AUSTRIAN JOURNAL OF STATISTICS
Volume 38 (2009), Number 2, 109–120

Effectivity of Modified Maximum Likelihood Estimators Using Selected Ranked Set Sampling Data

Tamanna Islam[1], Molla Rahman Shaibur[1], and S.S. Hossain[2]

[1]Iwate University, Japan
[2]University of Dhaka, Bangladesh

Abstract: This paper describes the modified maximum likelihood estimator (MMLE) of location and scale parameters based on selected ranked set sampling (SRSS) for normal, uniform and two-parameter exponential distributions. For these distributions, the MMLE of location and scale parameters for SRSS data were compared with the estimators of location and scale parameters for simple random sample (SRS) and ranked set sample (RSS). The MMLE based on SRSS data were found to be advantageous as compared to SRS and RSS estimators for the same number of measurements. The SRSS method with errors in ranking was also described. The minimum correlation between the actual and erroneous ranking was required for MMLE of SRSS to achieve better precision than usual SRS and RSS estimators. When the wrong assumption about the underlying distribution was present, the MMLE of the population mean based on SRSS was better than the RSS estimator of the population mean for all the cases considered.

Zusammenfassung: Diese Arbeit beschreibt die modifizierten Maximum Likelihood Schätzer (MMLE) des Lokation- und Skalenparameters basierend auf selektioniertes ranked set sampling (SRSS) für die Normal-, die Gleich- und die zweiparametrige Exponentialverteilung. Für diese Verteilungen werden die MMLE des Lokations- und Skalenparameters für SRSS Daten verglichen mit Schätzern für einfache Zufallsstichproben (SRS) und ranked set samples (RSS). Die MMLE basierend auf SRSS Daten scheinen verglichen mit den SRS und RSS Schätzern für dieselbe Anzahl von Beobachtungen Vorteile zu haben. Die SRSS Methode mit Fehlern im Ranking wird auch beschrieben. Die geringste Korrelation zwischen dem tatsächlichen und dem fehlerhaften Ranking wurde für den MMLE bei SRSS benötigt, um bessere Genauigkeit als die gewöhnlichen SRS und RSS Schätzer zu haben. Falls die falsche Annahme über die zugrunde liegende Verteilung gemacht wurde, war in allen betrachteten Fällen der MMLE des Populationsmittels basierend auf SRSS besser als der RSS Schätzer.

Keywords: Location and Scale Parameters, Errors in Ranking, Ranked Set Sampling.

1 Introduction

Sampling is the process of selecting some representative part of a population to estimate unknown characteristics by observing the selected part of the population. The assumptions to use the conventional sampling design are that the population of interest is finite and the sample size is also pre-determined. But there are some other situations where a population of interest may be difficult to measure and collection of information are expensive. For these situations, particularly in relation to environmental inquiries, different non-conventional sampling methods can be used. Ranked set sampling (RSS) is one of the sampling methods intended for use in environmental studies.

The RSS is used to increase the precision of estimated yield of a representative samples without the bias of researcher choice (McIntyre, 1952), but this proposal is not being used for a long time. Subsequently some properties (such as unbiasedness, variance and relative precision with respect to SRS) of RSS estimator of population mean (Zheng and Al-Saleh, 2002) and the necessary statistical theory (Takahasi and Wakimoto, 1968) have been established. Of late considerable development has been done in the field of RSS such as RSS has the MMLE for general parameters (Zheng and Al-Saleh, 2002) which has the same expression as the MLE for SRS. Zheng and Al-Saleh (2002) also showed that, the MMLE for the location parameter was always more efficient than the MLE using SRS and for the scale parameter, the MMLE was at least as efficient as the MLE using SRS, when the same sample size was used. They also found that in case of perfect judgement ranking, MMLE was relatively efficient than MLE based on RSS. When the judgement ranking was imperfect using simulation they also found that the MMLE was more robust than the MLE (Zheng and Al-Saleh, 2002).

At present forestry and range researchers are continuing to discover the effectiveness of the RSS. Recently, a number of parametric alternatives of the usual estimation have been suggested (Hossain and Muttlak, 1999; Bhoj, 2000). The SRSS method is an improvement of the RSS method, has the optimal estimators of location and scale parameters (Hossain and Muttlak, 2001).

In this paper, the MMLE of the location and scale parameters based on SRSS data were studied. These estimators were compared with usual estimators based on SRS and RSS data for normal, uniform and two-parameter exponential distributions. The MMLE for SRSS data was also discussed for the situation where errors in ranking might be occurred. And for studying the effect of errors in ranking, we used Dell and Clutter's model (Dell and Clutter, 1972). The minimum value of the correlation coefficient between the actual and erroneous ranking which is required for achieving better precision, with respect to the usual SRS estimator and RSS estimator, were also discussed.

