1.0031	1.1108	1.1141		0.50	1.0081	1.3257	1.3365
	0.75	1.0069	1.1655	1.1736		0.75	1.0201	1.5433	1.5743
	1.00	1.0123	1.2195	1.2345		1.00	1.0413	1.8203	1.8955
4	-1.00	1.0266	0.6612	0.6787	**15**	-1.00	1.0456	0.5817	0.6081
	-0.75	1.0138	0.7356	0.7457		-0.75	1.0219	0.6593	0.6736
	-0.50	1.0058	0.8164	0.8211		-0.50	1.0088	0.7538	0.7603
	-0.25	1.0014	0.9043	0.9055		-0.25	1.0021	0.8663	0.8681
	0.25	1.0014	1.1047	1.1062		0.25	1.0021	1.1603	1.1627
	0.50	1.0058	1.2196	1.2266		0.50	1.0088	1.3564	1.3683
	0.75	1.0138	1.3465	1.3650		0.75	1.0219	1.6044	1.6394
	1.00	1.0266	1.4877	1.5272		1.00	1.0456	1.9342	2.0223
6	-1.00	1.0340	0.6228	0.6439	**20**	-1.00	1.0479	0.5758	0.6033
	-0.75	1.0170	0.7009	0.7128		-0.75	1.0228	0.6525	0.6673
	-0.50	1.0070	0.7889	0.7944		-0.50	1.0091	0.7477	0.7544
	-0.25	1.0017	0.8881	0.8895		-0.25	1.0021	0.8624	0.8642
	0.25	1.0017	1.1271	1.1290		0.25	1.0021	1.1667	1.1691
	0.50	1.0070	1.2730	1.2819		0.50	1.0091	1.3731	1.3855
	0.75	1.0170	1.4428	1.4673		0.75	1.0228	1.6384	1.6757
	1.00	1.0340	1.6445	1.7004		1.00	1.0479	1.9999	2.0956
8	-1.00	1.0384	0.6047	0.6278	**25**	-1.00	1.0492	0.5725	0.6007
	-0.75	1.0189	0.6834	0.6963		-0.75	1.0233	0.6484	0.6635
	-0.50	1.0077	0.7746	0.7805		-0.50	1.0092	0.7440	0.7509
	-0.25	1.0018	0.8793	0.8809		-0.25	1.0021	0.8600	0.8619
	0.25	1.0018	1.1400	1.1421		0.25	1.0021	1.1706	1.1731
	0.50	1.0077	1.3047	1.3147		0.50	1.0092	1.3835	1.3963
	0.75	1.0189	1.5026	1.5309		0.75	1.0233	1.6600	1.6988
	1.00	1.0384	1.7475	1.8146		1.00	1.0492	2.0427	2.1434

We can also provided a ranked set sample of size n by each sample measurement of Y which is taken on the unit that has the maximum value for the X variable. So, let $Y_{[n]r}$ be concomitants of largest order statistics $X_{(n)r}$ of the rth sample for $r = 1, 2, \cdots, n$. Then we call the collection of observations $Y_{[n]1}, Y_{[n]2}, \cdots, Y_{[n]n}$ as the Upper ranked set sample (URSS). We can derive BLUE of θ_2 based on the observations URSS. From (1.5) the mean and variance of $Y_{[n]r}$, for $1 \leq r \leq n$, are given as

$$E[Y_{[n]r}] = \theta_2 \beta_n, \quad Var[Y_{[n]r}] = \theta_2^2 \lambda_n.$$

Note that, for $1 \leq r < s \leq n$, $Cov[Y_{[n]r}, Y_{[n]s}] = 0$. Now, a BLUE of θ_2 based on URSS is obtained as

$$\tilde{\theta}_2 = \frac{1}{n\beta_n} \sum_{r=1}^{n} Y_{[n]r}, \tag{2.6}$$

and its variance is given by

$$Var\left(\tilde{\theta}_2\right) = \frac{\theta_2^2 \lambda_n}{n\beta_n^2}. \tag{2.7}$$

