MISCELLANEOUS

ON SAMPLING WITH AND WITHOUT REPLACEMENT

By D. BASU

Indian Statistical Institute, Calcutta

SUMMARY. Certain aspects of sampling with or without replacement, with equal or unequal probabilities, are considered here in some details. Some comparisons have been made between the with and without replacement sampling schemes. When we are sampling with replacement the estimate should not depend on the number of times that any particular unit may appear in the sample. Thus, certain estimation procedures in current use are shown to be inefficient.

1. Introduction

Suppose a given population has N units. Let Y_j be some real valued characteristic of the j-th population unit ($j = 1, 2, ..., N$). Consider the problem of estimating the population mean

$$\bar{Y} = N^{-1} \Sigma \; Y_j.$$

Let

$$\sigma^2 = N^{-1} \; \Sigma \, (Y_j - \bar{Y})^2$$

be the population variance..

If we draw a sample of size n from the population with equal probabilities and with replacement then the variance of the sample mean is $n^{-1} \sigma^2$. If, on the other hand, we draw a sample of the same size n but this time without replacement then the variance of the sample mean is $n^{-1} \sigma^2 (N-n)(N-1)^{-1}$. Thus, it is usually claimed that sampling without replacement is better as it leads to an estimator of $\bar{Y}$ with a smaller variance. A little reflection, however, will show that this comparison between the two methods of sampling is not usually quite fair. Let us take a simple example. Suppose the units are villages and Y_j the number of households in the j-th village. Here the cost of selecting the sample villages from a given frame is negligible compared to the cost of travelling to the selected villages and ascertaining the exact number of households in the selected villages. Generally speaking, we need only consider the cost of measuring the Y-characteristics of the selected units—the small cost involved in selecting the units from a frame may be taken to be a part of the large overhead cost. In sampling with replacement it is then the number ν of distinct units appearing in the sample (and not the sample size n) that roughly determines the cost.

2. The distribution of ν

If ν be the number of distinct units appearing in a sample of size n drawn with equal probabilities and with replacements from a population with N units then it is clear that the distribution of ν depends only on n and N. It is not difficult to show (Feller p. 92) that

$$P(\nu = s) = N^{-n} \binom{N}{s} [1 - \binom{s}{1}(s-1)^n + \binom{s}{2}(s-2)^n \; ...]$$

where s runs from 1 to the smaller of n and N.

In terms of the 'differences of zeros' we may write the above in the more elegant form

$$P(\nu = s) = N^{-n} \binom{N}{s} \Delta^s 0^n \qquad \dots \quad (2.1)$$

where Δ is the usual difference operator with unit increments and $\Delta^s\, 0^n$ is to be interpreted as $\Delta^s\, x^n$ at $x = 0$.

From (2.1) we have the probability generating function of ν as

$$P_\nu(t) = E\, t^\nu = N^{-n} \sum_{s=0}^{N} \left[\binom{N}{s}\, t^s \Delta^s 0^n\right] = N^{-n} (1+t\Delta)^N\, 0^n. \qquad \dots \quad (2.2)$$

(Note that $\Delta^s\, 0^n = 0$ for $s = 0$ and $s > n$).

Writing $1+t$ for t in the probability generating function we have the factorial moment generating function of ν as

$$F_\nu(t) = N^{-n} (\mathcal{E}+t\Delta)^N 0^n \qquad \dots \quad (2.3)$$

where $\mathcal{E} = 1+\Delta =$ the usual increment operator with unit increments.

$$\therefore \quad E(\nu) = N^{-n} \binom{N}{1} \mathcal{E}^{N-1}\, \Delta\, 0^n = N^{-n} \binom{N}{1} (\mathcal{E}^N - \mathcal{E}^{N-1})\, 0^n = N\left[\, 1 - \left(\frac{N-1}{N}\right)^n\right]. \qquad \dots \quad (2.4)$$

$$\text{Also} \quad E\nu(\nu-1) = N^{-n} \binom{N}{2} 2\mathcal{E}^{N-2}\, \Delta^2\, 0^n = N^{-n}\;\; N(N-1)(\mathcal{E}^N - 2\mathcal{E}^{N-1} + \mathcal{E}^{N-2})\; 0^n$$

$$= N(N-1)\left[1 - 2\left(\frac{N-1}{N}\right)^n + \left(\frac{N-2}{N}\right)^n\right]$$

$$\therefore \quad V(\nu) = N\left(\frac{N-1}{N}\right)^n - N^2\left(\frac{N-1}{N}\right)^{2n} + N(N-1)\left(\frac{N-2}{N}\right)^n. \qquad \dots \quad (2.5)$$