The first step of the RSS procedure was to draw n random sets with n elements in each sample and ranked them (without actual measuring) with respect to the variable of interest. From the first set, the element with the smallest rank was chosen. From the second set, the element with the second smallest rank was chosen. The procedure was continued until the element with the largest rank from the nth sample was chosen. This procedure yielded a total number of n elements chosen to be measured, one from each sample. The cycle might be repeated m times until nm units were measured. These nm units formed the RSS data.

In the SRSS method instead of drawing n random sets of size n, only $k < n$ random sets of size n were drawn, and instead of measuring the ith smallest order statistic of the ith set, the n_ith smallest order statistic of the ith set was considered for measurement. The values of $n_1, \ldots, n_k$, $1 \leq n_1 < \cdots < n_k \leq n$, is required to determine beforehand. Determination of $n_1, \ldots, n_k$ has described extensively earlier (Hossain and Muttlak, 2001).

Present sampling method seems to be quite appealing because SRSS estimators are expected to be more precise than those obtained by SRS and RSS with the same number of measurements when the underlying distribution is known (Hossain and Muttlak, 2001).

2 MMLEs of Location and Scale Parameters Using SRSS Data

In a SRSS of size k, the measured elements are $x_{(n_1:n)1}, \ldots, x_{(n_k:n)k}$. For simplicity, these are denoted by $x_{(n_1)}, \ldots, x_{(n_k)}$. Now, let $x_{(n_i)}$, the n_ith smallest order statistic from the ith random sample of size n, be drawn from a distribution with location and scale parameters μ and σ, respectively. The likelihood function based on the SRSS can be written as (Balakrishnan and Cohen, 1991)

$$L = \prod_{i=1}^{k} \frac{n!}{(n_i-1)!(n-n_i)!} \{F(x_{(n_i)})\}^{n_i-1}\{1-F(x_{(n_i)})\}^{n-n_i} f(x_{(n_i)})\,.$$

The algebraic deduction of the MMLE for SRSS data for normal distribution is described below in short.

Let the sample be drawn from a normal distribution with pdf $f(x) = (2\pi\sigma^2)^{-1/2}$ $\exp\{-(x-\mu)^2/2\sigma^2\}$, hence the log-likelihood function can be written as

$$\log L = \sum_{i=1}^{k} \left\{-\log\sigma + \log f(z_{(n_i)}) + (n_i-1)\log F(z_{(n_i)}) + (n-n_i)\log(1-F(z_{(n_i)}))\right\}\,, \tag{1}$$

where

$$z_{(n_i)} = \frac{x_{(n_i)}-\mu}{\sigma}\,.$$

Equating the first derivatives with respect to μ and σ to zero and let $f'(z_{(n_i)}) = -z_{(n_i)} f(z_{(n_i)})$, we have the simplified form

$$\frac{1}{\sigma}\sum_{i=1}^{k}\left\{z_{(n_i)} - (n_i-1)\frac{f(z_{(n_i)})}{F(z_{(n_i)})} + (n-n_i)\frac{f(z_{(n_i)})}{1-F(z_{(n_i)})}\right\} = 0 \tag{2}$$

and

$$\frac{1}{\sigma}\sum_{i=1}^{k}\left\{1 - z^2_{(n_i)} + (n_i-1)\frac{f(z_{(n_i)})}{F(z_{(n_i)})}z_{(n_i)} - (n-n_i)\frac{f(z_{(n_i)})}{1-F(z_{(n_i)})}z_{(n_i)}\right\} = 0\,. \tag{3}$$

These equations do not admit explicit solutions. So, we expand the function $F(z_{(n_i)})$ appearing in (2) and (3) in a Taylor series around the point $F^{-1}(p_{(n_i)}) = \epsilon_{(n_i)}$. In general we know that, if f (function) can be differentiated n times at a, then we can define the nth Taylor polynomial for f about $x = a$ as

$$p_{(n_i)} = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^n(a)}{n!}(x-a)^n\,.$$

