RAJASTHAN UNIVERSITY STUDIES

IN

STATISTICS

1965

DEPARTMENT OF STATISTICS
UNIVERSITY OF RAJASTHAN
JAIPUR, India

CONTENTS.

* Please read "THEIR" instead of "ITS" in the first line of the heading of the article on page 1.

SOME ASPECTS OF QUEUES AND ITS ENGINEERING APPLICATIONS[1]

by

B. D. Tikkiwal

University of Rajasthan Jaipur

0.1 Summary

This paper presents first some management problems capable of being solved, in part or whole, by queueing theory. Then, it gives some aspects of queueing theory and its use in statistical analysis of the management problems with special reference to the historic problem of Telephony.

.1. Introduction

The scope of Operational Research is indicated from its following workable definition (cf. Ackoff, 1961; Beer, 1959).

'Operational research is the attack of modern science on problems of likelihood (accepting mischance) which arise in the management and control of men and machines, materials and money in their natural environment. Its special technique is to invent a strategy on control by measuring, comparing and predicting probable behaviour through scientific model of a situation.'

The theory of queues deals with the scientific models of the situations resulting in waiting lines of one kind or the other. It was first developed to solve the management problems in telephony. Following the pioneering work by Erlang of the Copenhagen Telephone Company in 1909, at the suggestion of F. Johansen, himself a pioneer contributor to the subject, fairly extensive research on the theory of queues has been done, specially after 1950. A systematic treatment of the theory from the point of view of stochastic processes

---

1 This paper was presented before the special seminar conducted by the Institution of Engineers on some selected topics on statistical techniques in modern management in December, 1964 at Hyderabad. The author is grateful to the organisers of the seminar, Mr. M. K. Rao and Prof, S. K. Ekambaram for their kind invitation.

is due to Kendall (1951, 1954). This paper gives some rudiments of queueing theory along with some aspects of its applications.

2. Some Management Problems

Before presenting the queueing theory, we first enumerate below some management problems where queueing theory can be usefully applied.

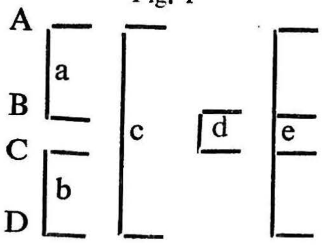
1. We find queues of patients requiring services at different doctors at out patient clinics.
2. There are queues of persons requiring services at different shops in different markets.
3. Articles passing alongwith conveyor belt in order to be packed in boxes.
4. The machines stopping intermittently and requiring attention of the mechanic before re-starting.
5. The air-crafts waiting to land or to take off at the air-ports.
6. The telephone calls requiring attention of the operators in a given area.

We discuss problems 5 and 6 in detail.

In the air-craft problem we assume that a particular air-port has several run-ways used for taking off and for landing. The air-ports are allowed to take off according to certain priority arrangements. Certain types of air crafts are allowed to take off in order of their arrivals whereas others such as Jets are given priority over the ordinary air crafts. The same is true about the priority arrangements, while allowing the air craft to land.

The telephony problem is a historic problem referred earlier in Sections O and 1. The figure I below gives a representative problem common to different telephony systems.

Fig. 1



This figure gives, for a particular area, a typical but sophisticated limited-availability telephonic system in which the line a is available to calls from the localities *A* and *B* only. The line *b* is available only to calls from *C* and *D*. The line *c* is available to calls from *A* and *D*. The line *d* is available to *B* and *C*. The line *e* is available to all localities.

3. The three Specifications of Queueing Theory

In all the above examples given in Section 2, we need to know (i) the arrival pattern (Input process) (ii) the queue-discipline (iii) the service-mechanism. A queue-system is completely specified by these three things. We shall refer these as three specifications of a queue-system. We discuss the three specifications in some detail, with reference to the historic problem of telephony, wherever possible without loss of generality.

(i) *The Arrival Pattern.*

The arrival pattern can be of many kinds. We discuss here only the Random (or Poissonian) pattern, which has found extensive use in telephony system.

The Random Pattern :

Let m calls be made in time t. Let $m \triangle t = t$. Let $\alpha \triangle t$ be the probability that a call is made in an interval $\triangle t$ an assumption due to Erlang. Let this event of the call being made in a particular interval be independent of the events of a call or no call in any other interval in (O, t). Let $P(r)$ denot the probability of r calls being made in (O, t). Then,

$$P(r) = \frac{m1}{r1\ m-r1} (\alpha \triangle t)^r (1 - \alpha \triangle t)^{m-r}$$

Therefore,

$$(3.1) \quad \lim_{m \to \infty} P(r) = \frac{\alpha^r\ t^r}{r1}\ e^{-\alpha t}$$

a well-known Poissonian frequency function. We know,

$$(3.2) \quad E(r) = \alpha\ t$$

$$(3.3) \quad \text{Var}\ (r) = \alpha\ t.$$

$$(3.4) \quad P(r \geqslant 2) = \sum_{r=2}^{\infty} \frac{\alpha^r t^r}{r!} e^{-\alpha t}$$

$$= (1 - \alpha t + \alpha^2 t^2 \ldots) \frac{\alpha^2 t^2}{2!} + \ldots$$

$$= O(t^2)$$

If t is small, Equation (3.4) will mean that the probability of getting two or more calls in a short time will also be small. Equation (3.2) means that, on an average, there are α calls per unit time. Equations (3.2) and (3.3) together give the coefficient of variation $= 100\sqrt{(\alpha t)}$. For $r=0$, Equation (3.1) gives the probability that the interval ξ, in which no call arrives, is less than or equal to t. Thus,

$$P(\xi \leqslant t) = e^{-\alpha t}.$$

Therefore,

$$(3.5) \quad f(t) = \lim_{\Delta t \to 0} \frac{e^{-\alpha t} - e^{-\alpha (t + \Delta t)}}{\Delta t}$$

$$= e^{-\alpha t} \lim_{\Delta t \to 0} \frac{1 - e^{-\alpha \Delta t}}{\Delta t}$$

$$= \alpha e^{-\alpha t}, \quad 0 \leqslant t \leqslant \infty$$

It may be noted that

$$(3.6) \quad \begin{cases} E(t) = \dfrac{1}{\alpha} \text{ and} \\ V(t) = \dfrac{1}{\alpha^2}. \end{cases}$$

Equation (3.6) means that, on an average, no calls come in $1/\alpha$ time, a result which is in accordance with the above result that, on an average, there are α calls per unit time. The coefficient of variation is 100 percent, the highest possible coefficient of variation.

Equation (3.5) can be interpreted as the frequency function for the interval from a fixed time to the time one call is received. Similarly, the frequency $fk(t)$ of the interval from a fixed time to the time k calls are received; is obtained as follows. From Equation (3.1),

$$P(r = k-1) = \frac{\alpha^{k-1} t^{k-1}}{(k-1)!} e^{-\alpha t}.$$

Therefore, the probability; that the interval ξ, in which $(k-1)$ calls arrive, is less than or equal to t; is given by

$$P(\xi \leqslant t) = \frac{\alpha^{k-1} t^{k-1}}{(k-1)!}$$

Thus,

$$(3.7)\; f(t) = \lim_{\Delta t \to 0} \frac{\alpha^{k-1} t^{k-1}}{(k-1)!} e^{-\alpha t} \quad \frac{1 - e^{-\alpha \Delta t}}{\Delta t}$$

$$= \alpha \frac{(\alpha t)^{k-1}}{(k-1)!} e^{-\alpha t}$$

We now present an alternative method. The laplace transform (Rainville, 1963) of Equation (3.5) is given by

$$L\left\{f(t)\right\} = \int_0^\infty e^{-st} \alpha e^{-\alpha t}\, dt.$$

$$= \frac{\alpha}{\alpha + s}$$

The interval from a fixed time to the time k calls are received is the sum of the k intervals, the i *th* interval being the interval from the time the $(i-1)$ *th* call is received to the time the i *th* call is received for $i = 1, 2, \ldots\ldots, k$; the time of zero call being referred as the fixed time. Since the k random intervals are statistically independent because of earlier assumption regarding the arrival of calls, the Laplace transform of $f_k(t)$ the frequency function of the interval from the fixed time to the time k calls are received is given by

$$L\left\{f_k(t)\right\} = \left(\frac{\alpha}{\alpha + s}\right)^k$$

Thus $f_k(t)$ is given by

$$(3.8)\; f_k(t) = \frac{\alpha (\alpha t)^{k-1} e^{-\alpha t}}{(k-1)!}, \; t \geqslant 0.$$

For $k = 1$, we get Equation (3.5) and for $k=\infty$

$$\lim_{k\to\infty} \frac{(\alpha t)^{k-1}}{(k-1)!} = 0, \; \alpha t < 1$$

being a term in the expansion of $e^{\alpha t}$. Therefore

$$\lim_{k\to\infty} f_k(t) = 0, \; \alpha t < 1.$$

Thus, for $k=\infty$, (3.8) gives

$$(3.9) \quad F_k(t) = \begin{cases} 0, & t < 1/\alpha \\ 1, & t \geqslant 1/\alpha \end{cases}$$

a result similar to that given by Kendall (1953).

It is clear from this result that for k large, the frequency function $f_k(t)$ is concentrated at the point $t=1/\alpha$, suggesting thereby that for k large, the random pattern behaves as the regular one in whieh calls are supposed to come at equally spaced intervals. Such an assumption of regular pattern does not normally hold good in telephony problems, but it is likely to hold good in problems like those of Conveyor-belt systems. In the regular pattern the coefficient of variation is zero, the lowest possible coefficient of variation. As in Erlangian service-time model, to be discussed later, we would like to have a frequency function which is flexible in nature for applicational purpose and whose coefficient of variation lies between 0 and 100. This suggests that we may have $f(t)$ as given in Equation (3.8) which includes the expression in Equation (3.5) as a special case. The coefficient of variation of this frequency function is (100/k).

(ii) *The queue discipline.*

It gives, in general, a procedure of selecting. for service, the customers (calls in case of telephony) waiting in queue. In telephony, the calls are generally dealt with in a random manner with respect to the order of their arrivals. However, there are situations where the procedure of 'First come, first-severed' or of 'Last come, first-served' is followed. There are also situations where neither of the procedures given above is followed e.g., in the air-crafts problem, where certain types of air-crafts are given priority over other types irrespective of the manner of their arrivals.

(iii) *The service Mechanism.*

There are three aspects of service-mechanism : (1) Service-time (2) Servic-capacity and (3) service-availability. The service capacity depends upon the number of servers available for service and the service availability is said to be complete if one server or the other is always available for service. In telephony, we normally have several operators for a given system and the service is continusly available. As regards service time, there are different service models. But we discuss here the random service model, a model which is most useful in telephony and which is similar to that in random pattern. In the discussions for random pattern; we know that $f(t)$ in Equation (3.5) also denotes the frequency function for the interval in which one call is completed. Thus, we may take the frequency function of the service-time for a call as

$$(3.10) \qquad g(t)=\beta e^{-\beta t};\ t \geqslant 0,\ \beta > 0.$$

On the lines of discussion in case of random pattern, we can in fact take a more general frequency function

$$g_1(t)=\frac{\beta\,(\beta t)^{k-1}}{(k-1)\,!}\,e^{-\beta t},\ t \geqslant 0;$$

known as Erlangian service-time frequency function (3). If there is only one server, we would like that the service-time, on an average, should be less than or equal to the average-interval in which a new customer would arrive otherwise the queue will grow indefinitely long. Thus, for any desirable situation, we have $\beta \leqslant \alpha$. This inequality is necessary for a system to be in statistical equilibrium, where the number of customers waiting in the queue oscillates in such a way that its mean and distribution remain constant over a long period. The traffic intensity I is, by definition, equal to β/α. Thus, for a system to be in statistical equilibrium we require $I \leqslant 1$.

