King Fahd University of Petroleum & Minerals

**DEPARTMENT OF MATHEMATICAL SCIENCES**

Technical Report Series

TR257

October 2000

# Random Ranked Set Samples with Imperfect Judgment Ranking

I. Rahimov, H.A. Muttlak

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

# RANDOM RANKED SET SAMPLES WITH IMPERFECT JUDGMENT RANKING

**I. Rahimov and H. A. Muttlak**
Department of Mathematical Sciences
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia

Random ranked set sampling (RRSS) as suggested by Rahimov and Muttlak (2000) assumed perfect ranking i.e. there is no errors in ranking the units with respect to the variable of interest. But in real life problems it is impossible to rank the units with no errors in ranking. Dell and Clutter (1972) pointed out that the errors in ranking would reduce the efficiency of the ranked set sampling (RSS) with fixed set size. In real life we may encounter problems that requiring random set size and (or) random number of replication. We suggested several unbiased estimators for the population mean using RRSS with errors in ranking estimator. The properties of these estimators are discussed and compare with usual RSS with errors in ranking. It turns out that under certain conditions, the newly suggested estimators are more efficient than the usual RSS with errors in ranking. Finally the asymptotic properties of the proposed estimators are discussed.

KEY WORDS: Discrete uniform distribution; errors in ranking; efficiency; random number of replications; random set size, relative precision.

## 1. INTRODUCTION

The ranked set sampling (RSS) proposed by McIntyre (1952) to estimate the pasture yields, has application in many other fields in particular in the field of environments and ecology. The RSS proven to be more efficient when units are difficult and costly to measure, but are easy and cheap to rank with respect to a variable of interest without actual measurement. One can often tell which shrub is the largest without measuring all the shrubs.

The RSS is based on fixed set size and number of replications. But in some applications we may encounter problems where either the set size or number of replications or both cannot be fixed. For real life example, see Muttlak and McDonald (1992).

Takahasi and Wakimoto (1968) independently proposed the same RSS method and supplied the mathematical theory that missing in McIntyre's (1952) suggestion. Dell and Clutter (1972) showed that errors in ranking reduce the efficiency of the RSS mean relative to the SRS mean. However, the RSS mean remains unbiased and more efficient than the SRS mean unless the ranking is so poor as to yield a random sample.

Several authors considered some modifications of the RSS method either to improve the efficiency of the estimators or (and) to reduce the errors in ranking and make the RSS method easier to implement in the field. Samawi et al. (1996) studied the properties of the extreme ranked set sampling (ERSS) in estimating the population mean. Muttlak (1997) considered using median ranked set sampling (MRSS) to estimate the population mean.

For classified and extensively reviewed work in the area of RSS, see Patil et al (1994) and Kaur et al (1995). Finally for bibliography in the area of RSS see Patil et al. (1999).

In this paper we provided a new direction of RSS via the notion of random ranked set sampling (RRSS) with errors in ranking. In Section 2 the idea of RRSS with errors in ranking is introduced in the case of number of replications are random and with fixed set size. The case of fixed number of replications with random set size is discussed in Sections 3. The general case of RRSS with errors in ranking with random set size and number of replications is considered in Section 4. The asymptotic properties of the estimator of the population mean suggested for the general case of RRSS with errors in ranking is established in Section 5.

## 2. RANDOM NUMBER OF REPLICATIONS WITH FIXED NUMBER OF CLASSES OR SET SIZE

Let $X_{11}, X_{12}, \ldots, X_{1n}; X_{21}, X_{22}, \ldots, X_{2n}; \ldots; X_{i1}, X_{i2}, \ldots, X_{in_i}; \ldots; X_{n1}$,

$X_{n2} \ldots, X_{nn}$ are independent and identically distributed random variables with cdf F(x), pdf f(x), mean, $\mu$ and variance $\sigma^2$. Let $v_1, v_2, \ldots$ be independent random variable taking values from $\Lambda = \{2, 3, \ldots\}$ and independent of the random variables $X_{ij}, i, j \geq 1$. First we randomly select $v_1$ sets of size n units, where $v_1 \in \Lambda$. We order the units within each set by judgment order i.e. there is errors in ordering the units within each set. Let $X_{i[1]}, X_{i[2]}, \ldots, X_{i[n]}$, i = 1, 2,..., $v_1$ be the judgment order statistics of $X_{i1}, X_{i2}, \ldots, X_{in}$ i = 1, 2, ..., $v_1$, which are written as $X_{i[j]}, j = 1, 2, \ldots, n$ to distinguish it from the actual order statistics $X_{i(j)}$. We select for measurement the first judgment order statistics $X^{(1)}_{1[1]}, X^{(1)}_{2[1]}, \ldots, X^{(1)}_{v_1[1]}$ and let $\bar{y}^{(n)}_{v_1} = \frac{1}{v_1}\sum_{i=1}^{v_1} X^{(1)}_{i[1]}$.

