CHAPTER 21

Ranked Set Sampling

21.1 INTRODUCTION

Ranked set sampling (RSS) was introduced by McIntyre (1952) to estimate the mean pasture and forage yield. The RSS is used when precise measurement of the variable of interest is difficult or expensive but the variable can be ranked easily without measuring the actual variable by an inexpensive method such as visual perception, judgment, and auxiliary information. For example, in estimating the mean height of trees in a forest, the heights of a small sample of two or three trees standing nearby can be ranked easily by visual inspection without measuring them. In estimating the number of bacterial cells per unit volume, we can rearrange two or three test tubes easily in the order of concentration using optical instruments without measuring exact values. In RSS, instead of selecting a single sample of size m, we select m sets of samples each of size m. In each of the sets, all the elements are ranked but only one is measured. Finally, an average of the m measured units is taken as an estimate of the population mean. The sample mean based on RSS is unbiased for the population mean regardless of the errors of ranking. The RSS mean is at least as precise as the sample mean of the simple random sampling with replacement (SRSWR) sampling scheme of the same size. Stokes (1980a,b, 1988) showed that RSS provides precise estimators for cumulative distribution function (cdf), population variance, and correlation coefficient.

21.2 RANKED SET SAMPLING BY SIMPLE RANDOM SAMPLING WITH REPLACEMENT METHOD

First, we choose a small number m (set size) such that one can easily rank m elements of the population with sufficient accuracy. Then the selection procedure of RSS is as follows: Select a sample of m^2 units from a population U by the SRSWR method. Allocate these m^2 units at random into m sets each of size m. Rank all the units in a set with respect to the values of the variable of interest y from 1 (minimum) to m (maximum) by a very inexpensive method such as eye inspection. No actual measurement is done

Survey Sampling Theory and Applications
ISBN 978-0-12-811848-1
http://dx.doi.org/10.1016/B978-0-12-811848-1.00021-2

at this stage. After the ranking has been completed, the unit holding rank 1 of the set 1, unit holding rank 2 of the set 2,…, and finally the unit holding rank m of the set m are measured accurately by using a suitable instrument. This completes a cycle of the sampling. The process is repeated for r cycles to obtain the desired sample of size $n = mr$ units. Thus in RSS, a total of m^2r units have been drawn from the population, but only mr of them are measured and the rest $mr(m-1)$ are discarded. These measured mr observations are called a "ranked set sample." Because the ordering of a large number of observations is difficult, increase in sample size $n(=mr)$ is done by increasing the number of cycles r. Let $y_{i1|k},\ldots, y_{ij|k},\ldots, y_{im|k}$ be the value of the variable of interest y of the ith set of elements of the kth cycle, $i = 1,\ldots, m$; $k = 1,\ldots, r$. Furthermore, let $y_{i(j)|k}$ be the smallest jth observation (order statistic) of $y_{i1|k},\ldots, y_{ij|k},\ldots, y_{im|k}$. From the ith set the ith order statistic is measured, i.e., the value of $y_{i(i)|k}$ is obtained. This can be represented as follows.

Set	**Cycle *k* after rearrangement of *y* values**					**Observed *y***
1	$y_{1(1)\|k}$	$\cdots$	$y_{1(i)\|k}$	$\cdots$	$y_{1(m)\|k}$	$y_{1(1)\|k}$
2	$y_{2(1)\|k}$	$\cdots$	$y_{2(i)\|k}$	$\cdots$	$y_{2(m)\|k}$	$y_{2(2)\|k}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
i	$y_{i(1)\|k}$	$\cdots$	$y_{i(i)\|k}$	$\cdots$	$y_{i(m)\|k}$	$y_{i(i)\|k}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
m	$y_{m(1)\|k}$	$\cdots$	$y_{m(i)\|k}$	$\cdots$	$y_{m(m)\|k}$	$y_{m(m)\|k}$

Thus we have the following data:

$$d = \left\{ \overbrace{y_{1(1)|1} \quad y_{2(2)|1} \quad \cdots \quad y_{m(m)|1}}^{\text{cycle } 1};\ldots;\overbrace{y_{1(1)|k} \quad y_{2(2)|k} \quad \cdots \quad y_{m(m)|k}}^{\text{cycle } k};\ldots; \right.$$

$$\left. \overbrace{y_{1(1)|r} \quad y_{2(2)|r} \quad \cdots \quad y_{m(m)|r}}^{\text{cycle } r} \right\}$$

21.2.1 A Fundamental Equality

Let $y_{i1|k},\ldots, y_{ij|k},\ldots, y_{im|k}$ be a random sample from a population with cdf $F(y)$ and probability density function (pdf) $f(y)$. Let the mean and variance of y be μ and σ^2, respectively. Then we have the following equalities:

$$\sum_{j=1}^{m} y_{ij|k} = \sum_{j=1}^{m} y_{i(j)|k} \tag{21.2.1}$$

and

$$\sum_{j=1}^{m} \left(y_{ij|k} - \mu\right)^2 = \sum_{j=1}^{m} \left(y_{i(j)|k} - \mu\right)^2 \tag{21.2.2}$$

Let $\mu_{(j)|m} = E(y_{i(j)|k})$ be the mean of the jth order statistic of a random sample of size m of the cycle k, which depends on m but is independent of the set i and the cycle k.

Eq. (21.2.1) yields

$$E\left(\frac{1}{m}\sum_{j=1}^{m} y_{ij|k}\right) = E\left(\frac{1}{m}\sum_{j=1}^{m} y_{i(j)|k}\right)$$

$$\text{i.e.,}\quad \mu = \frac{1}{m}\sum_{j=1}^{m} \mu_{(j)|m} \tag{21.2.3}$$

Similarly, Eq. (21.2.2) yields

$$\sum_{j=1}^{m} E\left(y_{ij|k} - \mu\right)^2 = \sum_{j=1}^{m} E\left(y_{i(j)|k} - \mu\right)^2$$

$$\text{i.e.,}\quad m\sigma^2 = \sum_{j=1}^{m}\left\{\sigma^2_{(j)|m} + \left(\mu_{(j)|m} - \mu\right)^2\right\}$$

where $\sigma^2_{(j)|m} =$ variance of $y_{i(j)|m}$.

$$\text{i.e.,}\quad \sigma^2 = \frac{1}{m}\sum_{j=1}^{m} \sigma^2_{(j)|m} + \frac{1}{m}\sum_{i=1}^{m}\left(\mu_{(j)|m} - \mu\right)^2 \tag{21.2.4}$$

21.2.2 Estimation of the Mean

Let $\overline{y}_{[m]|k} = \frac{1}{m}\sum_{i=1}^{m} y_{i(i)|k} =$ arithmetic mean of the m quantified values of the variable y for the cycle k and

$$\widehat{\mu}_{rss} = \frac{1}{r}\sum_{k=1}^{r} \overline{y}_{[m]|k} = \frac{1}{n}\sum_{k=1}^{r}\sum_{i=1}^{m} y_{i(i)|k} \tag{21.2.5}$$

is the mean of $n(=mr)$ quantified variables based on all the r cycles. The following theorem shows that the estimator $\widehat{\mu}_{rss}$ is unbiased for μ and possesses a lower variance than $\widehat{\mu}_{srs}$, the sample mean based on an SRSWR sample of the same size n.

Theorem 21.2.1

(i) $E(\widehat{\mu}_{rss}) = \mu$

(ii) $V(\widehat{\mu}_{rss}) = \dfrac{\sigma^2_{[m]}}{n} = \dfrac{1}{n}\left[\sigma^2 - \dfrac{1}{m}\sum_{i=1}^{m}\left(\mu_{(j)|m} - \mu\right)^2\right] \le \sigma^2/n$

(iii) An unbiased estimator of the variance of $V(\widehat{\mu}_{rss})$ is

$$\widehat{V}(\widehat{\mu}_{rss}) = \frac{1}{r(r-1)}\sum_{k=1}^{r}\left(\overline{y}_{[m]|k} - \widehat{\mu}_{rss}\right)^2.$$

where $\sigma^2_{[m]} = \dfrac{1}{m}\sum_{j=1}^{m}\sigma^2_{(j)|m}$.

Proof

(i) $E(\widehat{\mu}_{rss}) = \dfrac{1}{r}E\left(\sum_{k=1}^{r}\overline{y}_{[m]|k}\right)$

Eq. (21.2.1) yields

$$E\left(\overline{y}_{[m]|k}\right) = \frac{1}{m}\sum_{i=1}^{m}E\left(y_{i(i)|k}\right) = \frac{1}{m}\sum_{i=1}^{m}\mu_{(i)|m} = \mu$$

Hence, $E(\widehat{\mu}_{rss}) = \mu$.

(ii) Because $\overline{y}_{[m]|1}, \ldots, \overline{y}_{[m]|r}$ are iid random variables, we have

$$V(\widehat{\mu}_{rss}) = \frac{1}{r^2}\sum_{k=1}^{r}V\left(\overline{y}_{[m]|k}\right)$$
$$= \frac{V\left(\overline{y}_{[m]|k}\right)}{r} \qquad (21.2.6)$$

Now

$$V\left(\overline{y}_{[m]|k}\right) = \frac{1}{m^2}\left[\sum_{i=1}^{m}V\left(y_{i(i)|k}\right)\right]$$
$$= \frac{1}{m^2}\left[\sum_{i=1}^{m}\sigma^2_{(i)|m)}\right]$$
$$= \frac{1}{m}\left[\sigma^2 - \frac{1}{m}\sum_{i=1}^{m}\left(\mu_{(i)|m} - \mu\right)^2\right]$$

(using Eq. 21.2.4).

$$\text{Hence}\quad V(\widehat{\mu}_{rss}) = \frac{1}{n}\left[\sigma^2 - \frac{1}{m}\sum_{i=1}^{m}\left(\mu_{(i)|m} - \mu\right)^2\right] \le \sigma^2/n \qquad (21.2.7)$$

(iii) The result follows from the fact that $\overline{y}_{[m]|k}$ are distributed independently.

