Probability in the Engineering and Informational Sciences, **36**, 2022, 449–460.
doi:10.1017/S0269964820000625

RATIO ESTIMATION OF THE POPULATION MEAN USING AUXILIARY INFORMATION UNDER THE OPTIMAL SAMPLING DESIGN

CHUNXIAN LONG, WANGXUE CHEN, RUI YANG AND DONGSEN YAO
Department of Mathematics and Statistics, Jishou University, Jishou 416000, China
E-mail: chenwangxue2015@163.com

Cost-effective sampling design is a problem of major concern in some experiments especially when the measurement of the characteristic of interest is costly or painful or time-consuming. In this article, we investigate ratio-type estimators of the population mean of the study variable, involving either the first or the third quartile of the auxiliary variable, using ranked set sampling (RSS) and extreme ranked set sampling (ERSS) schemes. The properties of the estimators are obtained. The estimators in RSS and ERSS are compared to their counterparts in simple random sampling (SRS) for normal data. The numerical results show that the estimators in RSS and ERSS are significantly more efficient than their counterparts in SRS.

Keywords: auxiliary variable, extreme ranked set sampling, ranked set sampling, ratio estimator

1. INTRODUCTION

The auxiliary information associated with an auxiliary variable X that is correlated with the study variable Y can be employed to increase the efficiency of estimators. The most commonly used auxiliary information is the mean of the auxiliary variable. However, other auxiliary information associated with the variable X such as the median, coefficient of variation or correlation coefficient can be used to improve the efficiency of the estimators. Upadhyaya and Singh [19] suggested two ratio-type estimators for when the coefficient of variation and the coefficient of kurtosis of the auxiliary variable are available. Srivastava and Jhajj [15] suggested a class of estimators of the population mean.

The usual simple random sampling (SRS) ratio estimator of the population mean μ_Y of the study variable Y from a sample of size m is given by

$$\hat{\mu}_{Y_{\mathrm{SRS}}} = \mu_X \left(\frac{\overline{Y}_{\mathrm{SRS}}}{\overline{X}_{\mathrm{SRS}}} \right) \tag{1.1}$$

in which it is assumed that the population mean μ_X of the auxiliary variable X is known. Here, $\overline{X}_{\mathrm{SRS}} = (1/m)\sum_{i=1}^{m} X_i$ and $\overline{Y}_{\mathrm{SRS}} = (1/m)\sum_{i=1}^{m} Y_i$ are the sample means of X and

© The Author(s), 2020. Published by Cambridge University Press 0269-9648/19 $25.00

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

Y, respectively. The approximate mean square error (MSE) of $\hat{\mu}_{Y_{\mathrm{SRS}}}$ is given by

$$\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{SRS}}}) \approx \frac{1-\frac{m}{N}}{m}(\sigma_Y^2 + R^2\sigma_X^2 - 2R\sigma_{XY}), \tag{1.2}$$

where N is the population size, σ_X^2 and σ_Y^2 are the population variances of X and Y, respectively, $R = \mu_Y/\mu_X$ is the population ratio and σ_{XY} is the covariance of X and Y. For more details, see Cochran [7], Kadilar and Cingi [8] and Singh and Tailor [14].

Cost-effective sampling is a problem of major concern in some experiments especially when the measurement of the characteristic of interest is costly or painful or time-consuming. The method of ranked set sampling (RSS) provides an effective way to achieve observational economy in terms of precision achieved per unit of sampling. Initially, the concept of RSS was introduced by McIntyre [9] to estimate pasture yields. Takahasi and Wakimoto [18] established a very important statistical foundation for the theory of RSS. Later estimation of parameters of various commonly used distributions has been carried out using RSS. For more details, see Chen *et al.* [2–6], Qiu and Eftekharian [11] and Stokes [17]. However, ranking accuracy affects the efficiency of the estimator. When the set size is large, ranking error tends to occur. In order to reduce the error of ranking and keep optimality inherited in the original RSS procedure, Samawi *et al.* [13] introduced a modification of RSS called extreme RSS (ERSS).

There have also been some studies of parametric inference in ERSS. Qian *et al.* [10] studied the parameter estimation of the Pareto distribution. Yang *et al.* [20] studied the parameter estimation of the log-extended exponential-geometric distribution.

In this paper, we investigate ratio-type estimators of the population mean μ_Y of the study variable Y, involving either the first or the third quartile of the auxiliary variable X, using RSS and ERSS methods. In Section 2, the suggested estimators of μ_Y is given using SRS. In Sections 3 and 4, the suggested estimators of μ_Y given using RSS and ERSS, with their properties, are presented in detail. A comparison between the estimators will be presented in Section 5. Conclusions about the estimators are provided in Section 6.

2. ESTIMATION OF μ_Y IN SRS

Let $\bigcup_{i=1}^{m}(X_i, Y_i)$ be a bivariate simple random sample of m with probability density function (pdf) $f(x, y)$, cumulative distribution function (cdf) $F(x, y)$, means μ_X, μ_Y, variances σ_X^2, σ_Y^2 and correlation coefficient ρ.