3. Sampling cost considerations

If we assume that the variable part of the cost of sampling is proportional to the number of distinct units in the sample then we may compare the two methods of estimating $\bar{Y}$, as described in § 1, in the following manner. The expected sampling cost for a sample of size n with replacement is equal to the sampling cost for a sample of size $E\nu = N\left[\, 1 - \left(\frac{N-1}{N}\right)^n\,\right]$ without replacement (let us conveniently forget the fact that $E\nu$ is not necessarily an integer). The variances for the sample means for the two cases are then $n^{-1}\sigma^2$ and $(E\nu)^{-1}\sigma^2(N-E\nu)(N-1)^{-1}$ respectively. A little computation will show that the former is larger. Thus, from this comparison, sampling with replacement appears to be worse than sampling without replacement. This comparison between the two methods heavily depends on the assumption of linearity of the cost function and as such is not very satisfactory. For a different cost function sampling with replacement may appear to fare better than sampling without replacement. The issue that is raised in the next section is perhaps more pertinent to the problem.

4. Two estimators from a with replacement sample

If we draw a sample of size n with replacements and with equal probabilities and if ν be the number of distinct units appearing in the sample then the average Y-characteristic of the ν distinct units is also an unbiased estimator of $\bar{Y}$. We may enquire whether this estimator is better or worse than the average over all the n units. The variance of the former is

$$E\left(\frac{N-\nu}{N-1}\cdot\frac{\sigma^2}{\nu}\right) \qquad \ldots \quad (4.1)$$

whereas that of the latter is $n^{-1}\sigma^2$.

It is not possible to give a simple expression for (4.1). In the next section we shall give an indirect proof of the inequality.

$$E\left(\frac{N-\nu}{N-1}\cdot\frac{\sigma^2}{\nu}\right) < \frac{\sigma^2}{n} \quad (\text{if } n > 1) \qquad \ldots \quad (4.2)$$

Thus, the average Y-characteristics of the ν distinct units in the sample is a better estimator than the average over all the n units. This result may at first appear to be a little unfamiliar, even surprising. Let us take, for example, the familiar Binomial model where we draw n balls at random one by one and with replacements from an urn containing N identical balls Np of which are white, the rest being black. Here the sample observation consists of a sequence of n white or black balls. The number r of white balls in the sample then constitutes a complete[1] sufficient statistic for the unknown parameter p. (Note that in this situation the distribution of the sample depends only on p and not on N.) Hence r/n is the uniformly minimum variance unbiased estimator of p. Now suppose that the N balls are distinguishable from one another (as for example when the balls are villages) or suppose we put distinguishing marks on the balls drawn before they are replaced. The sample observation then is a sequence of n balls and there are N^n possible sample observations each having the same probability. (In the previous case the probability of getting a particular sample observation was $p^r(1-p)^{n-r}$ where r is the number of white balls in the sample). Here the sample observation is more detailed than in the previous case and actually contains more information about the parameter p. Now r is no longer a sufficient statistic. The ν distinct balls that came in the sample is a sufficient statistic and nothing less than this can be sufficient. Consider now a third kind of sample observation where for each of the n balls that are drawn we note down only its colour and the fact whether this particular ball was drawn before or not. Here the sample can be represented as a sequence of n whites and blacks with cross marks at ν places (ν a variable) to indicate at which draws we had the distinct balls. The sample observation now is more detailed than the first case and less so than the second. If ρ be the number of distinct white balls then the statistic (ρ, ν) is sufficient. The conditional expectation of r/n for fixed values of (ρ, ν) is ρ/ν and so by the Rao-Blackwell theorem ρ/ν is better than r/n.. Here the statistic (ρ, ν) though sufficient is not complete. This is obvious from the fact that the distribution of ν is independent of the parameter p. Thus

[1] The distribution of r is complete if there are at least $n+1$ admissible values for p. Thus if N be smaller than n then the distribution of r will not be complete.

we are unable[1] to prove that ρ/ν is the best unbiased estimator of p. The proof sketched above only demonstrators that, with the additional information of which are the distinct units, the standard estimator r/n is no longer the best estimator and that it is in fact worse than ρ/ν. In the next section we give a proof of inequality (4.2) in the general case.