Now we randomly select another $v_2$ sets of size n units each and order them by judgment order $X_{i[1]}, X_{i[2]}, \ldots, X_{i[n]}$, i = 1, 2, ..., $v_2$. We select the second judgment order statistics for measurement $X^{(2)}_{1[2]}, X^{(2)}_{2[2]}, \ldots, X^{(2)}_{v_1[2]}$ and denote $\bar{y}^{(n)}_{v_2} = \frac{1}{v_2}\sum_{i=1}^{v_2} X^{(2)}_{i[2]}$.

We repeat this process n times to get $\bar{y}^{(n)}_{v_1}, \bar{y}^{(n)}_{v_2}, \ldots, \bar{y}^{(n)}_{v_n}$. We propose the following as an estimator for the population mean μ

$$\bar{\bar{y}}_n = \frac{1}{n}\sum_{i=1}^{n}\bar{y}^{(n)}_{v_i} \tag{1}$$

The basic properties of the estimator $\bar{\bar{y}}_n$ are:

(i) $\bar{\bar{y}}_n$ is an unbiased estimator of population mean μ with variance

(ii) $Var(\bar{\bar{y}}_n) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{[i]n}\, E(\frac{1}{v_i})$, where $\sigma^2_{[i]n} = Var(X^{(n)}_{[i]j})$.

To prove the above two properties, we need the following results: If $X^{(n)}_{[r]}$ is the r[th] judgment order statistic with sample size n, then it is clear that $X^{(n)}_{[r]} \overset{d}{=} X^{(n)}_{i[r]}$. If $f^{(n)}_{[r]}(x)$ is the pdf of $X^{(n)}_{[r]}$, then as noted by Dell and Clutter (1972)

$$f(x) = \frac{1}{n}\sum_{i=1}^{n} f^{(n)}_{[i]}(x). \tag{2}$$

It follows from (2) that if we denote $\mu = E(X)$, $Var(X) = \sigma^2$, $\mu^{(n)}_{[i]} = E(X^{(n)}_{[i]})$ and $\sigma^2_{[i]n} = Var(X^{(n)}_{[i]})$, then $\sum_{i=1}^{n}\mu^{(n)}_{[i]} = n\mu$. Now it is easy to show (i), since $X^{(n)}_{i[r]}$, i = 1, 2, …, $v_r$ are i i d and $E(X^{(n)}_{i[r]}) = \mu_{[r]}$, then

$$E(\bar{y}^{(n)}_{v_i}) = E[E[\frac{1}{v_i}\sum_{j=1}^{v_i}X^{(n)}_{j[i]}|v_i]] = E[\frac{1}{v_i}\sum_{j=1}^{v_i}\mu^{(n)}_{[i]}] = \mu^{(n)}_{[i]}. \tag{3}$$

Using the result of equation (3), we obtain that

$$E(\bar{\bar{y}}_n) = E[\frac{1}{n}\sum_{i=1}^{n}\bar{y}^{(n)}_{v_i}] = \frac{1}{n}\sum_{i=1}^{n}E[\bar{y}^{(n)}_{v_i}] = \frac{1}{n}\sum_{i=1}^{n}\mu^{(n)}_{[i]} = \mu.$$

This shows that $\bar{\bar{y}}_n$ is unbiased estimator of μ.

To find the variance of $\bar{\bar{y}}_n$, first we need to find $Var(\bar{y}^{(n)}_{v_i})$. Using equation (3) and the total probability formula we have

$$Var(\bar{y}^{(n)}_{v_i})=\sum_{\ell=2}^{\infty}\frac{1}{\ell^2}E[\sum_{j=1}^{\ell}(X^{(n)}_{j[i]}-\mu^{(n)}_{[i]})]^2P(v_i=\ell)$$

$$Var(\bar{y}^{(n)}_{v_i})=\sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{j=1,m=1}^{\ell}E[X^{(n)}_{j[i]}-\mu^{(n)}_{[i]}][X^{(n)}_{m[i]}-\mu^{(n)}_{[i]}]P(v_i=\ell)$$

$$Var(\bar{y}^{(n)}_{v_i})=\sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{j=1}^{\ell}\sigma^2_{[i]n}P(v_i=\ell)=\sigma^2_{[i]n}E(\frac{1}{v_i}) \tag{5}$$