From (2.4) and (2.7) the efficiency of θ_2^* relative to $\tilde{\theta}_2$ is obtained as

$$e_3 = e\left(\tilde{\theta}_2 \mid \theta_2^*\right) = \frac{n\left[1 - \alpha\left(\frac{1-n}{3(n+1)}\right)\right]^2}{\sum_{r=1}^{n}\left[\frac{(1-\frac{\alpha(n-2r+1)}{3(n+1)})^2}{1-\frac{1}{3}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2}\right]\left[1 - \frac{1}{3}\left(\frac{\alpha(1-n)}{n+1}\right)^2\right]}.$$

Also, the efficiency of $\hat{\theta}_{2,\mathrm{RSS}}$ relative to $\tilde{\theta}_2$ is derived as

$$e_4 = e\left(\tilde{\theta}_2 \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \frac{\left[1 - \alpha\left(\frac{1-n}{3(n+1)}\right)\right]^2\left[1 - \frac{\alpha^2(n-1)}{9(n+1)}\right]}{1 - \frac{1}{3}\left(\frac{\alpha(1-n)}{n+1}\right)^2}.$$

We have computed the values of e_3 and e_4 for $n = 2(2)10(5)25$, and $\alpha = \pm.25, \pm.5, \pm.75, \pm1$ in Table 1. From the table, we can easily see that $\tilde{\theta}_2$ is relatively more efficient than θ_2^* and $\hat{\theta}_{2,\mathrm{RSS}}$ for $0 < \alpha \leq 1$. Also, e_3 and e_4 increases (decreases) with n and $0 < \alpha \leq 1$ $(-1 \leq \alpha < 0)$. Thus, we conclude that θ_2^* and $\hat{\theta}_{2,\mathrm{RSS}}$ are relatively more efficient than $\tilde{\theta}_2$ when $-1 \leq \alpha < 0$.

3. Estimators for θ_2 Based on ERSS and MERSS Methods

In this section, first we derive different estimators for θ_2 based on extreme ranked set sampling (ERSS) method with concomitant variable. This method introduced by Samawi *et al.* (1996) and can be described as follows:

Step 1. Select n random samples each of size n bivariate units from the population.

Step 2. If the sample size n is even, then select from $n/2$ samples the smallest ranked unit X together with the associated Y and from the other $n/2$ samples the largest ranked unit X together with the associated Y. This selected observations $(X_{(1)1}, Y_{[1]1})$, $(X_{(n)2}, Y_{[n]2})$, $(X_{(1)3}, Y_{[1]3})$, $\cdots$, $(X_{(1)n-1}, Y_{[1]n-1})$, $(X_{(n)n}, Y_{[n]n})$ can be denoted by ERSS_1.

Step 3. If n is odd then select from $(n-1)/2$ samples the smallest ranked unit X together with the associated Y and from the other $(n-1)/2$ samples the largest ranked unit X together with the associated Y and from one sample the median of the sample for actual measurement. In this case the selected observations $(X_{(1)1}, Y_{[1]1})$, $(X_{(n)2},\ Y_{[n]2})$, $(X_{(1)3},\ Y_{[1]3})$, $\cdots$, $(X_{(n)n-1},\ Y_{[n]n-1})$, $((X_{(1)n} + X_{(n)n})/2, (Y_{[1]n} + Y_{[n]n})/2)$ can be denoted ERSS_2 and $(X_{(1)1}, Y_{[1]1})$, $(X_{(n)2}, Y_{[n]2})$, $(X_{(1)3}, Y_{[1]3})$, $\cdots$, $(X_{(n)n-1}, Y_{[n]n-1})$, $(X_{((n+1)/2)n}, Y_{[(n+1)/2]n})$ can be denoted by ERSS_3.