21.2.3 Precision of the Ranked Set Sampling

The relative precision of $\widehat{\mu}_{rss}$ compared with $\widehat{\mu}_{srs}$, sample mean of an SRSWR sample of size n, is

$$RP = \frac{V(\widehat{\mu}_{srs})}{V(\widehat{\mu}_{rss})} = \frac{\sigma^2}{\sigma^2_{[m]}} \tag{21.2.8}$$

Example 21.2.1: Uniform Distribution

Let $X_1,\ldots, X_m$ be a random sample from a uniform distribution over (0,1). In this case $E(X_i) = \mu = 1/2$ and $\sigma^2 = V(X_i) = 1/12$. The distribution of the jth order statistic $X_{(j,m)}$ has the density

$$f_{(j,m)} = \begin{cases} \dfrac{\Gamma(m+1)}{\Gamma j\,\Gamma(m-j+1)} \quad x^{j-1}(1-x)^{m-j} & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases} \tag{21.2.9}$$

The variance of $X_{(j,m)} = \sigma^2_{(j)|m} = \dfrac{(m+1-j)j}{(m+1)^2(m+2)}$ and $\dfrac{1}{m^2}\sum_{j=1}^{m}\sigma^2_{(j)|m} = \dfrac{1}{6m(m+1)}$. Hence the relative efficiency of $\widehat{\mu}_{rss}$ with respect to $\widehat{\mu}_{srs}$ is

$$RP = \frac{V(\widehat{\mu}_{srs})}{V(\widehat{\mu}_{rss})} = \frac{1/(12mr)}{1/\{6m(m+1)r\}} = \frac{m+1}{2}$$

Example 21.2.2: Exponential Distribution

Following Takahasi and Wakimoto (1968), we find for the exponential distribution $\sigma^2_{(j)|m} = \sigma^2 \sum_{k=1}^{m} \dfrac{1}{(m-k+1)^2}$ and $\sigma^2_{[m]} = \dfrac{\sigma^2}{m}\sum_{j=1}^{m}\dfrac{1}{j}$. Hence the relative precision is

$$RP = 1 \Big/ \sum_{j=1}^{m}\frac{1}{j}$$

The exact expressions of the relative precisions of the most of distributions including gamma, normal, Weibull, and double exponential distribution are not simple. Takahasi and Wakimoto (1968) derived the relative precisions for specific values of m viz. 2, 3, 4, etc. McIntyre (1952) computed the values of RP for several populations and concluded that the upper bound of RP is $(m+1)/2$. Takahasi and Wakimoto (1968) proved that for the class of continuous distribution, the maximum value of RP is $(m+1)/2$ and the upper bound is attained by the uniform distribution.

However, the lower bound, in general, cannot be improved on zero. McIntyre (1952) reported that RP is very small for symmetric distributions.

21.2.4 Optimum Value of *m*

In a ranked set sampling we select samples in the form of r cycles, each consisting of m units. We need to find the optimum number of cycles r and the number of elements per cycles m, which minimize the variance of $\widehat{\mu}_{rss}$, keeping the total sample size $n(=mr)$ fixed. Let $\widehat{\mu}_{rss}(\widetilde{m},\widetilde{r})$ denote the mean of the ranked set sample of $\widetilde{r}$ cycles with $\widetilde{m}$ units so that $n=\widetilde{m}\widetilde{r}$. Then from the Theorem 21.2.1, we find the difference

$$\begin{aligned}
V\{\widehat{\mu}_{rss}(m,r)\}-V\{\widehat{\mu}_{rss}(\widetilde{m},\widetilde{r})\} &= \frac{1}{r^2}\sum_{k=1}^{r}V\left(\overline{y}_{[m]|k}\right)-\frac{1}{\widetilde{r}^2}\sum_{k=1}^{\widetilde{r}}V\left(\overline{y}_{[\widetilde{m}]|k}\right)\\
&= \frac{1}{rm^2}\left(\sum_{i=1}^{m}\sigma^2_{(i)|m}\right)-\frac{1}{\widetilde{r}\,\widetilde{m}^2}\left(\sum_{i=1}^{\widetilde{m}}\sigma^2_{(i)|\widetilde{m}}\right)\\
&= \frac{1}{n}\left(\sigma^2_{[m]}-\sigma^2_{[\widetilde{m}]}\right)
\end{aligned} \tag{21.2.10}$$

where $\sigma^2_{[m]}=\dfrac{1}{m}\displaystyle\sum_{i=1}^{m}\sigma^2_{(i,m)}$ and $\sigma^2_{[\widetilde{m}]}=\dfrac{1}{\widetilde{m}}\displaystyle\sum_{i=1}^{\widetilde{m}}\sigma^2_{(i,\widetilde{m})}$.

First we will prove the following result due to Takahasi and Wakimoto (1968).

Theorem 21.2.2

$$\sigma^2_{[m+1]}<\sigma^2_{[m]}$$

Proof

Let $X_1,\ldots,X_m$ be a random sample from a population with cdf $F(x)$ and density function $f(x)$, then the density function of jth order statistic $X_{(j,m)}$ is given by

$$f_{(j,m)}(x)=\frac{\Gamma(m+1)}{\Gamma j\,\Gamma(m-j+1)}F^{j-1}(x)\{1-F(x)\}^{m-j}f(x)\quad \text{for } j=1,\ldots,m \tag{21.2.11}$$

Similarly, the density function for the tth order statistic $X_{(t,m+1)}$, based on a random sample $(m+1)$ observation, is

$$f_{(t,m+1)}(x)=\frac{\Gamma(m+2)}{\Gamma t\,\Gamma(m-t+2)}F^{t-1}(x)\{1-F(x)\}^{m+1-t}f(x)\quad \text{for}\quad t=1,\ldots,m+1 \tag{21.2.12}$$

From Eqs. (21.2.11) and (21.2.12), we have

$$f_{(j,m)}(x) = \frac{m+1-j}{m+1} f_{(j,m+1)}(x) + \frac{j}{m+1} f_{(j+1,m+1)}(x) \tag{21.2.13}$$

Eq. (21.2.13) yields

$$\begin{aligned} \mu_{(j)|m} &= E\{X_{(j,m)}\} \\ &= \frac{m+1-j}{m+1} E\{X_{(j,m+1)}\} + \frac{j}{m+1} E\{X_{(j+1,m+1)}\} \\ &= \frac{m+1-j}{m+1} \mu_{(j)|m+1} + \frac{j}{m+1} \mu_{(j+1)|m+1} \end{aligned} \tag{21.2.14}$$

Similarly, the variance of $X_{(j,m)}$ is

$$\begin{aligned} \sigma^2_{(j,m)} &= E\left\{X^2_{(j,m)}\right\} - \mu^2_{(j,m)} \\ &= \frac{m+1-j}{m+1}\sigma^2_{(j)|m+1} + \frac{j}{m+1}\sigma^2_{(j+1)|m+1} + \left\{\frac{m+1-j}{m+1}\mu^2_{(j)|m+1}\right. \\ &\quad \left. + \frac{j}{m+1}\mu^2_{(j+1)|m+1} - \left(\frac{m+1-j}{m+1}\mu_{(j)|m+1} + \frac{j}{m+1}\mu_{(j+1)|m+1}\right)^2\right\} \end{aligned} \tag{21.2.15}$$

Eq. (21.2.15) yields

$$\begin{aligned} \sum_{j=1}^{m}\sigma^2_{(j)|m} &= \frac{1}{(m+1)}\left\{\sum_{j=1}^{m+1}(m+1-j)\sigma^2_{(j)|m+1)} + \sum_{j=1}^{m+1}(j-1)\sigma^2_{(j)|m+1}\right\} \\ &\quad + \frac{1}{m+1}\left\{\sum_{j=1}^{m}(m+1-j)\mu^2_{(j)|m+1} + \sum_{j=1}^{m} j\mu^2_{(j+1)|m+1)}\right. \\ &\quad \left. - \frac{1}{(m+1)}\sum_{j=1}^{m}\left((m+1-j)\mu_{(j)|m+1} + j\mu_{(j+1)|m+1)}\right)^2\right\} \\ &= \frac{m}{(m+1)}\sum_{j=1}^{m+1}\sigma^2_{(j)|m+1} \\ &\quad + \frac{1}{(m+1)^2}\sum_{j=1}^{m} j(m+1-j)\left(\mu_{(j+1)|m+1} - \mu_{(j)|m+1}\right)^2 \end{aligned}$$

Thus,

$$\sigma^2_{[m]} - \sigma^2_{[m+1]} = \frac{1}{m(m+1)^2} \sum_{j=1}^{m} j(m+1-j)\left(\mu_{(j+1)|m+1} - \mu_{(j)|m+1}\right)^2 > 0$$

This completes the proof of the theorem.

From the Theorem 21.2.2, we arrive at

Theorem 21.2.3

$V\{\widehat{\mu}_{rss}(m,r)\} - V\{\widehat{\mu}_{rss}(s,t)\} > 0$ for $s > m$ with $n = rm = st$

From Theorem 21.2.3, we note that efficiency $\widehat{\mu}_{rss}(m,r)$ increases with m when the total sample size n is fixed and maximum efficiency is achieved when $m = n$ and $r = 1$. However, a large $m(= n)$ is impractical because the cost of ordering a large number of elements will be expensive. In most practical cases m is taken as 3 or 4.

21.2.5 Optimum Allocation

In our description of RSS in Section 21.2, each of the elements holding rank 1, rank 2,..., and rank m were measured only once in each of the r cycles. Instead of measuring each of the elements holding rank 1, rank 2 ... and rank m an equal number of times (r), we may measure the element holding rank j, $r_j(\geq 1)$ times with $\sum_{j=1}^{m} r_j = n$. In this situation an unbiased estimator of the population mean μ is

$$\widehat{\mu} = \frac{1}{m}\sum_{j=1}^{m} \overline{y}_{(j)|m} \tag{21.2.16}$$

where $\overline{y}_{(j)|m} = \frac{1}{r_j}\sum_{k=1}^{r_j} y_{j(j)|k}$ and $y_{j(j)|k}$ represent the actual value of the jth order statistic of a random sample of m elements in the kth cycle. In case $\sigma_{(j)|m}$ is known, the optimum value of r_j is obtained using Neyman's optimum allocation as

$$r_{j0} = n\frac{\sigma_{(j)|m}}{\sum_{j=1}^{m} \sigma_{(j)|m}} \tag{21.2.17}$$

where $\sigma^2_{(j)|m} = V\left(y_{j(j)|k}\right)$.

