RANKED SET SAMPLING FOR ESTIMATING THE DIFFERENCE OF MEANS

Carlos N. Bouza[1], Universidad de La Habana

ABSTRACT
Ranked set sampling permits to use anciliary information to rank the units. The drawing of them using the ranks is at least equivalent to random sampling. In this paper this method is applied for deriving estimators of the difference.

Key words: ranked set sampling, accuracy, missing observations.

RESUMEN
El muestreo de conjunto previamente rankeado permite utilizar la información adicional para rankear las unidades. La selección de ellas es llevada a cabo usándoles y es cuando menos equivalente al uso del muestreo simple aleatorio. En este trabajo le usamos para obtener estimadores de la diferencia.

Palabras clave: muestreo de conjuntos, exactitud, observaciones perdidas.

AMS: 62D05.

1. INTRODUCTION

The aims of the researcher are to observe a convenient sample. It must have units representing the expected different behaviors of the variable of interest Y. They look for a *representative sample.* Statisticians have modeled these aims by stratifying, clustering, assigning different probabilities to the units, etc. The economic aspect is characterized by the sampling costs which should match with a fixed budget. The use of a sub-sampling procedure is common. Then a large sample s is selected but only some of the selected units, a sub-sample, are used for measuring Y. A less costly variable X is measured in s and Y only in the smaller sub-sample s*.

Ranked Set Sampling (rss) is based on the observation of a sub-sample m from a set of m independent samples of the same size. $s = s_1 \cup \ldots \cup s_m$ is the large sample. The sub-samples are selected independently by using Simple Random Sampling With Replacement (srswr) and their sizes are $|s_i| = m$ for any i = 1,...,m. Therefore $|s| = s^2$ and we have the measures of X as follows:

Subsample	*Variable*
1	$X_{11} \ldots X_{1m}$
2	$X_{21} \ldots X_{2m}$
⋮	⋮ … ⋮
m	$X_{m1} \ldots X_{mm}$

X_{ij} is the value of X in the j-th-sampled unit in s_i and $X_{i(t)}$ as the t-th order statistics (t-os). Then we may rearrange the samples as

Subsample	*Variable*
1	$\underline{X_{1(1)}} X_{1(2)} \ldots X_{1(m)}$
2	$X_{2(1)} \underline{X_{2(2)}} \ldots X_{2(m)}$
⋮	⋮ ⋮ … ⋮
m	$X_{m(1)} \ldots X_{m(2)} \underline{X_{m(m)}}$

The units with the underlined of are included in the sub-sample and Y is measured in them. The obtained sample s* is called Random set Sample and its size is m.

[1]**E-mail**:bouza@matcom.uh.cu

McIntyre (1952) recommended this method. He analyzed the economy due to its use without a mathematical support for his ideas. It was used for estimating pasture yields. He ranked a set of plots on the basis of a previous visual inspection. The subjective evaluation (X) were used for ranking the plots before measuring the yield (Y). The assumption that the ranking was perfect establishes that to measure Y in the unit with X-rank t in s_t^* is the tos of Y in this sample. This procedure is used in r cycles independently. The size of m is small and the size of the complete sample is n = rm.

In Section 2 we will discuss the efficiency of rss in comparison with srswr. In Section 3 we will develop estimators for the difference. A rss procedure is proposed and it is compared with an estimator used by Ware-Cunia (1962). Ratio and regression based estimators are worked out. The estimators are compared in Section 4. A Monte Carlo experiment illustrated the practical performance of the alternative estimators of the difference of population means.

2. SIMPLE RANDOM SAMPLING AND RANKED SET SAMPLING

Ranked set sample is based on the selection of m samples of size m. The procedure of selecting the ranked set sample is repeated r times independently and the total sample size is rm. Let us denote the os $Y_{i(i)}$ in a cycle j as $Y_{(i:m)j}$. An estimator of the population mean μ_Y is

$$\overline{Y}_{(m)r} = \frac{1}{mr}\sum_{i=1}^{m}\sum_{j=1}^{r} Y_{(i:m)j} = \frac{1}{m}\sum_{i=1}^{m} Y_{(i:m)} \qquad (2.1)$$

Halls-Dell (1966) and Cobby **et al** (1985) have used this method in other studies of pasture yields.

Takahasi-Wakimoto (1968) developed mathematical aspects of rss. The improvement of the efficiency of the proposed procedure dor estimating the mean with respect to the srswr was established under the assumption that the X-ranking and the Y-ranking coincide.

