AISM (2008) 60:301–318
DOI 10.1007/s10463-006-0088-y

Estimation of a parameter of Morgenstern type bivariate exponential distribution by ranked set sampling

Manoj Chacko · P. Yageen Thomas

Received: 12 December 2005 / Revised: 17 May 2006 / Published online: 11 November 2006

Abstract Ranked set sampling is applicable whenever ranking of a set of sampling units can be done easily by a judgement method or based on the measurement of an auxiliary variable on the units selected. In this work, we consider ranked set sampling, in which ranking of units are done based on measurements made on an easily and exactly measurable auxiliary variable X which is correlated with the study variable Y. We then estimate the mean of the study variate Y by the BLUE based on the measurements made on the units of the ranked set sampling regarding the study variable Y, when (X, Y) follows a Morgenstern type bivariate exponential distribution. We then consider unbalanced multistage ranked set sampling and estimate the mean of the study variate Y by the BLUE based on the observations made on the units of multistage ranked set sample regarding the study variable Y. Efficiency comparison is also made on all estimators considered in this work.

Keywords Ranked set sampling · Morgenstern type bivariate exponential distribution · Best linear unbiased estimator · Multistage ranked set sampling · Concomitants of order statistics

1 Introduction

The concept of ranked set sampling (RSS) was first introduced by McIntyre (1952) as a process of improving the precision of the sample mean as an

M. Chacko · P. Y. Thomas (✉)
Department of Statistics, University of Kerala, Trivandrum 695 581, India
e-mail: yageenthomas@gmail.com

M. Chacko
e-mail: manojchacko02@yahoo.co.in

estimator of the population mean. Ranked set sampling as described in McIntyre (1952) is applicable whenever ranking of a set of sampling units can be done easily by a judgement method (for a detailed discussion on the theory and applications of ranked set sampling see, Chen et al., 2004). Ranking by judgement method is not recommendable if the judgement method is too crude and is not powerful for ranking by discriminating the units of a moderately large sample. In certain situations, one may prefer exact measurement of some easily measurable variable associated with the study variable rather than ranking the units by a crude judgement method. Suppose the variable of interest say Y, is difficult or much expensive to measure, but an auxiliary variable X correlated with Y is readily measurable and can be ordered exactly. In this case as an alternative to McIntyre (1952) method of ranked set sampling, Stokes (1977) used an auxiliary variable for the ranking of the sampling units. The procedure of ranked set sampling described by Stokes (1977) using auxiliary variate is as follows: Choose n^2 independent units, arrange them randomly into n sets each with n units and observe the value of the auxiliary variable X on each of these units. In the first set, that unit for which the measurement on the auxiliary variable is the smallest is chosen. In the second set, that unit for which the measurement on the auxiliary variable is the second smallest is chosen. The procedure is repeated until in the last set, that unit for which the measurement on the auxiliary variable is the largest is chosen. The resulting new set of n units chosen by one from each set as described above is called the RSS defined by Stokes (1977). If $X_{(r)r}$ is the observation measured on the auxiliary variable X from the unit chosen from the rth set then we write $Y_{[r]r}$ to denote the corresponding measurement made on the study variable Y on this unit, then $Y_{[r]r}$, $r = 1, 2, \ldots, n$, form the ranked set sample. Clearly $Y_{[r]r}$ is the concomitant of the rth order statistic arising from the rth sample.

A striking example for the application of the ranked set sampling as proposed by Stokes (1977) is given in Bain (1978, p. 99), where the study variate Y represents the oil pollution of sea water and the auxiliary variable X represents the tar deposit in the nearby sea shore. Clearly collecting sea water sample and measuring the oil pollution in it is strenuous and expensive. However the prevalence of pollution in the sea water is much reflected by the tar deposit in the surrounding terminal sea shore. In this example ranking the pollution level of sea water based on the tar deposit in the sea shore is more natural and scientific than ranking it visually or by judgement method.

Stokes (1995) has considered the estimation of parameters of location-scale family of distributions using RSS. Lam et al. (1994, 1995) have obtained the BLUEs of location and scale parameters of exponential distribution and logistic distribution. The Fisher information contained in RSS have been discussed by Chen (2000) and Chen and Bai (2000). Stokes (1980) has considered the method of estimation of correlation coefficient of bivariate normal distribution using RSS. Modarres and Zheng (2004) have considered the problem of estimation of dependence parameter using RSS. Robust estimate of correlation coefficient for bivariate normal distribution have been developed by Zheng and Modarres (2006). Stokes (1977) has suggested the ranked set sample mean as an estimator for the mean of the study variate Y, when an auxiliary variable X is used for ranking the sample units, under the assumption that (X, Y) follows a bivariate normal distribution. Barnett and Moore (1997) have improved the estimator of Stokes (1977) by deriving the Best Linear Unbiased Estimator (BLUE) of the mean of the study variate Y, based on ranked set sample obtained on the study variate Y.

In this paper we are trying to estimate the mean of the population, under a situation where in measurement of observations are strenuous and expensive. Bain (1978, p. 99) has proposed an exponential distribution for the study variate Y, the oil pollution of the sea samples. Thus in this paper we assume a Morgenstern type bivariate exponential distribution (MTBED) corresponding to a bivariate random variable (X, Y), where X denote the auxiliary variable (such as tar deposit in the sea shore) and Y denote the study variable (such as the oil pollution in the sea water). A random variable (X, Y) follows MTBED if its probability density function (pdf) is given by (see, Kotz et al. 2000, p. 353)

$$f(x,y) = \begin{cases} \frac{1}{\theta_1\theta_2}\exp\left\{\frac{-x}{\theta_1}+\frac{-y}{\theta_2}\right\} & \left[1+\alpha\left(1-2\exp\left\{\frac{-x}{\theta_1}\right\}\right)\left(1-2\exp\left\{\frac{-y}{\theta_2}\right\}\right)\right], \\ & x>0, y>0; -1\leq\alpha\leq 1; \theta_1>0, \theta_2>0 \\ 0, \text{ otherwise.} & \end{cases} \tag{1}$$

In Sect. 2 of this paper we have derived an estimator θ_2^* of the parameter θ_2 involved in (1) using the ranked set sample mean. It may be noted that if (X, Y) has a MTBED as defined in (1) then the marginal distributions of both X and Y have exponential distributions and the pdf of Y is given by

$$f_Y(y) = \frac{1}{\theta_2}e^{-y/\theta_2}, \quad y>0, \theta_2>0. \tag{2}$$

The Cramer-Rao Lower Bound (CRLB) of any unbiased estimator of θ_2 based on a random sample of size n drawn from (2) is $\frac{\theta_2^2}{n}$. In Sect. 2, we have also shown that the variance of the proposed estimator θ_2^* is strictly less than the CRLB $\frac{\theta_2^2}{n}$ (associated with the marginal pdf (2)) for all $\alpha \in A$ where $A = [-1, 1] - \{0\}$. In this section, we have further made an efficiency comparison between θ_2^* and the maximum likelihood estimator (MLE) $\tilde{\theta}_2$ based on a random sample of size n arising from (1). In Sect. 3, we have derived the BLUE $\hat{\theta}_2$ of θ_2 involved in MTBED based on the ranked set sample and obtained the efficiency of $\hat{\theta}_2$ relative to $\tilde{\theta}_2$. In Sect. 4, we have considered the situation where we apply censoring and ranking on each sample and ultimately used a ranked set sample arising out of this procedure to estimate θ_2. In Sect. 5, we have derived the BLUE of θ_2 involved in (1) using unbalanced multistage ranked set sampling method. In this section, we have further analysed the efficiency of the estimator of θ_2 based on unbalanced MSRSS when compared with the $\tilde{\theta}_2$ the MLE of θ_2.