5. Proof of (4.2.)

Let there be N population units and let Y_j be the Y-characteristic of the j-th population unit ($j = 1, 2, \dots, N$). A sample $\mathbf{S}$ of size n is drawn one by one, with equal probabilities and with replacements. Let y_i be the observed Y-characteristic of the i-th sample unit ($i = 1, 2, \dots, n$). For each sample unit suppose we also note down its unit index (if a particular sample unit happens to be the j-th population unit then its unit-index is j). Let u_i be the unit-index of the i-th sample unit and let $\mathbf{x}_i = (y_i, u_i)$. We can then record the sample observation as

$$\mathbf{S} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$$

where the $\mathbf{x}_i$'s are independently and identically distributed random vectors.

Let ν be the number of distinct sample units and let $u_{(1)} < u_{(2)} < \dots < u_{(\nu)}$ be their unit-indices written in an ascending order. Let $y_{(i)}$ be the Y-characteristic of the sample unit with unit-index $u_{(i)}$ and let $\mathbf{x}_{(i)} = (y_{(i)}, u_{(i)})$ $i = 1, 2, \dots, \nu$.

Consider now the set of 'order-statistics'

$$\mathbf{T} = (\mathbf{x}_{(1)}, \mathbf{x}_{(2)} \dots, \mathbf{x}_{(\nu)}).$$

The usual estimator of $\bar{Y}$ (the population mean) is based on all the n observations and is

$$\bar{y} = \bar{y}(\mathbf{S}) = n^{-1} \sum_1^n y_i$$

whereas the estimator based on the ν distinct units is

$$\bar{y}^* = \bar{y}^*(\mathbf{T}) = \nu^{-1} \sum_1^\nu y_{(i)}.$$

Now, for fixed $\mathbf{T}$, the conditional distribution of $\mathbf{x}_i$ is concentrated at the ν points $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(\nu)}$ with equal probabilities at each of these points.

$$\therefore \quad E(y_i | \mathbf{T}) = \bar{y}^* \quad (i = 1, 2, \dots, n)$$

and hence

$$E(\bar{y} | \mathbf{T}) = \bar{y}^*.$$

Since $\bar{y}$ is an unbiased estimator of $\bar{Y}$, it follows at once that $\bar{y}^*$ is also unbiased. It also follows that, for any convex (downwards) loss function, $\bar{y}^*$ has a uniformly better risk function than $\bar{y}$. In particular

$$V(\bar{y}^*) \leqslant V(\bar{y})$$

the sign of equality holding only when $n = 1$.

Thus the inequality (4.2) is proved. We may note in passing that $\mathbf{T}$ is a sufficient statistic here though not a complete one. No uniformly best unbiased estimator of $\bar{Y}$ exists.

[1] If the parameter N is also unknown then it is easily demonstrated that the distribution of ν is complete. In this situtation we believe that ρ/ν is the best unbiased estimator of p.

6. The case when ν is fixed in advance

In the previous sections we fixed n and had ν as a random variable. Here we consider the situation where the number ν of distinct units in the sample is fixed in advance. We go on drawing samples one by one, with equal probabilities, and with replacements until we get n distinct units. The probability distribution of the number n of samples drawn may be obtained as follows. The event $n = k$ means that in the first $k-1$ draws there are exactly $\nu-1$ distinct units and that the k-th unit drawn is different from the $\nu-1$ distinct units that appeared in the first $k-1$ cases. Thus from (2.1) it follows that

$$P(n = k) = [N^{-(k-1)} \tbinom{N}{\nu-1} \Delta^{\nu-1} 0^{k-1}] \left(1-\frac{\nu-1}{N}\right) = \tbinom{N-1}{\nu-1} \Delta^{\nu-1}\left(\frac{x}{N}\right)^{k-1}\Bigg|_{x=0} \quad \ldots \quad (6.1)$$

where $$k = \nu, \nu+1, \ldots \text{ ad inf.}$$

From (6.1) it follows that the probability generating function of n is

$$P_n(t) = E\,t^n = \sum_{k=\nu}^{\infty} t^k\,P(n=k) = \tbinom{N-1}{\nu-1}\,t \sum_{k=\nu}^{\infty} \Delta^{\nu-1} \left(\frac{xt}{N}\right)^{k-1}\Bigg|_{x=0}$$

where Δ operates on x.