As the variables with the i-th rank have $E(Y_{(i:m)}) = \mu_{(i:m)}$, the expectation of the i-os, and $Var(Y_{(i:m)}) = \sigma^2_{(i:m)}$ is its variance. Then the unbiasedness of (2.1) follows easily, see Patil-Sinha-Taillie (1994), and

$$V\left(\overline{Y}_{(m)r}\right) = \frac{1}{m^2 r}\sum_{i=1}^{m}\sigma^2_{(i:m)} = \frac{1}{mr}\left(\sigma^2 - \sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2\right)$$

As n = mr is the total sample size we have that the sampling error in srswr is

$$V(\overline{Y}) = \frac{\sigma^2}{n}$$

(2.1) is more accurate. Takahasi-Wakimoto (1968) derived that rss and srswr are equivalent only if the distribution of the variable is uniform.

Dell-Clutter (1972) considered the case in which the ranking may be imperfect. They obtained that this estimator maintains its properties. Jonhson-Patil-Sinha (1993) applied rss to an important study of vegetation and Patil-Taillie (1993) discussed how related models permit to develop adequate inferences using rss. The interest in rss is close related with its usefulness in environmental surveys. Evaluating small sets of samples, which are ranked, diminishes the associated high costs, and an adequate representation of the existing units is achieved. Since the publication in the Volumen 12th of the Handbook of Statistics of the analysis of Patil-Sinha-Taillie (1994) on rss the number of publications on this subject has been increased considerably. We will cope with the problem of estimating the difference of two population means.

3. ESTIMATION OF THE DIFFERENCE

We are interested in estimating the difference between the means of two variables Y and Z. The auxiliary variable X permits to rank the units. An increase in X is related with an adequate response in Y and Z. This is a common situation in many applications. For example the ecologist wants to estimate the difference between the abundance of a pest after and before, a visual evaluation of the fields (X) permits to rank them, a

manager plans to estimate the consumption of his beverage in a population with respect to the competitor´s results, X may be the overall sales of the stores, etc.

A sample s of size m is selected by srswr for estimating $D = \mu_Y - \mu_Z$ and it is divided into three independent sub-samples

- s (1) = {j ∈ s| Y and Z are measured}, |s(1)| = m(1)
- s (2) = {j ∈ s| only Y is measured}, |s(2)| = m(2)
- s (3) = {j ∈ s| only Z are measured}, |s(3)| = m(3).

This model permits to reduce the sampling costs. Ware-Cunia (1962) derived a Minimum Variance estimator dealing with forest inventories. Pi-Erh Lin (1971) derived an optimal estimator of D when the variables are normal. We will use rss for estimating D and different estimators will be compared.

Take s as in Section 2 but each sub-sample s_i is partitioned into three sub-samples $s_i(j)$, j = 1,2,3, according with the above proposed scheme. For the first sample the ranking is made using an evaluation of the expected difference between Y and Z. For the other sub-samples the ranking is made as usual. Therefore we may compute for the sample where Y and Z are measured

$$d_1 = \bar{y}_{m(1)} - \bar{z}_{m(1)} = \frac{1}{m(1)r}\sum_{i=1}^{m(1)}\sum_{j=1}^{r} Y_{(i:m(1))j} - \frac{1}{m(1)r}\sum_{i=1}^{m(1)}\sum_{j=1}^{r} Z_{(i:m(1))j}$$

by using the structure of (2.1). It is unbiased for D and

$$V(d_1) = V(\bar{y}_{m(1)}) + V(\bar{z}_{m(1)}-) + Cov(\bar{y}_{m(1)}, \bar{z}_{m(1)}) = V_{y(1)} + V_{z(1)} + Cov_{y,z(1)}$$

where

$$V\left(\bar{y}_{(m)1}\right) = \frac{1}{rm(1)}\left(\sigma_y^2 - \sum_{i=1}^{m(1)}\left(\mu_{y(i:m(1))} - \mu_y\right)^2\right) = \frac{1}{rm(1)}\left(\sigma_y^2 - \Delta_{ym(1)}\right)$$

$$V\left(\bar{z}_{(m)1}\right) = \frac{1}{rm(1)}\left(\sigma_z^2 - \Delta_{zm(1)}\right)$$

Hence

$$V\left(d_1\right) = \frac{\sigma_y^2 + \sigma_z^2}{rm(1)} - \frac{1}{rm(1)}\left(\Delta_{ym(1)} + \Delta_{zm(1)}\right) - 2Cov_{y,z(1)}$$

Because the existence of errors in the ranking do not affect the propeties of (2.1).