2 Ranked set sample mean as an estimator of θ_2

Let (X, Y) be a bivariate random variable which follows a MTBED with pdf defined by (1). Suppose RSS in the sense of Stokes (1977) as explained in Sect. 1 is carried out. Let $X_{(r)r}$ be the observation measured on the auxiliary variate X in the rth unit of the RSS and let $Y_{[r]r}$ be the measurement made on the Y variate of the same unit, $r = 1, 2, \ldots, n$. Then clearly $Y_{[r]r}$ is distributed as the concomitant of rth order statistic of a random sample of n arising from (1). By using the expressions for means and variances of concomitants of order statistics arising from MTBED obtained by Scaria and Nair (1999), the mean and variance of $Y_{[r]r}$ for $1 \leq r \leq n$ are given below:

$$E[Y_{[r]r}] = \theta_2 \left[1 - \alpha \frac{n - 2r + 1}{2(n+1)}\right], \tag{3}$$

$$\mathrm{Var}[Y_{[r]r}] = \theta_2^2 \left[1 - \alpha \frac{n - 2r + 1}{2(n+1)} - \alpha^2 \frac{(n - 2r + 1)^2}{4(n+1)^2}\right]. \tag{4}$$

Since $Y_{[r]r}$ and $Y_{[s]s}$ for $r \neq s$ are measurements on Y made from two units involved in two independent samples we have

$$\mathrm{Cov}[Y_{[r]r}, Y_{[s]s}] = 0, \quad r \neq s. \tag{5}$$

In the following theorem we propose an estimator θ_2^* of θ_2 involved in (1) and prove that it is an unbiased estimator of θ_2.

Theorem 1 *Let $Y_{[r]r}$, $r = 1, 2, \ldots, n$, be the ranked set sample observations on a study variate Y obtained out of ranking made on an auxiliary variate X, when (X, Y) follows MTBED as defined in (1). Then the ranked set sample mean given by*

$$\theta_2^* = \frac{1}{n} \sum_{r=1}^{n} Y_{[r]r}$$

is an unbiased estimator of θ_2 and its variance is given by

$$\mathrm{Var}[\theta_2^*] = \frac{\theta_2^2}{n} \left[1 - \frac{\alpha^2}{4n} \sum_{r=1}^{n} \left(\frac{n - 2r + 1}{n+1}\right)^2\right].$$

Proof

$$E[\theta_2^*] = \frac{1}{n} \sum_{r=1}^{n} E[Y_{[r]r}].$$

On using (3) for $E[Y_{[r]r}]$ in the above equation we get

$$E[\theta_2^*] = \frac{1}{n}\sum_{r=1}^{n}\left[1 - \alpha\frac{n-2r+1}{2(n+1)}\right]\theta_2. \tag{6}$$

It is clear to note that

$$\sum_{r=1}^{n}(n-2r+1) = 0. \tag{7}$$

Applying (7) in (6) we get

$$E[\theta_2^*] = \theta_2.$$

Thus θ_2^* is an unbiased estimator of θ_2. The variance of θ_2^* is given by

$$\mathrm{Var}[\theta_2^*] = \frac{1}{n^2}\sum_{r=1}^{n}\mathrm{Var}(Y_{[r]r}).$$

Now using (4) and (7) in the above sum we get,

$$\mathrm{Var}[\theta_2^*] = \frac{\theta_2^2}{n}\left[1 - \frac{\alpha^2}{4n}\sum_{r=1}^{n}\left(\frac{n-2r+1}{n+1}\right)^2\right].$$

Thus the theorem is proved. □

Now we compare the variance of θ_2^* with the CRLB θ_2^2/n of any unbiased estimator of θ_2 involved in (2) which is the marginal distribution of Y in (1). If we write $e_1(\theta_2^*)$ to denote the ratio of θ_2^2/n with $\mathrm{Var}(\theta_2^*)$ then we have

$$e_1(\theta_2^*) = \frac{1}{\left[1 - \frac{\alpha^2}{4n}\sum_{r=1}^{n}\left(\frac{n-2r+1}{n+1}\right)^2\right]} \tag{8}$$

It is very trivial to note that

$$e_1(\theta_2^*) \geq 1.$$

Thus we conclude that there is some gain in efficiency on the estimator θ_2^* due to ranked set sampling. The reason for the above conclusion is that a ranked set sample always provides more information than simple random sample even if ranking is imperfect (see, Chen et al., 2004, p. 58). It is to be noted that $\mathrm{Var}(\theta_2^*)$ is a decreasing function of α^2 and hence the gain in efficiency of the estimator θ_2^* increases as $|\alpha|$ increases. Again on simplifying (8) we get

$$e_1(\theta_2^*) = \frac{1}{1 - \frac{\alpha^2}{4}[\frac{2}{3}(\frac{2+1/n}{1+1/n}) - 1]}.$$

Then

$$\lim_{n\to\infty} e_1(\theta_2^*) = \lim_{n\to\infty} \frac{1}{1 - \frac{\alpha^2}{4}[\frac{2}{3}(\frac{2+1/n}{1+1/n}) - 1]}$$
$$= \frac{1}{1 - \frac{\alpha^2}{12}}.$$

From the above relation it is clear that the maximum value for $e_1(\theta_2^*)$ is attained when $|\alpha| = 1$ and in this case $e_1(\theta_2^*)$ tends to 12/11.

