RATIO AND PRODUCT TYPE EXPONENTIAL ESTIMATORS FOR POPULATION MEAN USING RANKED SET SAMPLING


*Department of Applied Mathematics, Indian Institute of Technology (ISM) Dhanbad Jharkhand-826004, INDIA.
**Department of Mathematics and Computation, University of Havana, Havana, CUBA.

ABSTRACT
In this paper, we suggest an improved form of exponential ratio and product estimators using ranked set sampling (RSS) in order to estimate the population mean \( \bar{Y} \) of study variate \( Y \) using auxiliary variate \( X \). The mean-squared error (MSE) of the proposed class of estimators is obtained. To judge the merits of the suggested estimators over others a simulation study has been carried out.

KEYWORDS: Ranked set sampling, Auxiliary variate, Study variate, MSE.

MSC: 62D05

RESUMEN
En este trabajo sugerimos la forma mejorada exponencial de estimadores de razón y producto usando muestreo por conjuntos ordenados (RSS) para estimar la media poblacional \( \bar{Y} \) de una variable de estudio \( Y \) usando la variable auxiliar \( X \). El error cuadrático medio (MSE) de la clase de estimadores propuesta es obtenido. Para evaluar los méritos de los estimadores sugeridos respecto a otros se desarrolló un estudio de simulación.

1. INTRODUCTION

The ranked set sampling (RSS) is a method of sampling which provides more structure to the collected sample items and increases the amount of information present in the sample. The method of ranked set sampling (RSS) was first envisaged by McIntyre (1952) as a cost-efficient substitute to simple random sampling (SRS) for those circumstances where measurements are inconvenient or expensive to obtain but (judgment) ranking of units according to the variable of interests, say, \( Y \), is comparatively easy and cheap. It is known that the estimate of the population mean using RSS is more efficient than the one obtained using SRS. McIntyre (1952) and Takahasi and Wakimoto (1968) considered perfect ranking of the elements, that is, there are no errors in ranking the elements. Yet, in most circumstances, the ranking may not be done perfectly. Dell and Clutter (1972) demonstrated that the mean using the RSS is an unbiased estimator of the population mean, whether or not there are errors in ranking. Stokes (1977) considered the case where the ranking is done on the basis of a concomitant (auxiliary) variable \( X \) instead of judgment. We would expect the variable of interest will be highly correlated with the concomitant (auxiliary) variable. Stokes (1980) showed that the estimator of the variance based on RSS data is an asymptotically unbiased estimator of the population variance. Samawi and Muttlak (1996) deal the problem of estimating the population ratio of the two variables \( Y \) and \( X \) using RSS procedure. In addition, RSS has been investigated by many researchers like Al-Saleh and Al-Omary (2002), Wolfe(2004), Mandowara and Mehta (2013) Al-Omary and Bouza (2015).

The RSS has many statistical applications in agriculture, biology, environmental science, medical science etc. Let \( m \) random samples of size \( m \) bivariate units each and rank the bivariate units within each sample with respect to the auxiliary variate \( X \). Next, select the \( l^t \) smallest auxiliary variate \( X \) from the \( l^t \) sample for \( t = 1,2,3, ... m \) for actual measurement of the associated variate of interest \( Y \) with it. In this way, a total number of \( m \) measured bivariate units are obtained, one from each sample. The cycle may be repeated \( r \) times to get a sample of size \( n = rm \) bivariate units. These \( n = rm \) units built the RSS data. Note that we assume that the ranking of the variate \( X \) will be perfect while the ranking of the variate \( Y \) will be with errors, or at worst of a random order if the correlation between \( Y \) and \( X \) is close to zero. Also, note that in RSS, \( rm \) elements are identified, but only \( rm \) of them are quantified. So, comparing this sample with a

\*vishwak@googlemail.com, zeeshan008x52@gmail.com
simple random sample of size rm is reasonable. For more details about RSS, see Kaur, et al (1995).

We assume that ranking on the auxiliary variate, X, is perfect. The associated variate, Y, is then with an error unless the relation between X and Y is perfect. Let us denote \((X_i, Y_i)\) as the pair of the \(i^{th}\) order statistics of X and the associated element Y in the \(i^{th}\) cycle. Then the ranked set sample is

\[
(X_1(1), Y_1(1)) \ldots (X_1(m), Y_1(m)), (X_2(1), Y_2(1)) \ldots (X_2(m), Y_2(m)) \ldots (X_r(1), Y_r(1)) \ldots (X_r(m), Y_r(m)).
\]