Now using equation (5) and the independence of $v_i, i\geq 1$, we get

$$Var(\bar{\bar{y}}_n)=\frac{1}{n^2}\sum_{i=1}^{n}Var(\bar{y}^{(n)}_{v_i})=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{[i]n}E(\frac{1}{v_i}) \tag{6}$$

**Example 1.** Let $P(v_i=k)=\frac{1}{n-1}$, k = 2, 3, …, n. In this case we find that

$$Var(\bar{\bar{y}}_n)=\frac{1}{n^2(n-1)}\sum_{\ell=2}^{n}\frac{1}{\ell}\sum_{i=1}^{n}\sigma^2_{[i]n}. \tag{7}$$

From equation (2) it follows that

$$\sum_{i=1}^{n}\sigma^2_{[i]n}=n\sigma^2-\sum_{i=1}^{n}(\mu^{(n)}_{[i]}-\mu)^2, \tag{8}$$

an thus we have an alternative representation of

$$Var(\bar{\bar{y}}_n)=\frac{1}{n(n-1)}\sum_{\ell=2}^{n}\frac{1}{\ell}[\sigma^2-\frac{1}{n}\sum_{i=1}^{n}(\mu^{(n)}_{[i]}-\mu)^2]. \tag{9}$$

**Example 2.** Let $v_i, i \geq 1$, have the same binomial distribution with parameters n and p truncated at zero and one i.e. $P(v_i = k) = P(\xi = k | \xi > 1), k = 2,3,...,n$. In this case

$$P(v_i = k) = (1 - q^n - npq^{n-1})^{-1} \binom{n}{k} p^{k-1} q^{n-k}, k = 2,3,...,n, q = 1 - p$$

The variance of $Var(\overline{\overline{y}}_n)$ is

$$Var(\overline{\overline{y}}_n) = \frac{n! q^n}{n^2 (1 - q^n - npq^{n-1})} \sum_{i=1}^{n} \sigma_{[i]}^2 \sum_{\ell=2}^{n} \frac{1}{\ell! \ell (n - \ell)} \left( \frac{p}{q} \right)^{\ell} \quad (10)$$

In examples 1 and 2 we considered the case when the $v_i, i \geq 1$, are i.i.d. random variables. In the following two examples we will consider the case of different distributions for $v_i$.

**Example 3.** Let $v_i, i \geq 1$ have truncated binomial distribution with the parameters n and $p_i, 0 < p_i < 1$. In this case we obtain

$$E[\frac{1}{v_i}] = \frac{n! q_i^n}{1 - q_i^n - np_i q_i^{n-1}} \sum_{\ell=2}^{n} \frac{1}{\ell! (n - \ell)! \ell} \left( \frac{p_i}{q_i} \right)^{\ell}.$$

Thus the variance of $\overline{\overline{y}}_n$ can be written as

$$Var(\overline{\overline{y}}_n) = \frac{(n-1)!}{n} \sum_{i=1}^{n} \frac{q_i^n \sigma_{[i]n}^2 a_{ni}}{1 - q_i^n - np_i q_i^{n-1}}, \quad (11)$$

where $a_{in} = \sum_{\ell=2}^{n} \frac{1}{\ell! (n - \ell)! \ell} \left( \frac{p_i}{q_i} \right)^{\ell}$.

**Example 4.** It is not difficult to see that our model allows considering different configurations for the distribution of $v_i, i \geq 1$. As an example consider the case of two uniform distributions. Let $1 \leq r \leq n$ and $2 \leq m \leq n$ be two fixed numbers and

$v_i, 1 \le i \le r$ is uniformly distributed on $\{2,3,\ldots,m\}$ and $v_i, r+1 \le i \le n$, is uniformly distributed on $\{m+1, m+2, \ldots, n\}$. We find that

$$E\left\{\frac{1}{v_i}\right\} = \begin{cases} \dfrac{1}{m-1}\sum_{\ell=2}^{m}\dfrac{1}{\ell}, 1 \le i \le r \\ \dfrac{1}{n-m}\sum_{\ell=m+1}^{n}\dfrac{1}{\ell}, r+1 \le i \le n \end{cases}$$

Therefore the variance of $\overline{\overline{y}}_n$ can be found as

$$Var(\overline{\overline{y}}_n) = \frac{1}{n^2}\left\{\frac{1}{m-1}\sum_{\ell=2}^{m}\frac{1}{\ell}\sum_{i=1}^{r}\sigma^2_{[i]n} + \frac{1}{n-m}\sum_{\ell=m+1}^{n}\frac{1}{\ell}\sum_{i=m+1}^{n}\sigma^2_{[i]n}\right\}. \tag{12}$$

Note that the integers r and m are arbitrary numbers between 1 and n, one may choose them depending on the parent distribution i.e. the distribution of X to have smaller variance.