Let

$$f(x) = \frac{f(z)}{F(z)} \qquad \text{and} \qquad a = \epsilon\,.$$

The simplified form of the estimator of location parameter is found as follows (Balakrishnan and Cohen, 1991)

$$\hat{\mu}_M = C + L\sigma\,,$$

where

$$C = \frac{\sum_{i=1}^{k} x_{(n_i)} + \sum_{i=1}^{k}(n_i-1)x_{(n_i)}\beta_{(n_i)} + \sum_{i=1}^{k}(n-n_i)w_{(n_i)}x_{(n_i)}}{k + \sum_{i=1}^{k}(n_i-1)\beta_{(n_i)} + \sum_{i=1}^{k}(n-n_i)w_{(n_i)}}\,, \tag{4}$$

$$L = \frac{-\sum_{i=1}^{k}(n_i-1)\alpha_{(n_i)} + \sum_{i=1}^{k}(n-n_i)\gamma_{(n_i)}}{k + \sum_{i=1}^{k}(n_i-1)\beta_{(n_i)} + \sum_{i=1}^{k}(n-n_i)w_{(n_i)}}\,, \tag{5}$$

where

$$\begin{aligned}
\alpha_{(n_i)} &= \frac{f(\epsilon_{(n_i)})\left\{1 + \epsilon_{(n_i)}^2 + \frac{\epsilon_{(n_i)}f(\epsilon_{(n_i)})}{p_{(n_i)}}\right\}}{p_{(n_i)}}\,,\\
\beta_{(n_i)} &= \frac{f(\epsilon_{(n_i)})\left\{f(\epsilon_{(n_i)}) + p_{(n_i)}\epsilon_{(n_i)}\right\}}{p_{(n_i)}^2}\,,\\
\gamma_{(n_i)} &= \frac{f(\epsilon_{(n_i)})\left\{1 + \epsilon_{(n_i)}^2 - \frac{\epsilon_{(n_i)}f(\epsilon_{(n_i)})}{q_{(n_i)}}\right\}}{q_{(n_i)}}\,,\\
w_{(n_i)} &= \frac{f(\epsilon_{(n_i)})\{f(\epsilon_{(n_i)}) - q_{(n_i)}\epsilon_{(n_i)}\}}{q_{(n_i)}^2}\,.
\end{aligned}$$

The values of C and L in (4) and (5) have to be determined for given population distributions such as normal distribution.

The estimator of the scale parameter can be obtained by finding the positive root of the following equation

$$\begin{aligned}
&k\sigma_M^2 - \sigma_M\left\{-\sum_{i=1}^{k}(n_i-1)\alpha_{(n_i)}(x_{(n_i)}-\mu) + \sum_{i=1}^{k}(n-n_i)\gamma_{(n_i)}(x_{(n_i)}-\mu)\right\}\\
&-\left\{\sum_{i=1}^{k}(n_i-1)\beta_{(n_i)}(x_{(n_i)}-\mu)^2 + \sum_{i=1}^{k}(n-n_i)w_{(n_i)}(x_{(n_i)}-\mu)^2 + \sum_{i=1}^{k}(x_{(n_i)}-\mu)^2\right\}\\
&= 0\,.
\end{aligned}$$

3 Comparison of MMLE Using SRSS Data with Other Estimators

The usual estimators of the population mean μ and variance σ^2 based on SRS of size k is given by

$$\hat{\mu}_S = \frac{1}{k}\sum_{i=1}^{k} x_i \qquad \text{and} \qquad \hat{\sigma}_S^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_i - \bar{\mu}_k)^2\,, \tag{6}$$

respectively.

Let $\phi_\mu(M|S)$ and $\phi_\sigma(M|S)$ denote the relative precisions of $\hat{\mu}_M$ with respect to $\hat{\mu}_S$ and $\hat{\sigma}_M$ with respect to $\hat{\sigma}_S$, where M and S stand for MMLE for SRSS and SRS estimators.

The values of $\phi_\mu(M|S)$ and $\phi_\sigma(M|S)$ for various values of $n \leq 7$ $(k < n)$ for a normal, a uniform, and a two-parameter exponential distributions are presented in Table 1. The tabulated values show that high values of relative precision with respect to the usual SRS estimators of size k can be achieved by MMLE for SRSS data with the same number of measurements. For example, for $n = 6$ and $k = 2$, values of $\phi_\mu(M|S)$ are 2.296, 2.503, 2.303 and the values of $\phi_\sigma(M|S)$ are 1.646, 20.291, 2.214 for normal, uniform and two-parameter exponential distributions, respectively.

Table 1: Relative precision of MMLE of the population mean and standard deviation for SRSS data compared to the usual SRS estimators $\hat{\mu}_S$ and $\hat{\sigma}_S$.