The variance of $\widehat{\mu}$ with the optimum value of $r_j = r_{j0}$ is

$$V_{opt}(\widehat{\mu}) = \frac{\left(\sum_{j=1}^{m} \sigma_{(j)|m}\right)^2}{m^2 n} \tag{21.2.18}$$

21.2.5.1 Right-Tail Allocation Model

In practice, the implementation of the optimum allocation is not possible because $\sigma_{(j)|m}$ values are generally unknown. In this case it is of interest to study "near" optimal allocation where full knowledge of the value $\sigma_{(j)|m}$ is not required. It is well known for a positively skewed population that the variance of the order statistic increases with rank order: $\sigma^2_{(1)|m} \le \sigma^2_{(2)|m} \le \ldots \le \sigma^2_{(m)|m}$. For highly skewed distributions, the variance of the highest order statistic $\sigma^2_{(m)|m}$ is very high compared with the variances of the rest of the order statistics. Keeping this in mind, Kaur et al. (1997) proposed the following t-allocation

$$r' = r_1 = r_2 = \cdots = r_{m-1} = r_m/t \tag{21.2.19}$$

For $n = (m-1+t)r'$, the variance of $\widehat{\mu}$ under t-allocation is

$$\begin{aligned} V(\widehat{\mu}_t) = \frac{1}{m^2}\sum_{j=1}^{m}\frac{\sigma^2_{(j)|m}}{r_j} &= \frac{1}{m^2}\left(\sum_{j=1}^{m-1}\frac{\sigma^2_{(j)|m}}{r'} + \frac{\sigma^2_{(m)|m}}{tr'}\right) \\ &= \frac{1}{m^2 r'}\left(d' + \sigma^2_{(m)|m}\big/t\right) \end{aligned} \tag{21.2.20}$$

where $d' = \sum_{j=1}^{m-1} \sigma^2_{(j)|m}$.

The relative precision of $\widehat{\mu}_t$ compared with $\widehat{\mu}_{srs}$ is

$$RP_t = \frac{m^2 r'\sigma^2/n}{d' + \sigma^2_{(m)|m}\big/t} = \frac{m^2\sigma^2}{(m-1+t)\left(d' + \sigma^2_{(m)|m}\big/t\right)} \tag{21.2.21}$$

Under equal allocation $r_1 = r_2 = \cdots = r_m = r$, $\widehat{\mu}_t$ reduces to $\widehat{\mu}_e = \widehat{\mu} = \frac{1}{mr}\sum_{j=1}^{m}\sum_{k=1}^{r} y_{j(j)|k}$ and the relative precision of $\widehat{\mu}_t$ compared with equal allocation is

$$RP_{e,t} = \frac{m^2 r'\left(\frac{1}{m^2 r}\sum_{j=1}^{m}\sigma^2_{(j)|m}\right)}{\left(d' + \sigma^2_{(m)|m}\big/t\right)} = \frac{m^2}{(m-1+t)}\frac{\left(\frac{1}{m}\sum_{j=1}^{m}\sigma^2_{(j)|m}\right)}{\left(d' + \sigma^2_{(m)|m}\big/t\right)} \tag{21.2.22}$$

On maximizing Eq. (21.2.22) the optimum value of t is obtained as

$$t_{opt} = \sqrt{\frac{\sigma^2_{(m)|m}(m-1)}{d'}} \tag{21.2.23}$$

It can be shown that RP_t is a monotonically increasing function of t for $t=(1,t_{opt})$ and decreasing for $t=(t_{opt},\infty)$. On the basis of empirical studies based on of 225 positively skewed distributions belonging to different parametric populations including inverse Gaussian, reciprocal gamma, lognormal, Pareto and beta (type II) distribution, Kaur et al. (1997) provided guidelines for selection of the appropriate value of t. For details readers are referred to Kaur et al. (1997).

21.2.6 Judgment Ranking

In judgment ranking, each of the selected samples is ranked by a visual or judgment process, which may include the use of concomitant variable. Hence sometimes ranking may be imperfect. Let $y_{i\langle j\rangle|k}$ be the smallest jth "judgment order statistic" in the ith set of the cycle k. In case the judgment ranking is perfect, $y_{i\langle j\rangle|k}$ becomes equal to $y_{i(j)|k}$, otherwise if the judgment process is imperfect, we find $y_{i\langle j\rangle|k} \neq y_{i(j)|k}$. Here the selected data are $d^* = (y_{i\langle i\rangle|k};\ i=1,\ldots,\ m,\ k=1,\ldots,\ r)$.

21.2.6.1 Moment of the Judgment Order Statistic

For any constants c and p, the following relations hold

$$\sum_{i=1}^{m}\left(y_{ij|k}-c\right)^p = \sum_{i=1}^{m}\left(y_{i\langle j\rangle|k}-c\right)^p \quad \text{for } j=1,\ldots,m;\ k=1,\ldots,r \tag{21.2.24}$$

From Eq. (21.2.24) it follows that

$$m\,E\left(y_{ij|k}-c\right)^p = \sum_{i=1}^{m} E\left(y_{i\langle j\rangle|k}-c\right)^p$$

Substituting $p=1$, $p=2$, and $c=\mu$, we get

$$\mu = \frac{1}{m}\sum_{j=1}^{m}\mu_{\langle j\rangle|m} \quad \text{and } \sigma^2 = \frac{1}{m}\sum_{j=1}^{m}\sigma^2_{\langle j\rangle|m} + \frac{1}{m}\sum_{j=1}^{m}\tau^2_{\langle j\rangle|m} \tag{21.2.25}$$

where $\mu_{\langle j\rangle|m} = E\left(y_{i\langle j\rangle|k}\right)$, $\sigma^2_{\langle j\rangle|m} = V\left(y_{i\langle j\rangle|k}\right)$, and $\tau_{\langle j\rangle|m} = \mu_{\langle j\rangle|m} - \mu$.

Because (i) $\overline{y}_{\langle m\rangle|k} = \frac{1}{m}\sum_{i=1}^{m} y_{i\langle i\rangle|k}$ are independent random variables with mean $E\left(\overline{y}_{\langle m\rangle|k}\right) = \mu$ and variance $V\left(\overline{y}_{\langle m\rangle|k}\right) = \frac{1}{m}\sum_{i=1}^{m}\sigma^2_{\langle j\rangle|m}$, we find the following theorem analogous to Theorem 21.2.1.

Theorem 21.2.4

(i) $\widehat{\mu}_{\langle rss\rangle} = \frac{1}{r}\sum_{k=1}^{r}\overline{y}_{\langle m\rangle|k}$ is an unbiased estimator for μ.

(ii) The variance of $\widehat{\mu}_{\langle rss\rangle}$ is

$$V\left(\widehat{\mu}_{\langle rss\rangle}\right) = \frac{1}{n}\left[\sigma^2 - \frac{1}{m}\sum_{i=1}^{m}\left(\mu_{\langle j\rangle|m} - \mu\right)^2\right]$$

(iii) An unbiased estimator of the variance of $V\left(\widehat{\mu}_{\langle rss\rangle}\right)$ is

$$\widehat{V}\left(\widehat{\mu}_{\langle rss\rangle}\right) = \frac{1}{r(r-1)}\sum_{k=1}^{r}\left(\overline{y}_{\langle m\rangle|k} - \widehat{\mu}_{\langle rss\rangle}\right)^2$$

Consider the extreme situation where making a judgment is impossible. So, random ranking is assigned to all the elements of this set. In this case $y_{i\langle j\rangle|k}$ is a random sample from the original population and $\widehat{\mu}_{\langle rss\rangle}$ remains unbiased as it becomes equal to $\widehat{\mu}_{srs}$. In practice, ranking ability is expected to be between perfect and extreme random ranking. Actually, the error in ranking has little effect on the precision of the estimator. Dell and Clutter (1972) conducted simulation studies on the effect of errors in ranking of various populations including rectangular, exponential, and normal populations. They showed that the relative precision of $\widehat{\mu}_{\langle srs\rangle}$ with respect to $\widehat{\mu}_{srs}$ ranges between 1.05 and 3.00.

21.2.7 Estimation of Population Variance

It is well known that for simple random sampling, the sample variance is an unbiased estimator of the population variance σ^2. Stokes (1980a) showed that for RSS the sample variance

$$\widehat{\sigma}^2_{\langle rss\rangle} = \frac{1}{mr-1}\sum_{k=1}^{r}\sum_{j=1}^{m}\left(y_{j,\langle j\rangle|k} - \widehat{\mu}_{\langle rss\rangle}\right)^2$$

is not an unbiased estimator of σ^2.

Theorem 21.2.5

$\widehat{\sigma}^2_{\langle rss\rangle}$ is a biased estimator of σ^2 and the amount of bias of $\widehat{\sigma}^2_{\langle rss\rangle}$ is

$$\frac{1}{m(mr-1)}\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m)}-\mu\right)^2.$$

Proof

$$\begin{aligned}
E\left(\widehat{\sigma}^2_{\langle rss\rangle}\right) &= \frac{1}{mr-1}\left[\sum_{k=1}^{r}\sum_{j=1}^{m}E\left(y^2_{j,\langle j\rangle|k}\right)-mrE\left(\widehat{\mu}_{\langle rss\rangle}\right)^2\right]\\
&= \frac{1}{mr-1}\left[\sum_{k=1}^{r}\sum_{j=1}^{m}\left\{\sigma^2_{\langle j\rangle|m)}+\mu^2_{\langle j\rangle|m}\right\}-mr\left\{V\left(\widehat{\mu}_{\langle rss\rangle}\right)+\mu^2\right\}\right]\\
&= \frac{1}{mr-1}\left[r\sum_{j=1}^{m}\sigma^2_{\langle j\rangle|m)}-\frac{1}{m}\sum_{j=1}^{m}\sigma^2_{\langle j\rangle|m)}+r\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m}-\mu\right)^2\right]\\
&= \frac{1}{mr-1}\left[\frac{(mr-1)}{m}\sum_{j=1}^{m}\sigma^2_{\langle j\rangle|m)}+r\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m}-\mu\right)^2\right] \qquad (21.2.26)
\end{aligned}$$