To compare $\overline{\overline{y}}_n$ as an estimator for the population mean to the estimator proposed by Dell and Clutter (1972) which is denoted by $\hat{\mu}_{rss} = \frac{1}{n}\sum_{i=1}^{n}\overline{X}_{[i]}, \overline{X}_{[i]} = \frac{1}{k_i}\sum_{j=1}^{k_i}X_{[i]j}$ we need only to compare the variance of $\overline{\overline{y}}_n$ given in equation (6) to $Var(\hat{\mu}_{rss}) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{[i]n}/k_i$.

Clearly we need to know the distribution(s) of $v_i$. As a special case let assume that $k_i = k$ for i=, 1, 2, …, n. Then the variance of $\hat{\mu}_{rss}$ can be written as

$$Var(\hat{\mu}_{rss}) = \frac{1}{n^2 k}\sum_{i=1}^{n}\sigma^2_{[i]}. \tag{13}$$

It follows from (7) and (13) that to compare variances of $\overline{\overline{y}}_n$ and $\hat{\mu}_{rss}$ we need to compare

$A_n = \frac{1}{n-1}\sum_{\ell=2}^{n}\frac{1}{\ell}$ with 1/k. Here we can use

**Proposition 1.** Let $\varepsilon_k = A_n - \frac{1}{k}, k = 2,3,\ldots$, then

(i) $\varepsilon_k > 0$ for $k > \frac{n(n+1)}{2n+1}$.

(ii) $\varepsilon_k < 0$ for $k < \frac{\sqrt{2n^3 + 2n^2 + 1} - 2n - 1}{n-2}$

The relative precision of $\overline{\overline{y}}_n$ with respect to $\hat{\mu}_{rss}$ is

$$RP(\overline{\overline{y}}_n, \hat{\mu}_{rss}) = \frac{Var(\hat{\mu}_{rss})}{Var(\overline{\overline{y}}_n)} = [\frac{k}{n-1}\sum_{\ell=2}^{n}\frac{1}{\ell}]^{-1}. \tag{14}$$

## 3. FIXED NUMBER OF REPLICATIONS WITH RANDOM SET SIZE OR NUMBER OF CLASSES

Let the number of classes or the set size to be a random variable θ taking values from $\Lambda = \{2,3,\ldots\}$ and $k_1, k_2, \ldots, k_\theta$ are fixed integers. First we select $k_1$ sets of size θ units, order the units within each set and select from each set the first judgment order statistic for quantification. Let

$$\overline{y}_{k_1}^{(\theta)} = \frac{1}{k_1}\sum_{j=1}^{k_1} X_{j[1]}^{(\theta)}$$

Next we select $k_2$ sets of size θ units, order each set and select from each set the second smallest judgment order statistic for quantification. Then we have

$$\overline{y}_{k_2}^{(\theta)} = \frac{1}{k_2}\sum_{j=1}^{k_2} X_{j[2]}^{(\theta)}.$$

We repeat the above process θ times to obtain $\overline{y}_{k_1}^{(\theta)}, \overline{y}_{k_2}^{(\theta)}, \ldots, \overline{y}_{k_\theta}^{(\theta)}$. We propose

$$\overline{\overline{y}}_\theta = \frac{1}{\theta}\sum_{i=1}^{\theta} \overline{y}_{k_i}^{(\theta)} \tag{15}$$

as an estimator for the population mean μ. The properties $\overline{\overline{y}}_\theta$ are

(i) $\overline{\overline{y}}_\theta$ is an unbiased estimator for μ,

(ii) with variance $Var(\overline{\overline{y}}_{(\theta)}) = E[\frac{1}{\theta^2}\sum_{i=1}^{\theta}\frac{1}{k_i}\sigma^2_{[i]\theta}]$.

