On Ratio Estimation Using Extreme Ranked Set Samples

Hani M. Samawi

Received on March 20, 2001
Accepted for publication on May 28, 2002

Abstract

In many situations, the quantity estimated from a random sample is the ratio of two variables both of which vary from unit to unit. Also, ratio estimation is used to obtain increase in precision of estimating the population mean or total by taking advantage of the correlation between an auxiliary variable X and the variable of interest Y. In this paper two types of extreme ranked set sample (namely, ERSS₁ in case of even sample size and ERSS₂ in case of odd sample size, see Samawi et al., 1996) are used to estimate the ratio of the means of two populations. We show that using ERSS samples give approximately unbiased estimators of the population ratio, in case of symmetric populations, and it is more efficient than using SRS (simple random sample) for ratio estimation, using the same number of quantified units. Some inference using asymptotic results are given. Simulation is conducted to compare the efficiency of the estimators.

Key Words and Phrases: Simple random sample, ranked set sample, extreme ranked set sample, mean square error, ratio estimator.

1. Introduction

Ratio appears in many applications of the applied fields. For example, in a household survey, the average number of suits of clothes per adult male may be of interest. Example of this kind occur when the sampling unit (the household) comprises a group or cluster of elements (adult males) and our interest is in the population mean per element (see Cochran, 1977.) Also, it appears in other applications, for example, the ratio of loans for building purpose to total loans in a bank or the ratio of acres of wheat to total acres on a farm. Moreover, ratio estimators are used to obtain increased precision of estimating the population mean or total.

© 2002 by Yarmouk University, Irbid, Jordan.
* Department of Mathematics & Statistics, Qaboos University, Sultanate of Oman.
Samawi

Ranked set sampling procedure (RSS), as introduced by McIntyre (1952), involves randomly drawing \( n \) sets of \( n \) units each from an infinite population for which the unknown parameter is to be estimated. It is assumed that the units in each set can be ranked visually at no cost or very little cost. From the first set of \( n \) units the unit ranked lowest is measured. From the second set of \( n \) units the unit ranked second lowest is measured. The process is continued until the \( n-th \) ranked unit is measured from the \( n-th \) set of \( n \) units and this completes the cycle. Since in practice the number of units which are easily ranked cannot be more than four, when using RSS, this cycle could be repeated \( m \) times to obtain a larger sample size as needed (see Takahasi and Wakimoto, 1968.) This is because ranking, for example visually, more than four unit each time will introduce ranking error and hence this will reduce the advantages of using RSS.

Samawi and Muttilak (1996) improved the precision of ratio estimation by using RSS. However, Samawi et al. (1996) investigated a variety of extreme ranked set samples (ERSS) which have a potential of reducing the ranking errors since they do not need a complete ranking. Also, ERSS samples give more efficient estimators for the population mean than SRS in case of symmetric populations.

The ERSS procedure involves randomly drawing \( n \) sets of \( n \) units each from the infinite population for which the unknown parameter is to be estimated. We assume that the lowest or the largest units of this set can be detected visually with no cost or with a very little cost. From the first set of \( n \) units the lowest ranked unit is measured. From the second set of \( n \) units the largest ranked unit is measured. From the third set of \( n \) units the lowest ranked unit is measured, and so on. In this way we obtain the first \( n-1 \) measured units using the first \( n-1 \) sets. The choice of the \( n-th \) unit from the \( n-th \) (i.e., the last) set depends on whether \( n \) is even or odd.

(a) If \( n \) is even then the largest ranked unit is measured. Such a sample will be denoted by \( \text{ERSS}_a \). If \( n \) is odd then we have two options.

(b) For the measure of the \( n-th \) unit we take the average of the measures of the lowest and the largest units in the \( n-th \) set. Such a sample will be denoted by \( \text{ERSS}_b \).

(c) For the measure of the \( n-th \) unit we take the measure of the median. Such a sample will be denoted by \( \text{ERSS}_c \).

As indicated by Samawi et al. (1996), the choice (c) will be more difficult in applications than the choice (a) or (b). To obtain SRS we measure the first unit from each set. Also, this cycle can be repeated \( m \) times if needed.
On Ratio Estimation Using Extreme Ranked Set Samples

In this paper, we compare the efficiency of the estimates of the population ratio obtained by using SRS, RSS, ERSS_a and ERSS_e when the distribution of the population is assumed to be symmetric.