		Normal		Uniform		2-par. Exponential	
n	k	$\phi_\mu(M\|S)$	$\phi_\sigma(M\|S)$	$\phi_\mu(M\|S)$	$\phi_\sigma(M\|S)$	$\phi_\mu(M\|S)$	$\phi_\sigma(M\|S)$
3	2	1.625	1.667	2.186	9.283	1.525	2.240
4	2	1.647	1.405	2.469	14.043	2.060	1.013
	3	2.247	1.954	5.266	28.705	2.147	1.923
5	2	2.259	1.706	2.494	10.783	2.267	2.910
	3	2.399	1.955	5.413	23.372	2.473	2.145
	4	2.442	2.573	6.815	30.501	3.587	2.258
6	2	2.296	1.646	2.503	20.291	2.303	2.214
	3	2.429	2.021	5.978	39.772	2.738	2.223
	4	2.596	3.037	8.070	26.521	3.631	3.219
	5	2.775	3.416	10.685	20.210	3.753	2.509
7	2	2.373	1.752	2.834	16.319	2.933	2.585
	3	2.467	2.150	6.806	42.692	3.035	7.059
	4	2.697	3.135	10.012	10.437	3.984	2.578
	5	2.908	3.653	14.645	24.679	4.975	4.096
	6	2.957	4.510	18.137	19.640	5.143	3.012

For complete RSS data the estimators of location and scale parameters are

$$\hat{\mu}_k = \frac{1}{k}\sum_{i=1}^{k} x_{(i)} \qquad \text{and} \qquad \hat{\sigma}_k^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_i - \hat{\mu}_k)^2\,, \tag{7}$$

respectively.

The relative precisions of $\hat{\mu}_M$ and $\hat{\sigma}_M$ based on k measurements with respect to $\hat{\mu}_k$ and $\hat{\sigma}_k$ are defined as $\phi_\mu(M|R)$ and $\phi_\sigma(M|R)$, where M and R stand for MMLE for SRSS and RSS.

Table 2 shows that the values of $\phi_\mu(M|R)$ and $\phi_\sigma(M|R)$ for various values of $n \leq 7$ ($k < n$) for a normal, a uniform, and a two-parameter exponential distribution. The values of $\phi_\mu(M|R)$ and $\phi_\sigma(M|R)$ suggested that the MMLE for SRSS for all the cases considered were (no matter what the values of n and k) better than those based on RSS with same number of measurements. For example, for $n = 7$ and $k = 2$, values of $\phi_\mu(M|R)$ are 1.815, 1.752, 1.700, and the values of $\phi_\sigma(M|R)$ are 1.697, 7.631, 2.457 for normal, uniform and two-parameter exponential distributions, respectively.

Table 2: Relative precision of MMLE of the population mean and standard deviation for SRSS data compared to the usual RSS estimators $\hat{\mu}_R$ and $\hat{\sigma}_R$.

		Normal		Uniform		2-par. Exponential	
n	k	$\phi_\mu(M\|R)$	$\phi_\sigma(M\|R)$	$\phi_\mu(M\|R)$	$\phi_\sigma(M\|R)$	$\phi_\mu(M\|R)$	$\phi_\sigma(M\|R)$
3	2	1.226	1.473	1.458	8.834	1.143	2.162
4	2	1.248	1.331	1.625	9.883	1.274	1.006
	3	1.462	1.328	2.591	10.441	1.465	1.059
5	2	1.318	1.572	1.665	9.376	1.534	2.826
	3	2.067	1.656	2.700	8.866	1.554	2.098
	4	2.431	2.011	2.718	5.848	1.662	2.236
6	2	1.807	1.618	1.748	9.940	1.559	1.564
	3	2.382	1.323	2.967	10.629	1.710	1.966
	4	2.433	1.712	3.223	6.902	1.752	2.530
	5	2.656	2.610	3.526	10.452	1.899	2.175
7	2	1.815	1.697	1.752	7.631	1.700	2.457
	3	2.419	1.313	3.325	11.422	1.972	3.739
	4	2.662	1.951	4.014	8.624	1.919	2.038
	5	2.891	2.292	4.830	16.712	2.380	3.857
	6	2.913	2.393	5.163	13.885	3.403	1.971