To obtain Biases and Mean squared error, we consider

\[
T_y(i) = (\mu_y(i) - \mu_y), T_x(i) = (\mu_x(i) - \mu_x), T_{xy}(i) = (\mu_x(i) - \mu_x)(\mu_y(i) - \mu_y),
\]

\[
\sigma_{y(i)}^2 = E(Y(i) - \mu_y)^2, \sigma_{xy}^2 = E(X(i) - \mu_x)^2,
\]

and

\[
\sum_{i=1}^{n} T_x(i) = 0, \sum_{i=1}^{n} T_y(i) = 0,
\]

\[
\sum_{i=1}^{n} \sigma_{y(i)}^2 = n\sigma_y^2 - \sum_{i=1}^{n} T_{y(i)} \cdot \sum_{i=1}^{n} \sigma_{y(i)}^2 = n\sigma_y^2 - \sum_{i=1}^{n} T_{y(i)}^2,
\]

\[
\sum_{i=1}^{n} \sigma_{x(i)} = n\sigma_x - \sum_{i=1}^{n} T_{x(i)}.
\]

The sample mean of each variate based on RSS data and using the results obtained in Dell and Clutter (1972) can be defined as follows:

\[
\bar{X}(n) = \frac{1}{mr} \sum_{j=1}^{m} \sum_{i=1}^{n} X_{r(m)},
\]

\[
\bar{Y}[n] = \frac{1}{mr} \sum_{j=1}^{m} \sum_{i=1}^{n} Y_{r[m]},
\]

with variance

\[
Var(\bar{X}) = \frac{\sigma_x^2}{m} - \frac{1}{rm^2} \sum_{i=1}^{n} T_{x(i)}^2
\]

and co-variance

\[
Cov(\bar{X}, \bar{Y}) = \frac{\sigma_{xy}}{m} - \frac{1}{rm^2} \sum_{i=1}^{n} T_{xy(i)}
\]

Note that \(\bar{X}(i)\) and \(\bar{Y}(i)\) depend on order statistics from some specific distributions and these values can be found in Arnold et al (1992).

### 2. SOME EXISTING ESTIMATORS FOR THE POPULATION MEAN

For estimating the population mean \(\bar{Y}\), the usual ratio and product estimators for \(\bar{Y}\), respectively as

\[
\hat{\bar{Y}}_R = \bar{Y} \frac{\bar{X}}{\bar{X}},
\]

and their MSEs up to first degree of approximation are

\[
MSE(\hat{\bar{Y}}_R) = \frac{\sigma_y^2}{n} \left[\left(C_x^2 + C_y^2 - 2\rho C_x C_y\right)\right].
\]

\[
MSE(\hat{\bar{Y}}_P) = \frac{\sigma_y^2}{n} \left[\left(C_x^2 + C_y^2 + 2\rho C_x C_y\right)\right].
\]

Samawi and Muttlak (1996) approached ratio and product estimators under RSS as

\[
\hat{\bar{Y}}_{R_{rss}} = \bar{Y}_{[n]} \frac{\bar{X}_{[n]}}{\bar{X}},
\]

\[
\hat{\bar{Y}}_{P_{rss}} = \bar{Y}_{[n]} \frac{\bar{X}_{[n]}}{\bar{X}},
\]

and derived their MSEs to the first degree approximation as
\[
MSE(\hat{Y}_{r}) = \frac{\gamma^2}{m} \left( C_2^2 + C_γ^2 - 2 \rho C_x C_y \right) - \frac{1}{m} \left( \frac{\sum_{i=1}^{m} T_{x[i]}^2}{\mu_x^2} + \frac{\sum_{i=1}^{m} T_{y[i]}^2}{\mu_y^2} \right)
\]  
\[
MSE(\hat{Y}_{p}) = \frac{\gamma^2}{m} \left( C_2^2 + C_γ^2 + 2 \rho C_x C_y \right) - \frac{1}{m} \left( \frac{\sum_{i=1}^{m} T_{x[i]}^2}{\mu_x^2} + \frac{\sum_{i=1}^{m} T_{y[i]}^2}{\mu_y^2} \right)
\]  

For estimating the population mean \( \bar{Y} \), Bahl and Tuteja (1991) given the ratio and product type exponential estimators as

\[
\hat{Y}_{re} = \bar{y} \exp \left( \frac{x - \bar{x}}{x + \bar{x}} \right),
\]

\[
\hat{Y}_{pe} = \bar{y} \exp \left( \frac{x - \bar{x}}{x - \bar{y}} \right),
\]

and derived their MSEs to the first degree approximation as

\[
MSE(\hat{Y}_{re}) = \frac{\gamma^2}{n} \left( \frac{C_2^2 + C_γ^2 - \rho C_x C_y}{\bar{x} + \bar{y}} \right),
\]

\[
MSE(\hat{Y}_{pe}) = \frac{\gamma^2}{n} \left( \frac{C_2^2 + C_γ^2 + \rho C_x C_y}{\bar{x} + \bar{y}} \right),
\]