4. Some aspects of Statistical Analysis of Waiting-lines (Congestion)

We would like a situation where the customers do not wait long with minimum of efficient service. For this we need to know

(i) The mean-arrival time of the customer.

(ii) The mean-service time of the customer.

(iii) The mean-queue time the customer has to wait in the queue before he is served.

(iv) The average rate at which customers pass through a particular system.

(v) The proportion of time a server is busy. Assuming that there is only one server, we can easily find above quantities without using much of statistics as such. But if we are to find how the things will modify by addition of one server or more, the queueing theory comes to our rescue. If we assume (the assumption can be verified by actual collection of data) that for a given situation, the random model holds; then we know (i) and (ii) from theoretical discussions given in Section 3. The mean-queueing time is given by

(4.1) $$Y = \frac{I}{2\,(1-I)}\,\beta\,(1+c_s{}^2)$$

known as Pollaczek's formula (Kendall, 1951), with C_s as the coefficient of variation of servie-time. The average rate at which customers pass through a particular system in statistical equilibrium is $1/\alpha$. For this system, the proportion of time the server is busy is equal to the traffic intensity.

REFERENCES

1. Ackoff, R. L. (1961), "Progress in operational research", Vol. L New York-London : John Willey and Sons. Inc.
2. Bear, Stafford (1959), "Cybernatics and operational research", Operational Research Quarterly 10, 1-21.
3. Brockmayer, E. Hlastrom, H. L. and Jensen A. (1948), "The Life and Work of A. K. Erlang", Copen hagen Telephone Company, Copehagen.
4. Cox, D. R. (1955), "The statistical analysis of congestion", Jour. Royal Statist. Society, Vol. 115, series A, 324-335.
5. Cox, D. R. and Smith W. L. (1963), "Queues", New York : John Wiley & Sons Inc. Loudon : Methuen & Co., Ltd.,
6. Erlang, A. K. (1909), "Probability and Telephone calls", Nyt Tidsskr Mat. Vol 20, series B 33-39.

7. Fry, T. C. (1928), "Probability and its Engineering uses", New York: Van Nostrand.

8. Kendall, D. G. (1951), "Some problems in the theory of Queues", Jour. Royal Statist. Society Vol 13 series B, 1-185.

9. Kendall, D. G. (1953), "Stochastic processes occuring in the theory of queues and their analysis by means of the imbeded Markov chains", Ann. Math. Statist. Vol. 24, 338-354.

10. Rainville, E. D. (1963), "The Laplace transform", New York : Macmillan Company.

11. Saaty, T. L. (1961), "Elements of Queueing Theory with applications", Mcgraw Hill Book Co. Inc. (Gives extensive bibliography on Queueing theory).

Paper received : January, 1965

REGRESSION METHODS OF ESTIMATION[1]

By
B. D. Tikkiwal
UNIVERSITY OF RAJASTHAN,
JAIPUR, INDIA

1. Summary and Introduction

This paper includes the following topics under Regression Methods of estimation :

(i) The classical regression method of estimation;

(ii) Double sampling estimation procedure;

(iii) Successive Sampling.

The ratio method of estimation is a special case of the classical regression method and has not been dealt here. The classical regression method of estimation and double sampling estimation procedure are given in Sections 3 to 8 and successive sampling in Section 9. The lemmas, which are useful in deriving the theory of different methods of estimation, are given in Section 2 below.

2. Lemmas

Let $X_1, X_2, \ldots, X_p$ be p variates with their first and second moments finite. Without any loss of generality, let $E(X_i)=0$ for $i=1, 2, \ldots, p$. A linear bond between $X_1, X_2, \ldots, X_p$ is said to exist if $\Sigma C_i X_i=0$, $\Sigma C_i^2 \neq 0$. A system of r ($\leqslant p$) linear bonds between $X_1, X_2, \ldots, X_r$ is said to be complete, if all their bonds of the system are independent and any other bond of the system depends linearly on the r bonds. The rank $r(F)$ of the joint distribution $F(x_1, x_2, \ldots, x_p)$ of $X_1, X_2, \ldots, X_p$ is then $p-r$ by definition. When $r(F)=p$, the distribution is said to be non-degenerate. We give below lemmas 2.1 and its corollary 2.1 based on results of Lukomski (1939).

---

1. This paper is based on a course of five lectures delivered at the Summer School organised by the University of Kerala in 1965. The author is grateful to the organisers of the school for their kind invitation.

Lemma 2.1. The rank $r\ (F)$ of distribution $F(x_1, x_2, .., x_p)$ is equal to the rank of the covariance-matrix $\| \mu_{ij} \|$ where $\mu_{ij} = E(X_i X_j)$ for $i, j = 1, 2, ..., p$.

Proof ; Let $v = t_1 X_1 + t_2 X_2 + \dots, t_p X_p$ where $t_1, t_2, \dots, t_p$ are arbitrary real numbers not all equal to zero. Let

$$q = E(v^2) = \int\int\dots\int(\sum_{i=1}^{p} t_i x_i)^2 \, dF$$
$$= \sum_{i,j}\sum t_i t_j \mu_{ij}$$

Since $q \geqslant o$, it is minimum whenever $\Sigma t_i x_i = 0$ and vice-versa. Thus all the linear bonds of the system are obtained by setting $\partial q/\partial t_i = 0$ for $i = 1, 2, \dots, p$ and therefore by the equations $\Sigma_{j=1}^{p} \mu_{ij} t_j = 0$, $i = 1, 2, \dots, p$. If rank of convanance-matrix, which is also the coefficient-matrix of the p equations $\| \mu_{ij} \| = p - r$, then there are $p-(p-r)=r$ independent solutions for $t_1, t_2, \dots, t_p$ and therefore r independent linear bonds between $x_1, x_2 \dots, x_p$. This immediately gives the rank of the distribution as $p-r$, the same as that of the covariance-matrix.

Corollary 2.1. The distribution F is non-degenerate if and only if the covariance-matrix $\| \mu_{ij} \|$ is positive definite.

Proof : Let the distribution F may be non-degenerate, then the rank of $\| \mu_{ij} \| = r'$ for $i, j = 1, 2, \dots, r'$ and $r' = 1, 2, \dots, p$. Therefore, $| \mu_{ij} | > 0$ for $i, j = 1, 2, \dots, r'$ and $r' = 1, 2, \dots, p$. Hence, $\| \mu_{ij} \|$ is positive definite. The converse is obvious.

The following lemma recalls the necessary and sufficient conditions of Patterson (1950) in rigorous form (see Tikkiwal, 1960).

Lemma 2.2. Let there be a non-degenerate p - variate population with first and second moments finite. Let X_{ij} denote the outcome at the j th draw for $j=1, 2, \dots n_i$, to get a sample of size n_i for the i th variate for $i=1, 2, \dots p$. Let e_α be a linear unbiased estimator of μ_α the population mean for the α th variate based on $\Sigma\, n_i$ values at different draws. e_α is the minimum variance linear unbiased estimator (best estimator) among the class of linear unbiased estimators, if and only if

$$\text{Cov}\ (X_{ij}, e_\alpha)=k_{i\alpha}\ ,\ j=1, 2, \ldots, n_i\ ;\ i=1, 2, \ldots p.$$

Proof: Let $e_\alpha=\Sigma_{i=1}^{p} \Sigma_{j=1}^{n_i} w_{ij} X_{ij}$ where the weight w_{ij} depends upon the j th draw for the i th variate. Thus e_α is a linear estimator. Further

$$E\ (e_\alpha)=\Sigma_{i=1}^{p} \mu_i \sum_{j=1}^{n_i} w_{ij}.$$

If it is to be unbiased estimator of $\mu\ \alpha$,

$$\Sigma\ w_{ij}=\begin{cases} O\ , i \neq \alpha \\ 1\ , i = \alpha \end{cases}$$

Let

$$\phi=E\left[\sum_{i=1}^{p} \sum_{j=1}^{n_i} (w_{ij}\ X_{ij}) - \mu_\alpha\right]^2 - 2\ \Sigma\ k_{i\alpha} \sum_{j=1}^{n_i} w_{ij}$$

If e_α is to be the best estimator, $\partial\phi/\partial w_{ij}=0$
for different $(i\,j)$ and we must further verify whether

$$M=\left\| \frac{\partial^2\phi}{\partial w_{ij}\ \partial w_{i'j'}} \right\|$$

$(j=1, 2, \ldots, n_i\ ;\ j'=1, 2,\ n_i\ ';\ i, i'=1, 2, \ldots, p\ )$ is positive definite.
We note

$$\frac{\partial\varphi}{\partial w_{ij}} = 2\ \text{Cov}\ (X_{ij},\ e_\alpha) - 2\ k_{i\alpha}$$

and

$$\frac{\partial^2\phi}{\partial w_{ij}\,\partial w_{i'j'}} =2\ \text{Cov}\ (X_{ij},\ X_{i'j'}).$$

M is positive, definite if $\|\text{Cov}\ (X_{ij},X_{i'j'})\|$ is positive definite. By Corollary (2.1) this will be true if the joint distribution of $\Sigma\ n_i$ draws is non-degenerate. Since there is no linear bond between the draws for the same variate, the joint distribution of $\Sigma\ n_i$ draws will be non-degenerate provided the joint distribution of p variates is non-degenerate, which is true by assumption. Therefore, M is positive-definite. Since $\partial\phi/\partial w_{ij}=0$ for different i, j gives the conditions of the lemma, these conditions are necessary. Further when M

is positive definite, the conditions are also sufficient for ϕ to be minimum is clear from Mathematics of calculus. This completes the proof that e_α is the best estimator if and only if the conditions of the lemma are satisfied.

Corollary 2.2. If the weights w_{ij} are not to be estimated from the sample,

$$Var\ (e_\alpha) = K_{\alpha\alpha} = \text{Cov}\ (x_{\alpha j}, e_\alpha)$$

Proof. $Var\ (e_\alpha) = \text{Cov}\ [\ (\Sigma\Sigma w_{ij}\ x_{ij} - \mu_\alpha), e_\alpha]$

$$= \Sigma\ \Sigma\ w_{ij}\ \text{Cov}\ (x_{ij},\ e_\alpha)$$

$$= \Sigma\ k_{i\alpha}\ \Sigma\ w_{ij}$$

by Lemma 2.2. By using conditions of unbiasedness

$$Var\ (e_\alpha) = k_{\alpha\alpha} = \text{Cov}\ (X_{\alpha j},\ e_\alpha)$$

again by Lemma 2.2. The following lemma is due to Tikkiwal (1960).

Lemma 2.3. Let $\bar{x}$ and $\bar{y}$ be the sample means for the variates X and Y based on observations for the two variates on the same n units selected at random without replacement from the population of size N. Then

$$\text{Cov}\ (\bar{x}, \bar{y}) = \text{Cov}\ (x_i\ ,\ \bar{y}) = Cov\ (y_i\ ,\ \bar{x})$$

$$= \rho S_x\ S_y\ \frac{N - n}{Nn}$$

Where x_i and y_i denote the observations on i th unit for the variates X and Y and $S_z{}^2 = \sum_{i=1}^{N} (z_i - \mu_z)^2\ /\ (N-1)$, for $z = x,\ y$ and $\rho = \Sigma_{i=1}^{N}\ (x_i - \mu_x)\ (y_i - \mu_y)\ /\ (N-1)\ S_x\ S_y$, μ_z being equal to $\Sigma_{i=1}^{N} z_i\ /N$ for $z = x,\ y$.

If the sample mean for the variate X is based on n_x units and the sample mean for Y on n_y different units, then each of the covariance terms is given by the expression $-\rho S_x S_y / N$. This expression also determines Cov (x_i, y_j) for $i \neq j$.