Next we obtain the efficiency of θ_2^* by comparing the variance of θ_2^* with the asymptotic variance of MLE of θ_2 involved in MTBED. If (X, Y) follows a MTBED with pdf defined by (1), then

$$\frac{\partial \log f(x,y)}{\partial \theta_1} = \frac{1}{\theta_1}\left\{-1 + \frac{x}{\theta_1} + \frac{2\alpha x}{\theta_1}\frac{e^{-\frac{x}{\theta_1}}(2e^{-\frac{y}{\theta_2}} - 1)}{1 + \alpha(2e^{-\frac{x}{\theta_1}} - 1)(2e^{-\frac{y}{\theta_2}} - 1)}\right\}$$

and

$$\frac{\partial \log f(x,y)}{\partial \theta_2} = \frac{1}{\theta_2}\left\{-1 + \frac{y}{\theta_2} + \frac{2\alpha y}{\theta_2}\frac{e^{-\frac{y}{\theta_2}}(2e^{-\frac{x}{\theta_1}} - 1)}{1 + \alpha(2e^{-\frac{x}{\theta_1}} - 1)(2e^{-\frac{y}{\theta_2}} - 1)}\right\}.$$

Then we have

$$I_{\theta_1}(\alpha) = E\left(\frac{\partial \log f(x,y)}{\partial \theta_1}\right)^2,$$
$$= \frac{1}{\theta_1^2}\left\{1 + 4\alpha^2 \int_0^\infty\int_0^\infty \frac{u^2 e^{-3u}(2e^{-v} - 1)e^{-v}}{\{1 + \alpha(2e^{-u} - 1)(2e^{-v} - 1)\}}\mathrm{d}v\mathrm{d}u\right\},$$
$$I_{\theta_2}(\alpha) = E\left(\frac{\partial \log f(x,y)}{\partial \theta_2}\right)^2$$
$$= \frac{1}{\theta_2^2}\left\{1 + 4\alpha^2 \int_0^\infty\int_0^\infty \frac{v^2 e^{-3v}(2e^{-u} - 1)e^{-u}}{\{1 + \alpha(2e^{-u} - 1)(2e^{-v} - 1)\}}\mathrm{d}v\mathrm{d}u\right\}$$

and

$$I_{\theta_1\theta_2}(\alpha) = E\left(\frac{\partial^2 \log f(x,y)}{\partial\theta_1\partial\theta_2}\right)$$
$$= \frac{1}{\theta_1\theta_2}\left\{4\alpha\int_0^\infty\int_0^\infty \frac{uve^{-2u}e^{-2v}}{\{1+\alpha(2e^{-u}-1)(2e^{-v}-1)\}}\mathrm{d}v\mathrm{d}u\right\}.$$

Thus the Fisher information matrix associated with the random variable (X, Y) is given by

$$I(\alpha) = \begin{pmatrix} I_{\theta_1}(\alpha) & -I_{\theta_1\theta_2}(\alpha) \\ -I_{\theta_1\theta_2}(\alpha) & I_{\theta_2}(\alpha) \end{pmatrix}. \tag{9}$$

We have evaluated the values of $\theta_1^{-2}I_{\theta_1}(\alpha)$ and $\theta_1^{-1}\theta_2^{-1}I_{\theta_1\theta_2}(\alpha)$ numerically for $\alpha = \pm 0.25, \pm 0.50, \pm 0.75, \pm 1$ (clearly $\theta_1^{-2}I_{\theta_1}(\alpha) = \theta_2^{-2}I_{\theta_2}(\alpha)$) and are given below:

α	$\theta_1^{-2}I_{\theta_1}(\alpha)$	$\theta_1^{-1}\theta_2^{-1}I_{\theta_1\theta_2}(\alpha)$	α	$\theta_1^{-2}I_{\theta_1}(\alpha)$	$\theta_1^{-1}\theta_2^{-1}I_{\theta_1\theta_2}(\alpha)$
0.25	1.0062	0.0625	−0.25	1.0062	−0.0628
0.50	1.0254	0.1258	−0.50	1.0254	−0.1274
0.75	1.0596	0.1914	−0.75	1.0596	−0.1955
1.00	1.1148	0.2624	−1.00	1.1148	−0.2712

Thus from (9) the asymptotic variance of the MLE $\tilde{\theta_2}$ of θ_2 involved in MTBED based on a bivariate sample of size n is obtained as

$$\mathrm{Var}(\tilde{\theta_2}) = \frac{1}{n}I_{\theta_2}^{(-1)}(\alpha), \tag{10}$$

where $I_{\theta_2}^{(-1)}(\alpha)$ is the $(2,2)$th element of the inverse of $I(\alpha)$ given by (9).

We have obtained the efficiency $e(\theta_2^*|\tilde{\theta_2}) = \frac{\mathrm{Var}(\tilde{\theta_2})}{\mathrm{Var}(\theta_2^*)}$ of θ_2^* relative to $\tilde{\theta_2}$ for $n = 2(2)10(5)20$; $\alpha = \pm 0.25, \pm 0.50, \pm 0.75, \pm 1$ and are presented in table 1. From the table, one can easily see that θ_2^* is more efficient than $\tilde{\theta_2}$ and efficiency increases with n and $|\alpha|$ for $n \geq 4$.

3 Best linear unbiased estimator of θ_2

In this section we provide a better estimator of θ_2 than that of θ_2^* by deriving the BLUE $\hat{\theta}_2$ of θ_2 provided the parameter α is known. Let $X_{(r)r}$ be the observation measured on the auxiliary variate X in the rth unit of the RSS and let $Y_{[r]r}$ be the measurement made on the Y variate of the same unit, $r = 1, 2, \ldots, n$. Let

$$\xi_r = 1 - \alpha \frac{n-2r+1}{2(n+1)} \tag{11}$$

and

$$\delta_r = 1 - \alpha \frac{n-2r+1}{2(n+1)} - \alpha^2 \frac{(n-2r+1)^2}{4(n+1)^2}. \tag{12}$$

Using (11) in (3) and (12) in (4) we get $E[Y_{[r]r}] = \theta_2 \xi_r, 1 \leq r \leq n$, and $\mathrm{Var}[Y_{[r]r}] = \theta_2^2 \delta_r, 1 \leq r \leq n$. Clearly from (5), we have $\mathrm{Cov}[Y_{[r]r}, Y_{[s]s}] = 0$, $r, s = 1, 2, \ldots, n$ *and* $r \neq s$. Let $\mathbf{Y}_{[n]} = (Y_{[1]1}, Y_{[2]2}, \cdots, Y_{[n]n})'$ and if the parameter α involved in ξ_r and δ_r is known then proceeding as in David and Nagaraja (2003, p. 185) the BLUE $\hat{\theta}_2$ of θ_2 is obtained as

$$\hat{\theta}_2 = (\xi' G^{-1} \xi)^{-1} \xi' G^{-1} \mathbf{Y}_{[n]} \tag{13}$$

and

$$\mathrm{Var}(\hat{\theta}_2) = (\xi' G^{-1} \xi)^{-1} \theta_2^2, \tag{14}$$

where $\xi = (\xi_1, \xi_2, \cdots, \xi_n)'$ and $G = diag(\delta_1, \delta_2, \cdots, \delta_n)$. On substituting the values of ξ and G in (13) and (14) and simplifying we get

$$\hat{\theta}_2 = \frac{\sum_{r=1}^{n} (\xi_r/\delta_r) Y_{[r]r}}{\sum_{r=1}^{n} \xi_r^2/\delta_r} \tag{15}$$

and

$$\mathrm{Var}(\hat{\theta}_2) = \frac{1}{\sum_{r=1}^{n} \xi_r^2/\delta_r} \theta_2^2.$$