2. Samples And Estimating Population Ratio

For the \( k \)-th cycle, let

\[
(X_{1k}^*, Y_{1k}^*), (X_{2k}^*, Y_{2k}^*), ..., (X_{nk}^*, Y_{nk}^*), (X_{2k}, Y_{2k}), ..., (X_{2nk}, Y_{2nk}), ..., (X_{nk}, Y_{nk}),
\]

\[
(X_{n2k}, Y_{n2k}), ..., (X_{nmk}, Y_{nmk}),
\]

\( k = 1, 2, ..., m \), be \( n \) independent bivariate random samples each with sample size \( n \) and assume

that each element \((X_{jk}, Y_{jk})\), in the sample has the same bivariate distribution function \( F(x, y) \) with means \( \mu_x \) and \( \mu_y \), variances \( \sigma_x^2 \) and \( \sigma_y^2 \) and correlation coefficient \( \rho \). Assume that the ranking of the variable \( Y \) is perfect, while the ranking of the associated variable \( X \) will be with errors and will be at its worst if the correlation between \( X \) and \( Y \) is close to zero. For simplicity of notation we will assume that \((X_{jk}, Y_{jk})\), denotes the measure of the \((X_{jk}^*, Y_{jk}^*)\). Then according to our description \((X_{11k}, Y_{11k}), (X_{21k}, Y_{21k}), ..., (X_{nk}, Y_{nk})\) is the SRS. Then \((X_{1(1)k}, Y_{1(1)k}), (X_{2(2)k}, Y_{2(2)k}), ..., (X_{nk(n)k}, Y_{nk(n)k})\) denotes the RSS. If \( n \) is even then \((X_{1(1)k}, Y_{1(1)k}), (X_{2(2)k}, Y_{2(2)k}), ..., (X_{n-1[n-1]k}, Y_{n-1[n-1]k}), (X_{n[n]k}, Y_{n[n]k})\) denotes \( ERSS_a \). If \( n \) is odd then \((X_{1(1)k}, Y_{1(1)k}), (X_{2[n]k}, Y_{2[n]k}), ..., (X_{n-1[n-1]k}, Y_{n-1[n-1]k}), (X_{n(n+1)/2k}, Y_{n(n+1)/2k})\)

denotes \( ERSS_e \).

2.1 Definitions, notations and some related results

The following notation and results will be used throughout this paper.

For all \( i, i = 1, 2, ..., n \), let \( \mu_x = E(X_{ni}), \mu_y = E(Y_{ni}) \),

\[
\sigma_x^2 = Var(X_{ni}), \sigma_y^2 = Var(Y_{ni})
\]

\[
\mu_x(1) = E(X_{i(1)}), \mu_y(1) = E(Y_{i(1)}), \mu_x(n+1/2) = E(X_{i(n+1/2)}),
\]

\[
\mu_y(n+1/2) = E(Y_{i(n+1/2)}), \mu_x(n) = E(X_{i(n)}),
\]

817
Samawi

\[ \mu_Y(n) = \mathbb{E}(Y_{i(n)}) \], \sigma_{x(i)}^2 = \text{Var}(X_{i(n)}), \sigma_{y(i)}^2 = \text{Var}(Y_{i(n)}),
\]
\[ \sigma_x^2 = \text{Var}(X_{i(n)}), \sigma_y^2 = \text{Var}(Y_{i(n)}), \sigma_{x(y)}^{\frac{n+1}{2(i)}} = \text{Var}(X_{i(n)}^{\frac{n+1}{2}}), \]
\[ \sigma_{y(i)}^{\frac{n+1}{2}} = \text{Var}(Y_{i(n)}^{\frac{n+1}{2}}) \]

and \( \sigma_{x(y)}^{\frac{n+1}{2}} = \text{Var}(X_{i(n)}^{\frac{n+1}{2}}) \).