First we show (i). Since $E[\overline{y}_{k_i}^{(\theta)}] = \frac{1}{k_i}\sum_{j=1}^{k_i} E[X_{j[i]}^{(\theta)}] = E[\mu_{[i]}^{(\theta)}], i = 1,2,\ldots,\theta$, and

$\sum_{i=1}^{\theta}\mu_{[i]}^{(\theta)} = \theta\mu$ for any $\theta$, then

$$E(\overline{\overline{y}}_\theta) = E[\frac{1}{\theta}\sum_{i=1}^{\theta}E(\overline{y}_{k_i}^{(\theta)}|\theta)] = E[\frac{1}{\theta}\sum_{i=1}^{\theta}\mu_{[i]}^{(\theta)}] = \mu.$$

Thus $\overline{\overline{y}}_\theta$ is unbiased estimator for μ. Now we will find the variance of $\overline{\overline{y}}_\theta$.

$$Var(\overline{\overline{y}}_\theta) = E[\overline{\overline{y}}_\theta - \mu]^2 = E[\frac{1}{\theta}\sum_{i=1}^{\theta}\overline{y}_{k_i}^{(\theta)} - \frac{1}{\theta}\sum_{i=1}^{\theta}\mu_{[i]}^{(\theta)}]^2 = E[\frac{1}{\theta}\sum_{i=1}^{\theta}(\overline{y}_{k_i}^{(\theta)} - \mu_{[i]}^{(\theta)})]^2,$$

$$Var(\overline{\overline{y}}_\theta) = \sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{i,j=1}^{\ell}[(\overline{y}_{k_i}^{(\ell)} - \mu_{[i]}^{(\ell)})(\overline{y}_{k_j}^{(\ell)} - \mu_{[j]}^{(\ell)})]P(\theta = \ell),$$

$$Var(\overline{\overline{y}}_\theta) = \sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{i=1}^{\ell}Var(\overline{y}_{k_i}^{(\ell)})P(\theta = \ell).$$

Since for any fixed $\ell$

$$Var(\overline{y}_{k_i}^{(\ell)}) = \frac{1}{k_i}\sigma^2_{[i]\ell}, \tag{16}$$

we obtain (ii) from the last equation.

**Example 5.** Let θ be random variable having a Poisson distribution with λ > 0 truncated at zero and one, i.e. $p(\theta = m) = P(\xi = m|\xi > 1), m \geq 2$, where ξ is a Poisson random variable of the parameter λ. In this case we obtain that

$$Var(\overline{\overline{y}}_\theta) = \frac{1}{\ell^\lambda - 1 - \lambda} \sum_{m=2}^{\infty} \frac{\lambda^m}{m^2 m!} \sum_{i=1}^{m} \frac{1}{k_i} \sigma^2_{[i]m} .$$

# 4. RANDOM NUMBER OF REPLICATIONS WITH RANDOM SET SIZE OR NUMBER OF CLASSES

Let θ, $\nu_1, \nu_2, \ldots$ are independent random variables taking values from $\Lambda = \{2,3,\ldots\}$. First we draw randomly $\nu_1$ sets of size θ, order the units within each set, and select the first judgment order statistics for quantification from each set. Next we select $\nu_2$ sets of size θ, order the units within each set, and select the second judgment order statistics for quantification from each set. We repeat this process θ times. Denote the sample mean of each step by

$$\overline{y}^{(\theta)}_{\nu_i} = \frac{1}{\nu_i} \sum_{j=1}^{\nu_i} X^{(\theta)}_{j[i]}, i = 1,2,\ldots,\theta .$$

Then we propose

$$\overline{\tilde{y}}_\theta = \frac{1}{\theta} \sum_{i=1}^{\theta} \overline{y}^{(\theta)}_{\nu_i} , \tag{17}$$

as an estimator for μ. We have from (3) that $E(\overline{y}^{(\theta)}_{\nu_i}) = E[\mu^{(\theta)}_{[i]}]$. Using this fact and total probability formula we can show

$$E(\overline{\tilde{y}}_\theta) = \sum_{\ell=2}^{\infty} \frac{1}{\ell} \sum_{i=1}^{\ell} \mu^{(\ell)}_{[i]} P(\theta = \ell) = \mu ,$$

i.e. $\overline{\tilde{y}}_\theta$ is unbiased estimator for μ. Now we evaluate the variance of $\overline{\tilde{y}}_\theta$. Since θ, $\nu_1, \nu_2, \ldots$ are independent, using the fact that $\sum_{i=1}^{\ell} \mu^{(\ell)}_{[i]} = \ell \mu$, and total probability formula, we have

$$Var(\overline{\overline{y}}_\theta) = E[\overline{\overline{y}}_\theta - \mu]^2 = E[\frac{1}{\theta}\sum_{i=1}^{\theta}(\overline{y}_{v_i}^{(\theta)} - \mu_{[i]}^{(\theta)})]^2,$$

$$Var(\overline{\overline{y}}_\theta) = \sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{i,j=1}^{\ell} E[(\overline{y}_{v_i}^{(\ell)} - \mu_{[i]}^{(\ell)})(\overline{y}_{v_j}^{(\ell)} - \mu_{[j]}^{(\ell)})]P(\theta = \ell),$$

$$Var(\overline{\overline{y}}_\theta) = \sum_{\ell=2}^{\infty}\frac{1}{\ell^2}\sum_{i=1}^{\ell} Var(\overline{y}_{v_i}^{(\ell)})P(\theta = \ell).$$