Thus $\hat{\theta}_2$ as given in (15) can be explicitly written as $\hat{\theta}_2 = \sum_{r=1}^{n} a_r Y_{[r]r}$, where

$$a_r = \frac{\xi_r/\delta_r}{\sum_{r=1}^{n} \xi_r^2/\delta_r}, \quad r = 1, 2, \ldots, n.$$

Remark 1 As the association parameter α in (1) is involved in the BLUE $\hat{\theta}_2$ of θ_2 and its variance, an assumption that α is known may sometimes viewed as unrealistic. Hence when α is unknown, our recommendation is to compute the sample correlation coefficient q of the observations $(X_{(r)r}, Y_{[r]r}), r = 1, 2, \ldots, n$ and consider the model (1) for a value of α equal to α_0 given by

$$\alpha_0 = \begin{cases} -1 & \text{if } q < \frac{-1}{4} \\ 4q & \text{if } \frac{-1}{4} \leq q \leq \frac{1}{4} \\ 1 & \text{if } q > \frac{1}{4}, \end{cases}$$

as we know that the correlation coefficient between X and Y involved in the Morgenstern type bivariate exponential random vector is equal to $\frac{\alpha}{4}$.

Table 1 Efficiencies of the estimators θ_2^*, $\hat{\theta}_2$, $\hat{\theta}_2^{n(1)}$ and $\hat{\theta}_2^{1(1)}$relative to $\tilde{\theta_2}$ of θ_2 involved in Morgenstern type bivariate exponential distribution

n	α	$e(\theta_2^*)\|\tilde{\theta_2})$	$e(\hat{\theta}_2\|\tilde{\theta_2})$	$e(\hat{\theta}_2^{n(1)}\|\tilde{\theta_2})$	α	$e(\theta_2^*\|\tilde{\theta_2})$	$e(\hat{\theta}_2\|\tilde{\theta_2})$	$e(\hat{\theta}_2^{1(1)}\|\tilde{\theta_2})$
	0.25	0.9994	0.9994	1.0410	−0.25	0.9994	0.9994	1.0410
2	0.50	0.9970	0.9970	1.0795	−0.50	0.9974	0.9974	1.0800
	0.75	0.9911	0.9911	1.1130	−0.75	0.9926	0.9926	1.1147
	1.00	0.9767	0.9768	1.1349	−1.00	0.9806	0.9807	1.1394
	0.25	1.0008	1.0008	1.0782	−0.25	1.0008	1.0008	1.0782
4	0.50	1.0026	1.0027	1.1613	−0.50	1.0030	1.0031	1.1618
	0.75	1.0038	1.0044	1.2466	−0.75	1.0054	1.0060	1.2485
	1.00	0.9996	1.0019	1.3263	−1.00	1.0036	1.0059	1.3316
	0.25	1.0014	1.0014	1.0948	−0.25	1.0014	1.0014	1.0948
6	0.50	1.0051	1.0052	1.1994	−0.50	1.0055	1.0056	1.1998
	0.75	1.0094	1.0105	1.3111	−0.75	1.0109	1.0120	1.3131
	1.00	1.0097	1.0140	1.4224	−1.00	1.0137	1.0180	1.4281
	0.25	1.0018	1.0018	1.1042	−0.25	1.0018	1.0018	1.1042
8	0.50	1.0064	1.0066	1.2213	−0.50	1.0068	1.0070	1.2218
	0.75	1.0125	1.0139	1.3490	−0.75	1.0141	1.0154	1.3511
	1.00	1.0154	1.0211	1.4800	−1.00	1.0195	1.0252	1.4860
	0.25	1.0020	1.0020	1.1103	−0.25	1.0020	1.0020	1.1103
10	0.50	1.0073	1.0075	1.2355	−0.50	1.0077	1.0079	1.2360
	0.75	1.0145	1.0161	1.3739	−0.75	1.0161	1.0176	1.3760
	1.00	1.0191	1.0259	1.5184	−1.00	1.0232	1.0300	1.5245
	0.25	1.0023	1.0023	1.1189	−0.25	1.0023	1.0023	1.1189
15	0.50	1.0085	1.0088	1.2560	−0.50	1.0089	1.0092	1.2565
	0.75	1.0173	1.0192	1.4100	−0.75	1.0189	1.0208	1.4122
	1.00	1.0243	1.0328	1.5747	−1.00	1.0284	1.0370	1.5810
	0.25	1.0024	1.0024	1.1234	−0.25	1.0024	1.0024	1.1234
20	0.50	1.0091	1.0095	1.2669	−0.50	1.0095	1.0099	1.2674
	0.75	1.0188	1.0209	1.4295	−0.75	1.0204	1.0225	1.4317
	1.00	1.0270	1.0366	1.6054	−1.00	1.0311	1.0408	1.6118

We have computed the the ratio $e(\hat{\theta}_2|\tilde{\theta_2})= \frac{\mathrm{Var}(\tilde{\theta_2})}{\mathrm{Var}(\hat{\theta}_2)}$ as the efficiency of $\hat{\theta}_2$ relative to $\tilde{\theta_2}$ for $\alpha = \pm 0.25, \pm 0.5, \pm 0.75, \pm 1.0$ and $n = 2(2)10(5)20$ and are also given in Table 1. From the table, one can easily see that $\hat{\theta}_2$ is relatively more efficient than $\tilde{\theta_2}$. Further we observe from the table that $e(\hat{\theta}_2|\tilde{\theta_2})$ increases as n and $|\alpha|$ increases.