The population ratio of the two variables \( Y \) and \( X \) is defined by
\[ R = \frac{\mu_Y}{\mu_X}, \quad (2.1.1) \]

where \( \mu_x \) and \( \mu_y \) are the population means of the variables \( X \) and \( Y \) respectively. The ratio estimator using SRS can be defined by
\[ r_{SRS} = \frac{\overline{Y}}{\overline{X}}, \quad (2.1.2) \]

where \( \overline{X} = \frac{1}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n} X_{i(l)} \) and \( \overline{Y} = \frac{1}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n} Y_{i(l)} \) are the respective sample means. Hansen et al. (1953) showed that the variance of \( r_{SRS} \) can be approximated by
\[ \text{Var}(r_{SRS}) \approx \frac{R^2}{nm} \left( \frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - 2\rho \frac{\sigma_x}{\mu_x} \frac{\sigma_y}{\mu_y} \right), \quad (2.1.3) \]

where \( \rho = \frac{\text{Cov}(X_{i(l)}, Y_{i(l)})}{\sigma_x \sigma_y} \), \( \sigma_x \) and \( \mu_x \) and \( \sigma_y \) and \( \mu_y \) are the means and the standard deviations of the population of the variables \( X \) and \( Y \) respectively.

2.2 Ranking on the variable \( Y \)

As in Samawi and Muttalik (1996), assume that we can only rank the variable \( Y \) and that the ranking of \( Y \) is perfect while the ranking of \( X \) will be with error. Let \((X_{i(1)}^k, Y_{i(1)}^k), (X_{i(2)}^k, Y_{i(2)}^k), \ldots, (X_{i(n-1)}^k, Y_{i(n-1)}^k), (X_{i(n)}^k, Y_{i(n)}^k)\) \( \text{denotes } \) ERSS1_k, \( k = 1, 2, \ldots, m \).

The ratio estimate using ERSS1_k sample is defined by
On Ratio Estimation Using Extreme Ranked Set Samples

\[ r_{ERSS1_c} = \frac{\tilde{Y}_{(c)}}{\tilde{X}_{(c)}}. \]  

(2.2.1)

where

\[ \tilde{X}_{(c)} = \frac{1}{2} (\tilde{X}_{(1)} + \tilde{X}_{(n)}) \], \[ \tilde{X}_{(1)} = \frac{2}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n/2} X_{2l-1(l)k} \] and \[ \tilde{X}_{(n)} = \frac{2}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n/2} X_{2l(n)k} \].

\[ \tilde{Y}_{(c)} = \frac{1}{2} (\tilde{Y}_{(1)} + \tilde{Y}_{(n)}) \], \[ \tilde{Y}_{(1)} = \frac{2}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n/2} Y_{2l-1(l)k} \] and \[ \tilde{Y}_{(n)} = \frac{2}{nm} \sum_{k=1}^{m} \sum_{l=1}^{n/2} Y_{2l(n)k} \].

Replacing the notation (.) by [ ] when ranking with error, and using the bivariate Taylor series expansion (see, Bickel and Doksum, 1977) and the assumption of the symmetry of the distribution, the approximate variance of \( r_{ERSS1_c} \) can be written as

\[ Var(r_{ERSS1_c}) \approx \frac{R^2}{nm} \left( \frac{\sigma_{X(1)}^2}{\mu_x^2} + \frac{\sigma_{Y(1)}^2}{\mu_y^2} - 2 \frac{\sigma_{X(1)Y(1)}}{\mu_x \mu_y} \right). \]  

(2.2.2)

If \( n \) is odd then

\[ (X_{1(1)k}, Y_{1(1)k}), (X_{2(n)k}, Y_{2(n)k}), \ldots, (X_{n-l(n)k}, Y_{n-l(n)k}), (X_{n-l(n)k}, Y_{n-l(n)k}), \]

\[ k = 1, 2, \ldots, m \] denotes ERSS1_e. The ratio estimate using ERSS1_e data with errors in ranking for the variable \( X \) is defined by

\[ r_{ERSS1_e} = \frac{\tilde{Y}_{(c)}}{\tilde{X}_{(c)}}. \]  

(2.2.3)

where

\[ \tilde{X}_{(c)} = \frac{\sum_{k=1}^{m} (X_{1(1)k} + X_{2(n)k} + \ldots + X_{n-l(n)k} + X_{n-l(n)k})}{nm} \]

and

\[ \tilde{Y}_{(c)} = \frac{\sum_{k=1}^{m} (Y_{1(1)k} + Y_{2(n)k} + \ldots + Y_{n-l(n)k} + Y_{n-l(n)k})}{nm} \].