In the unmatched samples we will use the corresponding estimators of μ_y and μ_z:

$$\bar{y}_{(2)} = \frac{1}{rm(2)}\sum_{i=1}^{m(2)}\sum_{j=1}^{r} Y_{(t:m(1))j}$$

and

$$\bar{z}_{(3)} = \frac{1}{rm(3)}\sum_{i=1}^{m(3)}\sum_{j=1}^{r} Z_{(i:m(3))j}$$

with variances

$$V\left(\bar{y}_{(2)}\right)=\frac{1}{rm(2)}\left(\sigma_y^2-\sum_{i=1}^{m(2)}\left(\mu_{y(i:m(2))}-\mu_y\right)^2\right)=\frac{1}{rm(2)}\left(\sigma_y^2-\Delta_{ym(2))}\right)$$

$$V\left(\bar{z}_{(3)}\right)=\frac{1}{rm(3)}\left(\sigma_z^2-\sum_{i=1}^{m(3)}\left(\mu_{z(i:m(3))}-\mu_z\right)^2\right)=\frac{1}{rm(3)}\left(\sigma_z^2-\Delta_{z(m(3))}\right)$$

Then we estimate D by using

$$d_{rss}=d_1+\left(\bar{y}_{(2)}-\bar{z}_{(3)}\right)$$

with

$$V\left(d_{rss}\right)=\sigma_y^2\left(\frac{1}{rm(1)}+\frac{1}{rm(2)}\right)+\sigma_z^2\left(\frac{1}{rm(1)}+\frac{1}{rm(3)}\right)-\Delta-2Cov_{zy(1)}$$

where

$$\Delta=\Delta_y\left(\frac{1}{rm(1)}+\frac{1}{rm(2)}\right)+\Delta_z\left(\frac{1}{rm(1)}+\frac{1}{rm(3)}\right)$$

Ware-Cunia (1962) proposed the estimator

$$d_{VC}=a_1\bar{y}_1+a_2\bar{y}_2+b_1\bar{z}_1+b_3\bar{z}_3$$

the parameters are unknown buy they must satisfy that $a_1 + a_2 = -(b_1 + b_3) = 1$ for supporting the unbiasedness of it. Setting $a_1 = a$ and $b_1 = b$ the variance is

$$V(d_{VC})=\frac{\sigma_y^2\left(n_2a^2+n_1(1-a)^2\right)}{n_1n_2}+\frac{\sigma_z^2\left(n_3b^2+n_1(1-b)^2\right)}{n_1n_2}+\frac{2ab\sigma_{yz}}{n_1}$$

The optimum values of a and b are obtained by looking for a minimum variance estimator.

They are

$$a_0=\frac{\lambda n'\sigma_y+n_1\rho\sigma_z}{\sigma_y(\lambda n'+n_1\rho^2)} \tag{3.1}$$

$$b_0=\frac{\lambda n''\sigma_z+n_1\rho\sigma_y}{\sigma_{yz}(\lambda n''+n_1\rho^2)} \tag{3.2}$$

where

$$\lambda=\frac{n_1}{n_2},\quad n'=n_1+n_2,\quad n''=n_1+n_3\quad\text{and}\quad\rho=\sigma_{yz}/\sigma_y\sigma_z.$$

Using the approach based on Double Sampling and assuming that the ratios and regression coefficients are stable the knowledge of their values from previous experiences enables to develop ratio and regression estimators of D.

Taking $R_1 = \mu_z/\mu_y$ and $R_2 = \mu_y/\mu_z$

$$d_R = R_1(n_1\bar{y}_1 + n_2\bar{y}_2) - R_2\left((n_1\bar{z}_1 + n_3\bar{z}_3)\right) = R_1\bar{y}'_1 + R_2\bar{z}'_1$$

is unbiased because the involved means are unbiased. The variance is

$$V(d_R) = \frac{R_1^2\sigma_y^2}{n'} + \frac{R_2^2\sigma_z^2}{n''} - \frac{2R_1R_2\sigma_{yz}}{n'n''}$$