Since $\Delta^{\nu-1} x^r = 0$ for $r < \nu-1$ we have

$$P_n(t) = \tbinom{N-1}{\nu-1}\,t\Delta^{\nu-1}\left[\sum_{k=1}^{\infty}\left(\frac{xt}{N}\right)^{k-1}\right]_{x=0} = \tbinom{N-1}{\nu-1}\,t\Delta^{\nu-1}\left(1-\frac{xt}{N}\right)^{-1}\Bigg|_{x=0} \quad \ldots \quad (6.2)$$

In a like manner we have

$$E(n) = \tbinom{N-1}{\nu-1}\,\Delta^{\nu-1}\left(1-\frac{x}{N}\right)^{-2}\Bigg|_{x=0} \quad \ldots \quad (6.3)$$

If $\bar{y}$ be the average Y-characteristic of all the n observations then $\bar{y}$ is an unbiased estimator of $\bar{Y}$ with variance

$$\begin{aligned} V(\bar{y}) &= E\frac{\sigma^2}{n} = \sigma^2 \sum_{k=1}^{\infty} \frac{1}{k}\,P(n=k) \\ &= \sigma^2 \tbinom{N-1}{\nu-1}\,\Delta^{\nu-1}\sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{x}{N}\right)^{k-1}\Bigg|_{x=0} \\ &= \sigma^2 \tbinom{N-1}{\nu-1}\,\Delta^{\nu-1}\left[\frac{N}{x}\,\log\,\frac{N}{N-x}\right]_{x=0} \end{aligned} \quad \ldots \quad (6.4)$$

where, for $x = 0$, $\frac{N}{x}\log\frac{N}{N-x}$ is to be interpreted as 1.

If $\bar{y}^*$ be the average Y-characteristics of the ν distinct units then

$$V(\bar{y}^*) = \frac{N-\nu}{N-1}\,\frac{\sigma^2}{\nu}. \quad \ldots \quad (6.5)$$

That (6.5) is smaller than (6.4) may be proved in precisely the same way as we proved a similar result in §5.

7. Sampling with unequal probabilities

We consider now the more general situation where sampling is done with different probabilities attached to the different population units. Let P_j be the probability associated with the j-th population unit $(\Sigma P_j = 1)$. Suppose a sample of size n is taken with replacement and with the P_j's as the probabilities. Let us suppose that for the i-th sample unit we record its Y-characteristic y_i, its probability of selection p_i, and its unit-index $u_i (i = 1, 2, \ldots, n)$. Thus, the sample is

$$\mathbf{S} = (\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_n)$$

where

$$\mathbf{x}_i = (y_i, p_i, u_i) \quad (i = 1, 2, \ldots, n).$$

Clearly the $\mathbf{x}_i$'s are independently and identically distributed random vectors.

As in §5 let us define ν as the number of distinct x_i's and as before let $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \ldots \mathbf{x}_{(\nu)}$ be an arrangement of the ν distinct $\mathbf{x}_i$'s in ascending order of their unit-indices. Let

$$\mathbf{T} = (\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \ldots, \mathbf{x}_{(\nu)}).$$

It is easily seen that in this case also $\mathbf{T}$ is a sufficient statistic. Given $\mathbf{T}$, there are only

$$n! \; \Sigma' \; (\alpha_1! \quad \alpha_2! \quad \ldots. \quad \alpha_\nu!)^{-1}$$

(where the summation is taken over all positive integer α_i's such that $\alpha_1 + \ldots + \alpha_\nu = n$) values that $\mathbf{S}$ may take and the probability for each of these can be computed from the information $\mathbf{T}$ alone. Thus, any admissible estimator of $\bar{Y}$ must necessarily be a function of the statistic $\mathbf{T}$ alone. The usual estimator $\bar{y}$ of $\bar{Y}$ is based on all the n observations and is

$$\bar{y} = \bar{y}(\mathbf{S}) = \frac{1}{n} \sum_{1}^{n} \left(\frac{y_i}{Np_i} \right). \qquad \ldots \quad (7.1)$$

From the sufficiency of $\mathbf{T}$ it follows that

$$\bar{y}^* = E(\bar{y} | \mathbf{T}) = E \left(\frac{y_1}{Np_1} \middle| \mathbf{T} \right) \qquad \ldots \quad (7.2)$$

is a better unbiased estimator of $\bar{Y}$.