Now using Eq. (21.2.25) we find

$$E\left(\widehat{\sigma}^2_{\langle rss\rangle}\right)=\sigma^2+\frac{1}{m(mr-1)}\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m)}-\mu\right)^2$$

21.2.7.1 *Efficiency of* $\widehat{\sigma}^2_{\langle rss\rangle}$

The estimator $\widehat{\sigma}^2_{\langle rss\rangle}$ is asymptotically unbiased for σ^2 if either r or m becomes large. Eq. (21.2.25) shows that the upper bound of the bias of $\widehat{\sigma}^2_{\langle rss\rangle}$ is $\sigma^2/(mr-1)$. Stokes (1980a) reported that RSS may not always provide an efficient estimator for σ^2. The variance of $\widehat{\sigma}^2_{\langle rss\rangle}$ was given by Stokes (1980a) as

$$\begin{aligned}
V\left(\widehat{\sigma}^2_{\langle rss\rangle}\right) &= \frac{r}{(mr-1)^2}\left[\left(\frac{mr-1}{rm}\right)^2\sum_{j=1}^{m}\mu_{\langle 4j\rangle|m}+4\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m}-\mu\right)^2\sigma^2_{\langle j\rangle|m)}\right.\\
&+4\left(\frac{mr-1}{mr}\right)\sum_{j=1}^{m}\left(\mu_{\langle j\rangle|m}-\mu\right)\mu_{\langle 3j\rangle|m}+\frac{2r}{(mr)^2}\sum_{j\neq}^{m}\sum_{k=1}^{m}\sigma^2_{\langle j\rangle|m}\sigma^2_{\langle k\rangle|m}\\
&\left.+\frac{2(r-1)-(mr-1)^2}{(mr)^2}\sum_{j=1}^{m}\sigma^4_{\langle j\rangle|m}\right]
\end{aligned} \qquad (21.2.27)$$

where $\mu_{\langle pj\rangle|m)}=E(y_{j\langle j\rangle|m}-\mu_{\langle j\rangle|m)})^p$.

In case the ranking is totally imperfect in the sense that it is random, $\mu_{\langle pj\rangle|m} = E(y-\mu)^p = \mu_p$ and Eq. (21.2.27) yields

$$V\left(\widehat{\sigma}^2_{\langle rss\rangle}\right) = \frac{\mu_4}{mr} - \frac{(mr-3)}{mr(mr-1)}\sigma^4$$
$$= V(s^2)$$

where s^2 is the sample variance based on rm observations.

Stokes (1980a) proved that there exists N^* such that the relative precision of $\widehat{\sigma}^2_{\langle rss\rangle}$,

$$RP\left(\widehat{\sigma}^2_{\langle rss\rangle}\right) = \frac{V(s^2)}{V\left(\widehat{\sigma}^2_{\langle rss\rangle}\right)} \geq 1 \text{ if } mr > N^* \tag{21.2.28}$$

Equality in Eq. (21.2.28) holds if and only if $E(y_{j\langle j\rangle|m} - \mu)^2 = E(y-\mu)^2$ for all $j = 1, \ldots, m$. The increased precision can be realized by increasing either m or the number of cycles r. Because a substantial increase in m is impractical, the asymptotic results hold for large r.

21.2.8 Use of Concomitant Variables

Stokes (1977) considered the situation where the study variable y cannot be ordered easily but a related variable x that is observable and can be easily ordered accurately. For example, y be the internal characteristic of a patient or a laboratory animal, which is expensive, painful, or inconvenient to measure. In this case, RSS can be employed if judgment ordering can be accomplished on the y's by ordering some external characteristic x, which is correlated to y. So, in this case we have data of judgment ordering

$$d_J = \left\{ \overbrace{y_{1\langle 1\rangle|1} \quad y_{2\langle 2\rangle|1} \quad \cdots \quad y_{m\langle m\rangle|1}}^{\text{cycle } 1}; \ldots; \overbrace{y_{1\langle 1\rangle|k} \quad y_{2\langle 2\rangle|k} \quad \cdots \quad y_{m\langle m\rangle|k}}^{\text{cycle } k}; \ldots; \overbrace{y_{1\langle 1\rangle|r} \quad y_{2\langle 2\rangle|r} \quad \cdots \quad y_{m\langle m\rangle|r}}^{\text{cycle } r} \right\} \tag{21.2.29}$$

where $y_{i\langle j\rangle|k}$ be the smallest jth "judgment order statistic" of y corresponding to the smallest jth "order statistic" $x_{i(j)|k}$ of x in the ith set of the cycle k.

Let $\mu_{x(i)|m} = E(x_{i(i)|k})$, $\mu_{y(i)|m} = E(y_{i(i)|k})$ be the mean of the ith order statistic of x and y in a random sample of size m. The population means and variances of x and y will be denoted by μ_x,μ_y and σ_x^2, σ_y^2, respectively. The proposed estimator for the population mean μ_y is

$$\widehat{\mu}_{y\langle rss\rangle} = \frac{1}{r}\sum_{k=1}^{r} \overline{y}_{\langle m\rangle|k} \tag{21.2.30}$$

where $\overline{y}_{\langle m\rangle|k} = \dfrac{1}{m}\sum_{i=1}^{m} y_{i\langle i\rangle|k}$.

Stokes (1977) assumed that the regression of y on x is linear, i.e.,

$$y = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x) + \in \tag{21.2.31}$$

where x and $\in$ are independent, $E(\in|x) = 0$, $V(\in|x) = \sigma_y^2(1-\rho^2)$, and ρ is the correlation coefficient between x and y.

Then we have from Eq. (21.2.31)

$$y_{i\langle i\rangle|k} = \mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(x_{i(i)|k} - \mu_x\right) + \in_{i(i)|k}$$

Stokes further assumed that $\dfrac{y-\mu_y}{\sigma_y}$ and $\dfrac{x-\mu_x}{\sigma_x}$ have same distribution that holds for bivariate normal and bivariate Pareto distributions. Stokes (1977) derived the following theorem:

Theorem 21.2.6

(i) $E\left(\widehat{\mu}_{y\langle rss\rangle}\right) = \mu_y$

(ii) $V\left(\widehat{\mu}_{y\langle rss\rangle}\right) = \dfrac{1}{mr}\left[\sigma_y^2 - \dfrac{\rho^2}{m}\sum_{i=1}^{m}\tau^2_{y(i)|m}\right]$

where $\tau_{y(i)|m} = E(y_{(i)|m}) - \mu_y$.

Proof

(i)
$$E\left(\overline{y}_{\langle m\rangle|k}\right) = \frac{1}{m}\sum_{i=1}^{m} E\left(y_{i\langle i\rangle|k}\middle|x_{i(i)|k}\right)$$

$$= \mu_y + \rho\frac{\sigma_y}{\sigma_x}\frac{1}{m}\sum_{i=1}^{m} E\left(x_{i(i)|k} - \mu_x\right)$$

$$= \mu_y + \rho\frac{\sigma_y}{\sigma_x}\frac{1}{m}\sum_{i=1}^{m}\left(\mu_{x(i)|m} - \mu_x\right)$$

$$= \mu_y$$

$$
\begin{aligned}
\text{(ii)}\quad V\left(\widehat{\mu}_{y\langle rss\rangle}\right) &= \frac{V\left(\overline{y}_{\langle m\rangle|k}\right)}{r}\\
&= \frac{1}{m^2 r}\sum_{i=1}^{m}\left[E\left\{V\left(y_{i\langle i\rangle|k}\middle|x_{i(i)|k}\right)\right\} + V\left\{E\left(y_{i\langle i\rangle|k}\middle|x_{i(i)|k}\right)\right\}\right]\\
&= \frac{1}{m^2 r}\sum_{i=1}^{m}\left[\sigma_y^2(1-\rho^2) + \rho^2\frac{\sigma_y^2}{\sigma_x^2}V\left(\mu_{x(i)|m} - \mu_x\right)\right]
\end{aligned}
\tag{21.2.32}
$$

Now using the assumption that $\dfrac{y-\mu}{\sigma_y}$ and $\dfrac{x-\mu_x}{\sigma_x}$ have same distribution, we find

$$
\begin{aligned}
\sum_{i=1}^{m} V\left(\mu_{x(i)|m} - \mu_x\right) &= \frac{\sigma_x^2}{\sigma_y^2}\sum_{i=1}^{m} V\left(\mu_{y(i)|m} - \mu_y\right)\\
&= \frac{\sigma_x^2}{\sigma_y^2}\left(m\sigma_y^2 - \sum_{i=1}^{m}\tau_{y(i)|m}^2\right) \quad \text{(using Eq. 21.2.4)}
\end{aligned}
\tag{21.2.33}
$$

Part (ii) of theorem follows from Eqs. (21.2.32) and (21.2.33).

So, the relative efficiency of the ranking with the concomitant variable compared with the perfect ranking is

$$
\frac{V\left(\overline{y}_{(rss)}\right)}{V\left(\widehat{\mu}_{y\langle rss\rangle}\right)} = \frac{1 - \dfrac{1}{m}\displaystyle\sum_{i=1}^{m}\tau_{y(i)|m}^2}{1 - \dfrac{\rho^2}{m}\displaystyle\sum_{i=1}^{m}\tau_{y(i)|m}^2}
$$

Likewise, the efficiency of $\widehat{\mu}_{y\langle rss\rangle}$ with $\widehat{\mu}_{srs}$ is

$$
\frac{V\left(\overline{y}_{srs}\right)}{V\left(\widehat{\mu}_{y\langle rss\rangle}\right)} = \frac{1}{1 - \dfrac{\rho^2}{m}\displaystyle\sum_{i=1}^{m}\tau_{y(i)|m}^2}.
$$

Stokes (1977) used concomitant variable x for ranking y only. Yu and Lam (1997) used the variable x for estimation of μ_y. They proposed the following difference estimator when μ_x, the population mean of the concomitant variable x, is known.

$$
\widehat{\mu}_y^{(d)} = \widehat{\mu}_{y\langle rss\rangle} - B\left(\widehat{\mu}_{x(rss)} - \widehat{\mu}_x\right)
$$

where $\widehat{\mu}_{x(rss)} = \dfrac{1}{mr}\displaystyle\sum_{k=1}^{r}\sum_{i=1}^{m} x_{i(i)|k}$ and B is a constant to be determined.