4 Estimation of θ_2 based on censored ranked set sample

In the case of the example of pollution study on sea samples (see, Bain 1978, p. 99), sometimes if there is no tar deposit at the seashore then the correspondingly located sea sample will be censored and hence on these units the observations on Y is not measured. For ranking on X observations in a sample, the censored units are assumed to have distinct and consecutive lower ranks and the remaining units are ranked with the next higher ranks in a natural order. If in this censored scheme of ranked set sampling, k units are censored, then we may represent the ranked set sample observations on the study variate Y as

$\rho_1 Y_{[1]1}, \rho_2 Y_{[2]2}, \ldots, \rho_n Y_{[n]n}$ where

$$\rho_i = \begin{cases} 0 & \text{if the } i\text{th unit is censored} \\ 1 & \text{otherwise} \end{cases}$$

and hence $\sum_{i=1}^{n} \rho_i = n - k$. In this case the usual ranked set sample mean is equal to $\frac{\sum_{i=1}^{n} \rho_i Y_{[i]i}}{n-k}$. It may be noted that $\rho_i = 0$ need not occur in a natural order for $i = 1, 2, \ldots, n$. Hence if we write m_i, $i = 1, 2, \ldots, n-k$, as the integers such that $1 \leq m_1 < m_2 < \cdots < m_{n-k} \leq n$ and for which $\rho_{m_i} = 1$, then

$$E\left[\frac{\sum_{i=1}^{n} \rho_i Y_{[i]i}}{n-k}\right] = \theta_2 \left[1 - \frac{\alpha}{2(n+1)(n-k)} \sum_{i=1}^{n-k} (n - 2m_i + 1)\right].$$

Thus it is clear that the ranked set sample mean in the censored case is not an unbiased estimator of the population mean θ_2. However we can construct an unbiased estimator of θ_2 based on this mean. In the following theorem we have given the constructed unbiased estimator $\theta_2^*(k)$ of θ_2 based on the ranked set sample mean under censored situation and its variance.

Theorem 2 *Suppose that the random variable (X, Y) has a MTBED as defined in (1). Let $Y_{[m_i]m_i}$, $i = 1, 2, \ldots, n-k$, be the ranked set sample observations on the study variate Y resulting out of censoring and ranking applied on the auxiliary variable X. Then an unbiased estimator of θ_2 based on the ranked set sample mean $\frac{1}{n-k} \sum_{i=1}^{n-k} Y_{[m_i]m_i}$ is given by*

$$\theta_2^*(k) = \frac{2(n+1)}{\left[2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k} (n - 2m_i + 1)\right]} \sum_{i=1}^{n-k} Y_{[m_i]m_i}$$

and its variance is given by

$$\mathrm{Var}[\theta_2^*(k)] = \frac{4(n+1)^2 \theta_2^2}{\left[2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k} (n - 2m_i + 1)\right]^2} \sum_{i=1}^{n-k} \delta_{m_i},$$

where δ_{m_i} is as defined in (12).

Proof

$$
\begin{aligned}
&E[\theta_2^*(k)] \\
&= \frac{2(n+1)}{\left[2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k}(n-2m_i+1)\right]} \sum_{i=1}^{n-k} E[Y_{[m_i]m_i}] \\
&= \frac{2(n+1)}{\left[2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k}(n-2m_i+1)\right]} \sum_{i=1}^{n-k} \left[1 - \alpha \frac{n-2m_i+1}{2(n+1)}\right] \theta_2 \\
&= \frac{2(n+1)}{\left[2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k}(n-2m_i+1)\right]} \\
&\quad \times \left[n - k - \frac{\alpha}{2(n+1)} \sum_{i=1}^{n-k}(n-2m_i+1)\right] \theta_2 \\
&= \theta_2.
\end{aligned}
$$

Thus $\theta_2^*(k)$ is an unbiased estimator of θ_2. The variance of $\theta_2^*(k)$ is given by

$$
\begin{aligned}
\mathrm{Var}[\theta_2^*(k)] &= \frac{4(n+1)^2}{\left[(2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k}(n-2m_i+1)\right]^2} \sum_{i=1}^{n-k} \mathrm{Var}(Y_{[m_i]m_i}) \\
&= \frac{4(n+1)^2\, \theta_2^2}{\left[(2(n+1)(n-k) - \alpha \sum_{i=1}^{n-k}(n-2m_i+1)\right]^2} \sum_{i=1}^{n-k} \delta_{m_i},
\end{aligned}
$$

where δ_{m_i} is as defined in (12). Hence the theorem.

As a competitor of the estimator $\theta_2^*(k)$, we now propose the BLUE of θ_2 based on the censored ranked set sample, resulting out of ranking of observations on X.

Let $\mathbf{Y}_{[n]}(k) = (Y_{[m_1]m_1}, Y_{[m_2]m_2}, \ldots, Y_{[m_{n-k}]m_{n-k}})'$, then the mean vector and the dispersion matrix of $\mathbf{Y}_{[n]}(k)$ are given by

$$E[\mathbf{Y}_{[n]}(k)] = \theta_2\, \xi(k), \tag{16}$$

$$D[\mathbf{Y}_{[n]}(k)] = \theta_2^2\, G(k), \tag{17}$$

where $\xi(k) = (\xi_{m_1}, \xi_{m_2}, \ldots, \xi_{m_{n-k}})'$, $G(k) = \mathrm{diag}(\delta_{m_1}, \delta_{m_2}, \ldots, \delta_{m_{n-k}})$.

If the parameter α involved in $\xi(k)$ and $G(k)$ is known then (16) and (17) together defines a generalized Guass–Markov set up and hence the BLUE $\hat{\theta}_2(k)$ of θ_2 is obtained as

$$\hat{\theta}_2(k) = [(\xi(k))'(G(k))^{-1}\xi(k)]^{-1}(\xi(k))'(G(k))^{-1}\mathbf{Y}_{[n]}(k) \tag{18}$$

and

$$\mathrm{Var}(\hat{\theta}_2(k)) = [(\xi(k))'(G(k))^{-1}\xi(k)]^{-1}\theta_2^2. \tag{19}$$

On substituting the values of $\xi(k)$ and $G(k)$ in (18) and (19) and simplifying we get

$$\hat{\theta}_2(k) = \frac{\sum_{i=1}^{n-k}(\xi_{m_i}/\delta_{m_i})Y_{[m_i]m_i}}{\sum_{i=1}^{n-k}\xi_{m_i}^2/\delta_{m_i}} \tag{20}$$

and

$$\mathrm{Var}(\hat{\theta}_2(k)) = \frac{1}{\sum_{i=1}^{n-k}\xi_{m_i}^2/\delta_{m_i}}\theta_2^2.$$

□

Remark 2 In the expression for the BLUE's $\hat{\theta}_2$ given in (15) and $\hat{\theta}_2(k)$ given in (20) the quantities ξ_r and δ_r for $1 \leq r \leq n$ are all non-negative, and consequently the coefficients of ranked set sample observations are also non-negative. Thus the BLUE's $\hat{\theta}_2$ and $\hat{\theta}_2(k)$ of θ_2 are always non-negative. Thus unlike in certain situation of BLUE's where one encounters with inadmissible estimators, the estimators given by $\hat{\theta}_2$ and $\hat{\theta}_2(k)$ using ranked set sample are admissible estimators.

Remark 3 Since both the BLUE $\hat{\theta}_2(k)$ and the unbiased estimator $\theta_2^*(k)$ based on the censored ranked set sample utilize the distributional property of the parent distribution they lose the usual robustness property. Hence in this case the BLUE $\hat{\theta}_2(k)$ shall be considered as a more preferable estimators than $\theta_2^*(k)$.