It is not difficult to show that

$$P(\mathbf{x}_1 = \mathbf{x}_{(i)} | \mathbf{T}) = \frac{p_{(i)} \sum'' \frac{(n-1)!}{\alpha_1! \, \alpha_2! \ldots \alpha_\nu!} \, p_{(1)}^{\alpha_1} \, p_{(2)}^{\alpha_2} \cdots p_{(\nu)}^{\alpha_\nu}}{\sum' \frac{n!}{\alpha_1! \alpha_2! \ldots \alpha_\nu!} \, p_{(1)}^{\alpha_1} \, p_{(2)}^{\alpha_2} \cdots p_{(\nu)}^{\alpha_\nu}}$$

$$= c_i (\text{say}) \quad (i = 1, 2, \ldots, \nu)$$

where Σ' means summation over all integral α's such that

$$\alpha_1 + \alpha_2 + \ldots + \alpha_\nu = n \;\text{ and }\; \alpha_k > 0 \;\text{ for }\; k = 1, 2, \ldots, \nu$$

and Σ'' means summation over all integral α's such that

$$\alpha_1 + \alpha_2 + \ldots + \alpha_\nu = n - 1, \alpha_i \geqslant 0 \text{ and } \alpha_k > 0 \text{ for } k \neq i.$$

Thus

$$\bar{y}^* = E \left\{ \frac{y_1}{Np_1} \middle| \mathbf{T} \right\} = \sum_i c_i \frac{y_{(i)}}{Np_{(i)}}. \qquad \ldots \quad (7.3)$$

Unfortunately, it is rather troublesome to compute the c_i's. In the particular case where $n = 3$ and $\nu = 2$ we have

$$c_1 = (2p_{(1)} + p_{(2)}) / 3(p_{(1)} + p_{(2)})$$

and

$$c_2 = (p_{(1)} + 2p_{(2)}) / 3(p_{(1)} + p_{(2)}).$$

The estimator $\bar{y}^*$, though demonstrated to be superior to the usual estimator $\bar{y}$, cannot be of much use for large scale sample surveys. It is even more troublesome to estimate the variance of $\bar{y}^*$. An unbiased estimator for the variance of $\bar{y}$ is

$$\frac{1}{n(n-1)} \sum_{1}^{n} \left\{ \frac{y_i}{Np_i} - \bar{y} \right\}^2. \qquad \dots \quad (7.4)$$

The above will over-estimate the variance of $\bar{y}^*$ and so it will be on the safe side to take (7.4) as an estimator of the variance of $\bar{y}^*$.

8. The mean and variance of ν

In § 2 we have given the distribution of ν for the particular case where sampling is done with equal probabilities. In the unequal probability set-up the distribution of ν becomes very messy. Here we give expressions for the mean and variance of ν.

Let z_j be the characteristic function of the event that the sample of n units includes the j-th population unit.

Clearly

$$P(z_j = 1) = 1 - Q_j^n \quad (j = 1, 2, \dots, N)$$

where

$$Q_j = 1 - P_j.$$

Since

$$\nu = z_1 + z_2 + \dots, + z_N \qquad \dots \quad (8.1)$$

we have

$$E(\nu) = \sum_{1}^{N} (1 - Q_j^n). \qquad \dots \quad (8.2)$$

Also

$$V(\nu) = \Sigma \; V(z_j) + \sum_{i \neq j} \text{cov}\; (z_i, z_j). \qquad \dots \quad (8.3)$$