Using equation (5), we can write the variance of $\overline{\overline{y}}_\theta$ as

$$Var(\overline{\overline{y}}_\theta) = E[\frac{1}{\theta^2}\sum_{i=1}^{\theta}\sigma_{[i]\theta}^2 E(\frac{1}{v_i})]. \qquad (18)$$

**Example 6**. Let, as in Example 5, θ has a truncated Poisson distribution with λ > 0 and $v_i, i \geq 1$, have geometric distribution truncated at one with parameters $p_i, i \geq 1$. It is not difficult to see that in this case $E[\frac{1}{v_i}] = p_i q_i^{-2} T(q_i), T(q_i) = \sum_{\ell=2}^{\infty}\frac{1}{\ell}q_i^{\ell}$. Since $T'(q_i) = \frac{q_i}{1-q_i}$, we have $T(q_i) = \int_0^{q_i} T'(a)da = -[q_i + \ln p_i]$ and, consequent by

$$E[\frac{1}{v_i}] = -p_i q_i^{-2} T(q_i)[q_i + \ln p_i].$$

Thus we obtain from (18) the following formula for the variance of $\overline{\overline{y}}_\theta$

$$Var(\overline{\overline{y}}_\theta) = -(e^{\lambda} - 1 - \lambda)^{-1}\sum_{m=2}^{\infty}\frac{\lambda^m}{m^2 m!}\sum_{i=1}^{m}\sigma_{[i]m}^2 p_i q_i^2 [q_i + \ln p_i].$$

**Comparison**

Let θ = n be non-random and $v_1, v_2, \ldots$ be independent and identically distributed random variables taking values from $\Lambda = \{2, 3, \ldots\}$. In this case we obtain from (6)

$$Var(\overline{\overline{y}}_n) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_{[i]n}^2 E(\frac{1}{v_i}).$$

If $\hat{\mu}_{srs}$ is the simple random sampling (SRS) estimator for μ with sample size nk, then $Var(\hat{\mu}_{srs}) = \frac{\sigma^2}{nk}$. Using this result and the result of equation (8), we find the relative precision of $\overline{\overline{y}}_n$ with respect to $\hat{\mu}_{srs}$:

$$RP(\overline{\overline{y}}_n, \hat{\mu}_{srs}) = \frac{Var(\hat{\mu}_{srs})}{Var(\overline{\overline{y}}_n)} = \{kE(\frac{1}{v_i})[1 - \frac{1}{n\sigma^2}\sum_{i=1}^{n}(\mu_{[i]}^{(n)} - \mu)^2]\}^{-1} \quad (19)$$

If $v_i$ is random variable taking values from {2, 3, …, n) and uniformly distributed, then (19) will become

$$RP(\overline{\overline{y}}_n, \hat{\mu}_{srs}) = \{\frac{k}{n-1}\sum_{\ell=2}^{n}\frac{1}{\ell}[1 - \frac{1}{n\sigma^2}\sum_{i=1}^{n}(\mu_{[i]}^{(n)} - \mu)^2]\}^{-1}. \quad (20)$$

Let $\overline{\overline{y}}_{nk}$ be estimator with fixed number k quantification in each cycle, then the relative precision of $\overline{\overline{y}}_n$ with respect to $\overline{\overline{y}}_{nk}$ is $RP = \{kE(\frac{1}{v_i})\}^{-1}$.

## 5. ASYMPTOTIC PROPERTIES

We now consider the asymptotic properties of the estimator $\overline{\overline{y}}_n$ as $n \to \infty$. We will prove that under some natural conditions $\overline{\overline{y}}_n$ is asymptotically normal. We assume that $\sigma_{[i]n}^2 = Var(X_{[i]}^{(n)})$ and $\beta_{[i]n}^3 = E\left|X_{[i]}^{(n)} - \mu_{[i]}^{(n)}\right|^3$, where $\mu_{[i]}^{(n)} = E[X_{[i]}^{(n)}]$, are finite for each $i, n \geq 1$. Note that in the case of perfect ranking the above assumptions are satisfied if $\sigma^2 = Var(X)$ and $B^3 = E|X|^3$ are finite. We also need the following conditions

C1. $\lim_{n\to\infty}(nB_n)^{-3}\sum_{i=1}^{n}E[\frac{1}{v_i^2}]\beta_{[i]n}^3 = 0$;

where $B_n^2 = Var(\overline{\overline{y}}_n) = n^{-2}\sum_{i=1}^{n} E[\frac{1}{v_i}]\sigma_{[i]n}^2$.