When the regression estimator is used

$$\bar{y}_r = \bar{y}_1 + B_1(\bar{z}_3 - \bar{z}_1)$$

$$\bar{z}_r = \bar{z}_1 + B_2(\bar{y}_2 - \bar{y}_1)$$

where $B_1 = \rho\sigma_y/\sigma_z$ and $B_2 = \rho\sigma_z/\sigma_y$ and

$$d_r = \bar{y}_r - \bar{z}_r$$

Again the unbiasedness of the estimator follows from the knowledge of the involved parameters. Arranging conveniently the terms of $(d_r - D)^2$ after some algebraic manipulations we obtain that

$$V(d_r) = \frac{\sigma_y^2(n_3 + n'\rho^2)}{n_1n_3} + \frac{\sigma_z^2(n_2 + n''\rho^2)}{n_1n_2} - 2\frac{\sigma_{yz}(1+\rho^2)}{n_1}$$

4. COMPARISON OF THE ESTIMATORS

A theoretical comparison of the estimators is based on the analysis of their variances. The estimator d_{rss} is more accurate than d_{VC} when $V(d_{rss}) - V(d_{VC}) < 0$. Remembering that $n_h = rm$ (h = 1,2,3) we can write this difference as:

$$\sigma_y^2\left(\frac{n'-n_2a^2 - n_1(1-a)^2}{n_1n_2}\right) + \sigma_z^2\left(\frac{n''-n_3b^2 - n_1(1-b)^2}{n_1n_3}\right) - 2\left(Cov_{yz(1)} - \frac{n_1\sigma_{xy}}{n_1}\right) - \Delta$$

In this case the gain in accuracy of rss with respect to srswr is not expressed by a decomposition of the variance. The coefficients of the variance terms positive whenever a and/or – b belong to]0,1[. That is the case when we use the optimal values of them. The covariance term should be small in many cases. Therefore the possible gain in accuracy due to the use of rss depends on Δ.

A similar analysis of d_{rss} versus d_R yields the preference of the former when

$$V(d_{rss}) - V(d_R) = \sigma_y^2\left(\frac{n'}{n_1n_2} - \frac{R_1^2}{n'}\right) + \sigma_z^2\left(\frac{n''}{n_1n_3} - \frac{R_2^2}{n''}\right) - 2\left(Cov_{yz(1)} - \frac{n_1\sigma_{xy}}{n'n''}\right) - \Delta < 0$$

It is sufficient that

$$R_1 > \frac{n'}{\sqrt{n_1n_2}}$$

and

$$R_1 > \frac{n''}{\sqrt{n_1n_3}}$$

for preferring d_{rss}. In other cases the value of Δ must be analyzed.

The comparison with d_r establishes that if

$$V(d_{rss}) - V(d_r) = \sigma_y^2\left(\frac{n'}{n_1 n_2} - \frac{n_3 + n'\rho^2}{n_1 n_3}\right) + \sigma_z^2\left(\frac{n''}{n_1 n_3} - \frac{n_2 + n''\rho^2}{n_1 n_2}\right) - 2\left(Cov_{yz(1)} - \frac{\sigma_{zy}(1+\rho^2)}{n_1}\right) - \Delta < 0$$

we must prefer d_{rss}. Whenever

$$\rho > \sqrt{\frac{n_1 n_3}{n_2 n'}}$$

and

$$\rho > \sqrt{\frac{n_1 n_2}{n_3 n''}}$$

this decision is adequate. In other case the magnitude of Δ should be taken into account. The use of d_{VC} must be preferred to d_R if

$$V(d_{VC}) - V(d_R) = \sigma_y^2\left(\frac{n_2 a^2 + n_1(1-a)^2}{n_1 n_2} - \frac{R_1^2}{n'}\right) + \sigma_z^2\left(\frac{n_3 b^2 + n_1(1-b)^2}{n_1 n_3} - \frac{R_2^2}{n''}\right) - 2\sigma_{yz}\left(\frac{n_1}{n' n''} - \frac{ab}{n}\right) < 0$$

A sufficient condition for preferring d_{VC} is that

$$R_1 > \sqrt{\frac{a^2 n'^2 - (2a+1)n' n_1}{n_1 n_2}}$$

and

$$R_2 > \sqrt{\frac{b^2 n''^2 - (2b+1)n'' n_1}{n_1 n_3}}$$

in any case the covariance term is always negative.

The evaluation of the accuracy of d_{VC} with respect to d_r is based on the difference

$$V(d_{VC}) - V(d_R) = \sigma_y^2\left(\frac{n_2 a^2 + n_1(1-a)^2}{n_1 n_2} - \frac{n_3 + n'\rho^2}{n_1 n_3}\right) + \sigma_z^2\left(\frac{n_3 b^2 + n_1(1-b)^2}{n_1 n_3} - \frac{n_2 + n''\rho^2}{n_1 n_2}\right) - 2\sigma_{yz}\left(\frac{ab - 1 - \rho^2}{n_1}\right)$$

It is smaller than zero if

$$\rho > \sqrt{\frac{n_3\left(n_1(1-2a) + n' a^2 - n_2\right)}{n' n_2}}$$

and

$$\rho > \sqrt{\frac{n_2\left(n_1(1-2b) + n'' b^2 - n_3\right)}{n'' n_3}}$$

Note that the covariance term is positive.