Al-Omari [1] investigate ratio-type estimators of μ_Y of the study variable Y, involving either the first (q_1) or the third (q_3) quartile of the auxiliary variable X, using SRS methods. The suggested ratio estimators of μ_Y using SRS based on q_1 and q_3 of X, respectively, are given by

$$\hat{\mu}_{Y_{\mathrm{SRS1}}} = \bar{Y}_{\mathrm{SRS}}\left(\frac{\mu_X + q_1}{\bar{X}_{\mathrm{SRS}} + q_1}\right) \tag{2.1}$$

and

$$\hat{\mu}_{Y_{\mathrm{SRS3}}} = \bar{Y}_{\mathrm{SRS}}\left(\frac{\mu_X + q_3}{\bar{X}_{\mathrm{SRS}} + q_3}\right) \tag{2.2}$$

where $\overline{X}_{\mathrm{SRS}} = (1/m)\sum_{i=1}^{m} X_i$, $\overline{Y}_{\mathrm{SRS}} = (1/m)\sum_{i=1}^{m} Y_i$. For short, we write $\hat{\mu}_{Y_{\mathrm{SRS}k}}(k = 1, 3)$ for $\hat{\mu}_{Y_{\mathrm{SRS1}}}$ and $\hat{\mu}_{Y_{\mathrm{SRS3}}}$, respectively. Under the first-degree approximation of the Taylor series

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

method, the approximate MSE of $\hat{\mu}_{Y_{\mathrm{SRS}k}}$ is given by

$$\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{SRS}k}}) \approx \frac{1}{m}(\sigma_Y^2 + L_k^2\sigma_X^2 - 2L_k\rho\sigma_X\sigma_Y), \tag{2.3}$$

where $L_k = \mu_Y/(\mu_X + q_k)$.

3. ESTIMATION OF μ_Y IN RSS

An important advantage of RSS approach is that, it improves the efficiency of estimators of the population parameters by providing more representative sample from the target population. In RSS, one first draws m^2 units at random from the population and partitions them into m sets of m units. The m units in each set are ranked without making actual measurements. From the first set of m units, the unit ranked lowest is chosen for actual quantification. From the second set of m units, the unit ranked second lowest is measured. The process is continued until the unit ranked largest is measured from the mth set of m units. The whole process can be repeated k times if needed to obtain a sample of size km units.

Samawi and Muttlak [12] showed that ranking on variable X is more efficient than ranking on variable Y, so in this study, we will consider ranking on the auxiliary variable X. We investigate ratio-type estimators of the population mean μ_Y of the study variable Y, involving either the first or the third quartile of the auxiliary variable X, using RSS methods.

Let $\bigcup_{i=1}^{m}(X_{i(i:m)}, Y_{i[i:m]})$ be a bivariate ranked set sample of m, where $X_{i(i:m)}$ means that in the ith set, the ith order statistic in a random sample of size m is recorded, $Y_{i[i:m]}$ is the concomitant order statistic of $X_{i(i:m)}$. The suggested ratio estimators of μ_Y using RSS based on q_1 and q_3 of X, respectively, are given by

$$\hat{\mu}_{Y_{\mathrm{RSS1}}} = \bar{Y}_{\mathrm{RSS}}\left(\frac{\mu_X + q_1}{\bar{X}_{\mathrm{RSS}} + q_1}\right) \tag{3.1}$$

and

$$\hat{\mu}_{Y_{\mathrm{RSS3}}} = \bar{Y}_{\mathrm{RSS}}\left(\frac{\mu_X + q_3}{\bar{X}_{\mathrm{RSS}} + q_3}\right) \tag{3.2}$$

where $\bar{X}_{\mathrm{RSS}} = (1/m)\sum_{i=1}^{m} X_{i(i:m)}$ with

$$\mathrm{Var}(\bar{X}_{\mathrm{RSS}}) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2_{X_{(i:m)}} \tag{3.3}$$

and $\bar{Y}_{\mathrm{RSS}} = (1/m)\sum_{i=1}^{m} Y_{i[i:m]}$ with

$$\mathrm{Var}(\bar{Y}_{\mathrm{RSS}}) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2_{Y_{[i:m]}}. \tag{3.4}$$

The covariance of $\bar{X}_{\mathrm{RSS}}$ and $\bar{Y}_{\mathrm{RSS}}$ is

$$\mathrm{Cov}(\bar{X}_{\mathrm{RSS}}, \bar{Y}_{\mathrm{RSS}}) = \frac{1}{m^2}\sum_{i=1}^{m}\mathrm{Cov}(X_{i(i:m)}, Y_{i[i:m]}) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma_{XY_{[i:m]}}, \tag{3.5}$$

where $\sigma_{XY_{[i:m]}}$ is the covariance of $X_{i(i:m)}$ and $Y_{i[i:m]}$. Then, $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ has the following properties.