Proof. $$\text{Cov}(x_i, \bar{y}) = \frac{1}{n} E\left[(x_i - \mu_x) \sum_{j=1}^{n} (y_j - \mu_y)\right]$$

$$= \frac{1}{n}\left[E(x_i - \mu_x)(y_i - \mu_y) + \sum_{j \neq i}^{n} E(x_i - \mu_x)(y_j - \mu_y)\right]$$

$$= \frac{1}{n}\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) + \frac{2(n-1)}{N(N-1)} \sum_{i=1}^{N} \sum_{j>i} (x_i - \mu_x)(y_j - \mu_y)\right]$$

$$= \frac{1}{n}\left[\frac{1}{N}\Sigma (x_i - \mu_x)(y_i - \mu_y) - \frac{(n-1)}{N(N-1)} \Sigma (x_i - \mu_x)(y_i - \mu_y)\right]$$

$$= \frac{N-n}{nN(N-1)} \Sigma (x_i - \mu_x)(y_i - \mu_y)$$

$$= \frac{N-n}{nN} \rho S_x S_y$$

Similarly,

$$\text{Cov}(y_i, \bar{x}) = \frac{N-n}{Nn} \rho S_x S_y$$

since $\text{Cov}(\bar{x}, \bar{y}) = \frac{1}{n} \sum_{i=1}^{n} \text{Cov}(x_i, \bar{y})$;

$$\text{Cov}(\bar{x}, \bar{y}) = \text{Cov}(x_i, \bar{y}) = \text{Cov}(y_i, \bar{x})$$

$$= \frac{N-n}{Nn} \rho S_x S_y$$

Further, for $i \neq j$,

$$\text{Cov}(x_i, y_j) = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j>i}^{N} (x_i - \mu_x)(y_j - \mu_y)$$

$$= \frac{-1}{N(N-1)} \Sigma (x_i - \mu_x)(y_i - \mu_y)$$

$$\frac{-\rho S_x S_y}{N}.$$

Therefore, if the sample means $\bar{x}$ and $\bar{y}$ are based on different units,

$$\text{Cov}(\bar{x}, \bar{y}) = \text{Cov}(x_i, \bar{y}) = \text{Cov}(y_i, \bar{x}) = \text{Cov}(x_i, y_j)$$
$$= -\rho S_x S_y / N.$$

3. The classical regression method of estimation when the regression coefficient is known

In survey sampling we are often interested in estimating the mean value of the character Y of a particular population when the information on the auxiliary character X of the same population is already available or can be made available easily. The estimation procedure for the mean value of Y has been discussed by several authors for two different cases (1) When the population mean of the character X is already known and we obtain the information on Y through a sample of size n (2) When the population mean of the character X is not known and it is estimated through a random sample of size n' containing the sample for Y $(n' > n)$.

The estimation procedures used in cases (1) and (2) are known in literature respectively as regression estimation and double sampling estimation procedures. We shall present in this section and the following sections 4, 6 and 8 the first method of estimation which we shall refer as classical regression method of estimation. In the present section we present the case when regression coefficient is known. Let the regression estimator $\hat{Y}_r$ for this case be given by

$$Y_r = \bar{y}_n + \beta(\mu_x - \bar{x}_n) \qquad \text{.............(3.1)}$$

where $\bar{x}_n$ and $\bar{y}_n$ are the sample means based on a sample of size n drawn by the procedure of simple random sampling without replacement from the finite population of size N.

By Lemma 2.3,

$$\text{Cov}(x_i, \hat{Y}_r) = O \text{ for } i = l, 2, \ldots\ldots\ldots, N;$$
$$\text{Cov}(y_i, \hat{Y}_r) = \frac{N-n}{Nn} S_y^2 (1 - \rho^2).$$

By Lemma 2.2, when $|\rho| < 1$ so that the Bivariate finite population is nondegenerate, $\hat{Y}_r$ is the best estimator and by Corollary 2.2,

$$\text{Var}(\hat{Y}_r) = \text{Cov}(y_i, \hat{Y}_r)$$
$$= \frac{N-n}{Nn} S_y^2 (1 - \rho^2) \qquad \text{............(3.2)}$$

To obtain an unbiased estimator of variance, we note that $\hat{Y}_r$ is the sample mean of n observation of the form

$$y_i + \beta (\mu_x - x_i)$$

which when averaged over N different values of i, gives the population mean μ_y for the variate y. Applying the well known result of simple random sampling without replacement, the unbiased estimator $\hat{V}(\hat{Y}_r)$ of $Var(\hat{Y}_r)$ is given by

$$\hat{V} (\hat{Y}_r) = \frac{N-n}{Nn} \frac{1}{n-1} \left[\sum_{i=1}^{n} \left\{ y_i - \bar{y}_n - \beta (x_i - \bar{x}_n) \right\}^2 \right] \quad \text{...(3.3)}$$

In the following section we deal with the case when β is not known and so it is estimated from the sample. In this case, we assume that the sample of size n is drawn from a bivariate normal population with parameters μ_x, μ_y, σ_x, σ_y and ρ where $|\rho| < 1$ to ensure again the population to be non-degenerate.

4. The variance of the regression estimator and its unbiased estimator under the normality assumption when β is estimated from the sample.

Let, $b = [\Sigma (x_i - \bar{x}_n) (y_i - \bar{y}_n)] / \Sigma (x_i - \bar{x}_n)^2$ be the usual estimator of β from the sample. Let the regression estimator Y_r be given by

$$Y_r = \bar{y}_n + b (\mu_x - \bar{x}_n) \quad \text{......(4.1)}$$

The estimator Y_r is no more linear in x's and y's. Further

$$E (Y_r / b) = \mu_y$$

Therefore,

$$E (Y_r) = \mu_y$$

Thus Y_r is an unbiased estimator.

$$Var (Y_r) = E [E (Y_r - \mu_y)^2 | b]$$

and by Lemma 2.3; noting that it gives various covariance terms under normality assumption by putting $N = \infty$, $S_x = \sigma_x$, $S_y = \sigma_y$;

$$E [(Y_r - \mu_y)^2 | b] = E \left[\{ \bar{y}_n - \mu_y - b (\bar{x}_n - \mu_x) \}^2 | b \right]$$

$$= \frac{\sigma_y^2}{n} + b^2 \frac{\sigma_x^2}{n} - 2 b \rho \frac{\sigma_x \sigma_y}{n} .$$

Therefore,

$$V(Y_r) = \frac{\sigma_y^{\ 2}}{n} + \frac{\sigma_x^{\ 2}}{n} E(b^2) - 2\,\beta\,\rho \frac{\sigma_x\ \sigma_y}{n}.$$

Now,

$$E(b^2) = Var(b) + \beta^2$$

But, from Bi-variate normal theory

$$Var(b) = \sigma_y^{\ 2}(1-\rho^2)\, E\left[\frac{1}{\Sigma(x_i - \bar{x}_n)^2}\right].$$

Since under normality assumption

$$\frac{\Sigma(x_i - \bar{x}_n)^2}{\sigma_x^{\ 2}} \sim x^2_{n-1},$$

$$Var(b) = \frac{\sigma_y^{\ 2}}{\sigma_x^{\ 2}}(1-\rho^2)\, E\left(\frac{1}{x^2_{n-1}}\right),$$

and

$$E\left(\frac{1}{x^2_{n-1}}\right) = \frac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)} \int \frac{1}{x^2}(x^2)^{\frac{n-1}{2}-1} e^{\frac{-x^2}{2}}\, dx^2$$

$$= \frac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}\, 2^{\frac{n-3}{2}}\,\Gamma\left(\frac{n-3}{2}\right)$$

$$= \frac{1}{n-3}.$$

Therefore,

$$Var(b) = \frac{\sigma_y^2(1-\rho^2)}{\sigma_x^2(n-3)} \qquad \text{......(4.2)}$$

and

$$Var(Y_r) = \frac{\sigma_y^2(1-\rho^2)}{n}\left(1 + \frac{1}{n-3}\right) \qquad \text{......(4.3)}$$

To find an unbiased estimator of $Var(Y_r)$, let

$$Q = \Sigma_{i=1}^{n}\left[y_i - \bar{y}_n - b(x_i - \bar{x}_n)\right]^2$$

$$= \Sigma_{i=1}^{n}\left(y_i - \bar{y}_n\right)^2 - b^2\,\Sigma_{i=1}^{n}(x_i - \bar{x}_n)^2$$

$$= (1-r^2)\,\Sigma_{i=1}^{n}(y_i - \bar{y}_n)^2$$

where $r = b\frac{sx}{sy}$ is the usual estimator of ρ. Thus

$$E(Q) = E\sum_{i=1}^{n}(y_i - \bar{y}_n)^2$$

$$-EE\left\{b^2\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \middle| x_1, x_2, \ldots x_n\right\}$$

$$=(n-1)\sigma_y^2 - E\left[\Sigma(x_i - \bar{x}_n)^2\left\{\frac{\sigma_y^2(1-\rho^2)}{\Sigma(x_i - \bar{x}_n)^2} + \beta^2\right\}\right]$$

$$=(n-2)\,\sigma_y^2\,(1-\rho^2).$$

Hence the unbiased estimator of Var $(\hat{Y}_r)$ is

$$\frac{Q}{n(n-3)} \qquad \ldots\ldots\ldots\ldots(4.4)$$

So far we have discussed under the restrictive assumptions the estimation procedure for the population mean of Y when the population mean of X is known. We shall relax these assumptions and obtain the results in Section 6, when the population is finite and the regression coefficient is estimated from the sample. For this we shall require the results for the case when the population mean of X is not known under the same restrictive assumptions. We discuss these results in the following section.

5. Double sampling estimation procedure

Let there be a random sample of size n, drawn without replacement out of N population units as in Section 3, on which the information on both X and Y is obtained. Let us draw further $n' - n$ units in the same way to have further information on X alone and consider the following linear estimator $\hat{Y}_{ds}$ of μ_y known in the literature as Double Sampling Estimator

$$\hat{Y}_{ds} = \bar{y}_n + \beta\,(\bar{x}_n' - \bar{x}_n) \qquad \ldots\ldots\ldots\ldots(5.1)$$

$$E(\hat{Y}_{ds}) = \mu_y$$

and so $\hat{Y}_{ds}$ is a linear unbiased estimator. By Lemma 2.3,

$$\text{Cov}\,(x_i\,,\,\hat{Y}_{ds}) = \frac{N-n'}{Nn'}\,\rho\,S_x\,S_y,\; i = 1, 2, \ldots, n'$$

and

$$\text{Cov}\ (y_i\,,\ \hat{Y}_{ds}) = \frac{N-n}{N-n}\, S_y^2 + \beta\left(\frac{N-n'}{N\,n'} - \frac{N-n}{N\,n}\right)\rho S_x\, S_y$$

$$= S_y^2\left[\frac{N-n}{N\,n} + \rho^2\left(\frac{1}{n'} - \frac{1}{n}\right)\right],\ i = 1, 2\ldots\ldots n$$

So, by Lemma 2.2, $\hat{Y}_{ds}$ is the best estimator and by Corollary 2.2, its variance is given by

$$\text{Var}\ (\hat{Y}_{ds}) = \text{Cov}\ (y_i\ ,\ Yds\ )$$

$$= S_y^2\left[\frac{N-n}{Nn} + \rho^2\left(\frac{1}{n'} - \frac{1}{n}\right)\right] \qquad \ldots\ldots\ldots\ldots(5.2)$$

In general, β is to be estimated from the sample as in Section 4. Therefore, we consider the variance of the estimator

$$Y_{ds} = \bar{y}_n + b\,(\,\bar{x}_{n'} - \bar{x}_n\,) \qquad \ldots\ldots\ldots(5.3)$$

We assume that we are sampling from the bivariate normal population of Section 4. Then

$$E(Y_{ds}) = EE\,(Y_{ds}\,/b)$$

$$= E\,(\mu_y) = \mu_y.$$

Therefore, Y_{ds} is an unbiased estimator. Further

$$\text{Var}\ (Y_{ds}) = EE\,[\,(\,\bar{y}_n - \mu_y\,) + b\,(\,\bar{x}_n{}' - \bar{x}_n\,)\,]^2/b$$

$$= E\left[\frac{\sigma_y{}^2}{n} + b^2\left(\frac{1}{n'} + \frac{1}{n} - \frac{2}{n'}\right)\sigma_x{}^2 + 2b\rho\,\sigma_x\,\sigma_y\left(\frac{1}{n'} - \frac{1}{n}\right)\right]$$

by Lemma 2.3 as in Section 4. Substituting the value of $E\,(b^2)$ from Section 4 and then simplifying

$$\text{Var}\ (Y_{ds}) = \frac{n'-n}{n\,n'}\,\sigma_y^2\,(1-\rho^2)\left(1+\frac{1}{n-3}\right) + \frac{\sigma_y{}^2}{n'} \qquad \ldots\ldots\ldots\ldots(5.4)$$

The quantity $s_y{}^2 = \sum_{i=1}^{n} (y_i - \bar{y}_n)^2 \,/\, n-1$ is an unbiased estimater of $\sigma_y{}^2$. From Section 4, $Q/(n-2)$ is an unbiased estimator of $\sigma_y{}^2\,(1-\rho^2)$. Therefore, an unbiased estimator of Var $(\hat{Y}_{ds})$ is

$$\left(\frac{1}{n} - \frac{1}{n'}\right)\frac{Q}{n-3} + \frac{s_y{}^2}{n'} \qquad \ldots\ldots\ldots\ldots(5.5)$$

We shall return to the discussion of double sampling estimation procedure in Section 7.