4 MMLE for SRSS Data with Errors in Ranking

The usual RSS estimator and the MMLE for SRSS data of the population mean and the standard deviation are based on the ranking of the variable of interest. Usually the ranking is done subjectively by rankers and because of that an error may occur in such a ranking. The errors in ranking may have an effect on the estimates (Stokes, 1980). For SRSS, the ranking may not always be perfect, i.e., the n_ith smallest observation in the ith set measured by the SRSS method may not be the actual n_ith order statistic in the set of size n. Following the method in Dell and Clutter (1972) we call it the n_ith 'judgement order statistic'. Let $x_{[n_i]}$ denote the n_ith smallest 'judgement order statistic' from the ith random sample of size n. According to the approach of David and Levine (1972), we can suppose that, while ranking, an observer subjectively gives each element a value of a variable y. We assume that the regression of x on y is linear and that the values follow a bivariate normal distribution. Let ρ denote the correlation coefficient between x and y. Now the following can be obtained (Hossain and Muttlak, 2001)

$$\begin{aligned}\mathrm{E}(x_{[n_i]}) &= \mathrm{E}(\mathrm{E}(x|y_{(n_i)})) = \mu + \rho\sigma\delta_{n_i:n}\,,\\ \mathrm{var}(x_{[n_i]}) &= \mathrm{E}(\mathrm{var}(x|y_{(n_i)})) + \mathrm{var}(\mathrm{E}(x|y_{(n_i)})) = \sigma^2(1-\rho^2) + \rho^2\sigma^2\eta_{n_in_i:n}\,,\end{aligned}$$

where

$$\mathrm{E}(x_{(n_i)}) = \delta_{n_i:n}\,, \qquad \mathrm{cov}(x_{(n_i)}, x_{(n_j)}) = \eta_{n_in_i:n}$$

with $z_{(n_i)} = (x_{(n_i)} - \mu)/\sigma$.

The MMLE for SRSS data of the population mean μ using 'judgement order statistics' is given as

$$\mu^*_{[k]} = C + L\sigma\,,$$

where C and L are defined in (4) and (5). It can be shown that $\mu^*_{[k]}$ is an unbiased estimator of μ, and its variance is

$$\begin{aligned}\mathrm{var}(\mu^*_{[k]}) &= \frac{\displaystyle\sum_{i=1}^{k}\left\{1+(n_i-1)\beta_{(n_i)}+(n-n_i)w_{(n_i)}x_{(n_i)}\right\}^2 \mathrm{var}(x_{(n_i)})}{\left\{k+\displaystyle\sum_{i=1}^{k}(n_i-1)\beta_{(n_i)}+\sum_{i=1}^{k}(n-n_i)w_{(n_i)}\right\}^2}\\ &= \frac{\displaystyle\sum_{i=1}^{k}\left\{1+(n_i-1)\beta_{(n_i)}+(n-n_i)w_{(n_i)}\right\}^2\left\{\sigma^2(1-\rho^2)+\rho^2\sigma^2\eta_{n_in_i:n}\right\}}{\left\{k+\displaystyle\sum_{i=1}^{k}(n_i-1)\beta_{(n_i)}+\sum_{i=1}^{k}(n-n_i)w_{(n_i)}\right\}^2}\,.\end{aligned}$$

Denoting the relative precision of $\mu^*_{[k]}$ with respect to the SRS estimator $\hat{\mu}_S$ by $\phi_\mu(E|S)$ it is found that

$$\phi_\mu(E|S) = \frac{\left\{k+\displaystyle\sum_{i=1}^{k}(n_i-1)\beta_{(n_i)}+\sum_{i=1}^{k}(n-n_i)w_{(n_i)}\right\}^2}{k\displaystyle\sum_{i=1}^{k}\left\{1+(n_i-1)\beta_{(n_i)}+(n-n_i)w_{(n_i)}\right\}^2\left\{(1-\rho^2)+\rho^2\eta_{(n_i)(n_i):n}\right\}}\,,$$

where E stands for MMLE for SRSS with errors.

The value of $\phi_\mu(E|S)$ depends on the magnitude of ρ, because $\phi_\mu(E|S) = \psi(\rho^2)$, then

$$\psi(0) = \frac{\left\{k+\displaystyle\sum_{i=1}^{k}(n_i-1)\beta_{(n_i)}+\sum_{i=1}^{k}(n-n_i)w_{(n_i)}\right\}^2}{k\displaystyle\sum_{i=1}^{k}\left\{1+(n_i-1)\beta_{(n_i)}+(n-n_i)w_{(n_i)}\right\}^2}\,.$$

It can be shown that $\phi_\mu(E|S) \geq 1$ if $\rho^2 \geq \rho_0^2$, where $\psi(\rho_0^2) = 1$. The value of ρ_0^2 can be obtained as

$$\rho_0^2 = \frac{u - v}{w} ,$$

where

$$u = \left\{ k + \sum_{i=1}^{k} (n_i - 1)\beta_{(n_i)} + \sum_{i=1}^{k} (n - n_i) w_{(n_i)} \right\}^2 ,$$

$$v = k \sum_{i=1}^{k} \left\{ 1 + (n_i - 1)\beta_{(n_i)} + (n - n_i) w_{(n_i)} \right\}^2 ,$$

$$w = k \sum_{i=1}^{k} (\eta_{n_i n_i : n} - 1) \left\{ 1 + (n_i - 1)\beta_{(n_i)} + (n - n_i) w_{(n_i)} \right\}^2 .$$