5 Estimation of θ_2 based on unbalanced multistage ranked set sampling

Al-Saleh and Al-Kadiri (2000) have extended first the usual concept of RSS to double stage ranked set sampling (DSRSS) with an objective of increasing the precision of certain estimators of the population when compared with those obtained based on usual RSS or using random sampling. Al-Saleh and Al-Omari (2002) have further extended DSRSS to multistage ranked set sampling (MSRSS) and shown that there is increase in the precision of estimators obtained based on MSRSS when compared with those based on usual RSS and DSRSS. The MSRSS (in r stages)procedure is described as follows:

(1) Randomly select n^{r+1} sample units from the target population, where r is the number of stages of MSRSS.
(2) Allocate the n^{r+1} selected units randomly into n^{r-1}sets, each of size n^2.
(3) For each set in step (2), apply the procedure of ranked set sampling method to obtain a (judgment) ranked set, of size n; this step yields n^{r-1} (judgment) ranked sets, of size n each.
(4) Arrange n^{r-1} ranked sets of size n each, into n^{r-2} sets of n^2 units each and without doing any actual quantification, apply ranked set sampling method on each set to yield n^{r-2} second stage ranked sets of size n each.

(5) This process is continued, without any actual quantification, until we end up with the *rth* stage (judgment) ranked set of size n.
(6) Finally, the n identified elements in step (5) are now quantified for the variable of interest.

Instead of judgment method of ranking at each stage if there exist an auxiliary variate on which one can make measurement very easily and exactly and if the auxiliary variate is highly correlated with the variable of interest, then we can apply ranking based on these measurements to obtain the ranked set units at each stage of MSRSS. Then on the finally selected units one can make measurement on the variable of primary interest. In this section we deal with the MSRSS by assuming that the random variable (X, Y) has a MTBED as defined in (1), where Y is the variable of primary interest and X is an auxiliary variable. In Sect. 3, we have considered a method for estimating θ_2 using the $Y_{[r]r}$ measured on the study variate Y on the the unit having rth smallest value observed on the auxiliary variable X, of the rth sample $r = 1, 2, \ldots, n$, and hence the RSS considered there was balanced. Abo-Eleneen and Nagaraja (2002) have shown that in a bivariate sample of size n arising from MTBED the concomitant of largest order statistic possess the maximum Fisher information on θ_2 whenever $\alpha > 0$ and the concomitant of smallest order statistic possess the maximum Fisher information on θ_2 whenever $\alpha < 0$. Hence in this section, first we consider $\alpha > 0$ and carry out an unbalanced MSRSS with the help of measurements made on an auxiliary variate to choose the ranked set and then estimate θ_2 involved in MTBED based on the measurement made on the variable of primary interest. At each stage and from each set we choose an unit of a sample with the largest value on the auxiliary variable as the units of ranked sets with an objective of exploiting the maximum Fisher information on the ultimately chosen ranked set sample.

Let $U_i^{(r)}, i = 1, 2, \ldots, n$, be the units chosen by the (r stage) MSRSS. Since the measurement of auxiliary variable on each unit $U_i^{(r)}$ has the largest value, we may write $Y_{[n]i}^{(r)}$ to denote the value measured on the variable of primary interest on $U_i^{(r)}$, $i = 1, 2, \ldots, n$,. Then it is easy to see that each $Y_{[n]i}^{(r)}$ is the concomitant of the largest order statistic of n^r independently and identically distributed bivariate random variables with MTBED. Moreover $Y_{[n]i}^{(r)}$, $i = 1, 2, \ldots, n$, are also independently distributed with pdf given by (see, Scaria and Nair 1999)

$$f_{[n]i}^{(r)}(y;\alpha) = \frac{1}{\theta_2} e^{-\frac{y}{\theta_2}} \left\{ 1 + \alpha \left(\frac{n^r - 1}{n^r + 1} \right) \left(1 - 2e^{-\frac{y}{\theta_2}} \right) \right\}. \tag{21}$$

Thus the mean and variance of $Y_{[n]i}^{(r)}$ for $i = 1, 2, \ldots, n$, are given below:

$$E[Y_{[n]i}^{(r)}] = \theta_2 \left[1 + \frac{\alpha}{2} \left(\frac{n^r - 1}{n^r + 1} \right) \right], \tag{22}$$

$$\mathrm{Var}[Y_{[n]i}^{(r)}] = \theta_2^2 \left[1 + \frac{\alpha}{2} \left(\frac{n^r - 1}{n^r + 1} \right) - \frac{\alpha^2}{4} \left(\frac{n^r - 1}{n^r + 1} \right)^2 \right]. \tag{23}$$

If we denote

$$\xi_{n^r} = 1 + \frac{\alpha}{2} \left(\frac{n^r - 1}{n^r + 1} \right) \tag{24}$$

and

$$\delta_{n^r} = 1 + \frac{\alpha}{2} \left(\frac{n^r - 1}{n^r + 1} \right) - \frac{\alpha^2}{4} \left(\frac{n^r - 1}{n^r + 1} \right)^2, \tag{25}$$

then (22) and (23) can be written as

$$E[Y_{[n]i}^{(r)}] = \theta_2 \xi_{n^r} \tag{26}$$

and

$$\mathrm{Var}[Y_{[n]i}^{(r)}] = \theta_2^2 \delta_{n^r}. \tag{27}$$

Let $\mathbf{Y}_{[n]}^{(r)} = (Y_{[n]1}^{(r)}, Y_{[n]2}^{(r)}, \cdots, Y_{[n]n}^{(r)})'$, then by using (26) and (27) we get the mean vector and dispersion matrix of $\mathbf{Y}_{[n]}^{(r)}$ as

$$E[\mathbf{Y}_{[n]}^{(r)}] = \theta_2 \xi_{n^r} \mathbf{1} \tag{28}$$

and

$$D[\mathbf{Y}_{[n]}^{(r)}] = \theta_2^2 \delta_{n^r} \mathbf{I}, \tag{29}$$

where $\mathbf{1}$ is the column vector of n ones and $\mathbf{I}$ is a unit matrix of order n. If $\alpha > 0$ involved in ξ_{n^r} and δ_{n^r} is known then (28) and (29) together defines a generalized Gauss–Markov setup and hence the BLUE of θ_2 is obtained as

$$\hat{\theta}_2^{n(r)} = \frac{1}{n\xi_{n^r}} \sum_{i=1}^{n} Y_{[n]i}^{(r)} \tag{30}$$

with variance given by

$$\mathrm{Var}(\hat{\theta}_2^{n(r)}) = \frac{\delta_{n^r}}{n(\xi_{n^r})^2} \theta_2^2. \tag{31}$$

If we take $r = 1$ in the MSRSS method described above, then we get the usual single stage unbalanced RSS. Then the BLUE $\hat{\theta}_2^{n(1)}$ of θ_2 is given by

$$\hat{\theta}_2^{n(1)} = \frac{1}{n\xi_n} \sum_{i=1}^{n} Y_{[n]i},$$

with variance

$$\mathrm{Var}(\hat{\theta}_2^{n(1)}) = \frac{\delta_n}{n(\xi_n)^2} \theta_2^2,$$

where we write $Y_{[n]i}$ instead of $Y_{[n]i}^{(1)}$ and it represent the measurement on the variable of primary interest of the unit selected in the RSS. Also ξ_n and δ_n are obtained by putting $r = 1$ in (24) and (25), respectively.