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

THEOREM 1: *Under the first-degree approximation of the Taylor series method, the approximate MSE of* $\hat{\mu}_{Y_{RSSk}}$ *is given by*

$$MSE(\hat{\mu}_{Y_{RSSk}}) \approx \frac{1}{m^2}\sum_{i=1}^{m}(\sigma^2_{Y_{[i:m]}} + L_k^2\sigma^2_{X_{(i:m)}} - 2L_k\sigma_{XY_{[i:m]}}). \tag{3.6}$$

PROOF: Using Taylor series expansion, $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ can be approximated as

$$\hat{\mu}_{Y_{\mathrm{RSS}k}} \approx \bar{Y}_{\mathrm{RSS}} - L_k(\bar{X}_{\mathrm{RSS}} - \mu_X) + L_kV_k(\bar{X}_{\mathrm{RSS}} - \mu_X)^2 - V_k(\bar{X}_{\mathrm{RSS}} - \mu_X)(\bar{Y}_{\mathrm{RSS}} - \mu_Y), \tag{3.7}$$

where $V_k = 1/(\mu_X + q_k)$. In the first-degree of approximation, the estimator in (3.7) is given by

$$\hat{\mu}_{Y_{\mathrm{RSS}k}} \approx \bar{Y}_{\mathrm{RSS}} - L_k(\bar{X}_{\mathrm{RSS}} - \mu_X). \tag{3.8}$$

Then, use (3.3), (3.4), (3.5) and (3.8)

$$\begin{aligned}
\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{RSS}k}}) &= E(\hat{\mu}_{Y_{\mathrm{RSS}k}} - \mu_Y)^2 \\
&\approx E[\bar{Y}_{\mathrm{RSS}} - L_k(\bar{X}_{\mathrm{RSS}} - \mu_X) - \mu_Y]^2 \\
&= E(\bar{Y}_{\mathrm{RSS}} - \mu_Y)^2 + L_k^2E(\bar{X}_{\mathrm{RSS}} - \mu_X)^2 - 2L_kE(\bar{X}_{\mathrm{RSS}} - \mu_X)(\bar{Y}_{\mathrm{RSS}} - \mu_Y) \\
&= \mathrm{Var}(\bar{Y}_{\mathrm{RSS}}) + L_k^2\,\mathrm{Var}(\bar{X}_{\mathrm{RSS}}) - 2L_k\,\mathrm{Cov}(\bar{X}_{\mathrm{RSS}}, \bar{Y}_{\mathrm{RSS}}) \\
&= \frac{1}{m^2}\sum_{i=1}^{m}(\sigma^2_{Y_{[i:m]}} + L_k^2\sigma^2_{X_{(i:m)}} - 2L_k\sigma_{XY_{[i:m]}}).
\end{aligned}$$

This completes the proof of Theorem 1. ■

Suppose that the regression of Y on X is linear (see [16]), then $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ has the following properties.

COROLLARY 1: *Under the first-degree approximation of the Taylor series method,* $\hat{\mu}_{Y_{RSSk}}$ *satisfies*

$$Bias(\hat{\mu}_{Y_{RSSk}}) \approx 0 \tag{3.9}$$

and

$$MSE(\hat{\mu}_{Y_{RSSk}}) \approx \frac{1}{m^2}\left[m\sigma_Y^2(1-\rho^2) + (\rho\sigma_Y - L_k\sigma_X)^2\sum_{i=1}^{m}\sigma^2_{(i:m)}\right], \tag{3.10}$$

where $\sigma^2_{(i:m)}$ *is the variance of ith order statistic in random sample of size m from a standardized continuous distribution F with mean 0 and variance 1.*

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

PROOF: Since the regression of Y on X is linear (see Stokes [16]), we have

$$E(Y \mid X) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X - \mu_X) \tag{3.11}$$

and

$$\mathrm{Var}(Y \mid X) = \sigma_Y^2(1 - \rho^2). \tag{3.12}$$

Stokes [16] proved

$$\mathrm{Var}(X_{(i:m)}) = \sigma_X^2\sigma_{(i:m)}^2. \tag{3.13}$$

Use (3.11)–(3.13)

$$E(Y_{[i:m]}) = E[E(Y_{[i:m]} \mid X_{(i:m)})] = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}E(X_{_{(i:m)}} - \mu_X) \tag{3.14}$$

and

$$\begin{aligned}\mathrm{Var}(Y_{[i:m]}) &= E[\mathrm{Var}(Y_{[i:m]} \mid X_{(i:m)})] + \mathrm{Var}[E(Y_{[i:m]} \mid X_{(i:m)})] \\ &= \sigma_Y^2(1 - \rho^2) + \rho^2\frac{\sigma_Y^2}{\sigma_X^2}\,\mathrm{Var}(X_{(i:m)}) \\ &= \sigma_Y^2(1 - \rho^2) + \rho^2\sigma_Y^2\sigma_{(i:m)}^2. \end{aligned} \tag{3.15}$$