Now, if n is even then the estimator of the θ_2 using ERSS_1 is defined as

$$\hat{\theta}_{2,\text{ERSS}_1} = \frac{2}{n}\sum_{r=1}^{n/2}\left(Y_{[1]2r-1} + Y_{[n]2r}\right). \tag{3.1}$$

Since the elements in (3.1) are independent, so we have

$$Var\left(\hat{\theta}_{2,\text{ERSS}_1}\right) = \frac{\theta_2^2}{3n}\left[1 - \frac{1}{3}\left(\frac{\alpha(n-1)}{n+1}\right)^2\right]. \tag{3.2}$$

If n is odd then the estimators of θ_2 using ERSS_2 and ERSS_3 are obtained as

$$\hat{\theta}_{2,\text{ERSS}_2} = \frac{2\left(Y_{[1]1} + Y_{[n]2} + Y_{[1]3} + \cdots + Y_{[n]n-1} + (Y_{[1]n} + Y_{[n]n})/2\right)}{n}, \tag{3.3}$$

$$\hat{\theta}_{2,\text{ERSS}_3} = \frac{2\left(Y_{[1]1} + Y_{[n]2} + Y_{[1]3} + \cdots + Y_{[n]n-1} + Y_{e[(n+1)/2]n}\right)}{n}. \tag{3.4}$$

In the estimator $\hat{\theta}_{2,\text{ERSS}_2}$, it is easy to see that $Y_{[1]1}, Y_{[n]2}, Y_{[1]3}, \cdots, Y_{[n]n-1}$ are independent of $Y_{[1]n}$ and $Y_{[n]n}$, but the random variables $Y_{[1]n}$ and $Y_{[n]n}$ are dependent. Also, we can easily check that $Y_{[1]1}, Y_{[n]2}, Y_{[1]3}, \cdots, Y_{[n]n-1}$ and $Y_{[(n+1)/2]n}$ are all independent in $\hat{\theta}_{2,\text{ERSS}_3}$. So by using some simple algebra we have

$$Var\left(\hat{\theta}_{2,\text{ERSS}_2}\right) = \frac{\theta_2^2}{3n}\left[1 - \frac{\alpha^2(n-1)^3}{3n(n+1)^2} - \frac{1}{2n} + \frac{\alpha^2(2-n)}{6n(n+2)}\right], \tag{3.5}$$

$$Var\left(\hat{\theta}_{2,\text{ERSS}_3}\right) = \frac{\theta_2^2}{3n}\left[1 - \frac{\alpha^2(n-1)^3}{3n(n+1)^2}\right]. \tag{3.6}$$

Now, by using (2.2), (3.2), (3.5) and (3.6) the efficiency of $\hat{\theta}_{2,\text{RSS}}$ relative to the estimators $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$, respectively, are given by

$$e_5 = e\left(\hat{\theta}_{2,\mathrm{ERSS}_1} \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \frac{1-\frac{\alpha^2(n-1)}{9(n+1)}}{1-\frac{1}{3}\left(\frac{\alpha(n-1)}{n+1}\right)^2},$$

$$e_6 = e\left(\hat{\theta}_{2,\mathrm{ERSS}_2} \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \frac{1-\frac{\alpha^2(n-1)}{9(n+1)}}{1-\frac{\alpha^2(n-1)^3}{3n(n+1)^2}-\frac{1}{2n}+\frac{\alpha^2(2-n)}{6n(n+2)}},$$

$$e_7 = e\left(\hat{\theta}_{2,\mathrm{ERSS}_3} \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \frac{1-\frac{\alpha^2(n-1)}{9(n+1)}}{1-\frac{\alpha^2(n-1)^3}{3n(n+1)^2}}.$$

Note that $1 \leq e_i \leq 4/3$ for $i = 5, 6, 7$. Also, for fixed $n > 1$, e_i's increase in $|\alpha|$ and increase in n for fixed $|\alpha| > 0$. So we conclude that $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are more efficient than $\hat{\theta}_{2,RSS}$.

A1-Saleh and A1-Ananbeh (2007) proposed the concept of MERSS with concomitant variable for the estimation of the means of the bivariate normal distribution. Now, Suppose that the random vector (X, Y) has a MTBUD as defined in (1.2). The procedure of MERSS with concomitant variable in MTBUD is as follows:

Step 1. Select n units each of size n from MTBUD using SRS. Identify by judgment the minimum of each set with respect to the variable X.