6. The variance of the regression estimator Y_r for a finite population and its unbiased estimator

An approach to obtain the variance in this case is suggested from the following example giving the derivation of the well known variance covariance formulae for simple random sampling without replacement from a Bi-variate finite population.

Let n be the size of the sample drawn, from the finite Bivariate population of size N, with simple random sampling without replacement in order to estimate μ_x and μ_y the population means for the variates X and Y.

Let $\bar{x}_n$ and $\bar{y}_n$, the sample means based on sample size n, be the estimators of μ_x and μ_y. Let the Bivariate finite population itself be treated as a sample drawn from the Bivariate normal population with parameters μ'_x, μ'_y, σ'_x, σ'_y and ρ'

then

$$E(\bar{x}_n) = E(\mu_x) = \mu'_x \qquad \text{............(6.1)}$$

By definition, $\bar{x}_n$ is said to be an unbiased estimater of μ_x in the extended sense. Similarly $\bar{y}_n$ is also an unbiased estimator of μ_y in the extended sense. Further, let $\hat{V}(\bar{x}_n)$ the variance of $\bar{x}_n$ in the extended sense be defined as

$$\hat{V}(\bar{x}_n) = E(\bar{x}_n - \mu_x)^2 / 1, 2, \ldots, N.$$

Then

$$E[\hat{V}(\bar{x}_n)] = E[\bar{x}_n - \mu_x]^2$$

$$= E[\bar{x}_n - \mu'_x - (\mu_x - \mu'_x)]^2$$

$$= \sigma'^2_x \left[\frac{1}{n} + \frac{1}{N} - \frac{2}{N}\right]$$

$$= \frac{N-n}{N\,n}\,\sigma'^2_x \qquad \text{............(6.2)}$$

An unbiased estimator of $E\,\hat{V}(\bar{x}_n)$ based on the finite population is $\left[\frac{N-n}{Nn} \Sigma (x_i - \bar{x}_N)^2 / N-1)\right]$ the well known expression for the variance of the mean based on sample of size n drawn without replacement from the population of size N. The well known unbiased estimator of the variance is an unbiased estimator of the variance in the extended sense. We can similarly obtain the expression for the variance of $\bar{y}_n$ and covariance of $\bar{x}_n$ and $\bar{y}_n$.

We shall use the above approach, using the concept of super-population, to obtain the variance of the regression estimator in the case under study with obvious extensions of the terms unbiased in the extended sense and the variance in the extended sense to first and second moments of the estimators.

For the case under study, let the finite population of size N be regarded, for the study of X and Y, as a random sample from the Bi-variate normal population described in this section. Under this postulation, the expectation of Y_r of Section 4 is given by

$$E(Y_r) = E[\bar{y}_n + b(\mu_x - \bar{x}_n)/b]$$
$$= E(\mu'_y) = \mu'_y = E(\mu_y). \qquad \ldots\ldots\ldots\ldots(6.3)$$

Thus Y_r is an unbiased estimator, in the extended sense, of μ_y the finite population mean. Further,

$$E[\hat{V}(Y_r)] = E(Y_r - \mu_y)^2$$
$$= E[\bar{y}_n + b(\mu_x - \bar{x}_n) - \mu_y]^2$$
$$= E[\{\bar{y}_n + b(\mu_x - \bar{x}_n) - \mu'_y\} - (\mu_y - \mu'_y)]^2$$
$$= E\left[\bar{y}_n + b(\mu_x - \bar{x}_n) - \mu'_y\right]^2 + \frac{\sigma_y'^2}{N}$$
$$- 2\,\mathrm{Cov}(\bar{y}_n, \mu_y) - 2E[b(\mu_x - \bar{x}_n)(\mu_y - \mu'_y)]$$
$$= E[\bar{y}_n + b(\mu_x - \bar{x}_n) - \mu'_y]^2 - 2E\,b(\mu_x - \bar{x}_n)(\mu_y - \mu'_y) - \sigma_y'^2/N.$$

The first term on the right hand side is obtained by putting $n' = N$ in Equation (5.4). The second term

$$2E[bE\{(\mu_x - \bar{x}_n)(\mu_y - \mu'_y)/b\}]$$
$$= 2E(b\,O) = O$$

Therefore,

$$E\hat{V}(Y_r) = \frac{N-n}{Nn}\sigma_y'^2(1-\rho'^2)\left(1 + \frac{1}{n-3}\right) + \frac{\sigma_y'^2}{N} - \frac{\sigma_y'^2}{N}$$
$$= \frac{N-n}{Nn}\sigma_y'^2(1-\rho'^2)\left(1 + \frac{1}{n-3}\right) \qquad \ldots\ldots\ldots\ldots(6.4)$$

An unbiased estimator of $\hat{V}(Y_r)$ in the extended sense is

$$\frac{N-n}{Nn}\,\frac{Q}{n-3} \qquad \text{...........(6.5)}$$

where Q is the expression as given in Section 4. The expression in (6.5) reduces to the expression given in Equation (4.4) for $N=\infty$. We have seen above that the regression estimator Y_r, for the finite population, has the property of unbiasedness in the extended sense. Though it is desirable to have this property, one would like to know whether this estimator is, in fact, an unbised estimator of μ_y, the mean of the finite population. That, it is not so has been discussed by Cochran (1963) and others. It is shown by these authors that the bias in the estimator Y_r tends to zero for large n under specified assumptions. How large n we may take, that this bias becomes negligible can be worked out by using the unbiased regression estimator due to Micky (1959, Equation 3.14, p. 599). To describe this estimater, let b'_j be the value of the regression coefficient if the jth sample element is omitted, i.e.,

$$b'_j = \frac{\sum_{i=1}^{n}(x_i-\bar{x}_n)(y_i-\bar{y}_n)-\frac{n}{n-1}(x_j-\bar{x}_n)(y_j-\bar{y}_n)}{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2-\frac{n}{n-1}(x_j-\bar{x}_n)^2}$$

Let $\bar{b}=(\Sigma\, b'_j)/n$, then Micky's unbiased regression estimator of the population mean μ_y is

$$\bar{y}+\bar{b}(\mu_x-\bar{x}_n)-B \qquad \text{......(6.6)}$$

Where

$$B=\Big[(N-n)/Nn\Big]\Big[\sum_{j=1}^{n}x_j\,b'_j-n\,\bar{x}\,\bar{b}\Big].$$

For those large values of n, for which B is negligible and for which $\bar{b}\sim b$, Y_r will be approximately unbiased.

7. The variance of the double sampling estimator Y_{ds} for a finite population and its unbiased estimator.

We derive the various results in this section with the same postulation as in Section 6.

$$E(Y_{ds})=E[E(Y_{ds}/b)]$$
$$=E(\mu'_y)=\mu'_y=E(\mu_y)$$

Thus Y_{ds} is an unbiased estimator of μ_y in the extended sense. Further the variance of Y_{ds} in the extended sense is given by

$$V(Y_{ds}) = E(Y_{ds} - \mu_y)^2/1, 2, ...N$$

Therefore,

$$E[V(Y_{ds})] = E(Y_{ds} - \mu_y)^2$$
$$= E[\, Y_{ds} - \mu'_y - (\mu_y - \mu'_y)\,]^2$$
$$= Var\,(Y_{ds}) + V(\mu_y) - 2E(\bar{y}_n - \mu'_y)(\mu_y - \mu'_y)$$
$$-2E[\, b\,(\mu_y - \mu'_y)\,(\bar{x}_n' - \bar{x}_n)\,]$$
$$= Var\,(Y_{ds}) - \frac{\sigma'y^2}{n}$$
$$= \frac{n'-n}{nn'}\sigma'y^2\,(1-\rho'^2)\,(1+\frac{1}{n-3}) + \frac{\sigma'y^2}{n'} - \frac{\sigma'y^2}{N}$$

For $N=\infty$, the expression on the right hand side reduces to the expression in the Equation (5.4). An unbiased estimator of the above expression and therefore of $V(Y_{ds})$ in the extended sense is

$$\frac{n'-n}{nn'}\,\frac{Q}{n-3} + s_y{}^2\,\frac{N-n'}{Nn'}$$

which reduces to the expression in Equation (5.5) for $N=\infty$.

8. Stratified Regression and Double Sampling estimators

When regression coefficients in different strata are different and when sample sizes in different strata are fairly large, in each stratum we can take suitable regression or double sampling estimator, discussed in the various earlier sections, and then after assigning suitable weights we can get the estimator of the stratified population.

For example, let us see how the results of Section 3 generalise for stratified population.