Mclntyre [9] has obtained

$$E(\bar{X}_{\mathrm{RSS}}) = \mu_X. \tag{3.16}$$

Use (3.14) and (3.16)

$$E(\bar{Y}_{\mathrm{RSS}}) = \frac{1}{m}\sum_{i=1}^{m}E(Y_{[i:m]}) = \frac{1}{m}\sum_{i=1}^{m}\left[\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}E(X_{(i:m)} - \mu_X)\right] = \mu_Y. \tag{3.17}$$

Al-Omari [1] proved

$$\mathrm{Cov}(\bar{X}_{\mathrm{RSS}}, \bar{Y}_{\mathrm{RSS}}) = \rho\frac{\sigma_Y}{\sigma_X}\,\mathrm{Var}(\bar{X}_{\mathrm{RSS}}). \tag{3.18}$$

Use (3.13) and (3.18)

$$\mathrm{Cov}(\bar{X}_{\mathrm{RSS}}, \bar{Y}_{\mathrm{RSS}}) = \frac{1}{m^2}\rho\sigma_X\sigma_Y\sum_{i=1}^{m}\sigma_{(i:m)}^2. \tag{3.19}$$

Thus, use (3.8), (3.16) and (3.17)

$$\mathrm{Bias}(\hat{\mu}_{Y_{\mathrm{RSS}k}}) = E(\hat{\mu}_{Y_{\mathrm{RSS}k}}) - \mu_Y \approx E[\bar{Y}_{\mathrm{RSS}} - L_k(\bar{X}_{\mathrm{RSS}} - \mu_X)] - \mu_Y = 0.$$

Use (3.8), (3.13), (3.15) and (3.19)

$$\begin{aligned}\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{RSS}k}}) &= E(\hat{\mu}_{Y_{\mathrm{RSS}k}} - \mu_Y)^2 \\ &\approx \mathrm{Var}(\bar{Y}_{\mathrm{RSS}}) + L_k^2\,\mathrm{Var}(\bar{X}_{\mathrm{RSS}}) - 2L_k\,\mathrm{Cov}(\bar{X}_{\mathrm{RSS}}, \bar{Y}_{\mathrm{RSS}}) \\ &= \frac{1}{m^2}\left[m\sigma_Y^2(1 - \rho^2) + (\rho\sigma_Y - L_k\sigma_X)^2\sum_{i=1}^{m}\sigma_{(i:m)}^2\right].\end{aligned}$$

This completes the proof of Corollary 1. ■

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

4. ESTIMATION OF μ_Y IN ERSS

Take the sorting error into consideration, Samawi *et al.* [13] introduced a modification of RSS called extreme RSS (ERSS). The procedure of ERSS is described as follows: Randomly draw units of size m^2 from the population and randomly partition them into m sets of m units. If m is even, select the lowest ranked unit from each in $m/2$ sets for actual measurement. Select the largest ranked unit from each in the other sets for actual measurement. If m is odd, select the lowest ranked unit from each in $(m-1)/2$ sets for actual measurement. Select the largest ranked unit from each in $(m-1)/2$ sets for actual measurement. Select the $(m+1)/2$ ranked unit from the mth set for actual measurement. The whole process can be repeated k times if needed to obtain a sample of size km units.

In this section, we investigate ratio-type estimators of the population mean μ_Y of the study variable Y, involving either the first or the third quartile of the auxiliary variable X, using ERSS methods.

4.1. Estimation of μ_Y when m is odd

Let $\bigcup_{i=1}^{(m-1)/2}(X_{i(1:m)}, \quad Y_{i[1:m]}) \cup \bigcup_{i=(m+1)/2}^{m-1}(X_{i(m:m)}, \quad Y_{i[m:m]}) \cup (X_{m((m+1)/2:m)}, Y_{m[(m+1)/2:m]})$ be a bivariate extreme ranked set sample of m. The suggested ratio estimators of μ_Y using ERSS based on q_1 and q_3 of X, respectively, are given by

$$\hat{\mu}_{Y_{\mathrm{ERSS}O1}} = \bar{Y}_{\mathrm{ERSS}O}\left(\frac{\mu_X + q_1}{\bar{X}_{\mathrm{ERSS}O} + q_1}\right) \tag{4.1}$$

and

$$\hat{\mu}_{Y_{\mathrm{ERSS}O3}} = \bar{Y}_{\mathrm{ERSS}O}\left(\frac{\mu_X + q_3}{\bar{X}_{\mathrm{ERSS}O} + q_3}\right), \tag{4.2}$$

where $\bar{X}_{\mathrm{ERSS}O} = (1/m)(\sum_{i=1}^{(m-1)/2} X_{i(1:m)} + \sum_{i=(m+1)/2}^{m-1} X_{i(m:m)} + X_{m((m+1)/2:m)})$ with