819
Similarly, replacing the notation (.) by [.] when ranking with error, and using the assumption of the symmetry of the distribution, the approximate variance of \( r_{\text{ERSS}_c} \) can be written as

\[
\text{Var}(r_{\text{ERSS}_c}) \approx \frac{R^2}{nm} \left( \frac{(n-1)\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} + \frac{(n-1)\sigma_z^2}{\mu_x^2 \mu_y^2} \right)
\]

(2.2.4)

2.3 Ranking on the variable X

Now assume that we can rank the variable \( X \) only so that the ranking of \( X \) will be perfect while the ranking of \( Y \) will be with error. Let 

\[
(X_{1(1)}k, Y_{1(1)}k, X_{2(n)}k, Y_{2(n)}k, \ldots, X_{n-1(1)}k, Y_{n-1(1)}k, X_{n(n)}k, Y_{n(n)}k, k=1, 2, \ldots, m)
\]

denote ERSS2. The ratio estimate using ERSS2 data sample is defined by

\[
r_{\text{ERSS}_2} = \frac{\bar{Y}_{[a]}}{\bar{X}_{[a]}}
\]

(2.3.1)

where

\[
\bar{X}_{[a]} = \frac{1}{2} (\bar{X}_{(1)} + \bar{X}_{(n)}) \quad \bar{X}_{(1)} = \frac{2}{nm} \sum_{l=1}^{m} \sum_{i=1}^{n/2} X_{2l-1(1)k} \quad \bar{X}_{(n)} = \frac{2}{nm} \sum_{l=1}^{m} \sum_{i=1}^{n/2} X_{2l(n)k}
\]

\[
\bar{Y}_{[a]} = \frac{1}{2} (\bar{Y}_{[1]} + \bar{Y}_{[n]}) \quad \bar{Y}_{[1]} = \frac{2}{nm} \sum_{l=1}^{m} \sum_{i=1}^{n/2} Y_{2l-1(1)k} \quad \bar{Y}_{[n]} = \frac{2}{nm} \sum_{l=1}^{m} \sum_{i=1}^{n/2} Y_{2l(n)k}
\]

Again, replacing the notation (.) by [.] when ranking with error, and using the bivariate Taylor series expansion, the approximate variance of can be written as

\[
\text{Var}(r_{\text{ERSS}_2}) \approx \frac{R^2}{nm} \left( \frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - 2 \frac{\sigma_z^2}{\mu_x \mu_y} \right)
\]

(2.3.2)

If \( n \) is odd then

\[
(X_{1(n)k}, Y_{1(n)k}, X_{2(n)k}, Y_{2(n)k}, \ldots, X_{n-1(n)k}, Y_{n-1(n)k}, X_{n(n)k}, Y_{n(n)k}, k=1, 2, \ldots, m)
\]

denotes ERSS2. The ratio estimate using ERSS2 data with errors in ranking for the variable

\( X \) is defined by
On Ratio Estimation Using Extreme Ranked Set Samples

\[ r_{\text{ERSS2}_c} = \frac{\bar{Y}_c}{\bar{X}_c}, \]

(2.3.3)

where

\[ \bar{X}_c = \frac{\sum_{k=1}^{m} \left( X_1(k) + X_2(n) + \ldots + X_{n-1}(n) + X_{n(n+1)} \right)}{nm}, \]

and

\[ \bar{Y}_c = \frac{\sum_{k=1}^{m} \left( Y_1(k) + Y_2(n) + \ldots + Y_{n-1}(n) + Y_{n(n+1)} \right)}{nm}. \]

Similarly, the approximate variance of \( r_{\text{ERSS2}_c} \) can be written as

\[ \text{Var}(r_{\text{ERSS2}_c}) \approx \frac{R^2}{nm} \left( \frac{(n-1)\sigma_x^2 + \sigma_y^2}{2 \nu_x} + \frac{(n-1)\sigma_y^2 + \sigma_x^2}{2 \nu_y} - \frac{2(n-1)\sigma_x\sigma_y}{\nu_x^{3/2} \nu_y^{3/2}} \right) \]

(2.3.4)

Similar results for RSS samples can be found in Samawi and Muttlak (1996). Also, note that similar results can be obtained for the estimate of the population total and mean of the variable \( Y \), using the above ratio estimators.