The comparison of the two estimators based on Double Sampling principle yields that d_R should be preferred to d_r if

$$V(d_R) - V(d_r) = \sigma_y^2 \left(\frac{n_1 n_3 R_1^2 - n' n_3 \rho^2}{n' n_1 n_3} \right) + \sigma_z^2 \left(\frac{n_1 n_2 R_2^2 - n'' n_2 \rho^2}{n'' n_1 n_2} \right) - 2\sigma_{yz} \left(\frac{n_1}{n' n''} \frac{1-\rho^2}{n_1} \right)$$

is negative. Hence d_R is more accurate if

$$\rho > \sqrt{\frac{n_1 n_3 R_1^2 - n' n_3}{n'^2}}$$

and

$$\rho > \sqrt{\frac{n_1 n_2 R_2^2 - n'' n_{32}}{n''^2}}$$

because the term in covariance is expected to be generally negative.

In practice the parameters involved in the calculation of a_0, b_0, R_1, R_2, B_1 and B_2 are unknown. We will use the proposal of Bouza-Prabhu-Ajgaonkar (1993) for analyzing the behavior of d_{VC}, d_R and d_r by plugging in estimators of the parameters. The counterparts are

$$d'_{VC} = a' \bar{y}_1 (1 - a') \bar{y}_2 + b' \bar{z}_1 + (1 - b') \bar{z}_3$$

where a' and b' are calculated by sustituting the variances and the coefficient of correlation in (3.1) and (3.2) by

$$S^2_{y(j)} = \frac{\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2}{n_j - 1} \qquad \text{if} \qquad j = 1,2$$

$$S^2_{z(j)} = \frac{\sum_{i=1}^{n_j} (z_{ij} - \bar{z}_j)^2}{n_j - 1} \qquad \text{if} \qquad j = 1,3$$

$$\rho' = \frac{\dfrac{\sum_{i=1}^{n_1} (y_{i1} - \bar{y}_1)(y_{i1} - \bar{z}_1)}{n_j - 1}}{S_{y(1)} S_{z(1)}} = \frac{S_{yz(1)}}{S_{y(1)} S_{z(1)}}$$

For d_R we use instead of the population ratios

$$R'_1 = \frac{\bar{z}_1}{\bar{y}_1}$$

and

$$R'_2 = \frac{1}{R'_1}$$

then

$$d'_R = R'_1 \bar{y}' - R'_2 \bar{z}'$$

The Minimum Least Squares estimators estimate the regression coefficients

$$B'_1 = s_{xy}/S_y^2 \text{ and } B'_2 = s_{xy}/S_z^2 \text{ and } d'_r = \left(\bar{y}_1 + B'_r(\bar{z}_3 - \bar{z}_1)\right) - \left(\bar{z}_1 + B'_2(\bar{y}_2 - \bar{y}_1)\right).$$

Note that the use of rss does not depend on unknown parameters.

A simulation experiment was performed for evaluating the behavior of the estimators.

Three sample sizes n = 30, 90 and 300 were used and the sub-sample's size were equal to n/3. Both variables were generated using the same distribution. The distribution functions were an uniform in (0.2), U(0.2), an exponential with λ = 1, E(1) and a Normal, N(1,1). They have the same mean. The relative accuracy of the estimators was measured by

$$RP_t = \frac{\sum_{h=1}^{H} \left|\hat{D}_t - D\right|_h}{HD}$$

H = 1000 was the number of generated samples and t denotes the estimator. The number of cycles was r = 100.

The results for U(0.2) are given in Table 4. They suggest that d_{rss} was the best estimator but d'_R and d'_r have a similar behavior when the correlation is small. Note that the equivalence between (2.1) and the mean in srs derived by Takahasi. Wakimoto (1968) does not hold in this case. The increase in the correlation and the sample size have not a significant effect in the behavior of d_{rss} but it permits to obtain better estimates if d'_{VC} is used.

Table 4.1. Relative Accuracy of the estimators for a U(0.2).