$$\begin{aligned}\mathrm{Var}(\bar{X}_{\mathrm{ERSS}O}) &= \frac{1}{m^2}\left[\sum_{i=1}^{(m-1)/2}\mathrm{Var}(X_{i(1:m)}) + \sum_{i=(m+1)/2}^{m-1}\mathrm{Var}(X_{i(m:m)}) + \mathrm{Var}(X_{m((m+1)/2:m)})\right] \\ &= \frac{1}{m^2}\left[\frac{m-1}{2}(\sigma^2_{X(1:m)} + \sigma^2_{X(m:m)}) + \sigma^2_{X(\frac{m+1}{2}:m)}\right] \end{aligned} \tag{4.3}$$

and $\bar{Y}_{\mathrm{ERSS}O} = (1/m)(\sum_{i=1}^{(m-1)/2} Y_{i[1:m]} + \sum_{i=(m+1)/2}^{m-1} Y_{i[m:m]} + Y_{m[(m+1)/2:m]})$ with

$$\begin{aligned}\mathrm{Var}(\bar{Y}_{\mathrm{ERSS}O}) &= \frac{1}{m^2}\left[\sum_{i=1}^{(m-1)/2}\mathrm{Var}(Y_{i[1:m]}) + \sum_{i=(m+1)/2}^{m-1} Var(Y_{i[m:m]}) + Var(Y_{m[(m+1)/2:m]})\right] \\ &= \frac{1}{m^2}\left[\frac{m-1}{2}(\sigma^2_{Y[1:m]} + \sigma^2_{Y[m:m]}) + \sigma^2_{Y[(m+1)/2:m]}\right]. \end{aligned} \tag{4.4}$$

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

The covariance of $\bar{X}_{\mathrm{ERSSO}}$ and $\bar{Y}_{\mathrm{ERSSO}}$ is

$$\begin{aligned}\mathrm{Cov}(\bar{X}_{\mathrm{ERSSO}},\bar{Y}_{\mathrm{ERSSO}}) &= \frac{1}{m^2}\left[\sum_{i=1}^{(m-1)/2}\mathrm{Cov}(X_{i(1:m)},Y_{i[1:m]})+\sum_{i=(m+1)/2}^{m-1}\mathrm{Cov}(X_{i(m:m)},Y_{i[m:m]})\right.\\ &\quad\left.+\mathrm{Cov}(X_{m((m+1)/2:m)},Y_{m[(m+1)/2:m]})\right]\\ &= \frac{1}{m^2}\left[\frac{m-1}{2}(\sigma_{XY[1:m]}+\sigma_{XY[m:m]})+\sigma_{XY[(m+1)/2:m]}\right]. \qquad \textbf{(4.5)}\end{aligned}$$

Then, $\hat{\mu}_{Y_{\mathrm{ERSSO}k}}$ has the following properties.

THEOREM 2: *Under the first-degree approximation of the Taylor series method, the approximate MSE of* $\hat{\mu}_{Y_{ERSSOk}}$ *is given by*

$$\begin{aligned}MSE(\hat{\mu}_{Y_{ERSSOk}}) &\approx \frac{1}{m^2}\left\{\left[\frac{m-1}{2}(\sigma^2_{Y[1:m]}+\sigma^2_{Y[m:m]})+\sigma^2_{Y[(m+1)/2:m]}\right]\right.\\ &\quad+L_k^2\left[\frac{m-1}{2}(\sigma^2_{X(1:m)}+\sigma^2_{X(m:m)})+\sigma^2_{X((m+1)/2:m)}\right]\\ &\quad\left.-2L_k\left[\frac{m-1}{2}(\sigma_{XY[1:m]}+\sigma_{XY[m:m]})+\sigma_{XY[(m+1)/2:m]}\right]\right\}. \qquad \textbf{(4.6)}\end{aligned}$$

Since the proof procedure of this Theorem is similar to that of Theorem 1, it is omitted.

Suppose that the regression of Y on X is linear (see [16]) and the marginal distribution of X is symmetric, then $\hat{\mu}_{Y_{\mathrm{ERSSO}k}}$ has the following properties.

COROLLARY 2: *Under the first-degree approximation of the Taylor series method,* $\hat{\mu}_{Y_{ERSSOk}}$ *satisfies*

$$Bias(\hat{\mu}_{Y_{ERSSOk}}) \approx 0 \qquad \textbf{(4.7)}$$

and

$$MSE(\hat{\mu}_{Y_{ERSSOk}}) \approx \frac{1}{m^2}\{m\sigma_Y^2(1-\rho^2)+[(m-1)\sigma^2_{(m:m)}+\sigma^2_{((m+1)/2:m)}](\rho\sigma_Y-L_k\sigma_X)^2\}. \qquad \textbf{(4.8)}$$

Note that only when the marginal distribution of X is symmetric in ERSS, $\bar{X}_{\mathrm{ERSSO}}$ is an unbiased estimator of μ_X (see [13]). The other proof procedure is similar to that of Corollary 1. Hence, it is also omitted.