We have evaluated the ratio $e(\hat{\theta}_2^{n(1)}|\tilde{\theta_2}) = \frac{\mathrm{Var}(\tilde{\theta_2})}{\mathrm{Var}(\hat{\theta}_2^{n(1)})}$ for $\alpha = 0.25(0.25)1$; $n = 2(2)10(5)20$ as a measure of efficiency of our estimator $\hat{\theta}_2^{n(1)}$ relative to the MLE $\tilde{\theta_2}$ of θ_2 based on n observations and are also provided in Table 1. From the table, one can see that the efficiency increases with increase in α and n. Moreover the efficiency of the estimator $\hat{\theta}_2^{n(1)}$ is larger than the estimator θ_2^* based on RSS mean and the BLUE $\hat{\theta}_2$ based on usual RSS.

Al-Saleh (2004) has considered the steady-state RSS by letting r to $+\infty$. If we apply the steady-state RSS to the problem considered in this paper then the asymptotic distribution of $Y_{[n]i}^{(r)}$ is given by the pdf given by

$$f_{[n]i}^{(\infty)}(y;\alpha) = \frac{1}{\theta_2} e^{-\frac{y}{\theta_2}} \left\{1 + \alpha \left(1 - 2e^{-\frac{y}{\theta_2}}\right)\right\}. \tag{32}$$

From the definition of our unbalanced MSRSS it follows that $Y_{[n]i}^{(\infty)}$, $i = 1, 2, \ldots, n$, are independent and identically distributed random variables each with pdf as defined in (32). Then $Y_{[n]i}^{(\infty)}$, $i = 1, 2, \ldots, n$, may be regarded as unbalanced steady-state ranked set sample of size n. Then the mean and variance of $Y_{[n]i}^{(\infty)}$ for $i = 1, 2, \ldots, n$, are given below:

$$E[Y_{[n]i}^{(\infty)}] = \theta_2 \left[1 + \frac{\alpha}{2}\right]$$

and

$$\mathrm{Var}[Y_{[n]i}^{(\infty)}] = \theta_2^2 \left[1 + \frac{\alpha}{2} - \frac{\alpha^2}{4}\right].$$

Let $\mathbf{Y}_{[n]}^{(\infty)} = (Y_{[n]1}^{(\infty)}, Y_{[n]2}^{(\infty)}, \cdots, Y_{[n]n}^{(\infty)})'$. Then the BLUE $\hat{\theta}_2^{n(\infty)}$ based on $\mathbf{Y}_{[n]}^{(\infty)}$ and the variance of $\hat{\theta}_2^{n(\infty)}$ is obtained by taking the limit as $r \to \infty$ in (30) and (31), respectively and are given by

$$\hat{\theta}_2^{n(\infty)} = \frac{1}{n\left[1 + \frac{\alpha}{2}\right]} \sum_{i=1}^{n} Y_{[n]i}^{(\infty)}$$

and

$$\mathrm{Var}(\hat{\theta}_2^{n(\infty)}) = \frac{\left[1 + \frac{\alpha}{2} - \frac{\alpha^2}{4}\right]}{n\left[1 + \frac{\alpha}{2}\right]^2} \theta_2^2. \tag{33}$$

Table 2 Efficiencies of the estimators $\hat{\theta}_2^{n(\infty)}$ and $\hat{\theta}_2^{1(\infty)}$ relative to $\tilde{\theta_2}$

α	$e(\hat{\theta}_2^{n(\infty)}\|\tilde{\theta_2})$	α	$e(\hat{\theta}_2^{1(\infty)}\|\tilde{\theta_2})$
0.25	1.1382	−0.25	1.1382
0.50	1.3028	−0.50	1.3033
0.75	1.4943	−0.75	1.4960
1.00	1.7093	−1.00	1.7161

From (10) and (33) we get the efficiency of $\hat{\theta}_2^{n(\infty)}$ relative to $\tilde{\theta_2}$ by taking the ratio of $\mathrm{Var}(\tilde{\theta_2})$ with $\mathrm{Var}(\hat{\theta}_2^{n(\infty)})$ and is given by

$$\begin{aligned} e(\hat{\theta}_2^{n(\infty)}|\tilde{\theta_2}) &= \frac{\mathrm{Var}(\tilde{\theta_2})}{\mathrm{Var}(\hat{\theta}_2^{n(\infty)})}, \\ &= \frac{I_{\theta_2}^{(-1)}(\alpha)\left[1+\frac{\alpha}{2}\right]^2}{\left[1+\frac{\alpha}{2}-\frac{\alpha^2}{4}\right]}. \end{aligned}$$

Thus the efficiency $e(\hat{\theta}_2^{n(\infty)}|\tilde{\theta_2})$ is free of the sample size n. That is, for a fixed α, $e(\hat{\theta}_2^{n(\infty)}|\tilde{\theta_2})$ is a constant for all n. We have evaluated the value $e(\hat{\theta}_2^{n(\infty)}|\tilde{\theta_2})$ for $\alpha = 0.25(0.25)1$ and are presented in Table 2. From the table one can see that the efficiency of $\hat{\theta}_2^{n(\infty)}$ increases as α increases. moreover the estimator $\hat{\theta}_2^{n(\infty)}$ possess the highest efficiency amoung the other estimators of θ_2 proposed in this paper and the value of the efficiencies ranges from 1.1382 to 1.7093.

As mentioned earlier for MTBED the concomitant of smallest order statistic possess the maximum Fisher information on θ_2 whenever $\alpha < 0$. Therefore when $\alpha < 0$ we consider an unbalanced MSRSS in which at each stage and from each set we choose an unit of a sample with the smallest value on the auxiliary variable as the units of ranked sets with an objective of exploiting the maximum Fisher information on the ultimately chosen ranked set sample.