The estimator $\widehat{\mu}_y^{(d)}$ is unbiased for μ, as $\widehat{\mu}_{\langle rss\rangle}$ and $\widehat{\mu}_{x(rss)}$ are unbiased for μ_y and μ_x, respectively. The optimum value of B is obtained by minimizing

$$\begin{aligned}V\left(\widehat{\mu}_y^{(d)}\right) &= V\left(\widehat{\mu}_{y\langle rss\rangle}\right) + B^2 V\left(\widehat{\mu}_{x(rss)}\right) - 2BCov\left(\widehat{\mu}_{y\langle rss\rangle}, \widehat{\mu}_{x(rss)}\right)\\ &= V\left(\widehat{\mu}_{y\langle rss\rangle}\right) + B^2 V\left(\widehat{\mu}_{x(rss)}\right) - 2B\rho\frac{\sigma_y}{\sigma_x} V\left(\widehat{\mu}_{x(rss)}\right)\end{aligned} \quad (21.2.34)$$

where

$$\begin{aligned}V\left(\widehat{\mu}_{y\langle rss\rangle}\right) &= \frac{1}{mr}\left[\sigma_y^2 - \frac{\rho^2}{m}\sum_{i=1}^{m}\tau_{y(i)|m}^2\right] \quad \text{(using Theorem 21.2.6)}\\ &= \frac{\sigma_y^2}{mr}\left[1 - \frac{\rho^2}{m}\sum_{i=1}^{m}\left(\frac{\tau_{x(i)|m}}{\sigma_x}\right)^2\right],\end{aligned} \quad (21.2.35)$$

$$V\left(\widehat{\mu}_{x(rss)}\right) = \frac{1}{mr}\sigma_x^2\left[1 - \frac{1}{m}\sum_{i=1}^{m}\left(\frac{\tau_{x(i)|m}}{\sigma_x}\right)^2\right] \quad \text{(using Theorem 21.2.1)}$$

$$(21.2.36)$$

and $\tau_{x(i)|m} = \mu_{x(i)|m} - \mu_x$.

Eq. (21.2.34) yields the optimum value of B as

$$B^* = \rho\frac{\sigma_y}{\sigma_x}$$

A natural estimator of B^* is

$$\widehat{B} = \frac{\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)\left(y_{i\langle i\rangle|k} - \widehat{\mu}_{y\langle rss\rangle}\right)}{\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)^2}$$

Hence, RSS regression estimator of μ_y is defined as

$$\widehat{\mu}_{reg} = \widehat{\mu}_{y\langle rss\rangle} - \widehat{B}(\widehat{\mu}_{x(rss)} - \mu_x) \quad (21.2.37)$$

Theorem 21.2.7

Under the model (21.2.31),

(i) $\mathrm{E}\left(\widehat{\mu}_{reg}\right) = \mu_y$,

(ii) $\mathrm{V}\left(\widehat{\mu}_{reg}\right) = \frac{\sigma_y^2}{mr}(1-\rho^2)\left[1 + E\left(\frac{\overline{z}_{rss}^2}{s_z^2}\right)\right]$

where $\overline{z}_{rss} = \frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m} z_{i(i)|k}$, $s_z^2 = \frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m}\left(z_{i(i)|k} - \overline{z}_{rss}\right)^2$, and $z_{i(i)|k} = \frac{x_{i(i)|k} - \mu_x}{\sigma_x}$.

Proof

(i) $$\mathrm{E}\left(\widehat{\mu}_{reg}\right) = E\left\{\widehat{\mu}_{y\langle rss\rangle} - \widehat{B}\left(\widehat{\mu}_{x(rss)} - \mu_x\right)\right\}$$
$$= \mu_y - E\left[\left(\widehat{\mu}_{x(rss)} - \mu_x\right)\left\{E\left(\widehat{B}\middle|x\right)\right\}\right]$$
$$= \mu_y - BE\left(\widehat{\mu}_{x(rss)} - \mu_x\right)$$
$$= \mu_y$$

(ii) $$V\left(\widehat{\mu}_{reg}\right) = E\left(\widehat{\mu}_{reg} - \mu_y\right)^2$$
$$= E\left\{-\left(\widehat{B} - B\right)\left(\widehat{\mu}_{x(rss)} - \mu_x\right) + \overline{\in}_{rss}\right\}^2$$
$$\left(\text{where } \overline{\in}_{rss} = \frac{1}{rm}\sum_{k=1}^{r}\sum_{i=1}^{m}\in_{i(i)|k}\right)$$
$$= E\left[\left(\widehat{\mu}_{x(rss)} - \mu_x\right)^2 E\left\{\left(\widehat{B} - B\right)^2\middle|x\right\} + E(\overline{\in}_{rss})^2\right.$$
$$\left. - 2E\left\{\left(\widehat{B} - B\right)\left(\widehat{\mu}_{x(rss)} - \mu_x\right)\overline{\in}_{rss}|x\right\}\right]$$
$$= E\left[\left(\widehat{\mu}_{x(rss)} - \mu_x\right)^2 \frac{E\left\{\left(\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)\in_{i(i)|r}\right)^2\middle|x\right\}}{\left(\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)^2\right)^2} + E(\overline{\in}_{rss})^2\right]$$
$$= \sigma_y^2(1-\rho^2)E\left[\frac{\left(\widehat{\mu}_{x(rss)} - \mu_x\right)^2}{\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)^2} + \frac{1}{mr}\right]$$
$$= \frac{\sigma_y^2}{mr}(1-\rho^2)\left[1 + E\left\{\frac{\left(\widehat{\mu}_{x(rss)} - \mu_\times\right)^2}{\frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m}\left(x_{i(i)|k} - \widehat{\mu}_{x(rss)}\right)^2}\right\}\right]$$
$$= \frac{\sigma_y^2}{mr}(1-\rho^2)\left[1 + E\left(\frac{\overline{z}_{rss}^2}{s_z^2}\right)\right]$$

21.2.8.1 Relative Precision of $\widehat{\mu}_{reg}$

The relative precision of the RSS regression estimator with respect to the RSS naïve estimator $\widehat{\mu}_{(rss)}$ based on the model (21.2.31) is

$$RP\left(\widehat{\mu}_{reg}, \widehat{\mu}_{\langle rss\rangle}\right) = \frac{1 - \rho^2 E\left(\frac{1}{m}\sum_{i=1}^{n} z^2_{(i)r}\right)}{(1-\rho^2)\left[1 + E\left(\frac{\overline{z}^2_{rss}}{s^2_z}\right)\right]}$$

Because $\widehat{\mu}_{\langle rss\rangle}$ does not use any information on the concomitant variable x, a fairer comparison of the RSS regression estimator $\widehat{\mu}_{reg}$ is to compare with the naïve regression estimator based on SRSWR sampling defined by

$$\widehat{\mu}^*_{reg} = \overline{y}_{srs} - \widehat{\beta}(\overline{x}_{srs} - \mu_x)$$

where $\overline{x}_{srs}, \overline{y}_{srs}$, and $\widehat{\beta}$ denote the sample means of x,y and the sample regression coefficient based on an SRSWR sample of size $n(= mr)$. For the bivariate normal distribution $\widehat{\mu}^*_{reg}$ is unbiased for μ_y and the variance of $\widehat{\mu}^*_{reg}$ is $V\left(\widehat{\mu}^*_{reg}\right) = \frac{\sigma^2_y}{mr}(1-\rho^2)\left(1 + \frac{1}{mr-3}\right)$ (see Sukhatme & Sukhatme, 1970).

Hence the relative efficiency of $\widehat{\mu}_{reg}$ with respect to $\widehat{\mu}^*_{reg}$ is defined as follows:

$$RP\left(\widehat{\mu}_{reg}, \widehat{\mu}^*_{reg}\right) = \frac{V\left(\widehat{\mu}^*_{reg}\right)}{V\left(\widehat{\mu}_{reg}\right)}$$

$$= \frac{1 + \frac{1}{mr-3}}{1 + E\left(\frac{\overline{z}^2_{rss}}{s^2_z}\right)}$$

21.3 SIMPLE RANDOM SAMPLING WITHOUT REPLACEMENT

RSS for SRSWOR sampling was proposed by Patil et al. (1995). The method is described as follows. Consider a finite population $U = (U_1,\ldots, U_i,\ldots, U_N)$ of N identifiable units and let y_i be the value of the variable under study, y, for the ith unit. Let us assume $y_1 < y_2 < \cdots < y_N$. From the population U, a sample s of $n = m^2r$ units is selected by the SRSWOR method. The sample s is partitioned at random into mr sets each containing m distinct units. All the units in a set are ranked by some inexpensive procedure. The lowest rank holder is quantified in each of the first r sets:

$$y_{(1,m)1}, \ldots, y_{(1,m)i}, \ldots, y_{(1,m)r}$$

The second ranked unit is quantified for each of the next r sets to yield:

$$y_{(2,m)1}, \ldots, y_{(2,m)i}, \ldots, y_{(2,m)r}$$

The process is continued until the highest ranked unit is quantified for each of the last r sets:

$$y_{(m,m)1}, \ldots, y_{(m,m)i}, \ldots, y_{(m,m)r}$$

The ranked set estimator for the finite population mean $\overline{Y} = \sum_{i=1}^{N} y_i/N$ is given by

$$\widehat{\overline{Y}}_{rss} = \frac{1}{rm}\sum_{\alpha=1}^{r}\sum_{i=1}^{m} y_{(i,m)\alpha} \tag{21.3.1}$$

Let us consider a particular set s_α, consisting of m units. The set s_α may be considered as an SRSWOR sample of size m from the population U. Let us define the event

$$\{i \rightarrow t\} \tag{21.3.2}$$

to denote that the ith ranked unit in the set s_α is the tth ranked unit in the population U. We also define the probability of the event as

$$A_i^t = \Pr\{i \rightarrow t\} \tag{21.3.3}$$

Theorem 21.3.1

(i) $A_i^t = \dfrac{\binom{t-1}{i-1}\binom{N-t}{m-i}}{\binom{N}{m}}$

(ii) The inclusion probability of the unit t in a set is m/N.