$\hat{D}$	**0.1**	**N = 30** ρ **0.5**	**0.9**	**0.1**	**N = 90** ρ **0.5**	**0.9**	**0.1**	**N = 300** ρ **0.5**	**0.9**
d_{rss}	0.035	0.045	0.042	0.061	0.057	0.053	0.021	0.023	0.024
d'_{VC}	0.121	0.118	0.109	0.105	0.097	0.098	0.093	0.084	0.076
d'_R	0.056	0.057	0.051	0.063	0.067	0.060	0.076	0.069	0.064
d'_r	0.045	0.067	0.072	0.070	0.084	0.082	0.078	0.075	0.066

Table 4.2 presents the results of the Monte Carlo experiments for the E(1). The estimator based on rss is again the best alternative. The use of srs is the second best. The sample size plays an important role in the accuracy of the estimators. The ratio and regression estimators have a similar behavior.

Table 4.2. Relative Accuracy of the estimators for a E(1).

$\hat{D}$	**0.1**	**N = 30** ρ **0.5**	**0.9**	**0.1**	**N = 90** ρ **0.5**	**0.9**	**0.1**	**N = 300** ρ **0.5**	**0.9**
d_{rss}	0.081	0.080	0.079	0.067	0.070	0.064	0.058	0.055	0.060
d'_{VC}	0.094	0.096	0.093	0.079	0.077	0.078	0.069	0.065	0.058
d'_R	0.104	0.098	0.094	0.100	0.098	0.087	0.091	0.086	0.082
d'_r	0.106	0.100	0.098	0.099	0.084	0.085	0.081	0.076	0.072

When the variables are normal d_{VC} is the best alternative but d_{rss} and d'_r are close to it.

Table 3. Relative Accuracy of the estimators for a N(1,1).

$\hat{D}$	0.1	N = 30 ρ 0.5	0.9	0.1	N = 90 ρ 0.5	0.9	0.1	N = 300 ρ 0.5	0.9
d_{rss}	0.063	0.065	0.057	0.041	0.060	0.054	0.033	0.036	0.031
d_{VC}	0.008	0.008	0.013	0.012	0.012	0.009	0.007	0.004	0.004
d'_R	0.087	0.087	0.089	0.061	0.063	0.089	0.087	0.083	0.082
d'_r	0.064	0.066	0.067	0.072	0.081	0.080	0.079	0.075	0.068

Therefore d_{rss} seems to be the best alternative when the distribution is unknown and d_{VC} under the normality of the variables.

ACKNOWLEDGEMENTS

The computation of the Monte Carlo experiments and the writing of this paper were developed while the author was visiting the Laboratorio de Investigaciones y Asesoría Estadística of Universidad Veracruzana, México.

REFERENCES

BOUZA, C.N. and S.G. PRABHU-AJGAONKAR (1993): "Estimation of the difference of population means with missing observations", **Biometrical J**. 35, 245-252.

COBBY, J.M.; M.S. RIDOUT; P.J. BASSET and R.V. LARGE (1985): "An investigation into the use of ranked set sampling on grass and grass-clover swards", **Grass and Forage Sc**., 20, 257-263.

DELL, T.R. and J.L. CLUTTER (1972): "Ranked-set Sampling theory with order statistics background", **Biometrics**, 28, 545-555.

HALLS, L.K. and T.R. DELL (1994): "Trial of ranked set sampling for forage yields", **Forst Sc**., 12, 22-26.

JOHNSON, G.D; G.P. PATIL and A.K. SINHA (1993): "Ranked set sampling for vegetation research", **Abstracta Botanica** 17, 87-102.

Mc INTYRE, G.A. (1952): "A method of unbiased selective sampling using ranked sets", **Australian J. Agric. Res**., 3, 385-390.

PATIL, G.P.; A.K. SINHA and C. TAILLIE (1993): "Observational economy of ranked set sampling: comparison with the regression estimator", **Enviromentics**, 4, 399-412.

_____ (1994): "Ranked set sampling", Handbook of Statistics, 12 (Eds. G.P. Patil and C.R. Rao), 167-200. N. Holland/Elsevier Publishers, New York-Amsterdam.

PI-EHR, LIN (1971): "Estimation procedures for the difference of means with missing observations", **J. of the Amer. Stat. Assoc**., 41, 517-529.

TAKAHASI, K. and K. WAKIMOTO (1968): "On unbiased estimates of the population mean based on the sample stratified by means of ordering", **Ann. Inst. Statis. Math**., 20, 1-31.

WARE, K. and T. CUNIA (1962): "Continuous forest inventory with partial replacement of samples", **Forest Sc. Monographs**, 3.