Let $Y_{[1]i}^{(r)}$, $i = 1, 2, \ldots, n$, be the value measured on the variable of primary interest on the units selected at the rth stage of the unbalanced MSRSS. Then it is easily to see that each $Y_{[1]i}^{(r)}$ is the concomitant of the smallest order statistic of n^r independently and identically distributed bivariate random variables with MTBED. Moreover $Y_{[1]i}^{(r)}, i = 1, 2, \ldots, n$, are also independently distributed with pdf given by

$$f_{[1]i}^{(r)}(y;\alpha) = \frac{1}{\theta_2}e^{-\frac{y}{\theta_2}}\left\{1-\alpha\left(\frac{n^r-1}{n^r+1}\right)\left(1-2e^{-\frac{y}{\theta_2}}\right)\right\}. \tag{34}$$

Clearly from (21) and (34) we have

$$f_{[1]i}^{(r)}(y;\alpha) = f_{[n]i}^{(r)}(y;-\alpha), \tag{35}$$

and hence $E(Y^{(r)}_{[n]i})$ for $\alpha > 0$ and $E(Y^{(r)}_{[1]i})$ for $\alpha < 0$ are identically equal. Similarly $\mathrm{Var}(Y^{(r)}_{[n]i})$ for $\alpha > 0$ and $Var(Y^{(r)}_{[1]i})$ for $\alpha < 0$ are identically equal. Consequently if $\hat{\theta}_2^{1(1)}$ is the BLUE of θ_2, involved in MTBED for $\alpha < 0$, based on the unbalanced MSRSS observations $Y^{(r)}_{[1]i}$, $i = 1, 2, \ldots, n$,, then the coefficients of $Y^{(r)}_{[1]i}$, $i = 1, 2, \ldots, n$, in the BLUE $\hat{\theta}_2^{1(1)}$ for $\alpha < 0$ is same as the coefficients of $Y^{(r)}_{[n]i}$, $i = 1, 2, \ldots, n$, in the BLUE $\hat{\theta}_2^{n(r)}$ for $\alpha > 0$. Further we have $\mathrm{Var}(\hat{\theta}_2^{1(1)}) = \mathrm{Var}(\hat{\theta}_2^{n(r)})$ and hence $\mathrm{Var}(\hat{\theta}_2^{1(1)}) = \mathrm{Var}(\hat{\theta}_2^{n(1)})$ and $\mathrm{Var}(\hat{\theta}_2^{1(\infty)}) = Var(\hat{\theta}_2^{n(\infty)})$, where $\hat{\theta}_2^{1(1)}$ is the BLUE of θ_2, for $\alpha < 0$ based on the usual unbalanced single stage RSS observations $Y_{[1]i}$, $i = 1, 2, \ldots, n$, and $\hat{\theta}_2^{1(\infty)}$ is the BLUE of θ_2, for $\alpha < 0$ based on the unbalanced steady-state RSS observations $Y^{(\infty)}_{[1]i}$, $i = 1, 2, \ldots, n$,. We have obtained the efficiency $e(\hat{\theta}_2^{1(1)}|\tilde{\theta_2})$ of the BLUE $\hat{\theta}_2^{1(1)}$ relative to $\tilde{\theta_2}$, the MLE of θ_2 for $\alpha = -0.25, -0.5, -0.75, -1$; $n = 2(2)10(5)20$ and are incorporated in Table 1. Similarly as in the case of $\hat{\theta}_2^{n(\infty)}$ for a fixed α the efficiency $e(\hat{\theta}_2^{1(\infty)}|\tilde{\theta_2})$ is same for all n. We have evaluated $e(\hat{\theta}_2^{1(\infty)}|\tilde{\theta_2})$ for $\alpha = -0.25, -0.5, -0.75, -1$ and are incorporated in Table 2. From the table, one can see that efficiency increases as $|\alpha|$ increases and the value of efficiency ranges from 1.1382 to 1.7161.

Remark 4 If (X, Y) follows an MTBED with pdf defined in (1), then the correlation coefficient between X and Y is given by

$$\mathrm{Corr}(X, Y) = \frac{\alpha}{4}, \quad -1 \leq \alpha \leq 1.$$

Clearly when $|\alpha|$ goes to 1, correlation coefficient between X and Y is high. Thus, using the ranks of Y to induce the ranks of X becomes more accurate. Thus when $|\alpha|$ is large (that is α tends to ± 1) we see that the ranked set sample obtained based on the ranking made on X becomes more informative for making inference on θ_2 than the case with small values of $|\alpha|$. From the table we notice that for a given sample size the efficiencies of all estimators increase as $|\alpha|$ increase. Consequently we note that more information on θ_2 can be extracted from the ranked set sample when $|\alpha|$ is large subject to $|\alpha| \leq 1$. Thus we conclude that concomitant ranking is more effective in estimating θ_2 when the absolute value of the association parameter α is large (that is when α tends to ± 1).

Acknowledgements The authors are highly grateful to the referee for the constructive comments which helped much in the improvement of the revised version of the paper.

References

Abo-Eleneen, Z. A., Nagaraja, H. N. (2002). Fisher information in an order statistic and its concomitant. *Annals of the Institute of Statistical Mathematics*, *54*, 667–680.

Al-Saleh, M. F. (2004). Steady-state ranked set sampling and parametric inference. *Journal of Statistical Planning and Inference*, *123*, 83–95.

Al-Saleh, M. F., Al-Kadiri, M. (2000). Double ranked set sampling. *Statistics and Probability Letters*, *48*, 205–212.

Al-Saleh, M. F., Al-Omari, A. (2002). Multistage ranked set sampling. *Journal of Statistical Planning and Inference*, *102*, 273–286.

Bain, L. J. (1978). *Statistical analysis of reliability and life testing models: theory and methods*. New York: Marcel Dekker.

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

Chen, Z. (2000). The efficiency of ranked-set sampling relative to simple random sampling under multi-parameter families. *Statistica Sinica*, *10*, 247–263.

Chen, Z. and Bai, Z. (2000). The optimal ranked set sampling scheme for parametric families. *Sankhya Series A*, *46*, 178–192.

Chen, Z., Bai, Z., Sinha, B. K. (2004). *Lecture notes in statistics, ranked set sampling, theory and applications*. New York: Springer.

David, H. A., Nagaraja, H. N (2003). *Order statistics*. New York: Wiley.

Kotz, S., Balakrishnan, N., Johnson, N. L. (2000). *Distributions in statistics: continuous multivariate distributions*. New York: Wiley.

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

Lam, K., Sinha, B. K., Wu, Z. (1995). Estimation of location and scale parameters of a logistic distribution using ranked set sample. In: H. N. Nagaraja, P. K. Sen D. F. Morrison (Eds) *Statistical theory and applications: papers in honor of Herbert A. David*. New-York: Springer.

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

Modarres, R., Zheng, G. (2004). Maximum likelihood estimation of dependence parameter using ranked set sampling. *Statistics and Probability Letters*, *68*, 315–323.

Scaria, J., Nair, N. 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). Inference on the correlation coefficient in bivariate normal populations from ranked Set Samples. *Journal of the American Statistical Association*, *75*, 989–995.

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

Zheng, G., Modarres, R. (2006). A robust estimate of correlation coefficient for bivariate normal distribution using ranked set sampling. *Journal of Statistical Planning and Inference*, *136*, 298–309.