Step 2. Repeat step 1, but for the maximum.

Note that the $2n$ pairs of set $\{(X_{(1)r}, Y_{[1]r}), (X_{(n)r}, Y_{[n]r}); r = 1, 2, \cdots, n\}$ that are obtained using the above procedure, are independent but not identically distributed. An unbiased estimator of θ_2 based on MERSS is given by

$$\hat{\theta}_{2,\mathrm{MERSS}} = \frac{1}{n}\sum_{r=1}^{n}\left(Y_{[1]r} + Y_{[n]r}\right). \tag{3.7}$$

and its variance is

$$Var\left(\hat{\theta}_{2,\mathrm{MERSS}}\right) = \frac{\theta_2^2}{6n}\left[1-\frac{1}{3}\left(\frac{\alpha(n-1)}{n+1}\right)^2\right]. \tag{3.8}$$

The efficiency of $\hat{\theta}_{2,\mathrm{RSS}}$ relative to $\hat{\theta}_{2,\mathrm{MERSS}}$ is given by

$$e_8 = e\left(\hat{\theta}_{2,\mathrm{MERSS}} \mid \hat{\theta}_{2,\mathrm{RSS}}\right) = \frac{1-\frac{\alpha^2(n-1)}{9(n+1)}}{\frac{1}{2}-\frac{1}{6}\left(\frac{\alpha(n-1)}{n+1}\right)^2}.$$

Note that $1 \le e_8 \le 8/3$. Thus, we conclude that $\hat{\theta}_{2,\mathrm{MERSS}}$ is more efficient than $\hat{\theta}_{2,\mathrm{RSS}}$. Also, the efficiency of $\hat{\theta}_{2,\mathrm{MERSS}}$ relative to θ_2^* and $\tilde{\theta}_2$ are given by

$$\begin{aligned} e_9 &= e\left(\theta_2^* \mid \hat{\theta}_{2,\mathrm{MERSS}}\right) \\ &= \left(\frac{1}{2n} - \frac{1}{6n}(\frac{\alpha(n-1)}{n+1})^2\right)\left(\sum_{r=1}^{n}\left[\frac{\left(1-\frac{\alpha(n-2r+1)}{3(n+1)}\right)^2}{1-\frac{1}{3}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2}\right]\right), \\ e_{10} &= e\left(\tilde{\theta}_2 \mid \hat{\theta}_{2,\mathrm{MERSS}}\right) = \frac{\left(\frac{1}{2}-\frac{1}{6}\left(\frac{\alpha(n-1)}{n+1}\right)^2\right)\left[1-\alpha\left(\frac{-n+1}{3(n+1)}\right)\right]^2}{1-\frac{1}{3}\left(\frac{\alpha(-n+1)}{n+1}\right)^2}. \end{aligned}$$

Finally, the efficiency of θ_2^* relative to the estimators $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are given by

$$e_{11} = e\left(\hat{\theta}_{2,\mathrm{ERSS}_1} \mid \theta_2^*\right) = \frac{1}{\left(\frac{1}{n}-\frac{1}{3n}\left(\frac{\alpha(n-1)}{n+1}\right)^2\right)\left(\sum_{r=1}^{n}\left[\frac{\left(1-\frac{\alpha(n-2r+1)}{3(n+1)}\right)^2}{1-\frac{1}{3}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2}\right]\right)},$$

$$\begin{aligned} e_{12} &= e\left(\hat{\theta}_{2,\mathrm{ERSS}_2} \mid \theta_2^*\right) \\ &= \frac{1}{\left(\frac{1}{n}-\frac{\alpha^2(n-1)^3}{3n^2(n+1)^2}-\frac{1}{2n^2}+\frac{\alpha^2(2-n)}{6n^2(n+2)}\right)\left(\sum_{r=1}^{n}\left[\frac{\left(1-\frac{\alpha(n-2r+1)}{3(n+1)}\right)^2}{1-\frac{1}{3}(\frac{\alpha(n-2r+1)}{n+1})^2}\right]\right)}, \end{aligned}$$

$$e_{13} = e\left(\hat{\theta}_{2,\mathrm{ERSS}_3} \mid \theta_2^*\right) = \frac{1}{\left(\frac{1}{n}-\frac{\alpha^2(n-1)^3}{3n^2(n+1)^2}\right)\left(\sum_{r=1}^{n}\left[\frac{\left(1-\frac{\alpha(n-2r+1)}{3(n+1)}\right)^2}{1-\frac{1}{3}\left(\frac{\alpha(n-2r+1)}{n+1}\right)^2}\right]\right)}.$$