C2. $\underset{n\to\infty}{Lim}(nB_n)^{-2}\sum_{i=1}^{n} E\left|\frac{1}{v_i} - E(\frac{1}{v_i})\right|\sigma_{[i]n}^2 = 0;$

C3. $\underset{n\to\infty}{Lim}(nB_n)^{-2}\underset{1\le i\le n}{Max}\,\sigma_{[i]n}^2 = 0$.

**Theorem**

If conditions C1 – C3 are satisfied, then $B_n^{-1}(\overline{\overline{y}}_n - \mu)$ is asymptotically normal i.e. as $n \to \infty$

$$\sup_x \left| P\{B_n^{-1}(\overline{\overline{y}}_n - \mu) \le x\} - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt \right| \to 0 .$$

Note that condition (C1) is a version of so called Lyapuhov condition for asymptotic normality of a sum of independent random variables. Condition (C2) may be considered as a condition for the variance of $1/v_i$. In fact it is hard to expect asymptotic normality of $\overline{\overline{y}}_n$ without a condition of "smallity" of deviations of $1/v_i$ from their expectations.

**Proof.** We denote $S_n = B_n^{-1}(\overline{\overline{y}}_n - \mu)$. Using the independence of the random variables $v_i, X_{[i]j}^{(n)}; i, j, n \ge 1$, by the total probability formula we obtain the following formula for the characteristic function of $S_n$:

$$E[e^{itS_n}] = E[\prod_{i=1}^{n} f_{in}^{v_i}(\frac{t}{nB_n v_i})], \tag{21}$$

where $f_{in}(t) = E[e^{itv_{i1}}]$, $v_{ij}^{(n)} = X_{[i]j}^{(n)} - \mu_{[i]}^{(n)}, j \ge 1$. We now consider

$$\varphi_n(t) = \prod_{i=1}^{n} f_{in}^{v_i}(\frac{t}{nB_n v_i})$$

for arbitrary but fixed $t \in R$. First we prove that as $n \to \infty$

$$I_1=\sum_{i=1}^{n}\nu_i[f_{in}(\frac{t}{nB_n\nu_i})-1]\xrightarrow{P}-\frac{t^2}{2} \qquad (22)$$

where P means convergence in probability. To do it we use the following representation for $f_{in}(t)$

$$f_{in}(t)=1-\frac{t^2}{2}\sigma_{[i]n}^2-i\frac{t^3}{6}[E[\nu_{ij}^{(n)^3}]+\varepsilon_{in}(t)] \qquad (23)$$

for $j\geq 1$, where $|\varepsilon_{in}(t)|\leq 3\beta_{[i]n}^3$ and $\varepsilon_{in}(t)\to 0$ as $t\to 0$. If we use (23) it is not difficult to see that

$$I_1=-\frac{t^2}{2}+\frac{t^2}{2n^2B_n^2}\sum_{i=1}^{n}(E[\frac{1}{\nu_i}]-\frac{1}{\nu_i})\sigma_{[i]n}^2+\frac{it^3}{6}I_2, \qquad (24)$$

where $I_2=\frac{1}{n^3B_n^3}\sum_{i=1}^{n}\frac{1}{\nu_i^2}[E(\nu_{ij}^{(n)^3})+\varepsilon_{in}(\frac{t}{nB_n\nu_i})]$. Using the Chebyshev inequality and condition (C2) we obtain that as $n\to\infty$ the second term in (24) converges in probability to zero. Now we consider $I_2$, taking into the account the estimate for $\varepsilon_{in}(t)$ and using the Chebyshev inequality again we have for any δ

$$P\{|I_2|>\delta\}\leq\frac{4}{\delta n^3B_n^3}\sum_{i=1}^{n}E[\frac{1}{\nu_i^2}]\beta_{[i]n}^3.$$

This together with condition (C1) yields that the last term of (24) is converges in probability to zero as $n\to\infty$ for each fixed t. Thus relation (22) is proved.