Proof

(i) The event $\{i \rightarrow t\}$ will hold if exactly $i - 1$ units are selected from the smallest $t - 1$ members of the population U and $m - i$ must be selected from the $N - t$ units of the population holding rank greater than t. So, probability of the event is A_i^t.

(ii) The inclusion probability of the unit t in a set is

$$\sum_{i=1}^{m} A_i^t = \sum_{i=1}^{m} \frac{\binom{t-1}{i-1}\binom{N-t}{m-i}}{\binom{N}{m}}$$

$$= m/N$$

Let us define the event that the ith ranked unit from the set s_α has rank t in the population and jth ranked unit from the set $s_\beta(\beta \neq \alpha)$ has rank l in the population as

$$\{i \rightarrow t, j \rightarrow l\}$$

The probability of the event $\{i \rightarrow t, j \rightarrow l\}$ is denoted by

$$B_{ij}^{t,l} = \Pr\{i \rightarrow t, j \rightarrow l\}$$

Let $\mathbf{B}_{ij}$ be the $N \times N$ matrix with $B_{ij}^{t,l}$ as its (t,l) component. Clearly, $\mathbf{B}_{ij} = \mathbf{B}_{ji}^T$ as $B_{ij}^{t,l} = B_{ji}^{l,t}$. Patil et al. (1995) derived the following theorem.

Theorem 21.3.2

(i)

$$B_{ij}^{t,l} = \begin{cases} \displaystyle\sum_{\lambda=0}^{m-i} \frac{\binom{t-1}{i-1}\binom{l-t-1}{\lambda}\binom{N-l}{m-i-\lambda}\binom{l-1-i-\lambda}{j-1}\binom{N-l-m+i+\lambda}{m-j}}{\binom{N}{m,m}} & \text{for } t < l \\ 0 & \text{for t} = l \\ B_{ji}^{l,t} & \text{for t} > l \end{cases}$$

(ii) The probability of inclusion of tth unit in a set α and lth in a set β is

$$\frac{m^2}{N(N-1)}(1-\delta_{tl})$$

where $\binom{N}{m,m} = \dfrac{N!}{m!m!(N-2m)!}$, λ is the number of units in set α that lies between y_t and y_l, $\delta_{tl} = 1$ if $t = l$, and $\delta_{tl} = 0$ if $t \neq l$.

Proof

(i) The proof follows from the similar argument in the Theorem 21.3.1. In fact λ must satisfy the following restrictions:

$$0 \le \lambda \le l - t - 1; \ \ 2m - i - j + l - N \le \lambda \le m - i; \ \ \lambda \le l - i - j$$

(ii) Probability of inclusion of the tth unit in set α and lth unit in set the β is

$$\sum_{i=1}^{m} \sum_{j=1}^{m} B_{ij}^{tl} = \frac{m^2}{N(N-1)} \text{ for } l > t \tag{21.3.4}$$

and for $t = l$, $\sum_{i=1}^{m} \sum_{j=1}^{m} B_{ij}^{tt} = 0$ because $B_{ij}^{tt} = 0$ for $t = 1, \ldots, N$.

Theorem 21.3.3

(i) $\mu_{(i,m)} = E\left[y_{(i,m)\alpha}\right] = \mathbf{A}_i^T \mathbf{y}$

(ii) $\sigma_{(i,m)}^2 = V\left[y_{(i,m)\alpha}\right] = \mathbf{A}_i^T \mathbf{y}^{(2)} - \left(\mathbf{A}_i^T \mathbf{y}\right)^2$

(iii) $\sigma_{(ij,m)} = Cov\left[y_{(i,m)\alpha}, y_{(j,m)\beta}\right] = \mathbf{y}^T\left(\mathbf{B}_{ij} - \mathbf{A}_i \mathbf{A}_i^T\right)\mathbf{y}$

where $\mathbf{A}_i^T = \left(A_i^1, \ldots, A_i^t, \ldots, A_i^N\right)$ and $\mathbf{y}^{(2)} = \left(y_1^2, \ldots, y_i^2, \ldots, y_N^2\right)^T$.

Proof

(i)

$$\begin{aligned}
\mu_{(i,m)} = E\left[y_{(i,m)\alpha}\right] &= \sum_{t=1}^{N} y_t \Pr\left(y_{(i,m)\alpha} = y_t\right) \\
&= \sum_{t=1}^{N} y_t \Pr(\{i \rightarrow t\}) \\
&= \sum_{t=1}^{N} y_t A_i^t \\
&= \mathbf{A}_i^T \mathbf{y}
\end{aligned} \tag{21.3.5}$$

(ii)

$$\begin{aligned}
\sigma_{(i,m)}^2 &= V\left[y_{(i,m)\alpha}\right] = E\left[y_{(i,m)\alpha}\right]^2 - \mu_{(i,m)}^2 \\
&= \sum_{t=1}^{N} y_t^2 \Pr\left(y_{(i,m)j} = y_t\right) - \mu_{(i,m)}^2 \\
&= \sum_{t=1}^{N} y_t^2 A_i^t - \mu_{(i,m)}^2 \\
&= \mathbf{A}_i^T \mathbf{y}^{(2)} - \left(\mathbf{A}_i^T \mathbf{y}\right)^2
\end{aligned}$$

(iii) $$\begin{aligned}\sigma_{(ij,m)} &= Cov\left[y_{(i,m)\alpha}, y_{(j,m)\beta}\right]\\ &= E\left[y_{(i,m)\alpha}y_{(j,m)\beta}\right] - \mu_{(i,m)}\mu_{(j,m)}\\ &= \sum_{t=1}^{N}\sum_{l=1}^{N} y_t y_l B_{ij}^{tl} - \mu_{(i,m)}\mu_{(j,m)}\\ &= \mathbf{y}^T\mathbf{B}_{ij}\mathbf{y} - \left(\mathbf{A}_i^T\mathbf{y}\right)\left(\mathbf{A}_j^T\mathbf{y}\right)\\ &= \mathbf{y}^T\left(\mathbf{B}_{ij} - \mathbf{A}_i\mathbf{A}_i^T\right)\mathbf{y}\end{aligned}$$

Theorem 21.3.4

(i) $\widehat{\overline{Y}}_{rss} = \frac{1}{rm}\sum_{\alpha=1}^{r}\sum_{i=1}^{m} y_{(i,m)\alpha}$ is unbiased for $\overline{Y}$

(ii) $V\left(\widehat{\overline{Y}}_{rss}\right) = \left\{\frac{1}{rm}\;\frac{N-1-mr}{N-1}\sigma_y^2 - \frac{1}{m}(\mathbf{y}-\boldsymbol{\mu})^T\left(\sum_{i=1}^{m}\mathbf{B}_{ii}\right)(\mathbf{y}-\boldsymbol{\mu})\right\}$

where $\sigma_y^2 = \sum_{i=1}^{N}\left(y_i - \overline{Y}\right)^2/N$, $\boldsymbol{\mu}^T = (\mu,\ldots\mu)$ and $\mu = \overline{Y}$.

Proof

(i) $$\begin{aligned}E\left(\widehat{\overline{Y}}_{rss}\right) &= \frac{1}{rm}\sum_{\alpha=1}^{r}\sum_{i=1}^{m} E\left(y_{(i,m)\alpha}\right)\\ &= \frac{1}{rm}\sum_{\alpha=1}^{r}\sum_{i=1}^{m}\mathbf{A}_i^T\mathbf{y}\\ &= \frac{1}{rm}\sum_{\alpha=1}^{r}\sum_{i=1}^{m}\sum_{t=1}^{N} A_i^t y_t\\ &= \frac{1}{m}\sum_{t=1}^{N}\left(\sum_{i=1}^{m} A_i^t\right)y_t\\ &= \frac{1}{N}\sum_{t=1}^{N} y_t \quad \text{(using Theorem 21.3.1)}\\ &= \overline{Y} = \mu\end{aligned}$$

(ii) $$V\left(\widehat{\overline{Y}}_{rss}\right)=\frac{1}{(rm)^2}V\left(\sum_{\alpha=1}^{r}\sum_{i=1}^{m}y_{(i,m)\alpha}\right)$$

$$=\frac{1}{(rm)^2}\left\{\sum_{\alpha=1}^{r}V\left(\sum_{i=1}^{m}y_{(i,m)\alpha}\right)\right.$$

$$\left.+\sum_{\alpha\neq}^{r}\sum_{\alpha'=1}^{r}Cov\left(\sum_{i=1}^{m}y_{(i,m)\alpha},\sum_{i=1}^{m}y_{(i,m)\alpha'}\right)\right\}$$

Now, writing $V\left(y_{(i,m)\alpha}\right)=\sigma^2_{(i,m)}$, $Cov\left(y_{(i,m)\alpha},y_{(j,m)\alpha}\right)=\sigma^2_{(ij,m)}$ for $i\neq j$, $Cov\left(y_{(i,m)\alpha},y_{(i,m)\alpha'}\right)=\sigma^2_{(ii,m)}$ for $\alpha\neq\alpha'$, and $Cov\left(y_{(i,m)\alpha},y_{(j,m)\alpha'}\right)=\sigma^2_{(ij,m)}$ for $\alpha\neq\alpha'$, we have

$$V\left(\widehat{\overline{Y}}_{rss}\right)=\frac{1}{(rm)^2}\left\{\sum_{\alpha=1}^{r}\left(\sum_{i=1}^{m}\sigma^2_{(i,m)}+\sum_{i\neq}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)}\right)\right.$$

$$\left.+\sum_{\alpha\neq}^{r}\sum_{\alpha'=1}^{r}\left(\sum_{i=1}^{m}\sigma_{(ii,m)}+\sum_{i\neq}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)}\right)\right\}$$

$$=\frac{1}{(rm)^2}\left\{r\left(\sum_{i=1}^{m}\sigma^2_{(i,m)}+\sum_{i\neq}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)}\right)\right.$$

$$\left.+r(r-1)\left(\sum_{i=1}^{m}\sigma_{(ii,m)}+\sum_{i\neq}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)}\right)\right\}$$

$$=\frac{1}{rm^2}\left(\sum_{i=1}^{m}\sigma^2_{(i,m)}+r\sum_{i=1}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)}-\sum_{i=1}^{m}\sigma_{(ii,m)}\right)\qquad(21.3.6)$$