4.2. Estimation of μ_Y when m is even

Let $\bigcup_{i=1}^{m/2}(X_{i(1:m)},Y_{i[1:m]})\cup\bigcup_{i=(m+2)/2}^{m}(X_{i(m:m)},Y_{i[m:m]})$ be a bivariate extreme ranked set sample of m. The suggested ratio estimators of μ_Y using ERSS based on q_1 and q_3 of X,

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

respectively, are given by

$$\hat{\mu}_{Y_{\mathrm{ERSSE}1}} = \bar{Y}_{\mathrm{ERSS}E}\left(\frac{\mu_X + q_1}{\bar{X}_{\mathrm{ERSS}E} + q_1}\right) \tag{4.9}$$

and

$$\hat{\mu}_{Y_{\mathrm{ERSSE}3}} = \bar{Y}_{\mathrm{ERSS}E}\left(\frac{\mu_X + q_3}{\bar{X}_{\mathrm{ERSS}E} + q_3}\right) \tag{4.10}$$

where $\bar{X}_{\mathrm{ERSS}E} = (1/m)(\sum_{i=1}^{m/2} X_{i(1:m)} + \sum_{i=(m+2)/2}^{m} X_{i(m:m)})$ and $\bar{Y}_{\mathrm{ERSS}E} = (1/m)(\sum_{i=1}^{m/2} Y_{i[1:m]} + \sum_{i=(m+2)/2}^{m} Y_{i[m:m]})$. Then, $\hat{\mu}_{Y_{\mathrm{ERSS}Ek}}$ has the following properties.

THEOREM 3: *Under the first-degree approximation of the Taylor series method, the approximate MSE of* $\hat{\mu}_{Y_{ERSSOk}}$ *is given by*

$$\begin{aligned} MSE(\hat{\mu}_{Y_{ERSSEk}}) \approx \frac{1}{2m}[&(\sigma^2_{Y[1:m]} + \sigma^2_{Y[m:m]}) + L_k^2(\sigma^2_{X(1:m)} + \sigma^2_{X(m:m)}) \\ &- 2L_k(\sigma_{XY[1:m]} + \sigma_{XY[m:m]})]. \end{aligned} \tag{4.11}$$

Suppose that the regression of Y on X is linear (see [16]) and the marginal distribution of X is symmetric, then $\hat{\mu}_{Y_{\mathrm{ERSS}Ek}}$ has the following properties.

COROLLARY 3: *Under the first-degree approximation of the Taylor series method,* $\hat{\mu}_{Y_{ERSSEk}}$ *satisfies*

$$Bias(\hat{\mu}_{Y_{ERSSEk}}) \approx 0 \tag{4.12}$$

and

$$MSE(\hat{\mu}_{Y_{ERSSEk}}) \approx \frac{1}{m}[\sigma_Y^2(1-\rho^2) + \sigma^2_{(m:m)}(\rho\sigma_Y - L_k\sigma_X)^2]. \tag{4.13}$$

Since the proof procedures of Theorem 3 and Corollary 3 are, respectively, similar to that of Theorem 2 and Corollary 2, they are all omitted.

5. NUMERICAL COMPARISON

In this subsection, we will compare the relative efficiencies of the above estimators. The samples were generated from a bivariate normal distribution $N(\mu_X,\ \mu_Y, \sigma_X^2,\ \sigma_Y^2,\ \rho)$ with pdf

$$\begin{aligned} f(x,y) = &\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \\ &\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}. \end{aligned}$$

The efficiencies of $\hat{\mu}_{Y_{\mathrm{SRS}k}}$ with respect to (w.r.t.) $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ and $\hat{\mu}_{Y_{\mathrm{SRS}k}}$ with respect to (w.r.t.) $\hat{\mu}_{Y_{\mathrm{ERSSQ}k}}$ are, respectively, denoted as

$$\mathrm{eff}(\hat{\mu}_{Y_{\mathrm{RSS}k}},\ \hat{\mu}_{Y_{\mathrm{SRS}k}}) = \frac{\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{SRS}k}})}{\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{RSS}k}})} \tag{5.1}$$

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

TABLE 1. The efficiency of $\hat{\mu}_{Y_{\mathrm{RSS1}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS1}}}$ for $\rho < 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
−0.95	1.8848	2.2641	2.6644	3.0582	3.4263
−0.90	1.8141	2.2371	2.5816	2.9059	3.2421
−0.85	1.8157	2.1715	2.4981	2.7777	3.1015
−0.80	1.7594	2.1040	2.4199	2.6836	2.9569
−0.75	1.7423	2.0505	2.3312	2.6029	2.8432
−0.70	1.7352	2.0103	2.2701	2.5220	2.7408

TABLE 2. The efficiency of $\hat{\mu}_{Y_{\mathrm{RSS1}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS1}}}$ for $\rho > 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
0.95	1.2283	1.2949	1.3402	1.3724	1.4085
0.90	1.1856	1.2308	1.2623	1.2877	1.3052
0.85	1.1724	1.2152	1.2484	1.2689	1.2870
0.80	1.1779	1.2183	1.2470	1.2719	1.2907
0.75	1.1782	1.2232	1.2573	1.2796	1.2984
0.70	1.1878	1.2351	1.2664	1.2943	1.3148