Now,

$$V(z_j) = Q_j^n (1 - Q_j^n)$$

and

$$\begin{aligned} \text{cov}(z_i, z_j) &= P(z_i = z_j = 1) - P(z_i = 1) P(z_j = 1) \\ &= (1 - Q_i^n - Q_j^n + Q_{ij}^n) - (1 - Q_i^n)(1 - Q_j^n) \\ &= -(Q_i^n \; Q_j^n - Q_{ij}^n) \end{aligned}$$

where

$$Q_{ij} = 1 - P_i - P_j$$

$$\therefore \quad V(\nu) = \Sigma \; Q^n(1 - Q^n) - \sum_{i \neq j} (Q^n Q^n - Q_{ij}^n) = \Sigma \; Q_j^n - (\Sigma \, Q_i^n)^2 + \sum_{i \neq j} Q_{ij}^n. \qquad \dots \quad (8.4)$$

9. Unequal probabilities and without replacement

Now let us consider the case of sampling without replacement and with different probabilities. As in §7 let P_j be the probability attached to the j-th population unit. Let y_i and p_i be the Y-characteristic and the probability corresponding to the i-th sample unit $(i = 1, 2, \dots, n)$. Writing $\mathbf{x}_i = (y_i, p_i)$ we can record the sample observation[1] as

$$\mathbf{S} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n). \qquad \dots \quad (9.1)$$

[1] Here we need not record the unit indices of the sample units as we know that they must be all different. Unless we have some additional information about the population units it appears that it is impossible to utilise any information about the sample unit-indices to improve on any estimator of $\bar{Y}$.

Now, let us order the $\mathbf{x}_i$'s by some method, say, in ascending order of the y_i's and for $\mathbf{x}_i$'s with equal y_i's, in ascending order of their p_i's. Let $\mathbf{x}_{(i)}$ be the i-th order statistic and let

$$\mathbf{T} = (\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \ldots, \mathbf{x}_{(n)}) \qquad \ldots \quad (9.2)$$

be the set of order statistics.

For given $\mathbf{T}$, there are $n!$ or less (in the case where some of the $\mathbf{x}_{(i)}$'s are the same) values that $\mathbf{S}$ may take and the conditional probability for each of these may be computed from the information $\mathbf{T}$ alone. Thus, in this case also $\mathbf{T}$ is a sufficient statistic. Hence no estimator that is not a function of $\mathbf{T}$ alone can be admissible. Any estimator that makes use of the order in which the sample was drawn can be uniformly improved upon by its conditional expectation given $\mathbf{T}$.

For example, consider the particular case of $n = 2$. We may estimate $\hat{Y}$ from the first component of $\mathbf{S}$, i.e., from $\mathbf{x}_1$. This estimator is, of course, $y_1 | Np_1$.

Now $$P(\mathbf{x}_1 = \mathbf{x}_{(1)} | \mathbf{T}) = \frac{p_{(1)}\, p_{(2)}/(1-p_{(1)})}{p_{(1)}\, p_{(2)}/(1-p_{(1)}) + p_{(2)}\, p_{(1)}/(1-p_{(2)})} = \frac{1-p_{(2)}}{2-p_{(1)}-p_{(2)}} \qquad \ldots \quad (9.3)$$

and similarly

$$P(\mathbf{x}_1 = \mathbf{x}_{(2)} | \mathbf{T}) = \frac{1-p_{(1)}}{2-p_{(1)}-p_{(2)}} \qquad \ldots \quad (9.4)$$

$$\therefore \; E\left(\frac{y_1}{Np_1}\middle| \mathbf{T}\right) = \left[(1-p_{(2)})\,\frac{y_{(1)}}{Np_{(1)}} + (1-p_{(1)})\frac{y_{(2)}}{Np_{(2)}}\right] \times$$
$$\times (2-p_{(1)}-p_{(2)})^{-1}. \qquad \ldots \quad (9.5)$$

The estimator (9.5) is uniformly better than $y_1 | Np_1$.

Since $\mathbf{T}$ is not a complete sufficient statistic, we cannot prove that (9.5) is the uniformly best estimator. For further discussion about the problem dealt with in this section one may refer to (Murthy, 1957).

Acknowledgement

The author wishes to thank Mr. S. Raja Rao who suggested the problem in 1952 and then collaborated with the author in obtaining the results contained in § 2 to § 6.

References

Feller, W. (1957) : *An introduction to probability theory and its applications*, John Wiley & Sons New York.

Fraser, D. A. S. (1957) : *Non parametric methods in statistics*, John Wiley & Sons New York.

Murthy, M. N. (1957) : Ordered and unordered estimators in sampling without replacement. *Sankhyā*, 18, 379.

Paper received : January, 1958.