3. Inference About \( R \) Using Asymptotic Results

Assume that \( X \) and \( Y \) have finite second moments. Since the ratio is a function of the means of \( X \) and \( Y \), i.e., \( R = \frac{\mu_y}{\mu_x} \), \( R \) has at least two bounded derivatives of all types in some neighborhood of \( (\mu_x, \mu_y) \), provided that \( \mu_x \neq 0 \). The following two propositions follow directly from the definition of \( r_{\text{ERSS}_i} \), and the well known bivariate Delta method (see, Bickel and Doksum, 1977.)

Proposition 1. \( E(r_{\text{ERSS}_i}) = \frac{\mu_y}{\mu_x} + O((nm)^{-1}) \), \( s = a \) or \( c \) and \( i = 1 \) or \( 2 \). Hence, \( r_{\text{ERSS}_i} \) is asymptotically unbiased estimate for \( R \), as \( m \to \infty \), when \( n \) is fixed.
Proposition 2. \( \frac{(r_{\text{ESSS}_i} - R)}{\sqrt{\text{Var}(r_{\text{ESSS}_i})}} \), \( s = a \) or \( c \) and \( i = 1 \) or \( 2 \) converges in distribution to a standard normal random variable, as \( m \to \infty \), when \( n \) is fixed.

By propositions (1) and (2), an approximate 100(1 - \( \alpha \))% confidence interval for \( R \) is

\[
\{ r_{\text{ESSS}_i} \pm z_{\alpha/2} \sqrt{\text{Var}(r_{\text{ESSS}_i})} \},
\]

where \( z_{\alpha/2} \) is the upper 100(\( \alpha / 2 \))% quantile of the standard normal distribution and \( \text{Var}(r_{\text{ESSS}_i}) \) is any consistent estimate for \( \text{Var}(r_{\text{ESSS}_i}) \) based on ERSS data. For example:

\[
\text{Var}(\bar{X}_{(n)}) = \frac{1}{2nm} (S^2_{x(l)} + S^2_{x(n)}), \text{ where}
\]

\[
S^2_{x(l)} = \frac{1}{m(n/2 - 1)} \sum_{k=1}^{m} \sum_{l=1}^{n/2} (X_{2l-1(l)} - \bar{X}_{(l)})^2
\]

and

\[
S^2_{x(n)} = \frac{1}{m(n/2 - 1)} \sum_{k=1}^{m} \sum_{l=1}^{n/2} (X_{2l-1(n)} - \bar{X}_{(n)})^2.
\]

Also,

\[
\text{Var}(\bar{Y}_{(n)}) = \frac{1}{2nm} (S^2_{y(l)} + S^2_{y(n)}), \text{ where}
\]

\[
S^2_{y(l)} = \frac{1}{m(n/2 - 1)} \sum_{k=1}^{m} \sum_{l=1}^{n/2} (Y_{2l-1(l)} - \bar{Y}_{(l)})^2
\]

and

\[
S^2_{y(n)} = \frac{1}{m(n/2 - 1)} \sum_{k=1}^{m} \sum_{l=1}^{n/2} (Y_{2l-1(n)} - \bar{Y}_{(n)})^2. \text{ Similar formulas can be derived}
\]

for the case of \([.]\).

In the same way we can have,
On Ratio Estimation Using Extreme Ranked Set Samples

\[ C\hat{r}(\overline{Y}_{(a)}, \overline{X}_{(a)}) = \frac{1}{2nm} \left( S_{Y(1)x(1)} + S_{y(n)x(n)}^2 \right), \text{ where} \]

\[ S_{Y(1)x(1)} = \frac{1}{m(n/2)-1} \sum_{k=1}^{m} \sum_{i=1}^{n/2} (Y_{2i-1(l)} - \overline{Y}_{(1)}) (X_{2i-1(l)} - \overline{X}_{(1)})^2 \]

and

\[ S_{y(n)x(n)} = \frac{1}{m(n/2)-1} \sum_{k=1}^{m} \sum_{i=1}^{n/2} (Y_{2i(n)} - \overline{Y}_{(n)}) (X_{2i(n)} - \overline{X}_{(n)})^2. \]