TABLE 3. The efficiency of $\hat{\mu}_{Y_{\mathrm{RSS3}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS3}}}$ for $\rho < 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
−0.95	1.8445	2.2537	2.6061	3.0156	3.3499
−0.90	1.8204	2.1752	2.5125	2.8230	3.1250
−0.85	1.7815	2.0977	2.4044	2.6693	2.9507
−0.80	1.7419	2.0387	2.2983	2.5289	2.7625
−0.75	1.7132	1.9741	2.1911	2.4222	2.6157
−0.70	1.6446	1.9009	2.1263	2.3038	2.4799

TABLE 4. The efficiency of $\hat{\mu}_{Y_{\mathrm{RSS3}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS3}}}$ for $\rho > 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
0.95	1.0411	1.0514	1.0597	1.0599	1.0629
0.90	1.0047	1.0063	1.0079	1.0083	1.0090
0.85	1.0001	1.0020	1.0033	1.0045	1.0067
0.80	1.0041	1.0055	1.0055	1.0057	1.0058
0.75	1.0122	1.0145	1.0159	1.0177	1.0191
0.70	1.0224	1.0269	1.0289	1.0313	1.0339

and

$$\mathrm{eff}(\hat{\mu}_{Y_{\mathrm{ERSSQ}k}},\ \hat{\mu}_{Y_{\mathrm{SRS}k}}) = \frac{\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{SRS}k}})}{\mathrm{MSE}(\hat{\mu}_{Y_{\mathrm{ERSSQ}k}})}, \tag{5.2}$$

where $Q = O$ or E. In the simulation, we considered that $\mu_X = 2$, $\mu_Y = 4$, $\sigma_X^2 = \sigma_Y^2 = 1$ and different values of ρ. The results involving $q_1 = 1.32551$ and $q_3 = 2.67449$ are presented in Tables 1–8.

From Tables 1–4, we conclude the following:

(1) For different values of m, ρ and quartile, $\mathrm{eff}(\hat{\mu}_{Y_{\mathrm{RSS}k}},\ \hat{\mu}_{Y_{\mathrm{SRS}k}}) > 1$, which means $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ is more efficient $\hat{\mu}_{Y_{\mathrm{SRS}k}}$.

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

TABLE 5. The efficiency of $\hat{\mu}_{Y_{\mathrm{ERSSQ1}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS1}}}$ for $\rho < 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
−0.95	1.8641	1.9812	2.3305	2.3536	2.6293
−0.90	1.8266	1.9546	2.2514	2.2626	2.5555
−0.85	1.8172	1.9208	2.2340	2.1592	2.4757
−0.80	1.7604	1.8435	2.1185	2.1404	2.4028
−0.75	1.7215	1.8874	2.0859	2.1392	2.3179
−0.70	1.7420	1.8294	2.0637	2.0292	2.2454

TABLE 6. The efficiency of $\hat{\mu}_{Y_{\mathrm{ERSSQ1}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS1}}}$ for $\rho > 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
0.95	1.2311	1.2469	1.3024	1.3058	1.3430
0.90	1.1826	1.1952	1.2319	1.2401	1.2602
0.85	1.1703	1.1837	1.2200	1.2226	1.2424
0.80	1.1740	1.1905	1.2237	1.2185	1.2454
0.75	1.1802	1.1826	1.2332	1.2339	1.2507
0.70	1.1883	1.2001	1.2396	1.2400	1.2689

TABLE 7. The efficiency of $\hat{\mu}_{Y_{\mathrm{ERSSQ3}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS3}}}$ for $\rho < 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
−0.95	1.8536	1.9564	2.2992	2.3000	2.6028
−0.90	1.8006	1.9254	2.2089	2.2099	2.4336
−0.85	1.7423	1.8537	2.1519	2.1617	2.3677
−0.80	1.7110	1.8064	2.0767	2.0944	2.2828
−0.75	1.6672	1.7672	1.9908	1.9972	2.2160
−0.70	1.6588	1.7360	1.9475	1.9504	2.1163

TABLE 8. The efficiency of $\hat{\mu}_{Y_{\mathrm{ERSSQ3}}}$ w.r.t. $\hat{\mu}_{Y_{\mathrm{SRS3}}}$ for $\rho > 0$

ρ	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
0.95	1.0411	1.0431	1.0525	1.0536	1.0556
0.90	1.0047	1.0033	1.0052	1.0063	1.0074
0.85	1.0011	1.0015	1.0021	1.0032	1.0067
0.80	1.0041	1.0041	1.0044	1.0050	1.0058
0.75	1.0126	1.0129	1.0147	1.0149	1.0158
0.70	1.0218	1.0238	1.0265	1.0269	1.0283

(2) For fixed ρ, the efficiency of $\hat{\mu}_{Y_{\mathrm{RSS}k}}$ increases as m increases.
(3) The efficient for $\rho < 0$ is greater than it is for $\rho > 0$.
(4) In conclusion, the estimators of μ_Y in RSS are more efficient than that of μ_Y in SRS.