3. PROPOSED ESTIMATORS FOR POPULATION MEAN

We define the following ratio and product type exponential estimators for \( \bar{Y} \) under RSS, respectively as

\[
\hat{Y}_{r}^{rss} = \bar{y}_{[n]} \exp \left( \frac{x - \bar{x}_{[n]}}{x + \bar{x}_{[n]}} \right),
\]

\[
\hat{Y}_{p}^{rss} = \bar{y}_{[n]} \exp \left( \frac{x - \bar{x}_{[n]}}{x - \bar{y}_{[n]}} \right),
\]

here we have ranked auxiliary variate and thus, there is an induced rank in study variate. the induced rank on the study variate will be perfect if the correlation between the variate is perfect otherwise it will be worst if there will be no correlation.(the worst case will not affect our problem since it has already been proven by Dell and Clutter (1972). Therefore, the MSE of \( \hat{Y}_{r}^{rss} \) and \( \hat{Y}_{p}^{rss} \) using bi-variate Taylor series expansion given as

\[
MSE(\hat{Y}_{r}^{rss}) = \frac{\gamma^2}{m} \left( \frac{C_2^2}{4} + C_γ^2 - \rho C_x C_y \right) - \frac{1}{m} \left( \frac{\sum_{i=1}^{m} T_{x[i]}^2}{4 \bar{x}^2} + \frac{\sum_{i=1}^{m} T_{y[i]}^2}{\bar{y}^2} \right)
\]

\[
MSE(\hat{Y}_{p}^{rss}) = \frac{\gamma^2}{m} \left( \frac{C_2^2}{4} + C_γ^2 + \rho C_x C_y \right) - \frac{1}{m} \left( \frac{\sum_{i=1}^{m} T_{x[i]}^2}{4 \bar{x}^2} + \frac{\sum_{i=1}^{m} T_{y[i]}^2}{\bar{y}^2} \right)
\]

Proposition: Let \( W_{x[i]} = \frac{x(i) - \bar{x}}{\mu_x} \) and \( W_{y[i]} = \frac{y(i) - \bar{y}}{\mu_y} \) and also using the result from Dell and Clutter (1972) the above equation may be written as

\[
MSE(\hat{Y}_{r}^{rss}) = \frac{\gamma^2}{m} \left( \frac{C_2^2}{4} + C_γ^2 - \rho C_x C_y \right) - \frac{1}{m} \left( \sum_{i=1}^{m} W_{x[i]}^2 \frac{4}{4} + \sum_{i=1}^{m} W_{y[i]}^2 - 2 \sum_{i=1}^{m} \frac{W_{x[i]} W_{y[i]}}{2} \right)
\]

\[
= \frac{\gamma^2}{m} \left( \frac{C_2^2}{4} + C_γ^2 - \rho C_x C_y \right) - \frac{1}{m} \left( \sum_{i=1}^{m} \frac{W_{x[i]}^2}{2} - W_{y[i]}^2 \right)
\]

It is clear that \( \sum_{i=1}^{m} \left( \frac{W_{x[i]} - W_{y[i]}}{2} \right) \) is greater than zero. Hence

\[
MSE(\hat{Y}_{r}^{rss}) \leq MSE(\hat{Y}_{r}).
\]

Also, it can be proved in similar ways that

\[
MSE(\hat{Y}_{p}^{rss}) \leq MSE(\hat{Y}_{p}).
\]

3.1 Generalized Exponential Estimators Using RSS

We propose a ratio-cum-product type exponential estimators using RSS as
\[ \hat{\varphi}_{G}^{\text{rss}} = \bar{y} \left[ \frac{1}{\tau_{(m)}} \left( 1 + \left( \frac{X}{\tau_{(m)}} \right)^{2} \right) \right]^{-1}, \]

where \( \alpha \) is some suitable real number whose values make the minimum MSE of \( \hat{\varphi}_{G}^{\text{rss}} \). It can also be noticed that for \( \alpha = 1 \) and \( \alpha = -1 \) the above equation becomes Bahl and Tuteja (1991) usual ratio and product exponential estimators respectively.