The RSS estimator of population mean in the presence of errors in ranking is given by Dell and Clutter (1972) as

$$\hat{\mu}_{[k]} = \frac{1}{k} \sum_{i=1}^{k} x_{[i]} ,$$

with variance

$$\operatorname{var}(\hat{\mu}_{[k]}) = \frac{\sigma^2}{k^2} \sum_{i=1}^{k} ((1 - \rho^2) + \rho^2 \eta_{ii:k}) .$$

The relative precision of μ_k^* with respect to the RSS estimator $\hat{\mu}_k$ is

$$\phi_\mu(E|R) = \frac{\frac{1}{k^2} \sum_{i=1}^{k} \left((1 - \rho^2) + \rho^2 \eta_{ii:k} \right) \left\{ k + \sum_{i=1}^{k} (n_i - 1)\beta_{(n_i)} + \sum_{i=1}^{k} (n - n_i) w_{(n_i)} \right\}^2}{\sum_{i=1}^{k} \left\{ 1 + (n_i - 1)\beta_{(n_i)} + (n - n_i) w_{(n_i)} \right\}^2 \left\{ (1 - \rho^2) + \rho^2 \eta_{n_i n_i : n} \right\}} .$$

It can be shown that $\phi_\mu(E|R) \geq 1$ if $\rho^2 \geq \rho_1^2$ where

$$\rho_1^2 = \frac{C_1 - C_2}{D_1 - D_2}$$

with

$$C_1 = \frac{1}{k} \left\{ k + \sum_{i=1}^{k} (n_i - 1)\beta_{(n_i)} + \sum_{i=1}^{k} (n - n_i) w_{(n_i)} \right\}^2 ,$$

$$C_2 = \sum_{i=1}^{k} \left\{ 1 + \sum_{i=1}^{k} (n_i - 1)\beta_{(n_i)} + \sum_{i=1}^{k} (n - n_i) w_{(n_i)} \right\}^2 ,$$

$$D_1 = \frac{1}{k^2} \sum_{i=1}^{k} (\eta_{ii:k} - 1) \left\{ k + \sum_{i=1}^{k} (n_i - 1)\beta_{(n_i)} + \sum_{i=1}^{k} (n - n_i) w_{(n_i)} \right\}^2 ,$$

$$D_2 = \sum_{i=1}^{k} (\eta_{n_i n_i : n} - 1) \left\{ 1 + (n_i - 1)\beta_{(n_i)} + (n - n_i) w_{(n_i)} \right\}^2 .$$

Table 3: Relative precision $\phi_\mu(E|S)$ of MMLE of the population mean for SRSS data compared to the usual SRS estimator $\hat{\mu}_S$ when errors occur in ranking.

		ρ								
n	k	0	0.25	0.4	0.5	0.7	0.8	0.9	0.95	$\|\rho_0\|$
3	2	1.00	1.58	1.69	1.74	1.78	1.85	2.05	2.11	0.178
4	2	0.987	1.842	1.851	2.071	2.097	2.291	2.385	2.409	0.207
	3	1.000	1.871	1.924	2.084	2.154	2.333	2.516	2.551	0.21
5	2	1.000	2.678	2.889	2.901	3.112	3.329	3.595	3.668	0.318
	3	0.812	1.995	2.042	2.068	2.188	2.190	2.522	2.635	0.225
	4	0.912	2.210	2.303	2.398	2.348	2.580	2.77	2.800	0.253
6	2	1.000	2.174	2.251	2.310	2.508	2.675	2.721	2.845	0.235
	3	0.928	2.186	2.260	2.276	2.292	2.458	2.625	2.838	0.245
	4	1.000	3.29	3.415	3.435	3.723	4.075	4.085	4.261	0.362
	5	1.000	2.778	2.465	2.677	2.796	3.033	3.165	3.233	0.290
7	2	1.000	2.164	2.219	2.266	2.421	2.701	2.717	2.794	0.247
	3	0.895	2.311	2.372	2.393	2.644	2.675	2.731	2.888	0.253
	4	0.981	2.501	2.845	2.911	3.033	3.294	3.342	3.441	0.291
	5	1.000	2.759	2.911	3.133	3.210	3.229	3.238	3.258	0.278
	6	1.000	3.484	3.633	3.701	3.847	4.214	4.313	4.578	0.379

Table 4: Relative precision $\phi_\mu(E|R)$ of MMLE of the population mean for SRSS data compared to the usual RSS estimator $\hat{\mu}_R$ when errors occur in ranking.