Let the population consist of k strata. Let x_{ij} and y_{ij} denote the values of the unit, for X and Y, drawn at the i th draw in the j th stratum. Then, with obvious notations, we take the estimator $Y_{r\,(st)}$ of $\Sigma_{j=1}^{k}\left(\frac{N_j}{N}\mu_{yj}\right)$ as given by

$$Y_{r\,(st)} = \Sigma_{j=1}^{k}\left[\, \bar{y}_{n_j} + \beta_j\,(\mu_{xj} - \bar{x}_{n_j})\,\right]\frac{N_j}{N} \qquad ......(8.1)$$

$$E\,(\,Y_{r\,(st)}\,)=\sum_{j=1}^{k}\left(\frac{N_j}{N}\,\mu_{yj}\right),\text{ and}$$

$$Var\,(\,Y_{r\,(st)}\,)=\frac{1}{N^2}\sum_{j=1}^{k}\left[\,N_j^2\;\frac{N_j-n_j}{N_j\,n_j}\,S^2_{yj}\,(1-\rho_j^2)\,\right]\quad\ldots(8.2)$$

An unbiased estimator of Var $Y_{r\,(st)}$. is

$$\frac{1}{N^2}\sum_{j=1}^{k}N_j^2\,\frac{N_j-n_j}{N_j\,n_j}\,\frac{1}{n_j-1}$$

$$\left[\sum_{j=1}^{n_j}\left\{y_{ij}-\bar{y}_{n_j}-\beta_j\,(x_{ij}-\bar{x}_{n_j})\right\}^2\right]\quad\ldots\ldots\ldots\ldots(8.3)$$

When the sample sizes in different strata are small and when the situation suggests that $\beta j=\beta$ for all j, a combined regression estimator or a combined double sampling estimator will be better than the type of estimator considered above. This is seen from the discussion below regarding the combined regression estimator. Let the combined regression estimator $\hat{Y}_{rc}$ of $\sum_{j=1}^{k}\left(\frac{N_j}{N}\,\mu_{yj}\right)$ be

$$\hat{Y}_{rc}=\sum_{j=1}^{k}\frac{N_j}{N}\left\{\bar{y}_{nj}+\beta\,(\mu_{xj}-\bar{x}_{n_j})\right\}\quad\ldots\ldots\ldots\ldots(8.4)$$

$$E\,(\hat{Y}_{rc})=\sum_{j=1}^{k}\left(\frac{N_j}{N}\,\mu_{yj}\right).$$

Therefore, $\hat{Y}_{rc}$ is an unbiased estimator of the mean of the stratified population. Further

$$\text{Var}\,(\hat{Y}_{rc})=\sum_{j=1}^{k}\left[\frac{N_j^2}{N^2}\;\frac{N_j-n_j}{N_j\,n_j}\;\frac{1}{N_j-1}\right.$$

$$\left.\sum_{i=1}^{N_j}\left\{y_{ij}+\beta\,(\mu_{xj}-x_{ij})-\mu_{yj}\right\}^2\right]\quad\ldots\ldots\ldots\ldots(8.5)$$

noting that $\bar{y}_{n_j}+\beta\,(\mu_{xj}-\bar{x}_{n_j})$

can be regarded as the mean of n_j observations randomly drawn from a finite population consisting of N_j observations

$$y_{ij}+\beta\,(\mu_{xj}-x_{ij})$$

The value of β, where Var $(\hat{Y}_{rc})$ is minimum, is given by

$$\frac{\partial\ \text{Var}\,(\hat{Y}_{rc})}{\partial\beta}=0,$$

or by,

$$\sum_{j=1}^{k}\left[\frac{N_j^{\,2}}{N^2}\,\frac{N_j - n_j}{N_j\, n_j}\,\frac{1}{N_j - 1}\sum_{i=1}^{N_j}(\mu_{xj} - x_{ij})\left\{y_{ij} + \beta\,(\mu_{xj} - x_{ij}) - \mu_{yj}\right\}\right] = 0.$$

Therefore,

$$\beta\left[\sum_{j=1}^{k}\left\{\frac{N_j^{\,2}}{N^2}\,\frac{N_j - n_j}{N_j\, n_j}\,\frac{1}{N_j - 1}\sum_{i=1}^{N_j}(x_{ij} - \mu_{xj})^2\right\}\right]$$

$$= \sum_{j=1}^{k}\left\{\frac{N_j^{\,2}}{N^2}\,\frac{N_j - n_j}{N_j\, n_j}\,\frac{1}{N_j - 1}\sum_{i=1}^{N_j}(y_{ij} - \mu_{yj})(x_{ij} - \mu_{xj})\right\},$$

or,

$$\beta = \frac{\sum_{j=1}^{k}\frac{N_j^{\,2}}{N^2}\,\frac{N_j - n_j}{N_j\, n_j}\,S_{xyj}}{\sum_{j=1}^{k}\frac{N_j^{\,2}}{N^2}\,\frac{N_j - n_j}{N_j\, n_j}\,S^2_{xj}} \qquad \text{...........(8.6)}$$

when S_{xyj} and S^2_{xj} are defined in a way similar to the quantities defined in Lemma 2.3. If $n_j = nN_j/N$, as in Proportional sampling,

$$\beta = \frac{\sum_{j=1}^{k} N_j\, S_{xyj}}{\sum_{j=1}^{k} N_j\, S^2_{xj}} \qquad \text{...........(8.7)}$$

Further if $N_j \doteq N_j - 1$, then

$$\beta = \frac{\sum_{j=1}^{k}\sum_{i=1}^{N_j}(x_{ij} - \mu_{xj})(y_{ij} - \mu_{yj})}{\sum_{j=1}^{k}\sum_{i=1}^{N_j}(x_{ij} - \mu_{xj})^2} \qquad \text{...........(8.8)}$$

From the theory of simple random sampling without replacement, an estimator of Var ($\hat{Y}_{rc}$) for a given β is

An asymptotically unbiased estimator of $E[\hat{V}(\hat{Y}_2)]$ and therefore of $\hat{V}(\hat{Y}_2)$ in the extended sense is

$$s_2^2 \left[\frac{\hat{\phi}_2}{n_2''} - \frac{1}{N}\right] \qquad \ldots\ldots(9.3)$$

where

$$s_2^2 = \sum_{j=1}^{n_2} \left(y_{2j} - \bar{y}_{2n_2}\right)^2 \Big/ (n_2 - 1),$$

$\bar{y}_{2n_2}$ being the mean of n_2 units on the second occasion.

We shall prove the theorem in part up to Equation (9.1). The remaining part can be proved on an approach similar to one adopted in Section 6 after first obtaining the variance of $\hat{Y}_2$ under normality assumption.

Proof. If y_{1j} is an observation on one of the n_2' units, then by Lemma 2.3,

$$\text{Cov}(y_{1j}, \hat{Y}_2) = (1-\phi_2)\left[\rho S_1 S_2 \frac{N-n_2'}{Nn_2'} + \rho\frac{S_2}{S_1} S_1^2 \left(\frac{N-n_1}{Nn_1} - \frac{N-n_2'}{Nn_2'}\right)\right] - \phi_2 \frac{\rho S_1 S_2}{N}$$

$$= \rho S_1 S_2 \left[\frac{1-\phi_2}{n_1} - \frac{1}{N}\right] \qquad \ldots\ldots(9.4)$$

When y_1j is an observation on one of the n_1'' units, then

$$\text{Cov}[y_1\text{j}, \hat{Y}_2] = (1-\phi_2)\left[-\frac{\rho S_1 S_2}{N} + \rho\frac{S_2}{S_1}\left(\frac{N-n_1}{Nn_1} - \frac{1}{N}\right)S_1^2\right] - \phi_2 \rho \frac{S_1 S_2}{N}$$

$$= \rho S_1 S_2 \left[\frac{1-\phi_2}{n_1} - \frac{1}{N}\right] \qquad \ldots\ldots(9.5)$$

when y_2 j is an observation on one of the n_2' units, then

$$\text{Cov}(y_2\text{j}, \hat{Y}_2) = (1-\phi_2)\left[S_2^2 \frac{N-n_2'}{Nn_2'} + \rho\frac{S_2}{S_1}\left(\frac{N-n_1}{Nn_1} - \frac{N-n_2'}{Nn_2'}\right) \rho S_2 S_1\right] - \phi_2 \frac{S_2^2}{N}$$

$$= (1-\phi_2) S_2^2 \left[\frac{1-\rho^2}{n_2'} + \frac{\rho^2}{n_1}\right] - \frac{S_2^2}{N}$$

$$= \phi_2 \frac{S_2^2}{n_2''} - \frac{S_2^2}{N}$$

$$= S_2^2 \left[\frac{\phi_2}{n_2''} - \frac{1}{N}\right] \qquad \ldots\ldots(9.6)$$

When y_{2j} is an observation on one of the n_2'' units, then

$$\text{Cov}\ (y_{2j}, \hat{Y}_2) = (1-\phi_2)\left[-\frac{S_2^2}{N} + \rho\ S_2^2\left(-\frac{1}{N} + \frac{1}{N}\right)\right] + \phi_2 \frac{N - n_2''}{Nn_2''} S_2^2$$

$$= S_2^2 \left[\frac{\phi_2}{n_2''} - \frac{1}{N}\right] \qquad \text{......(9.7)}$$

The equations (9.4) and (9.5) show that *Cov* $(y_{1j}, \hat{Y}_2)$ is independence of j for all $j=1, 2, ..., n_1$. The equations (9.6) and (9.7) show that Cov $(y_{2j}, \hat{Y}_2)$ is independent of j for all $j=1, 2, ..., n_2$. Therefore, by Lemma 2.2, $\hat{Y}_2$ is the best estimator of μ_2. Further, that the variance of $\hat{Y}_2$ is given by Equation (9.1) is seen either from Equation (9.6) or from Equation (9.7) in view of the Corollary 2.2.

REFERENCES

1. Cochran, W. G. (1963); "Sampling techniques", Wiley publication.

2. Lukomsky, J. (1939); "On some properties of Multidimensional distributions", Ann. Math. Statist.. 10. 236-246.

3. Micky, M. R. (1959) "Some finite Population unbiased ratio and regression estimators", Jour. Amer. State Assoc., 54, 594-612.

4. Patterson, H. D. (1950); "Sampling on successive occasions with partial replacement of units", J. R. Statist. Ass., 37, 95-102.

5. Prabhu Ajgaonkar, S. G., (1962); "Some aspects of the class of linear estimation with special reference to successive sampling", Ph.D. Thesis, Karnatak University, India.

6. Sukhatme, P. V. (1953); "Sampling theory of surveys with applications", New Delhi: Indian Soc. of Agric. Stat., Ames, Iowa: Iowa State college Press, U.S.A.

7. Tikkiwal, B. D. (1950); "Estimation by successive sampling", Presented at the fourth annual meeting of the Indian society of Agricultural statistics and thesis entitled 'Theory of successive sampling' for Diploma, I.C.A.R., New Delhi, 1–17.

8 .... (1953); "On the recurrence formula in sampling on succesive occasions", J. Ind. Soc. Agric. Stat. 5, 96-99.

9. .... (1955); "Multiphase sampling on successive occasions", Ph. D. Thesis, Faculty of North Carolina State College, U.S.A.

10. .... (1956); "A further contribution to the theory of univariate sampling on successive occasions", J. Ind. Soc. Agric. Stat., 8, 84-90

11. .... (1958); "An examination of the effect of matched sampling on the efficiency of estimations in the theory of successive sampling", J. Ind. Soc. Agric. Stat., 10, 16–21

12. Tikkiwal, B. D. (1960); "Classical regression and double sampling estimation", J. Roy. Stat. Soc., B, 22, 131-138.

13. .... (1963); "A note on two-stage sampling on successive occasions", Sankhya, series A, 26, 97–100.

14. Yates, F. (1949); "Sampling methods for censuses and surveys", Charles Griffin and Company, Ltd., London.

SOME ASPECTS OF SAMPLING WITH VARYING PROBABILITIES FROM A FINITE POPULATION

By

B. D. Tikkiwal
University of Rajasthan

0. INTRODUCTION AND SUMMARY

Hansen and Hurwitz (1943) suggested a stratified two-stage sampling design, for estimating the population mean or total, in which the primary units within a given stratum are selected with probability proportional to their sizes and with replacement. The second stage units within a selected primary unit are selected with simple random sampling without replacement so as to make the design self-weighting in each of the strata or over all strata taken together. Later, these authors (1949); while determining the optimum probabilities of selection in the above two-stage design for a given stratum, where primary units are selected with varying probabilities and with replacement; mentioned the need of developing the theory further when primary units are selected with varying probabilities and without replacement.

Narain (1951) on one hand and Horvitz and Thompson (1952) on the other, apparently working independently, developed the theory further for sampling with varying probabilities and without replacement. Narain obtained the variance of the two-stage sampling estimator with the same number of two-stage sampling units in each of the selected primary units and also examined the conditions necessary to make this estimator unbiased. Though, the main estimator and its variance, given by Horvitz and Thompson in their equations (6) and (9), can be obtained with slight modification in the estimator and its variance given by Narain, the two authors together presented the theory in a way suggestive of further extension. They noted that there can be a number of classes of linear estimators and mentioned three such estimators by way of illustration.

Koop (1961) and Prabhu Ajgaonkar and Tikkiwal (1961) independently examined all the possible classes of linear estimators. Prabhu Ajgaonkar and Tikkiwal pointed out that their are only seven classes of linear estimators where as the classes of linear estimators considered by Koop are in effect only six.