Again using Taylor series expansion we get the MSE of \( \hat{\varphi}_{G}^{\text{rss}} \) as

\[ \text{MSE}(\hat{\varphi}_{G}^{\text{rss}}) = \frac{\sigma_{\hat{\varphi}_{G}^{\text{rss}}}}{\sigma_{\varphi}} \left( \frac{1}{m^2} \left[ \sum_{i=1}^{m} \frac{x_{i}^{2} + y_{i}^{2}}{m^2} \right] + \frac{\rho \sigma_{x} \sigma_{y}}{m^2} \right)^{2} \]

In order to get the minimum MSE we differentiate the above equation (25) by \( \alpha \) and equate it with 0. Hence we get optimum value of \( \alpha \) as

\[ \alpha_{\text{opt}} = 2 \left( \frac{\rho \sigma_{x} \sigma_{y}}{m \sigma_{Y}^{2}} \right) \left( \frac{1}{m^2} \sum_{i=1}^{m} \frac{x_{i}^{2} + y_{i}^{2}}{m^2} \right) \]

Using the above result we get the minimum MSE of \( \hat{\varphi}_{G}^{\text{rss}} \) as

\[ \text{MSE}(\hat{\varphi}_{G}^{\text{rss}})_{\text{min}} = \frac{\sigma_{\hat{\varphi}_{G}^{\text{rss}}}}{\sigma_{\varphi}} \left( \frac{1}{m^2} \sum_{i=1}^{m} \frac{x_{i}^{2} + y_{i}^{2}}{m^2} \right) + \frac{\rho \sigma_{x} \sigma_{y}}{m^2} \left( \frac{1}{m^2} \sum_{i=1}^{m} \frac{x_{i}^{2} + y_{i}^{2}}{m^2} \right) \]

### 4. A SIMULATION STUDY

To illustrate how one can gain the insight in the application or the properties of the proposed estimator a computer simulation was conducted. Bivariate random observations were generated from a bivariate normal distribution with parameters \( \mu_{x}, \mu_{y}, \sigma_{x}, \sigma_{y} \) and correlation coefficient \( \rho \). The sampling method explained above is used to pick an RSS data with sets of size \( m \) and \( r \) repeated cycles to get a RSS of size \( mr \).

A sample of size \( mr \) bivariate units is randomly chosen from the population (we refer to these data as SRS data). The simulation was performed with \( m = 3, 4, 5 \) and with \( r = 3 \) and 6 (i.e., with total sample sizes of 9, 12, 15, 18, 24 and 30) for the RSS and SRS data sets. Here, we have ranked the auxiliary variate \( X \) which induces ranking in study variate \( Y \) (ranking on \( Y \) will be perfect if \( \rho = 1 \) or will be with errors in ranking if \( \rho < 1 \)). Using R software we have conducted 5,000 replications for estimates of the means and mean square errors. Results of these simulations are summarized by the percentage relative efficiencies of the estimators using the formula.

\[ \text{PRE} \left[ *, \hat{\varphi}_{G}^{\text{rss}} \right] = \frac{\text{MSE}(\hat{\varphi}_{G}^{\text{rss}})}{\text{MSE}(\varphi)} \times 100 \]

where, \(* = \hat{\varphi}_{G}^{\text{rss}}, \hat{\varphi}_{G}^{\text{rss}} \).

<table>
<thead>
<tr>
<th>( r )</th>
<th>( m )</th>
<th>( \rho = 0.5 )</th>
<th>( \rho = 0.6 )</th>
<th>( \rho = 0.7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( \hat{\varphi}_{G}^{\text{rss}} )</td>
<td>( \hat{\varphi}_{G}^{\text{rss}} )</td>
<td>( \hat{\varphi}_{G}^{\text{rss}} )</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>100.00</td>
<td>180.96</td>
<td>197.60</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>100.00</td>
<td>195.96</td>
<td>197.16</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>100.00</td>
<td>88.08</td>
<td>102.37</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>100.00</td>
<td>137.81</td>
<td>148.54</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>100.00</td>
<td>227.19</td>
<td>228.51</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>100.00</td>
<td>459.04</td>
<td>2552.20</td>
</tr>
</tbody>
</table>

Table I: Percentage Relative Efficiencies (PREs) of different estimators of \( \bar{Y} \) with respect to \( \bar{Y}_{R} \).
Finally from Tables 1 and 2 we can conclude that the proposed estimators $\hat{y}_{rss}^{R}$, $\hat{y}_{rss}^{P}$ and $\hat{y}_{rss}^{G}$ are more appropriate estimators than the existing popular estimators $\hat{y}_{r}^{R}$, $\hat{y}_{r}^{P}$, $\hat{y}_{r}^{G}$ and $\hat{y}_{P}^{rss}$ has appreciable efficiency.

Acknowledgement: The authors are thankful to the editor and the learned referees for their valuable comments and suggestions towards the improvement of the paper.

RECEIVED JUNE, 2016
REVISED SEPTEMBER, 2016

270
REFERENCES