		ρ								
n	k	0	0.25	0.4	0.5	0.7	0.8	0.9	0.95	$\|\rho_0\|$
3	2	0.991	0.990	1.069	1.074	1.098	1.165	1.176	1.279	0.094
4	2	0.887	0.968	1.075	1.092	1.116	1.153	1.209	1.296	0.091
	3	0.611	0.789	0.832	0.872	0.881	0.929	0.986	1.026	0.088
5	2	0.801	0.826	0.879	0.932	0.988	0.992	1.019	1.045	0.093
	3	0.712	0.769	0.811	0.825	0.891	0.906	0.996	1.019	0.082
	4	1.000	1.242	1.269	1.352	1.376	1.386	1.406	1.429	0.089
6	2	0.763	0.781	0.816	0.905	0.926	0.946	1.016	1.092	0.081
	3	0.751	0.746	0.812	0.816	0.845	0.879	1.529	1.586	0.083
	4	0.723	0.777	0.792	0.826	0.832	0.912	0.956	0.976	0.088
	5	1.000	1.788	1.792	1.801	1.886	1.892	1.902	1.982	0.085
7	2	1.000	1.782	1.796	1.06	1.882	1.892	1.906	1.983	0.085
	3	0.695	0.716	0.776	0.816	0.862	0.889	0.935	1.011	0.079
	4	0.771	0.779	0.788	0.862	0.856	0.873	0.968	1.015	0.081
	5	1.000	1.319	1.386	1.472	1.489	1.546	1.576	1.645	0.083
	6	1.000	1.302	1.356	1.363	1.436	1.636	1.675	1.859	0.097

Values of $\phi_\mu(E|S)$ for various values of ρ, n, and k showed that values of $\phi_\mu(E|S)$ increased as the magnitudes of ρ increased. The values of $|\rho_0|$ for various n and k are also given in Table 3. Table 4 gives the values of $\phi_\mu(E|R)$ for various values of ρ, and $|\rho_1|$ for various n and k. Considering Table 4, it can be seen that for the larger values of ρ, values of $\phi_\mu(E|R)$ were also large.

5 Sensitivity of MMLE for SRSS of the Population Mean

In this section we studied the effects of any possible misspecification of the assumed distribution. Using the assumed distributions normal, uniform and two-parameter exponential values of $\phi_\mu(A|R)$ are computed for various true distributions (rectangular, normal, two-parameter exponential and exponential) and the values are given in Table 5 and 6 for various n and k.

Table 5: Values of the relative precision $\phi_\mu(A|R)$ for various true distributions and for the rectangular and normal as assumed distributions.

		True Distribution			
n	k	Rectangular	Normal	2-par. Exponential	Exponential
Assumed Distribution: Rectangular					
3	2	1.885	1.492	1.319	1.368
4	2	2.080	1.581	1.454	1.561
	3	2.163	2.017	1.505	1.781
5	2	2.291	1.927	1.561	1.475
	3	2.307	2.465	1.663	1.601
	4	2.481	2.620	1.806	1.655
6	2	2.382	2.148	1.786	1.436
	3	2.406	2.670	2.178	1.722
	4	2.493	2.932	2.246	1.695
	5	2.687	3.066	2.336	1.660
7	2	2.515	2.313	2.078	3.258
	3	2.613	2.781	2.212	2.084
	4	2.691	3.169	2.561	2.751
	5	2.785	3.571	2.708	2.743
	6	3.839	5.969	3.454	4.313
Assumed Distribution: Normal					
3	2	0.789	1.193	0.971	1.068
4	2	0.875	1.484	0.990	1.104
	3	0.924	1.512	1.051	1.110
5	2	0.876	1.730	1.319	1.171
	3	0.951	1.782	1.222	1.274
	4	0.959	1.808	1.486	1.306
6	2	0.881	1.828	1.336	1.336
	3	0.960	1.787	1.302	1.354
	4	1.032	1.941	1.559	1.379
	5	1.039	1.973	1.651	1.542
7	2	0.910	1.928	1.351	1.413
	3	0.980	2.863	1.411	1.573
	4	1.029	2.073	1.665	1.442
	5	1.059	2.100	1.782	1.622
	6	1.075	2.211	1.971	2.504

Table 6: Values of the relative precision $\phi_\mu(A|R)$ for various true distributions and for the two-parameter exponential distribution as assumed distribution.