Furthermore, the efficiency of $\tilde{\theta}_2$ relative to the estimators $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are given by

$$\begin{aligned} e_{14} &= e\left(\hat{\theta}_{2,\mathrm{ERSS}_1} \mid \tilde{\theta}_2\right) = \frac{1}{\left[1-\alpha\left(\frac{-n+1}{3(n+1)}\right)\right]^2}, \\ e_{15} &= e\left(\hat{\theta}_{2,\mathrm{ERSS}_2} \mid \tilde{\theta}_2\right) = \frac{1-\frac{1}{3}\left(\frac{\alpha(-n+1)}{n+1}\right)^2}{\left[1-\alpha\left(\frac{-n+1}{3(n+1)}\right)\right]^2\left[1-\frac{\alpha^2(n-1)^3}{3n(n+1)^2}-\frac{1}{2n}+\frac{\alpha^2(2-n)}{6n(n+2)}\right]}, \end{aligned}$$

$$e_{16} = e\left(\hat{\theta}_{2,\mathrm{ERSS}_3} \mid \tilde{\theta}_2\right) = \frac{1-\frac{1}{3}\left(\frac{\alpha(-n+1)}{n+1}\right)^2}{\left[1-\alpha\left(\frac{-n+1}{3(n+1)}\right)\right]^2\left[1-\frac{\alpha^2(n-1)^3}{3n(n+1)^2}\right]}.$$

We have computed the values of e_j for $j = 9, 10, \cdots, 16$ with $\alpha = \pm.25, \pm.5, \pm.75, \pm1$, $n = 5(5)20$ and these are given in Table 2. From Table 2, we may conclude:

(a) The efficiencies of $\hat{\theta}_{2,\mathrm{MERSS}}$ relative to θ_2^* and $\tilde{\theta}_2$ are less than 1 for $n \geq 5$. So, $\hat{\theta}_{2,\mathrm{MERSS}}$ is relatively more efficient than θ_2^* and $\tilde{\theta}_2$.

(b) The efficiencies of θ_2^* relative to the estimators $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are more than 1 for $n \geq 5$. Thus, $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are relatively more efficient than θ_2^*.

(c) The efficiencies of $\tilde{\theta}_2$ relative to the estimators $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ are more than (less than) 1 for $-1 \leq \alpha < 0$ $(0 < \alpha \leq 1)$ and $n \geq 5$. Thus $\tilde{\theta}_2$ is relatively more efficient than $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{ERSS}_2}$ and $\hat{\theta}_{2,\mathrm{ERSS}_3}$ when $0 < \alpha \leq 1$.

4. An Application

Recently, a biological study on purslane plants (portulaca oleracea) are done in the research and studied center of Persian Gulf University. The results of this study show that the shoot height of the plant is a correlated character with the shoot diameter. Now, we consider a bivariate data set from the 256 plant data such that the first component X represents the shoot height in centimeter, and the second components Y represents the shoot diameter in centimeter. Clearly the shoot height can be measured very easily but the shoot diameter is difficult to measure. Under the assumption that (X, Y) follows MTBUD, we select 8 random samples with size 8 from 256 plant data and rank the sampling units of each sample according to the X variate (shoot height). Now, we measure the raked set sample observations $Y_{[r]r}$ corresponding to $X_{(r)r}$. The obtained RSS, ERSS_1 and MERSS observations are given in Table 3. Since the sample correlation coefficient is $q > 1/3$. Therefore, an estimate for α is 1 (see Remark 1).