Now,

$$\begin{aligned}
\sum_{i=1}^{m}\sigma^2_{(i,m)} &= \sum_{i=1}^{m} E\left\{y_{(im)r}-\mu_{(i,m)}\right\}^2\\
&= \sum_{i=1}^{m} E\left(y_{(im)r}-\mu\right)^2-\sum_{i=1}^{m}\left(\mu_{(i,m)}-\mu\right)^2\\
&= \sum_{i=1}^{m}\sum_{t=1}^{N} A_i^t(y_t-\mu)^2-\sum_{i=1}^{m}\left(\mu_{(i,m)}-\mu\right)^2\\
&= \sum_{t=1}^{N}(y_t-\mu)^2\sum_{i=1}^{m}A_i^t-\sum_{i=1}^{m}\left(\mu_{(i,m)}-\mu\right)^2\\
&= m\sigma_y^2-\sum_{i=1}^{m}\left(\mu_{(i,m)}-\mu\right)^2 \qquad (21.3.7)\\
&\quad \text{(using Theorem 21.3.1)}
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{m}\sigma_{(ij,m)} &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{E\left(y_{(i,m)\alpha}y_{(j,m)\beta}\right)-\mu_{(i,m)}\mu_{(j,m)}\right\}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\sum_{t=1}^{N}\sum_{l=1}^{N} y_t y_l B_{ij}^{tl}-\mu_{(i,m)}\mu_{(j,m)}\right)\\
&= \sum_{t\neq}^{N}\sum_{l=1}^{N} y_t y_l\left(\sum_{i=1}^{m}\sum_{j=1}^{m}B_{ij}^{tl}\right)-m^2\mu^2 \quad \left(\text{since } B_{ij}^{tt}=0\right)\\
&= \frac{m^2}{N(N-1)}\sum_{t\neq}^{N}\sum_{l=1}^{N}y_t y_l-m^2\mu^2 \quad \text{(using Eq. 21.3.5)}\\
&= -\frac{m^2}{(N-1)}\sigma_y^2 \qquad (21.3.8)
\end{aligned}$$

Substituting Eqs. (21.3.7) and (21.3.8) in Eq. (21.3.6), we find

$$V\left(\widehat{\overline{Y}}_{rss}\right) = \frac{r}{(rm)^2}\left(\frac{m(N-1-mr)}{N-1}\sigma_y^2 - \sum_{i=1}^{m}\left(\mu_{(i,m)}-\mu\right)^2 - \sum_{i=1}^{m}\sigma_{(ii,m)}\right) \tag{21.3.9}$$

We note that $\sum_{i=1}^{m}\sigma_{(ii,m)}$ remains unchanged if the population is centered, i.e., if y_i is replaced by $y_i - \mu$ and $\mu_{(i,m)}$ is replaced by $\mu_{(i,m)} - \mu$. Thus

$$\begin{aligned}\sigma_{(ii,m)} &= \mathbf{y}^T\mathbf{B}_{ii}\mathbf{y} - \mu_{(i,m)}^2 \\ &= (\mathbf{y}-\boldsymbol{\mu})^T\mathbf{B}_{ii}(\mathbf{y}-\boldsymbol{\mu}) - \left(\mu_{(i,m)}-\mu\right)^2 \end{aligned} \tag{21.3.10}$$

Substituting Eq. (21.3.10) in Eq. (21.3.9), we obtain

$$V\left(\widehat{\overline{Y}}_{rss}\right) = \frac{1}{rm}\left(\frac{N-1-mr}{N-1}\sigma_y^2 - \frac{1}{m}(\mathbf{y}-\boldsymbol{\mu})^T\left(\sum_{i=1}^{m}\mathbf{B}_{ii}\right)(\mathbf{y}-\boldsymbol{\mu})\right) \tag{21.3.11}$$

21.3.1 Relative Precision

The variance of the sample mean $\overline{y}_{wor}$ based on a sample of size mr selected by the SRSWOR method is

$$V\left(\overline{y}_{wor}\right) = \frac{N-mr}{mr(N-1)}\sigma_y^2$$

So, the relative precision of RSS compared with the sample mean of an SRSWOR sample based on the same sample size mr is

$$\begin{aligned} RP_{wor} &= \frac{V\left(\overline{y}_{wor}\right)}{V\left(\widehat{\overline{Y}}_{rss}\right)} \\ &= \frac{1}{1-\dfrac{N-m}{N-mr}\tau} \end{aligned} \tag{21.3.12}$$

where

$$\tau = \frac{1}{N-m} + \frac{N-1}{N-m}\frac{\gamma}{\sigma_y^2} \quad \text{and} \quad \gamma = \frac{1}{m}(\mathbf{y}-\boldsymbol{\mu})^T\left(\sum_{i=1}^{m}\mathbf{B}_{ii}\right)(\mathbf{y}-\boldsymbol{\mu}).$$

The expression for the relative savings is

$$\begin{aligned} RS_{wor} &= 1 - 1/RP_{wor} \\ &= \frac{1 - f/r}{1 - f}\tau \end{aligned} \tag{21.3.13}$$

where $f = mr/N = n/N$.

The expression (21.3.12) indicates that the relative precision depends on the replication factors r, m, γ, and σ_y^2. The relative savings is a monotonic increasing function of r for a given value of m.

21.4 SIZE-BIASED PROBABILITY OF SELECTION

Cox (1969) proposed the harmonic mean as an estimator of the mean of a random variable y, if the sample is selected from a population with probability proportional to y. For example, y may be lake size if we are using point-intercept sampling for lakes in a region (Muttlak and McDonald, 1990), fiber length when using point-intercept sampling for fibers in yarn (Cox, 1969), or particle width if we are using line-intercept sampling to estimate the number of particles in a region (McDonald, 1980). Let $f(y)$ be the pdf of y. If each unit has a probability of selection proportional to its size y, then Cox (1969) showed that the pdf of the observed sizes is

$$g(y) = \frac{yf(y)}{\mu}, \quad y > 0 \tag{21.4.1}$$

where $g(y)$ is called the weighted (or length-biased) pdf and μ is the mean of the unweighted density $f(y)$.

Cox (1969) derived the following results:

$$\text{(i) } E_g\left(\frac{1}{y}\right) = \frac{1}{\mu} \text{ (ii) } E_g(\widetilde{\mu}) = \frac{1}{\mu} \text{ and (iii) } V_g(\widetilde{\mu}) = \frac{1}{n}\left(\frac{\mu\mu_{-1} - 1}{\mu^2}\right) \tag{21.4.2}$$

where E_g denotes expectation with respect to the weighted density $g(y)$, $\widetilde{\mu} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i}$ and $\mu_{-1} = E\left(\frac{1}{y}\right)$.

Cox (1969) has shown that for large n, the harmonic mean, $\widehat{\mu}_h = \left(\frac{1}{\widetilde{\mu}}\right)$, is asymptotically normal with mean μ and variance

$$\frac{\mu^2(\mu\,\mu_{-1} - 1)}{n} \tag{21.4.3}$$

provided $V_g(\widetilde{\mu})$ is finite.

Muttlak and McDonald (1990) applied size-biased probability selection in RSS. In this method, initially m^2 units are selected with size-biased probability of selection with density function $f(y)$. The selected m^2 units are partitioned at random into m sets each of size m. Hence, the pdf of each of the selected m^2 observed y values is $g(y)$, given in Eq. (21.4.1). All the units in a set are ranked by some inexpensive procedure. The lowest rank holder is quantified for the first set, the second ranked holder is quantified for the next set, and the process is continued until the highest rank unit is quantified for the last mth set. The cycle is repeated r times. Let $y_{i(i)|k}$ be the value of the variate of interest y based on ith ranked unit of the ith set of the kth cycle, $i = 1,\ldots, m$; $k = 1,\ldots, r$.

Consider

$$\widetilde{\mu}_k = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{y_{i(i)|k}} \tag{21.4.4}$$

as an estimator of $1/\mu$.

We can write

$$g(y) = \frac{1}{m}\sum_{i=1}^{m} g_{m,i}(y) \tag{21.4.5}$$

where $g_{m,i}(y) = \dfrac{\Gamma(m+1)}{\Gamma(m)\Gamma(m-i+1))}[G(y)]^{i-1}[1-G(y)]^{m-i}g(y)$ and $G(y) = \int_{-\infty}^{y} g(y)dy$.

Eq. (21.4.5) yields

$$\begin{aligned} E(g(y)) &= \frac{1}{m}\sum_{i=1}^{m} E\left(\frac{1}{y_{i(i)|k}}\right) \\ \text{i.e., } \frac{1}{\mu} &= \frac{1}{m}\sum_{i=1}^{m}\mu_{-(i)|m} \end{aligned} \tag{21.4.6}$$

where $E\left(\dfrac{1}{y_{i(i)|k}}\right) = \mu_{-(i)|m}$.

Hence from Theorem 21.2.1, we get

$$E(\widetilde{\mu}_k) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{\mu_{-(i)|m}} = \frac{1}{\mu} \text{ and } V(\widetilde{\mu}_k) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2_{-(i)|m} \le V\left(\frac{1}{y}\right)\Big/ m \tag{21.4.7}$$

where $\sigma^2_{-(i)|m} = V\left(\dfrac{1}{y_{i(i)|k}}\right)$.

Eq. (21.4.7) leads to the following theorem:

Theorem 21.4.1

(i) $\widetilde{\mu}_{\bullet} = \frac{1}{r}\sum_{k=1}^{r}\widetilde{\mu}_k = \frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m}\frac{1}{y_{i(i)|k}}$ is an unbiased estimator of $\frac{1}{\mu}$

(ii) $V(\widetilde{\mu}_{\bullet}) \leq V(\widetilde{\mu}_{mr})$

where $\widetilde{\mu}_{mr} = \frac{1}{mr}\sum_{i=1}^{mr}\frac{1}{y_i}$ and $y_1,\ldots,\ y_{mr}$ is a size-biased sample from a population with density of $f(y)$.