The main estimator, for sampling with varying probabilities and without replacement, considered by Horvitz and Thompson lies in their T_2—class where the weight assigned to a particular draw depends upon the out-come at the draw, i.e., if the *i th* unit comes at a particular draw, then a weight β_i is associated with that draw for $i=1,2,\ldots N$, N being the number of sampling units in the population.

Horvitz and Thompson's estimator can be used, with some modification, where there is sampling with varying probabilities and with replacement. With this modification, Prabhu Ajgaonkar and Tikkiwal (1961) gave a unified theory for sampling with varying probabilities and with or without replacement.

Horvitz and Thompson also gave the estimator of the variance of their estimator. This estimator was shown to assume negative values by Yates and Grundy (1953) who gave, in turn, another estimator which was also shown to have negative values by Durbin (1953). Simultaneously with Yates and Grundy (1953), Sen (1953) gave the same estimator of variance and showed that this estimator is always positive under some sampling schemes. Hartley and Rao (1962) also examined the positive character of this estimator. From these studies, it seems that Sen, Yates and Grundy's estimator assumes negative values much less frequently than the estimator given by Horvitz and Thompson, making the latter less desirable than the former. Therefore, Sen (1953) suggested some modification in Horvitz and Thompson's estimator, making it sometimes more desirable than Sen, Yates and Grundy's estimator. Apart from the difficulty of getting over some times negative estimator of the variance, there is also difficulty, when the sample size is greater than two, in computation of probabilities of inclusion of population units in singles or in pairs. Such probabilities occur in the estimator of the population total and also in the estimator of the variance of the estimator. Several authors have tried to resolve this difficulty but not with much success. Therefore, Rao, Hartley and Cochran (1962) has suggested a slightly modified estimator, based on different sampling scheme, which is claimed to be only slightly less efficient than Horvitz and Thompson's estimator with a well known classical sampling scheme considered by Hartley and Rao (1962). The modified estimator has an always positive estimator of its variance.

From what is said in the beginning of this section, it is clear that sampling with varying probabilities and with replacement was first considered by Hansen and Hurwitz (1943) in a restricted manner for a two-stage design. The authors did not give specifically the estimation procedure for estimating the variances of their various estimators, even while considering an actual sampling problem in section 7 of their paper. The estimation procedure for sampling with replacement seems to have been given first, for a two-stage design by Sukhatme and Narain (1952) and for any multi-stage design by Durbin (1953) who, in fact, discussed the general theory of multi-stage sampling with varying probabilities and with or without replacement. The estimation procedure for two-stage sampling with varying probabilities and without replacement was first discussed by Horvitz and Thompson (1952).

The paper gives in Section 1 a unified general discussion, of T_2-class estimators for sampling with varying probabilities and with or without replacement, due to Prabhu Ajgaonkar and Tikkiwal. Horvitz and Thompson's estimator in T_2-class and its variance for sampling without replacement is obtained as a special case in Section 2. Then, Section 3 gives three estimators of the variance and a discussion of their relative merits. In Section 4, certain probabilities of inclusion; of the population units in singles and pairs, required in the estimator and in its variance as well as in the estimator of the variance for sampling without replacement; are calculated. Section 5 gives the results, for sampling without replacement, due to Rao, Hartley and Cochran (1962) to overcome the difficulty in calculation of the probabilities in general for $n>2$. Section 6 gives the estimator and its variance together with the estimator of the variance for sampling with replacement. The relative efficiency of the two sampling systems with and without replacements is discussed in Section 7. Section 8 gives the general theory of multi-stage sampling with varying probabilities and with or without replacement. It is shown by the author that the various estimators due to Sukhatme and Narain, Durbin and Ecimovic belong to T_2-class.

1. A General Discussion of T_2-Class Estimators.

The problem frequently faced in surveys is one that of estimation of totals (or of population means) for different characteristics of the population at a given time or at different times. The estimators generally adopted are linear functions of the observations on the units in the sample drawn from the population. Such linear functions can be obtained by associating with a particular unit a weight dependent either (i) on the draw, at which the particular unit occurs, or (ii) on the unit itself or (iii) on the sample or (iv) on any combination of these three. The estimators, which are linear functions of above types are said to belong to T–class and in particular of type (ii) to T_2–class. In what follows, we confine ourselves to the problem of estimation of T, the total for a particular characteristic of the population.

When a sample of size n is drawn from the population with varying probabilities and with or without replacement in order to estimate the total T, an estimator $\hat{T}$ of T in T–class is given by

$$\hat{T} = \sum_{r=1}^{n} \beta_{ro}^{s_n} X_r$$

where X_r denotes the random out-come at the rth draw and $\beta_{ro}^{s_n}$ denote the weight to be associated with X_r for a given sample s_n. An estimator in T_2-class is obtained by taking $\beta_{ro}^{s_n} = \beta_i$ if $X_r = x_i$, for any s_n, x_i denoting the value of the i th unit in the population for $i=1,2,.. N$. Let P_{ir} denote the probability of i th unit of the population occuring at the r th draw. For an estimator in T_2-class,

$$E(\hat{T}) = \sum_{r=1}^{n} \left(\sum_{i=1}^{N} x_i \ \beta_i \ P_{ir} \right)$$

$$= \sum_{i=1}^{N} \left[x_i \ \beta_i \left(\sum_{r=1}^{n} P_{ir} \right) \right]$$

$$= \sum_{i=1}^{N} x_i \ \beta_i \ k_i \ ;$$

where

$$k_i = \sum_{r=1}^{n} P_{ir}$$

In order that $\hat{T}$ be an unbiased estimator; we must have $\beta_i = 1/k_i$, a unique value for β_i, for all $i=1, 2, \ldots, N$. Thus, there is only one unbiased estimator in T_2-class and therefore it is also the minimum-variance-linear unbiased estimator. Let this minimum-variance-linear-unbiased estimator in T_2-class be denoted by

$$(1.1)\quad \hat{T}_2 = \sum_{r=1}^{n} \beta_{ro} X_r ,$$

where $\beta_{ro} = 1/k_i$, when $X_r = x_i$. Also,

$$(1.2)\quad E(\hat{T}_2)^2 = \sum_{i=1}^{N} \frac{x_i^2}{k_i} + \sum_{r\neq s}^{n} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{x_i\, x_j}{k_i\, k_j} P\,(ir, js)$$

where $P\,(ir, js)$ denotes the probobility of selecting x_i and x_j units at *rth* and *sth* draws respectively. The variance of $\hat{T}_2$ can now be calculated easily, when there is sampling with or without replacement. We present results for sampling without replacement in Sections 2 to 5, and for sampling with replacement in Sections 6 and 7.

2. Sampling with Varying Probabilities and without Replacement.

For sampling without replacement, let $P_{ir} = p_{ir}$ and $k_i = \pi_i$. It may be noted that π_i is the probability of including *ith* unit in the sample for the present case. Since $p\,(ir, js) = 0$ for $i = j$, from equation (1.2)

$$E\,(\hat{T}_2)^2 = \sum_{i=1}^{N} \frac{x_i^2}{\pi_i} + \sum_{i\neq j}^{N} \frac{x_i\, x_j}{\pi_i\, \pi_j}\, \pi_{ij}$$

where π_{ij} is the probability of i *th* and j *th* units being included in the sample. Hence, if $\hat{T}_{ht}$ denotes the estimator for sampling without replacement; its variance $V\,(\hat{T}_{ht})$ is given by.

$$(2.1)\quad V(\hat{T}_{ht}) = \sum_{i=1}^{N} x_i^2 \,(1-\pi_i)\,/\,\pi_i + \sum_{i\neq j}^{N} \frac{x_i\, x_j}{\pi_i\, \pi_j}\,(\pi_{ij} - \pi_i\, \pi_j)$$

a result, in the present form, due to Horvitz and Thompson (1952, equ. 9, p. 670).

3. The Three Unbiased Estimators of the Varianee of $\hat{T}_{ht}$ and their Relative Merits.

The unbiased estimator, of the variance of $\hat{T}_{ht}$, due to Horvitz and Thompson is

$$(3.1)\quad \hat{T}_{ht}^2 - (\text{ an unbiased estimator of } T^2)$$

$$= \hat{T}_{ht}^2 - \sum_{r=1}^{n} X_r^2 \,\beta_{ro} - \sum_{r\neq r'}^{n} X_{rr'}$$

where $X_{rr'} = x_i\, x_j \,/\, \pi_{ij}$ if x_i and x_j units occur at the *rth* and *r'th* draws respectively.

In order to obtain the alternate unbiased estimator of the variance of $\hat{T}_{ht}$ due to Sen (1953) and Yates and Grundy (1953), we first present the results due to Horvitz and Thompson (1952) in the form of following lemma.

Lema 3.1.

(i) $\sum_{i=1}^{N} \pi_i = n$

(ii) $\sum_{j(\neq i)}^{N} \pi_{ij} = (n-1)\, \pi_i$

Proof

(i) $\sum_{i=1}^{N} \pi_i = \sum_{i=1}^{N} \sum_{r=1}^{n} p_{ir}$

$$= \sum_{r=1}^{n} \sum_{i=1}^{N} p_{ir} = n$$

noting that $\sum_{i=1}^{N} p_{ir} = 1$ for any given *r*.

(ii) Let $p\,(j,s\,|\,i,r)$ denote the conditional probability of jth unit occuring at sth draw given that ith unit has occured at the rth draw. Then

$$\sum_{j(\neq i)}^{N} \pi_{ij}$$

$$= \sum_{j(\neq i)}^{N} \sum_{r=1}^{n} p_{ir} \sum_{s(\neq r)}^{n} p\,(j,s\,|\,i,r)$$

$$= \sum_{r=1}^{n} p_{ir} \sum_{s(\neq r)}^{n} \sum_{j(\neq i)}^{N} p\,(j,s\,|\,i,r)$$

$$= \sum_{r=1}^{n} p_{ir}\,(n-1)$$

$$= (n-1)\,\pi_i.$$

Corollary 3.1.

$$\sum_{j(\neq i)} (\pi_{ij} - \pi_i\,\pi_j)$$

$$= (n-1)\,\pi_i - \pi_i\,(n-\pi_i)$$

$$= \pi_i^2 - \pi_i\,.$$

Using the corollary in (2.1)

$$V(\hat{T}_{ht}) = \sum_{i=1}^{N} \frac{x_i^2}{\pi_i^2} \sum_{j(\neq i)}^{N} (-\pi_{ij} + \pi_i\,\pi_j) + \sum_{i\neq j}^{N} \frac{x_i\,x_j}{\pi_i\,\pi_j} (\pi_{ij} - \pi_i\,\pi_j)$$

$$= \sum_{i=1}^{N} \sum_{j>i}^{N} \left(\frac{x_i}{\pi_i} - \frac{x_i}{\pi_j}\right)^2 (\pi_i\,\pi_j - \pi_{ij}).$$

Therefore, the alternate unbiased estimator due to Sen, Yates and Grundy is

$$3.2)\quad \sum_{r=1}^{n} \sum_{s>r}^{n} C_{rs}$$

where

$$C_{rs} = (\pi_i\,\pi_j - \pi_{ij})\,/\,\pi_{ij}\cdot (x_i\,/\,\pi_i - x_j\,/\,\pi_j)^2,$$

If, i th and j th units occur at r th and s th draws in respective order.

Yates and Grundy (1953) in their psuedo empirical study (Table 3, p. 258) and also in the study of Horvitz and Thompson's population of 20 city blocks noted that Horvitz and Thompson's estimator gives negative values where as their estimator does not. Durbin (1953, p. 264), showed that Sen, Yates and Grundy's estimator can also assume negative values. Sen (1953, Theorems 1 and 2) showed that this estimator of variance is always positive, when sampling is done according to Midzuno's scheme for any n described by Horvitz and Thompson as sampling scheme 1 and also according to Horvitz and Thompson's scheme 2 given for $n=2$ (1952, pp. 678–679). Hartley and Rao (1962, p. 364) proved the result for Horvitz and Thopson's scheme 2 under restrictive conditions perhaps being unaware of the result due to Sen. These studies indicate that Sen, Yates and Grundy's estimator is likely to assume negative values much less often than Horvitz and Thompson's estimator.