Therefore,

\[ V\hat{r}(r_{ERSS}) \approx \frac{S_{x(a)}^2}{nm} \left( \frac{S_{x[a]}^2}{\overline{X}_{[a]}^2} + \frac{S_{y(a)}^2}{\overline{Y}_{(a)}^2} - 2 \frac{S_{x[a]y(a)}}{\overline{X}_{[a]} \overline{Y}_{(a)}} \right), \text{ where} \]

\[ S_{x[a]}^2 = \frac{1}{2} (S_{x(1)}^2 + S_{x(n)}^2), \]

\[ S_{y(a)}^2 = \frac{1}{2} (S_{y(1)}^2 + S_{y(n)}^2) \quad \text{and} \quad S_{x[a]y(a)} = \frac{1}{2} (S_{y(1)x(1)} + S_{y(n)x(n)}). \]

Similar results can be derived for the other ERSS samples. Also, similar results for RSS samples can be found in Samawi and Muttak (1996). Furthermore, we can use this asymptotic result to test the null hypothesis \( H_0 : C \mu_y = \mu_x \), where \( C \) is real number constant,

i.e \( H_0 : \frac{\mu_y}{\mu_x} = C \) vs for example an alternative \( H_1 : \frac{\mu_y}{\mu_x} \neq C \). In this case we reject \( H_0 \) if \( C \in [r_{ERSS} \pm z_{\alpha/2} \sqrt{Va(r_{ERSS})}] \), for large \( m \). Also, the test based on the SRS rejects \( H_0 \) if \( C \in [r_{SRS} \pm z_{\alpha/2} \sqrt{Va(r_{SRS})}] \), for large \( m \). Without loss of generality assume the one sided alternative hypothesis \( H_1 : R > C \) and let \( C_i > C > 0 \), then the asymptotic powers for the ERSS based test and SRS based test respectively are

\[ 1 - \Phi \left( z_{\alpha/2} + \frac{C - C_i}{\sqrt{Va(r_{ERSS})}} \right) \quad \text{and} \quad 1 - \Phi \left( z_{\alpha/2} + \frac{C - C_i}{\sqrt{Va(r_{SRS})}} \right). \]

823
Since $C_i < 1$ and $\sqrt{\text{Var}(r_{\text{EASSI}_i})} < \sqrt{\text{Var}(r_{\text{SRS}})}$, it follows that, ERSS based test is more powerful than SRS based test asymptotically.

4. Simulation Study

We conducted a computer simulation to study the behavior of the ratio estimators using SRS, RSS and ERSS. Bivariate random observation were generated from a bivariate normal distribution with parameters $\mu_x, \mu_y, \sigma^2_x, \sigma^2_y$ and correlation coefficient $\rho$. The sampling method of section 1 is used to pick SRS, RSS and ERSS data with sets of size $nm$. We investigated the performance of the ratio estimates for $n=4, 5, 6, 7, 10$ and $m=2$ and 4. The ratio of the population means were estimated from the SRS, RSS and ERSS data sets. Using 5,000 replications, estimates of the means and the mean square errors for the ratio estimates were computed.

We consider ranking on either variable $Y$ or $X$. Results of these simulations are summarized by the relative efficiency of the estimators of $R$ for different values of correlation coefficient $\rho = -0.99, -0.95, 0.00, 0.25, 0.50, 0.70, 0.95$ and 0.99. The simulation results are given in Table 1 for ranking on the variable $X$ and in Table 2 for ranking on the variable $Y$. The efficiency of the ratio estimator is defined by

$$RE(r_{\text{SRS}}, r_{\text{EASSI}_i}) = \frac{\text{MSE}(r_{\text{SRS}})}{\text{MSE}(r_{\text{EASSI}_i})}$$

4.1. Result of the simulation study

The values obtained by simulation are given in Tables 1 and 2.