It follows from representation

$$\ln(1+z)=z-\frac{z^2}{2}+\frac{z^3}{3}-\ldots,|z|<1$$

That $|\ln(1+z)-z|\leq|z|^2$ for any complex z such that $|z|\leq 1/2$. Note that here $\ln(z)$ denotes the principal value of the logarithm, i.e., $\ln(z)=\ln|z|+i\arg z$,

$-\pi < \arg z \leq \pi$. To use the last estimate for $\varphi_n(t)$ we need to prove that for sufficiently large n and each fixed t

$$\max_{1\leq i\leq n}\left|f_{in}\left(\frac{t}{nB_n v_i}\right)-1\right| \leq \frac{1}{2} \tag{25}$$

Now we use the representation

$$f_{in}(t)=1-\frac{t^2}{2}[\sigma_{[i]n}^2+\varepsilon_{in}^{(1)}(t)], \tag{26}$$

where $\left|\varepsilon_{in}^{(1)}(t)\right| \leq 3\sigma_{[i]n}^2$. From (26) we get for any $i,n\geq 1$

$$\left|f_{in}\left(\frac{t}{nB_n v_i}\right)-1\right| \leq 2t^2\frac{\sigma_{[i]n}^2}{(nB_n v_i)^2}. \tag{27}$$

It follows from the last inequality that under the condition (C3) as $n\to\infty$ for each fixed t

$$\max_{1\leq i\leq n}\left|f_{in}\left(\frac{t}{nB_n v_i}\right)-1\right| \to 0. \tag{28}$$

Consequently (25) holds for sufficiently large n and for each fixed t.

Now we will prove that as $n\to\infty$

$$I_3=\sum_{i=1}^{n} v_i\left(f_{in}\left(\frac{t}{nB_n v_i}\right)-1\right)^2 \xrightarrow{P} 0. \tag{29}$$

To show that we use (27) and obtain the following estimate

$$|I_3| \leq \theta_n^{(1)}\theta_n^{(2)}.$$

where $\theta_n^{(1)}=\dfrac{2t^2}{n^2B_n^2}\sum_{i=1}^{n}\dfrac{1}{v_i}\sigma_{[i]n}^2$ and $\theta_n^{(2)}=\dfrac{2t^2}{n^2B_n^2}\max_{1\leq i\leq n}\sigma_{[i]n}^2$. Note that here $E[\theta_n^{(1)}]=2t^2$. Thus for any $\delta>0$

$$P\{|I_3|>\delta\} \leq P\{\theta_n^{(1)}\theta_n^{(2)}>\delta\} \leq 2t^2\theta_n^{(2)}$$

As we can see (29) follows from this and condition (C3). From (22) and (29) we conclude that

$$\varphi_n(t) = e^{\ln(\varphi_n(t)} \xrightarrow{\ p\ } e^{-t^2/2},$$

as $n \to \infty$. Since $|\varphi_n(t)| \leq 1, n \geq 1$ by dominated convergence theorem (see Shiryaev (1996), Theorem 3, p.187 and remark on p. 258), we obtain from the last that $E[\varphi_n(t)] \to e^{-t^2/2}$, $n \to \infty$, which completes the prove of the theorem.

## REFERENCES


Dell, D. R. and Clutter, J. L. (1972). "Ranked Set Sampling Theory with Order Statistics Background," *Biometrics*, 28, 545-55.

Kaur, A., Patil, G. P., Sinha, B. K. and Taillie, C. (1995), "Ranked Set Sampling: an Annotated Bibliography," *Environmental and Ecological Statistics*, 2, 25-54.

McIntyre, G. A. (1952), "A Method of Unbiased Selective Sampling, using Ranked Sets," *Australian Journal of Agricultural Research*, 3, 385-390.

Muttlak, H. A. and McDonald, L. L. (1992), "Ranked Set Sampling and the Line Intercept Method: A more Efficient procedure," *Biometrical Journal*, 34, 329-46.

Muttlak, H. A. (1997), "Median Ranked Set Sampling," J*ournal of Applied Statistical Science*, 6(4), 245-55.

Patil, G. P., Sinha, A. K. and Taillie, C. (1994), "Ranked Set Sampling," *A Handbook of Statistics*, vol. 12, pp167-200.

Patil, G. P., Sinha, A. K. and Taillie, C. (1999), "Ranked Set Sampling: A Bibliography," *Environmental and Ecological Statistics*, 6, 91-98.

Rahimov, I. And Muttlak, H. A., (2000), "Random Ranked Set Samples," Technical Report no. 256, King Fahd University of Petroleum & Minerals,.

Samawi, H. M., Ahmed, M. S. and Abu-Dayyeh (1996), "Estimating the Population Mean using Extreme Ranked Set Sampling," *Biometrical Journal*, 38(5), 577-86.

Shiryaev, A. N. (1996), *Probability, 2end*: Springer, New York.

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,* 21, 249-55.