The computed values of $\hat{\theta}_{2,\mathrm{RSS}}$, $\hat{\theta}_{2,\mathrm{ERSS}_1}$, $\hat{\theta}_{2,\mathrm{MERSS}}$ are 2.9075, 2.9625, and 2.9737, and their estimated variances are 0.3212, 0.2913, and 0.1467, respectively. We can find that the estimated values for θ_2 based on different samplings are close.

Acknowledgements

We are grateful to the editor and referees for their helpful comments and suggestions which improved the presentation of the paper.

Table 2: The values of e_j for $j = 9, 10, \cdots, 16$ in MTBUD

n	α	e_9	e_{10}	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}
5	-1.00	0.4741	0.3024	1.0545	1.1708	1.0190	1.6530	1.8354	1.5974
	-0.75	0.4857	0.3472	1.0293	1.1432	1.0109	1.4400	1.5994	1.4142
	-0.50	0.4937	0.3950	1.0126	1.1249	1.0049	1.2656	1.4060	1.2559
	-0.25	0.4984	0.4459	1.0031	1.1145	1.0012	1.1211	1.2456	1.1190
	0.25	0.4984	0.5570	1.0031	1.1145	1.0012	0.8975	0.9971	0.8958
	0.50	0.4937	0.6172	1.0126	1.1249	1.0049	0.8100	0.8998	0.8038
	0.75	0.4857	0.6805	1.0293	1.1432	1.0109	0.7346	0.8160	0.7215
	1.00	0.4741	0.7469	1.0545	1.1708	1.0190	0.6694	0.7432	0.6469
10	-1.00	0.4449	0.2644	1.1237	1.1828	1.0924	1.8906	1.9900	1.8378
	-0.75	0.4700	0.3163	1.0636	1.1195	1.0485	1.5804	1.6635	1.5580
	-0.50	0.4870	0.3729	1.0266	1.0806	1.0206	1.3407	1.4112	1.3328
	-0.25	0.4968	0.4341	1.0064	1.0594	1.0050	1.1516	1.2123	1.1500
	0.25	0.4968	0.5705	1.0064	1.0594	1.0050	0.8764	0.9225	0.8751
	0.50	0.4870	0.6456	1.0266	1.0806	1.0206	0.7744	0.8151	0.7698
	0.75	0.4700	0.7254	1.0636	1.1195	1.0485	0.6892	0.7254	0.6794
	1.00	0.4449	0.8099	1.1237	1.1828	1.0924	0.6173	0.6497	0.6001
15	-1.00	0.4312	0.2508	1.1592	1.1992	1.1333	1.9930	2.0617	1.9485
	-0.75	0.4629	0.3051	1.0801	1.1173	1.0681	1.6384	1.6948	1.6202
	-0.50	0.4839	0.3648	1.0331	1.0687	1.0284	1.3706	1.4178	1.3644
	-0.25	0.4960	0.4297	1.0079	1.0427	1.0068	1.1634	1.2036	1.1622
	0.25	0.4960	0.5755	1.0079	1.0427	1.0068	0.8686	0.8986	0.8677
	0.50	0.4839	0.6564	1.0331	1.0687	1.0284	0.7616	0.7879	0.7582
	0.75	0.4629	0.7426	1.0801	1.1173	1.0681	0.6732	0.6964	0.6658
	1.00	0.4312	0.8342	1.1592	1.1992	1.1333	0.5993	0.6200	0.5859
20	-1.00	0.4235	0.2438	1.1805	1.2107	1.1587	2.0501	2.1026	2.0123
	-0.75	0.4588	0.2993	1.0896	1.1176	1.0795	1.6700	1.7128	1.6550
	-0.50	0.4822	0.3605	1.0368	1.0633	1.0330	1.3866	1.4222	1.3816
	-0.25	0.4956	0.4274	1.0088	1.0346	1.0079	1.1697	1.1997	1.1687
	0.25	0.4956	0.5782	1.0088	1.0346	1.0079	0.8646	0.8868	0.8639
	0.50	0.4822	0.6621	1.0368	1.0633	1.0330	0.7551	0.7744	0.7523
	0.75	0.4588	0.7517	1.0896	1.1176	1.0795	0.6650	0.6821	0.6591
	1.00	0.4235	0.8470	1.1805	1.2107	1.1587	0.5902	0.6054	0.5794