**Table 1.** The relative efficiency of the estimators of $R$ when ranking on the variable $X$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>m(n)=</th>
<th>2(4)</th>
<th>2(5)</th>
<th>2(6)</th>
<th>2(7)</th>
<th>2(10)</th>
<th>4(4)</th>
<th>4(5)</th>
<th>4(6)</th>
<th>4(7)</th>
<th>4(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td></td>
<td>2.16</td>
<td>2.60</td>
<td>2.66</td>
<td>2.80</td>
<td>2.92</td>
<td>2.04</td>
<td>2.55</td>
<td>2.29</td>
<td>2.60</td>
<td>2.75</td>
</tr>
<tr>
<td>0.95</td>
<td></td>
<td>1.97</td>
<td>2.20</td>
<td>2.14</td>
<td>2.15</td>
<td>2.16</td>
<td>1.75</td>
<td>2.04</td>
<td>1.83</td>
<td>1.98</td>
<td>1.99</td>
</tr>
<tr>
<td>0.90</td>
<td></td>
<td>1.74</td>
<td>1.75</td>
<td>1.88</td>
<td>2.01</td>
<td>1.84</td>
<td>1.63</td>
<td>1.73</td>
<td>1.81</td>
<td>1.82</td>
<td>1.82</td>
</tr>
<tr>
<td>0.70</td>
<td></td>
<td>1.57</td>
<td>1.71</td>
<td>1.65</td>
<td>1.73</td>
<td>1.66</td>
<td>1.47</td>
<td>1.58</td>
<td>1.52</td>
<td>1.61</td>
<td>1.60</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>1.54</td>
<td>1.83</td>
<td>1.64</td>
<td>1.70</td>
<td>1.65</td>
<td>1.49</td>
<td>1.58</td>
<td>1.62</td>
<td>1.63</td>
<td>1.58</td>
</tr>
</tbody>
</table>
On Ratio Estimation Using Extreme Ranked Set Samples

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0.25</th>
<th>1.74</th>
<th>1.87</th>
<th>1.73</th>
<th>1.73</th>
<th>1.78</th>
<th>1.56</th>
<th>1.65</th>
<th>1.54</th>
<th>1.80</th>
<th>1.69</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>1.93</td>
<td>1.96</td>
<td>1.83</td>
<td>2.22</td>
<td>1.98</td>
<td>1.68</td>
<td>1.80</td>
<td>1.81</td>
<td>1.88</td>
<td>1.85</td>
<td></td>
</tr>
<tr>
<td>-0.95</td>
<td>2.55</td>
<td>2.87</td>
<td>2.77</td>
<td>2.84</td>
<td>2.99</td>
<td>2.14</td>
<td>2.48</td>
<td>2.47</td>
<td>2.71</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>-0.99</td>
<td>2.38</td>
<td>2.82</td>
<td>2.78</td>
<td>3.29</td>
<td>3.05</td>
<td>2.02</td>
<td>2.64</td>
<td>2.49</td>
<td>2.94</td>
<td>3.09</td>
<td></td>
</tr>
</tbody>
</table>

RSS data sets

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0.99</th>
<th>2.45</th>
<th>3.01</th>
<th>3.39</th>
<th>3.52</th>
<th>4.39</th>
<th>2.40</th>
<th>2.68</th>
<th>3.04</th>
<th>3.31</th>
<th>3.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>2.16</td>
<td>2.34</td>
<td>2.51</td>
<td>2.56</td>
<td>2.82</td>
<td>1.91</td>
<td>2.20</td>
<td>2.08</td>
<td>2.31</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>1.91</td>
<td>2.05</td>
<td>2.10</td>
<td>2.18</td>
<td>2.30</td>
<td>1.73</td>
<td>1.91</td>
<td>1.98</td>
<td>2.05</td>
<td>2.14</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>1.71</td>
<td>1.85</td>
<td>1.83</td>
<td>1.90</td>
<td>2.00</td>
<td>1.60</td>
<td>1.61</td>
<td>1.72</td>
<td>1.75</td>
<td>1.81</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>1.68</td>
<td>1.91</td>
<td>1.86</td>
<td>1.98</td>
<td>1.87</td>
<td>1.56</td>
<td>1.73</td>
<td>1.74</td>
<td>1.73</td>
<td>1.88</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>1.87</td>
<td>2.07</td>
<td>1.94</td>
<td>2.01</td>
<td>2.19</td>
<td>1.69</td>
<td>1.86</td>
<td>1.78</td>
<td>2.01</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>2.10</td>
<td>2.11</td>
<td>2.13</td>
<td>2.54</td>
<td>2.33</td>
<td>1.86</td>
<td>1.92</td>
<td>1.98</td>
<td>2.07</td>
<td>2.29</td>
<td></td>
</tr>
<tr>
<td>-0.95</td>
<td>2.82</td>
<td>3.23</td>
<td>3.67</td>
<td>3.67</td>
<td>4.84</td>
<td>2.48</td>
<td>2.91</td>
<td>3.36</td>
<td>3.54</td>
<td>4.66</td>
<td></td>
</tr>
<tr>
<td>-0.99</td>
<td>2.80</td>
<td>3.35</td>
<td>3.76</td>
<td>4.31</td>
<td>5.20</td>
<td>2.30</td>
<td>2.96</td>
<td>3.39</td>
<td>4.01</td>
<td>5.38</td>
<td></td>
</tr>
</tbody>
</table>