Following Cox (1969), the harmonic mean $1/\widetilde{\mu}_{\bullet}$ based on the RSS is approximately unbiased for the population mean μ. In addition, one would expect that the variance of $1/\widetilde{\mu}_{\bullet}$ is smaller than that of $1/\widetilde{\mu}_{mr}$ because the variance of $\widetilde{\mu}_{\bullet}$ is smaller than that of $\widetilde{\mu}_{mr}$. Muttlak and McDonald (1990) supported correctness of this inequality by using computer simulations.

21.5 CONCLUDING REMARKS

RSS is useful for the situations where quantification of an element is difficult but elements can be ranked fairly by eye or some other method with negligible cost. RSS provides an unbiased and efficient estimate for the population mean; even the ranking is imperfect. RSS does not provide unbiased estimates of the population variance. Unlike estimation of the mean, RSS does not provide appreciable gain in efficiency for estimation of variance. However, RSS is found to be most beneficial when estimates of both mean and variances are needed (Stokes, 1980a). RSS for estimating the population mean under SRSWOR sampling was proposed by Patil et al. (1995). However, the expression of variance was not simple as it involves a quantity $\gamma = (\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Gamma}(\mathbf{y}-\boldsymbol{\mu})/m$, where $\boldsymbol{\Gamma} = \left(\sum_{i=1}^{m}\mathbf{B}_{ii}\right)$ is an $N \times N$ matrix whose entries are a function of the population size N and the set size m but independent of the study variable y. A computer program for computation of $\boldsymbol{\Gamma}$ was provided by Patil et al. (1995). Several extensions to the theories and applications of RSS have been proposed; that is, RSS with imperfect ranking, RSS by ranking with a concomitant variable, and RSS with multivariate samples and random selection method were considered by Stokes (1977, 1980a), Patil et al. (1993), and Li et al. (1999), among others. Application of RSS in Mann–Whitney Wilcoxon test was provided by Bonn and Wolfe (1992, 1994) while an RSS version of the sign test was provided by Koti and Babu (1996). RSS for density estimation was investigated by Chen (1999), and RSS for parametric families was considered by Shen (1994) and Sinha et al. (1996).

21.6 EXERCISES

21.6.1 Let m sets of samples each of size m be selected independently from a population by SRSWR method. A sample $s_1,\ldots, s_r$ of $r(<m)$ sets is selected at random from the selected m sets by SRSWOR method. The units of the selected samples $s_1,\ldots, s_r$ are ordered with respect to the value of the study variable y. Let a subset of integers $k_1,\ldots, k_r$ be selected from the set of integers $1,\ldots, n$ by SRSWOR method. Let $y_{(i,k_i)}$ be the k_ith ordered statistic from the set s_i, $i = 1,\ldots, r$.

(a) Show that

(i) $\widehat{\mu}_{mrss} = \frac{1}{r}\sum_{i=1}^{r} y_{(i,k_i)}$ is an unbiased estimator of μ, the population mean of y.

(ii) $$Var(\widehat{\mu}_{mrss}) = \frac{Q}{mr}\,\frac{m-r}{m-1} + \frac{\sum_{j=1}^{m}\sigma^2_{j:m}}{mr}$$

where $Q = \sum_{j=1}^{m}\left(\mu_{j:m}-\mu\right)^2$, $\mu_{j:m} = E(y_{(i,j)})$, and $\sigma^2_{j:m} = Var\left(y_{(i,j)}\right)$.

(b) Let $\overline{y}_m$ be the sample mean of size m selected by SRSWR method. Show that a necessary condition for $Var(\widehat{\mu}_{mrss}) < Var\left(\overline{y}_m\right)$ is $r > m^2/(2m-1)$ (Li et al., 1999).

21.6.2 Consider a finite population $U = \{1,\ldots, N\}$ of N units. Let y_i and x_i be the values of the study and auxiliary variables y and x, respectively. From the population U, m independent samples $s_1,\ldots, s_m$ each of size m are selected by SRSWOR method. The selected units in $s_j(j = 1,\ldots, m)$ are ranked with respect to the x values. Let $y_{j\langle j\rangle}$ be the judgment jth order statistic corresponding to $x_{j(j)}$, jth order statistic of x for the sample s_j, $j = 1,\ldots, m$. Thus, we get the following ranked set sample $rss = (y_{1\langle 1\rangle},\ldots, y_{j\langle j\rangle},\ldots, y_{m\langle m\rangle})$. The selection of ranked set sample is repeated r times (cycles). The ranked set sample based on the kth cycle is denoted by $S = \bigcup_{k=1}^{r} S_k$, where $S_k = (y_{1\langle 1\rangle|k},\ldots, y_{j\langle j\rangle|k},\ldots, y_{m\langle m\rangle|k})$, $k = 1,\ldots, r$. Let S^* be the set of distinct units in S.

(a) Show that the inclusion probabilities for the ith unit, i and $t(\neq i)$ units are, respectively,

$$\pi_i = 1 - \prod_{j=1}^{m}\left(1-\alpha_{ij}\right)^r \text{ and}$$

$$\pi_{it} = 1 - \prod_{j=1}^{m}\left(1-\alpha_{ij}\right)^r - \prod_{j=1}^{m}\left(1-\alpha_{tj}\right)^r + \prod_{j=1}^{m}\left(1-\alpha_{ij}-\alpha_{tj}\right)^r$$

where $\alpha_{ij} = \binom{i-1}{j-1}\binom{N-i}{m-j}\Big/\binom{N}{m}$

(b) Let δ_i be the number of times ith unit appears in S. Show that $E(\delta_i) = mr/N$.

(c) Let $T_{yk} = \frac{1}{m}\sum_{j=1}^{m} y_{j\langle j\rangle|k}$ and $T_y = \frac{1}{r}\sum_{k=1}^{r} T_{yk}$. Show that

(i) $E(T_y) = \overline{Y}$

(ii) The variance of T_y is

$$V(T_y) = \frac{1}{mr}\sigma_y^2 - \frac{1}{rm^2}\sum_{j=1}^{m}\left(\mu_{y\langle j\rangle} - \overline{Y}\right)^2$$

and

(iii) An unbiased estimator of $V(T_y)$ is

$$\widehat{V}(T_y) = \frac{1}{r(r-1)}\sum_{j=1}^{r}(T_{yr} - T_y)^2$$

where $\overline{Y} = \sum_{i=1}^{N} y_i/N$, $\sigma_y^2 = \sum_{i\in U}\left(y_i - \overline{Y}\right)^2/N$,

and $\mu_{y\langle j\rangle} = E(y_{j\langle j\rangle|k})$ (Jozani and Johnson, 2011).

21.6.3 Let $y_{i(i)|k}$, $i = 1,\ldots, m$; $k = 1,\ldots, r$, be a ranked set sample of size mr selected from a population with cdf F with r cycles each of which consists of m sets of size m each. The empirical distribution function (edf) based on the ranked set sample is denoted by $F^*(t) = \frac{1}{mr}\sum_{k=1}^{r}\sum_{i=1}^{m}\Delta\left(y_{i(i)|k} - t\right)$, where $\Delta(u) = 1$ if $u \le 0$ and $\Delta(u) = 0$ if $u > 0$. Show that

(i) $F^*(t)$ is an unbiased estimator of $F(t)$

(ii) $Var(F^*(t)) = \frac{1}{rm^2}\sum_{j=1}^{m} F_{(j)}\left(1 - F_{(j)}\right)$, where $F_{(j)}$ is the distribution function of $y_{i(i)|k}$ and

(iii) $Var(F^*(t)) \le Var\left(\widehat{F}(t)\right)$, where $\widehat{F}(t)$ is the edf of a random sample of size mr (Stokes and Sager, 1988).

21.6.4 Let a ranked set sample of size $n(=mr)$ be selected from a population with r cycles each of which consists of m sets of size m each. Let $y_{i\langle i\rangle|k}$ ($x_{i(i)|k}$) denote ith judgment ordering (order statistic) in the ith set of the cycle r for the study (auxiliary) variable $y(x)$.

Consider the ratio estimator $\widehat{\mu}_{rRSS} = \dfrac{\overline{y}_{\langle n\rangle}}{\overline{x}_{[n]}}\mu_x$ of the population mean μ_y of the study variable y, where $\overline{y}_{\langle n\rangle} = \dfrac{1}{rm}\sum_{k=1}^{r}\sum_{i=1}^{m} y_{i\langle i\rangle|k}$, $\overline{x}_{[n]} = \dfrac{1}{rm}\sum_{k=1}^{r}\sum_{i=1}^{m} x_{i(i)|k}$, and μ_x is the population mean of the auxiliary variable x, which is assumed to be known. Show that the bias and MSE of $\widehat{\mu}_{rRSS}$ are as follows:

$$B(\widehat{\mu}_{rRSS}) \cong \mu_y\left[\frac{1}{mr}\left(C_x^2 - C_{yx}\right) - \left(W_{x(i)}^2 - W_{yx(i)}\right)\right],$$

$$M(\widehat{\mu}_{rRSS}) \cong \frac{1}{mr}\left(\sigma_y^2 - 2R\sigma_{yx} + R^2\sigma_x^2\right) - \frac{1}{m^2 r}\left(\sum_{i=1}^{m}\tau_{y\langle i\rangle}^2 - 2R\sum_{i=1}^{m}\tau_{yx(i)}^2 + R^2\sum_{i=1}^{m}\tau_{x(i)}^2\right)$$

where $C_x(C_y) =$ population CV of $x(y)$, $C_{yx} = \rho C_x C_y$, $\rho =$ correlation coefficient between y and x, $W_{x(i)}^2 = \dfrac{1}{m^2 r\mu_x^2}\sum_{i=1}^{m}\tau_{x(i)}^2$, $W_{yx(i)} = \dfrac{1}{m^2 r\mu_x\mu_y}\sum_{i=1}^{m}\tau_{yx(i)}$, $\tau_{x(i)} = E(x_{i(i)|k}) - \mu_x$, $\tau_{y\langle i\rangle} = E(y_{i\langle i\rangle|k}) - \mu_y$, $\tau_{yx(i)} = E(y_{i\langle i\rangle|k} - \mu_y)(x_{i(i)|k} - \mu_x)$, $\sigma_y^2\left(\sigma_x^2\right) =$ population variance of $y(x)$, and $R = \mu_y/\mu_x$.

(Kadilar et al., 2009).