Sen (1953) suggested a third estimator $\hat{V}_s$; which is essentially the same as Horvitz and Thompson's estimator $\hat{V}_{ht}$ when it assumes positive values with the difference that for negative values of the latter estimator, it takes the value zero. Thus,

$$\hat{V}_s = \begin{cases} \hat{V}_{ht} , & \hat{V}_{ht} \geqslant 0 \\ 0 , & \hat{V}_{ht} < 0 . \end{cases}$$

Further

$$\begin{aligned} E(\hat{V}_s) &= E(\hat{V}_{ht}) + k \\ &= V(\hat{T}_{ht}) + k \end{aligned}$$

where $k \geqslant 0$. Thus $\hat{V}_s$ is a positively biased estimator of $V(\hat{T}_{ht})$. Further

$$E\left[\hat{V}_{ht} - V(\hat{T}_{ht}) \right]^2 = E\left[\hat{V}_s - V(\hat{T}_{ht})\right]^2 + c$$

where c is some non-negative quantity. Therefore,

$$\text{(3.3)} \qquad V(\hat{V}_{ht}) = V(\hat{V}_s) + k^2 + c.$$

Thus, the second moment of $\hat{V}_s$ about $V(\hat{T}_{ht})$ is always less than or equal to the varience of $\hat{V}_{ht}$.

If this second moment is also less than the variance of the estimator of the variance due to Sen, Yates and Grundy, one may prefer to use Sen's modified estimator even when it is biased. On this criterion, one would use Sen's modified estimator to estimate the variance of the population total in table II of Sen (1954, p. 125) given by him to compare the relative efficiencies of the three estimators.

4. Calculation of Probabilities of Inclusion.

Let $\pi_i(n)$ and $\pi_{ij}(n)$ denote the probabilites of inclusion of *i th* unit and (i, j) *th* units in the sample of size n. Then for Midzuno's scheme of sampling; where the probability of *i th* unit $(i=1,2,...,N)$ being drawn at the first draw is p_i and at a subsequent draw is $1/(N-1)$ given that some other units have occured at the previous draws ;

$$p_{i_1} = p_i \text{ and for } r > 1,$$

$$p_{ir} = (1-p_i) \prod_{s=2}^{r-1} \left(\frac{N-s}{N-s+1} \right) \cdot \frac{1}{N-r+1}$$

$$= \frac{(1-p_i)}{N-1};$$

Therefore, we get

$$(4.1) \qquad \pi_i(n) = \sum_{r=1}^{n} p_{ir} = p_i + \frac{n-1}{N-1}(1-p_i).$$

Also,

(4.2) $\pi_{ij}(n)$ = (Prob. of *ith* unit or *jth* unit occuring at the first draw and the remaining unit occuring at a subsequent draw) + (Prob. of *ith* and *jth* units occuring after the first draw)

$$= (p_i + p_j)\frac{n-1}{N-1} + (1-p_i-p_j)\frac{(n-1)(n-2)}{(N-1)(N-2)},$$

since $\pi_i(n) = n/N$ and $\pi_{ij}(n) = n(n-1)/N(N-1)$ for simple random sampling without replacemet. It may be noted that Midzuno's scheme of sampling gives to a given sample the probability proportional to the total size of the sample provided p_i is proportional to the size of the *i th* unit.

For Horvitz and Thompson's sampling scheme where $n=2$ and where $p(j, 2/i, 1) = p_j / (1-p_i)$ for all i, j ;

$$(4.3) \qquad \pi_i(2) = p_i\left(1 + S - \frac{p_i}{1-p_i}\right)$$

where $S = \sum_{j=1}^{N} \left[p_j \,/\, (1 - p_j) \right]$; and

$$\pi_{ij}(2) = p_i\, p_j \left(\frac{1}{1-p_i} + \frac{1}{1-p_j}\right).$$

The experession in (4.1), (4.2) and (4.3) are due to Horvitz and Thompson (1952).

For $n \geqslant 2$, the recurrence relations are given by Singh (1954, equations on pp. 49–50) for a generalised sampling scheme which is the same as Horvitz and Thompson's scheme for $n=2$. But the expressions are complicated and therefore the computations for $n>2$ and for moderately large N are difficult for such a sampling scheme.

Hartley and Rao (1962), while considering an alternate scheme of sampling to resolve the above difficulty, has commented on the work of Des Raj (1956) that his assumption, $x_i = \alpha + \beta y_i$, y_i being an auxiliary variate, in finding those values of π_{ij} which minimise the variance of $\hat{T}_{ht}$ in (2.1); nullify the utility of his results, because, in that case the regression estimator is the best estimator. The argument is not enought to justify the comment. For, from the works of Tikkiwal (1960) and Prabhu Ajgaonkar (1962), the regression estimator can be shown to be the best estimator in T_1-class where as the estimator considered by Des Raj is the best estimator in T_2-class. Therefore, a direct comparison of the variances of the two estimators is necessary in order to examine the relative efficiency of the two estimators.

5. Unequal Probability Sampling without Replacement due to Rao, Hartley and Cochran.

When the clusters are randomly formed; it is shown (Sukhatme, sec. 6a. 2, p. 242) that for estimating population mean (or total), cluster sampling is as efficient as simple random sampling without replacement. That, the result is true also for simple random sampling with replacement, is easily seen. Similarly, if the population is divided into k strata at random and if there is proportional sampling within each stratum, then it is easy to see that stratified sampling has the same efficiency as simple random sampling with or without replacement. However, if the sample sizes in different strata are chosen arbitrarily, then, the relative efficiency either way will depend upon the sizes of the strata and the sample sizes in them.

In the light of above analysis, we consider the sampling procedure with varying probabilities and without replacement due to Rao, Hartley and Cochran (1962). In order to draw a sample of size n according to this sampling procedure, the population under study is divided at random into n strata, the i *th* stratum being of size N_i for $i=1,2,\ldots,n$ so that $\sum_{i=1}^{n} N_i = N$, the size of the population.

Let x_{i_j}, for $j=1,2,\ldots,N_i$, denote the value of i_j *th* unit falling in the i *th* stratum. One unit from each stratum is selected so as to have the ultimate sample of size n. Let the conditional probability of i_j *th* unit being selected from the i *th* stratum be p_{i_j} / π_i, where

$$\pi_i = \sum_{j=1}^{N_i} p_{i_j} \text{ and } \sum_{i=1}^{n} \pi_i = 1.$$

Then, from (1.1) the minimum-variance-linear-unbiased estimator of T_i, the total of i *th* stratum is

$$\beta^i_{10} X^i_1 \text{ where } \beta^i_{10} = \pi_i / p_{ij} \text{ if } X^i_1 = X_{i_j}.$$

Therefore, an unbiased estimator of T is given by

$$(5.1) \qquad \hat{T}_{rhc} = \sum_{i=1}^{n} \beta^i_{10} X^i_1 .$$

Further,

$$V(\hat{T}_{rhc}) = E\left[\sum_{i=1}^{n} (\beta^i_{10} X^i_1 - T_i) \right]^2$$

$$= E_1 \sum_{i=1}^{n} E_2 (\beta^i_{10} X^i_1 - T_i)^2$$

where E_2 is the conditional expectation for a given set of n strata and E_1 is the expectation when strata are randomly formed. Therefore,

$$V(\hat{T}_{rhc}) = E_1 \sum_{i=1}^{n} \sum_{j=1}^{N_i} \left(\frac{x_{i_j} \pi_i}{p_{i_j}} - T_i \right)^2 . \frac{p_{i_j}}{\pi_i}$$

$$= E_1 \sum_{i=1}^{n} \left[\sum_{j=1}^{N_i} \sum_{j'(\neq j)}^{N_i} \frac{x^2_{i_j}}{p_{i_j}} p_{i_{j'}} - \sum_{j \neq j'}^{N_i} x_{i_j} x_{i_{j'}} \right].$$

When the strata are randomly formed, the probability that ij *th* and ij' *th* units are included in the i *th* stratum of size N_i is $N_i\,(N_i-1)/N\,(N-1)$. Therefore

$$(5.2)\quad V(\hat{T}_{rhc}) = \sum_{i=1}^{n} \frac{N_i\,(N_i-1)}{N\,(N-1)} \left[\sum_{j=1}^{N} \frac{x_j^2}{p_j}\,(1-p_j) - \sum_{j\neq j'}^{N} x_j\, x_{j'} \right.$$

$$= \frac{\sum_{i=1}^{n} N_i^2 - N}{N\,(N-1)} \left(\sum_{j=1}^{N} \frac{x_j^2}{p_j} - T^2 \right)$$

a result due to Rao et al (1962, equ. 5, p. 484). The variance of $\hat{T}_{rhc}$ is minimum when all N_i's are equal, giving $N_i=N/n$ for $i=1, 2,\ldots,n$. Thus, N must be a multiple of n for variance of $\hat{T}_{rhc}$ to be minimum. When N is not a multiple of n, let $N=nR+k$, where $O<k<n$ and R is a positive integer. Then we may choose $N_i=R+1$, $i=1, 2,\ldots, k$ and $N_i=R$, $i>k$, for random stratification. The variance of $\hat{T}_{rhc}$, so obtained, is shown by Rao et al to be close to the minimum value for $k=1, n-1$. In order to obtain an unbiased estimator of the variance of $\hat{T}_{rhc}$, we note

$$E\left(\sum_{i=1}^{n} (\beta_{10}\, X_1^i)^2 / \pi_i - \hat{T}_{rhc}^2 \right)$$

$$= \sum_{i=1}^{n} E_1 \sum_{j=1}^{N_i} (x_{ij}^2 / p_{ij}) - V(\hat{T}_{rhc}) - T^2$$

$$= \left(\sum_{j=1}^{N} \frac{x_j^2}{p_j} - T^2 \right) - V(\hat{T}_{rhc})$$

$$= \left(\frac{N\,(N-1)}{\Sigma N_i^2 - N} - 1 \right) V(\hat{T}_{rhc})$$

$$= \frac{N^2 - \Sigma N_i^2}{\Sigma N_i^2 - N}\, V(\hat{T}_{rhc}).$$

Therefore, an unbiased estimator of the variance of $\hat{T}_{rhc}$ is

$$(5.3)\quad \frac{\Sigma N_i^2 - N}{N^2 - \Sigma N_i^2} \left[\sum_{i=1}^{n} (\beta_{10}\, X_1^i)^2 / \pi_i - \hat{T}_{rhc}^2 \right]$$

an expression similar to that given by Rao, Hartley and Cochran (1962), equ. 15, p. 485). When N_i's are all equal, selection of one unit from each random stratum amounts to proportional allocation of n units to n strata. If the analogy with simple random sampling is to hold good, in this case, the estimator $\hat{T}_{rhc}$ should be as efficient as random sampling with varying probabilities and with or without replacement. It is not so for sampling with varying probabilities and with replacement is seen by noting, in view of equation (6.2), that when N_i's are all equal,

$$(5.4) \qquad V(\hat{T}_{rhc}) = \frac{N-n}{N-1} V(\hat{T}_2) w$$

the same as equation (9) of Rao et al. The analogy is not true even in case of sampiing with varying probabilities and without replacement in the light of the work of Gupta (1964) showing that the estimator in this case is sometimes less and some times more efficient than the estimator due to Rao et al depending upon the population.