Table 3: Obtained RSS, $ERSS_1$ and MERSS observations

	r	1	2	3	4	5	6	7	8
RSS	$X_{(r)r}$	10.37	7.25	8.33	9.00	8.87	11.37	10.50	8.50
	$Y_{[r]r}$	1.37	1.27	1.10	1.15	1.72	1.75	1.57	1.70
$ERSS_1$	$X_{(1)2r-1}$	10.37	4.00	7.12	4.25				
	$Y_{[1]2r-1}$	1.37	1.40	1.32	1.12				
	$X_{(n)2r}$	10.00	11.25	12.50	8.50				
	$Y_{[n]2r}$	1.57	1.57	1.80	1.70				
MERSS	$X_{(1)r}$	10.37	6.70	4.00	6.75	7.12	7.16	4.25	5.12
	$Y_{[1]r}$	1.37	1.72	1.40	1.22	1.32	1.70	1.12	0.97
	$X_{(n)r}$	15.00	10.00	15.37	11.25	9.50	12.50	11.83	8.50
	$Y_{[n]r}$	1.60	1.57	1.75	1.57	1.45	1.80	1.53	1.70

References

Al-Saleh, M. F. and Ananbeh, A. (2005). Estimating the correlation coefficient in a bivariate normal distribution using moving extreme ranked set sampling with a concomitant variable. *Journal of the Korean Statistical Society* **34**, 125-140.

Al-Saleh, M. F. and Ananbeh, A. (2007). Estimation of the means of the bivariate normal distribution using moving extreme ranked set sampling with concomitant variable. *Statistical Papers* **48**, 179-195.

Al-Saleh, M. F. and Diab, Y. A. (2009). Estimation of the parameters of Downton's bivariate exponential distribution using ranked set sampling scheme. *Journal of Statistical Planning Inference* **139**, 277-286.

A1-Odat, M. T. and A1-Saleh, M. F. (2001). A variation of ranked set sampling. *Journal of Applied Statistical Science* **10**, 137-246.

Barnett, V. and Moore, K. (1997). Best linear unbiased estimates in ranked set sampling with particular reference to imperfect ordering. *Journal of Applied Statistics* **24**, 697-710.

Chacko, M. and Thomas, Y. (2008). Estimation of a parameter of Morgenstern type bivariate exponential by ranked set sampling. *Annals of the Institute of Statistical Mathematics* **60**, 273-300.

David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, 3rd edition. Wiley, New York.

Lam, K., Sinha, B. K. and Wu, Z. (1994). Estimation of parameters in two-parameter exponential distribution using ranked set sampling. *Annals of the Institute of Statistical Mathematics* **46**, 723-736.

McIntyre, G. (1952). A method for unbiased selective sampling using ranked set sampling. *Australian Journal of Agricultural Research* **3**, 385-390.

Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. *Mitteilungsblatt für Mathematische Statistik* **8**, 234-235.

Samawi, H., Ahmad, M. and Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Joumal* **30**, 577-586.

Scaria, J. and Nair, U. (1999). On concomitants of order statistics from Morgenstern family. *Biometrical Journal* **41**, 483-489.

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics - Theory and Methods* **6**, 1207-1211.

Stokes, S. L. (1980). Inferences on the correlation coefficient in bivariate normal populations from ranked set samples. *Journal of American Statistical Association* **75**, 989-995.

Stokes, S. L. (1995). Parametric ranked set sampling. *Annals of the Institute of Statistical Mathematics* **47**, 465-482.

Received February 1, 2011; accepted July 4, 2011.

Saeid Tahmasebi
Department of Statistics
Persian Gulf University
Boushehr 75168, Iran
stahmasby@yahoo.com

Ali Akbar Jafari
Department of Statistics
Yazd University
Yazd, 89195-741, Iran
aajafari@yazduni.ac.ir