		True Distribution			
n	k	Rectangular	Normal	2-par. Exponential	Exponential
		Assumed Distribution: 2-par. exponential			
3	2	1.002	1.042	0.937	1.076
4	2	1.147	1.140	1.098	1.162
	3	1.175	1.172	1.112	1.183
5	2	1.169	1.171	1.127	1.232
	3	1.208	1.226	1.167	1.280
	4	1.210	1.308	1.225	1.304
6	2	1.195	1.265	1.135	1.352
	3	1.259	1.254	1.223	1.471
	4	1.218	1.350	1.233	1.484
	5	1.343	1.389	1.603	1.512
7	2	1.196	1.267	1.255	1.435
	3	1.353	1.277	1.230	1.538
	4	1.297	1.415	1.259	1.719
	5	1.350	1.424	1.896	2.114
	6	1.444	1.571	2.128	2.668

6 Discussion and Conclusions

When the number of measurements was small and the cost of measurements was high, instead of SRS and RSS a precise estimator could be obtained by using SRSS method. For this, we compared MMLE for SRSS with SRS and RSS estimators based on the same numbers of measurements. The results of the present research revealed that, for all the cases considered in this study, the MMLE for SRSS were more efficient than the estimators of SRS obtained from the same number of measurements (Table 1). So for all cases considered, a generous gain in precision was realized by using MMLE for SRSS estimators with $k < n$ measurements.

The values of $\phi_\mu(M|R)$ and $\phi_\sigma(M|R)$ in Table 2 suggested that for all the cases considered here (no matter what the values of n or k are) the MMLE for SRSS were more efficient than the RSS estimators obtained from the same number of measurements for normal, uniform and two-parameter exponential distributions.

Looking at the values in Table 1 and 2, it can be observed that using the same sample size the values of relative precisions were greater than 1 for all the distributions considered. It also can be seen that, for the same to number of measurements k, the values of relative precisions (for population mean) increased as set size n increased.

In the presence of errors in ranking, MMLE for SRSS performed better than the SRS and RSS estimators when the actual ranking was highly correlated with the 'judgment ranking' (Table 3 and 4). Table 3 also showed that, in favor of MMLE for SRSS to perform better than SRS, the required lower limit of the correlation coefficient between the actual ranking and the 'judgment ranking' can be less than 0.2. Table 4 showed that, on behalf of MMLE for SRSS to perform better than RSS, the same limit can be less than 0.24. The values of relative precisions for all n and k increased as the values of ρ increased. This indicated that lesser was the extent of errors in ranking, better the MMLE for SRSS performed.

From Table 5 and 6, we can see that, even if the assumption about the underlying distribution is wrong, the MMLE of the population mean for SRSS data performed better than the RSS estimator of the population mean for the cases considered.

Acknowledgements

The authors wish to express appreciation to the Editor in Chief Herwig Friedl and Ahmad Ismail Mustafa for improving the manuscript.

References

Balakrishnan, N., and Cohen, C. A. (1991). *Order statistics and inference*. San Diego: Academic Press, Inc.

Bhoj, D. (2000). New ranked set sampling for one-parameter family of distributions. *Biometrical Journal*, *42*, 647-658.

David, H. A., and Levine, D. N. (1972). Ranked set sampling in the presence of judgment error. *Biometrics*, *28*, 553-555.

Dell, T. R., and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, *28*, 545-553.

Hossain, S. S., and Muttlak, H. A. (1999). Paired ranked set sampling: a more efficient procedure. *Environmetrics*, *10*, 195-212.

Hossain, S. S., and Muttlak, H. A. (2001). Selected ranked set sampling. *Australian and New Zealand Journal of Statistics*, *43*, 311-325.

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

Stokes, S. L. (1980). Estimation of variance using judgment ordered ranked set samples. *Biometrics*, *36*, 35-42.

Takahasi, K., and 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*, 1-31.

Zheng, G., and Al-Saleh, F. M. (2002). Modified maximum likelihood estimators based on ranked set sampling. *Annals of the Institute of Statistical Mathematics*, *54*, 642-658.

Authors' addresses:

Molla Rahman Shaibur and Tamanna Islam
Department of Agro-bioscience
Iwate University
Morioka 020-8550, Japan
E-mail: shaibur75@yahoo.com

S. S. Hossain
Institute of Statistical Research and Training
University of Dhaka
Dhaka 1000, Bangladesh