From Tables 5–8, we conclude the following:

(1) For different values of m, ρ and quartile, $\mathrm{eff}(\hat{\mu}_{Y_{\mathrm{ERSSQ}k}}, \hat{\mu}_{Y_{\mathrm{SRS}k}}) > 1$, which means $\hat{\mu}_{Y_{\mathrm{ERSSQ}k}}$ is more efficient $\hat{\mu}_{Y_{\mathrm{SRS}k}}$.

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

(2) For fixed ρ, the efficiency of $\hat{\mu}_{Y_{\mathrm{ERSSQ}k}}$ increases as m increases.

(3) The efficient for $\rho < 0$ is greater than it is for $\rho > 0$.

(4) In conclusion, the estimators of μ_Y in ERSS are more efficient than that of μ_Y in SRS.

(5) In conclusion, the estimators of μ_Y in RSS and ERSS are more efficient than that of μ_Y in SRS.

6. CONCLUSIONS

In this article, we investigate ratio-type estimators of the population mean of the study variable, involving either the first or the third quartile of the auxiliary variable, using RSS and ERSS schemes. The properties of the estimators are obtained. The estimators in RSS and ERSS are compared to their counterparts in SRS for normal data. The numerical results show that the estimators in RSS and ERSS are significantly more efficient than their counterparts in SRS. The numerical results also show that the estimators in RSS are significantly more efficient than their counterparts in ERSS. However, the procedure of ERSS is easier to rank than the RSS procedure.

Acknowledgments

The authors thank the Editor in Chief, an associate editor and reviewers for their valuable comments and suggestions to improve the paper. This research was supported by the National Science Foundation of China (Grant No.11901236), the Scientific Research Fund of Hunan Provincial Science and the Technology Department (Grant No. 2019JJ50479) and the Scientific Research Fund of Hunan Provincial Education Department (Grant No.18B322) and Young Core Teacher Foundation of Hunan Province (Grant No. [2020]43).

References

1. Al-Omari, A.I. (2012). Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling. *Statistics and Probability Letters* 82(11): 1883–1890.
2. Chen, Z.H., Bai, Z.D., & Sinha, B.K. (2003). *Ranked set sampling: theorem and applications*. New York: Springer.
3. Chen, W.X., Xie, M.Y., & Wu, M. (2013). Parametric estimation for the scale parameter for scale distributions using moving extremes ranked set sampling. *Statistics and Probability Letters* 83(9): 2060–2066.
4. Chen, W.X., Xie, M.Y., & Wu, M. (2016). Modified maximum likelihood estimator of scale parameter using moving extremes ranked set sampling. *Communications in Statistics – Simulation and Computation* 45(6): 2232–2240.
5. Chen, W.X., Tian, Y., & Xie, M.Y. (2018). The global minimum variance unbiased estimator of the parameter for a truncated parameter family under the optimal ranked set sampling. *Journal of Statistical Computation and Simulation* 88(17): 3399–3414.
6. Chen, W.X., Yang, R., Yao, D.S., & Long, C.X. (2019). Pareto parameters estimation using moving extremes ranked set sampling. Statistical Papers. https://doi.org/10.1007/s00362-019-01132-9.
7. Cochran, W.G. (1997). *Sampling technique*, 3rd ed. New York: Wiley and Sons.
8. Kadilar, C. & Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation* 151(3): 893–902.
9. McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research* 3(4): 385–390.
10. Qian, W.S., Chen, W.X., & He, X.F. (2019). Parameter estimation for the Pareto distribution based on ranked set sampling. *Statistical Papers*. doi:10.1007/s00362-019-01102-1.
11. Qiu, G.X & Eftekharian, A. (2020). Extropy information of maximum and minimum ranked set sampling with unequal samples. *Communications in Statistics – Theory and Methods*. doi:10.1080/03610926.2019.1678640.

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press

12. Samawi, H.M. & Muttlak, H.A. (1996). Estimation of ratio using rank set sampling. *The Biometrical Journal* 38(6): 753–764.
13. Samawi, H.M., Ahmed, M.S., & Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Journal* 38(5): 577–586.
14. Singh, H.P. & Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition* 6:555–560.
15. Srivastava, K. & Jhajj, H.S. (1981). A class on estimators of the population mean in survey sampling using auxiliary information. *Biometrika* 68(1): 341–343.
16. Stokes, L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics – Theory and Methods* 6(12): 1207–1211.
17. Stokes, L. (1995). Parametric ranked set sampling. *Annals of the Institute of Statistical Mathematics* 47(3): 465–482.
18. Takahasi, K. & 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): 1–31.
19. Upadhyaya, L.N. & Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal* 41(5): 627–636.
20. Yang, R., Chen, W.X., Yao, D.S., Long, C., Dong, Y., & Shen, B. (2020). The efficiency of ranked set sampling design for parameter estimation for the log-extended exponential-geometric distribution. *Iranian Journal of Science and Technology, Transactions A: Science* 44(2): 497–507.

https://doi.org/10.1017/S0269964820000625 Published online by Cambridge University Press