Tables 1 and 2 show the efficiency of ratio estimation using ERSS for different values of \( \rho \) and sample sizes, ranking the variable \( X \) and \( Y \) respectively. They indicate that ranking on variable \( X \) is better than ranking on variable \( Y \).

Table 2. The relative efficiency of the estimator when ranking the variable \( Y \)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( m(n) )</th>
<th>2(4)</th>
<th>2(5)</th>
<th>2(6)</th>
<th>2(7)</th>
<th>2(10)</th>
<th>4(4)</th>
<th>4(5)</th>
<th>4(6)</th>
<th>4(7)</th>
<th>4(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.99</td>
<td>2.00</td>
<td>2.14</td>
<td>2.35</td>
<td>2.53</td>
<td>2.37</td>
<td>1.93</td>
<td>2.16</td>
<td>2.06</td>
<td>2.35</td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>1.47</td>
<td>1.57</td>
<td>1.46</td>
<td>1.52</td>
<td>1.53</td>
<td>1.42</td>
<td>1.37</td>
<td>1.35</td>
<td>1.43</td>
<td>1.52</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>1.25</td>
<td>1.27</td>
<td>1.24</td>
<td>1.32</td>
<td>1.24</td>
<td>1.15</td>
<td>1.26</td>
<td>1.17</td>
<td>1.18</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>1.00</td>
<td>1.08</td>
<td>1.06</td>
<td>1.02</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.06</td>
<td>1.00</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

825
When ranking the variable $X$, as $\rho$ decreases from 0.99 to 0.50 the relative efficiency of the ratio estimates using ERSS with respect to SRS decreased in most cases. However, it is not the case when $\rho$ decreases from 0.50 to zero. Also, for negative $\rho$ the relative efficiency of the ratio estimate is higher than for positive $\rho$.

Moreover, the relative efficiency of the ratio estimate using RSS is higher than using either SRS or ERSS for the same $\rho$ and the same sample size $m(n)$ (see Table 1 and 2.) However, in practice we notice that it is easy to rank the extremes (maximum and/or the minimum) even for $n>4$. Thus, when $n>4$ and $m$ is small ERSS could be more efficient and practical than RSS. For example, the relative efficiency for the ratio estimate using ERSS when $\rho=0.99$, $m=2$, $n=10$ and ranking on the variable $X$ is 2.92 (see table 1) but for using RSS when $m=4$ and $n=5$ is 2.68.

Finally, we recommend using ERSS for estimating the ratio or any population parameter based on the ratio estimate because it is more efficient than using SRS for symmetric underlying distribution and more practical than RSS. Also, whenever ERSS is used to estimate the population ratio, it is recommended to rank on that variable which is in the denominator of the ratio estimate.
On Ratio Estimation Using Extreme Ranked Set Samples

هاني سماوي

ملخص

في كثير من الأوضاع الكمية التي يراد تقديرها من العينة العشوائية في النسبة بين متغيرين، أيضًا
النسبة تستخدم لزيادة الدقة عندما تكون الوسط الحسابي أو المجموع لمجموعة ما بالاستفادة من
العلاقة بين المتغيرين. في هذا البحث، نموذج من العينات المرتبة المتطرفة استخدم في تقدير النسبة
بين الوسطين المسجلين للمجتمع الذي يوجد دراسة، النوع الأول عندما يكون حجم العينة زوجي والنوع
الثاني عندما يكون حجم العينة فردي، وسوف نثبت أن استخدام العينات المرتبة المتطرفة سوف يعني
تقديرات غير منهجية تقريباً للنسب عندما يكون المجتمع متماثل. وأيضًا سوف نثبت أنه يمكن تقييم
أكثر دقة من استخدام العينة العشوائية البسيطة، وسوف نعطي بعض الاستقراء حول النسب وبعض
المحاكاة للتقييمات باستخدام الحاسوب.

References