6. Sampling with Varying Probabilities and with Replacement.

For sampling with replacement, let $P_{ir} = p'_{ir}$ and $k_i = \pi'_i$. Then,

$$\sum_{r \neq s}^{n} P(ir, js) = \pi'_i \pi'_j - \sum_{r=1}^{n} p'_{ir} p'_{jr}$$

Therefore, from equation (1.2), if $\hat{T}_{2w}$ denotes the estimator for sampling with replacement,

$$(6.1) \qquad V(\hat{T}_{2w}) = \sum_{i=1}^{N} \frac{x_i^2}{\pi'_i} + \sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j \left(1 - \frac{\sum_{r=1}^{n} p'_{ir} p'_{jr}}{\pi'_i \pi'_j}\right) - T^2$$

$$= \sum_{i=1}^{N} \frac{x_i^2}{\pi'^2_i} (\pi'_i - p'^2_{ir}) - \sum_{i \neq j} \frac{x_i x_j}{\pi'_i \pi'_j} \left(\sum_{r=1}^{n} p'_{ir} p'_{jr} \right)$$

as given by Prabhu Ajgaonkar and Tikkiwal (1961).

In many practical situations, $p'_{ir} = p'_i$ for $r = 1,2,\ldots,n$ and therefore $\pi'_i = np_i'$.

Substituting these values in (6.1)

$$(6.2) \quad V(\hat{T}_{2w}) = \frac{1}{n}\left(\sum_{i=1}^{N} \frac{x_i^2}{p_i} - T^2 \right).$$

In order to obtain an unbiased estimator of the variance of $\hat{T}_{2w}$, we note

$$E\left(\sum_{r=1}^{n} \beta^2_{ro}\, x_r^2 - \frac{\hat{T}^2_{2w}}{n} \right)$$

$$= \sum_{i=1}^{N} \frac{x_i^2}{np_i'} - \frac{1}{n}\left[\text{Var}\,(\hat{T}_{2w}) - T^2 \right]$$

$$= \left(\frac{n-1}{n}\right) \text{Var}\,(\hat{T}_{2w}).$$

Therefore, an unbiased estimator of the variance of $\hat{T}_{2w}$ is

$$(6.3) \quad \frac{1}{n-1}\left(n \sum_{r=1}^{n} \beta_{ro}^2\, x_r^2 - \hat{T}^2_{2w} \right).$$

7. The Relative Efficiency of the two Sampling Systems with and without Replacements.

We compare the efficiency of the two sampling systems under the assumption that $\pi_i = \pi'_i$. Then, from equations (2.1) and (6.1),

$$(7.1) \quad V(\hat{T}_{ht}) - V(\hat{T}_{2w}) = - \sum_{i=1}^{N} \frac{x_i^2}{\pi_i^2} \sum_{r \neq s}^{n} p'_{ir}\, p'_{is} +$$

$$\sum_{i \neq j}^{n} \frac{x_i\, x_j}{\pi_i\, \pi_j} \left(- \sum_{r \neq s}^{n} p'_{ir}\, p'_{js} + \pi_{ij} \right);$$

a quadratic form which is negative definite when there are equal probabilities of selection in the two sampling systems and when, in the population, x_i's are not all equal. This shows that sampling with replacement cannot be more efficient than sampling without replacement for equal probabilities of selection, a well known classical result. When there is varying probabilities of selection, no general statement can be made except that we note that the quadratic form cannot be positive definite. According to Durbin (1953, sec. 4, p. 266), the quadratic expression in (7.1) is usually negative, but it is easy to invent cases in which it is positive. This statement needs further examination.

8. Multi-stage sampling with Varying Probabilities and with or without Replacement.

We first discuss the theory of two-stage sampling in detail and then indicate its generalisation to multi-stage sampling.

Let x_{ij} denote the value of *j th* second-stage unit in the *i th* first-stage unit for $i=1,2,...,N$ and $j=1,2,...,M_i$. Let us estimate the population total $T = \sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}$. For this, let us draw n first-stage units and m_i second-stage units from the M_i second-stage units of the *i th* first-stage unit, if it is selected at a particular first-stage draw, for $i=1,2,...,N$. Thus, in sampling with replacement; where a given first-stage unit, say *i th*, may occur more than once, say t times, in the sample; t independent samples of size m_i second-stage units will be drawn out of M_i second-stage units in the *i th* first-stage unit. Let X_{rs} denote the variate value of the unit drawn at the *s th* second-stage draw after the *r th* first-stage draw for $r=1,2,...n$ and $s=1,2,...,$ m_i if *i th* first-stage unit occurs at the *r th* draw. Let p_{js}^{ir} denote the probability of *i th* first-stage unit occuring at the *r th* first-stage draw and then *j th* second-stage unit, in the *i th* first-stage unit, occuring at the *s th* second-stage draw. Let p_{ir} denote the probability of *i th* first-stage unit occuring at the *r th* first-stage draw. Let $p(js \mid ir)$ denote the conditional probability of *j th* second-stage unit, in the *i th* first-stage unit, occuring at the *s th* second-stage draw having occured *i th* first-stage unit at the *r th* first-stage draw. Let $\hat{T}_2$, an estimator in T_2–class, of population total T, be given by

$$(8.1) \qquad \hat{T}_{2,\,ms} = \Sigma\,\Sigma\, \beta_{so}^{r}\, x_{rs},$$

where $\beta_{so}^{r} = \beta_j^{i}$ if *i th* first-stage unit is selected at the *r th* first-stage draw and then *j th* second-stage unit, in the *i th* first-staee unit, is selected at the *s th* second-stage draw. Now,

$$E(\hat{T}_{2ms}) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} (x_{ij}\, \beta_j^{i}) \left(\sum_{r=1}^{n} \sum_{s=1}^{m_i} p_{js}^{ir} \right)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{M_i} x_{ij}\, \beta_j^{i}\, \pi_j^{i},$$

where $\pi_j^i = \sum_{r=1}^{n} \sum_{s=1}^{m_i} p_{js}^{ir}.$

If this is to be unbiased estimator of the population total T,

$$\beta_j^i = \frac{1}{\pi_j^i}$$

a unique value for β_j^i. Thus, there is only one unbiased estimator in T_2–class and therefore it is minimum-variance-linear-unbiased estimator. The variance and its estimator, for the minimum-variance-linear-unbiased estimator in (8.1) for sampling with or without replacement or both at the two stages of sampling, can be obtained as in Section 1.

In order to discuss the various results obtained by Durbin (1953) and others on multi-stage sampling, let $p\,(js|ir)$ be independent of r. This assumption would mean that the sampling system adopted at the second-stage is same for all first-stage draws. It is reasonable to adopt a uniform system of sampling at the second-stage. There does not seem to be any two-stage sampling system in use which does not satisfy the above assumption. This assumption is in fact true for any known multi-stage sampling system. Therefore, this assumption is a practical assumption.

Under the above practical assumption,

$$\pi_j^i = \sum_{r=1}^{n} \sum_{s=1}^{m_i} p_{ir}\, p\,(js \mid ir)$$

$$= \pi_i \sum_{s=1}^{m_i} p\,(js \mid ir).$$

Therefore, the estimator in (8.1) is equivalent to

$$(8.2) \qquad \hat{T}_{2,ms} = \sum_{r=1}^{n} \beta_{ro} X_r$$

where $\beta_{ro} = 1/\pi_i$, π_i being the probability of first-stage unit being included in the sample and where $X_r = \hat{T}_i$ the estimator of the total of *ith* first-stage unit; when *i th* first-stage unit occurs at the *r th* first-stage draw. If there are more than two stages, that the estimator can be put in the form in (8.2) is now obvious. In this form the result was given for the first-time by Durbin (1953). Thus under the above practical assumption, Durbin's estimator lies in T_2–class.

The estimator given by Sukhatme and Narain (1952); for sampling with replacement in their case (ii) where m_i second-stage units are selected each time, with simple random sampling without replacement, from the total M_i second-stage units in the *i th* first-stage unit, whenever it is selected in n first-stage draws; is a special case of Durbin's form and therefore it also belongs to T_2-class. As regards the other estimator given by these authors for their case (i) where, if a given first-stage unit, say *i th*, is selected t times in n draws, then, a sample of tm_i second-stage units is drawn out of M_i second-stage units with simple random samling without replacement; it is seen that this estimator also belongs to T_2-class by noting

$$\pi_j^i = \sum_t \left\{ \binom{n}{t} p_i^t \, (1-p_i)^{n-t} \right\} \frac{tm_i}{M_i}$$

$$= np_i \frac{m_i}{M_i};$$

since, by the authors assumption $p_{ir} = p_i$ for all r for sampling with replacement at the first-stage.

The above two estimators for two different types of sampling have been discussed in detail by the first author, Sukhatme (1953, sec. 8.2, p. 358 and sec. 8.9, p. 379).

The various unbiased estimators considered by Ecimovic (1956) are obtained by him from the unbiased estimator, of the total given in his equation (1), when there is simple random sampling without replacement at all the three stages. Since this unbiased estimator is also special case of Durbin's form, the various estimators due to Ecimovic belong to T_2- class.

Hansen and Hurvitz (1943, p. 341) and later Horvitz and Thompson (1952, pp 673-675) have considered a self-weighting design where m_i second-stage units are so chosen that the weight associated to the unit x_{ij} is constant for all i,j. There estimators for such a design are easily obtained by taking $\pi_j^i = t$, a constant quantity. Thus the estimators considered by these authors also belong to T_2- class.

Theoretically, it is possible to have sampling systems where p $(js|ir)$ is not independent of r. In such a case, the estimators in

(8.1) and (8.2) are not identical. In order to choose between the two estimators, one has to compare the efficiencies of the two estimators for the sampling system adopted. Since such sampling systems are not in use, we confine our attention to the case where $p\ (js|ir)$ is independent of r and therefore the estinators in (8.1) and (8.2) are identical. We proceed to find the variance of $\hat{T}_{2,ms}$ in (8.2) and an estimator of its variance, when there is sampling with or without replacement. Let

$$E\left[\hat{T}_{2,ms}\,|r=1,2,...,n\right] = \hat{T}_{2,ms}$$

an estimator of T when there is only one stage of sampling. Then, for sampling with or without replacement;

$$V(\hat{T}_{2,ms}) = E\left[\hat{T}_{2,ms} - \hat{T}'_{2,ms} + \hat{T}'_{2,ms} - T\right]^2$$

$$= V(\hat{T}'_{2,ms}) + E\left[V(\hat{T}_{2,ms}\,|\,r=1,2,...,n)\right]$$

$$(8.3) \qquad = V(\hat{T}'_{2,ms}) + \sum_{i=1}^{N}\left(\frac{V(\hat{T}_i\,|\,i)}{\pi_i}\right)$$

The estimation of the variance, in more than one stage of sampling, for sampling with or without replacement does not present any new problem other than already considered in Sections 3 and 6. In order to see this, let $\hat{V}_{wo}$ denote some unbiased estimator of the variance in case of sampling without replacement, such as one of those given in Section 3. Let $\hat{V}_w$ denote some unbiased estimator of the variance, in case of sampling with replacement, such as one given in Section 6. Let s_i^2 be an unbiased estimator of $V(\hat{T}_i\,/\,i)\,/\pi_i^2$.

Let $x_i \equiv \hat{T}_i$ in $\hat{V}_{wo}$ and $\hat{V}_w$ and let $s_r = \pi_i\, s_i^2$ whenever *i th* first-stage unit occurs at the *r th* first-stage draw. Then,

$$(8.4) \qquad E\left[\hat{V}_{wo} + \sum_{i=1}^{n} s_r\right] = \text{Var}\,(\hat{T}_{2,ms})$$

and

$$(8.5) \qquad E(\hat{V}_w) = V(\hat{T}_{2,ms})$$

is easily noted. Thus the expressions within brackets on the left in

(8.4) and (8.5) give unbiased estimators of the variances for sampling without and with replacements.

Yates (1949, sec. 7.17, pp 226-227) has given a rule for estimation of the variance in multi-stage sampling when there is simple random sampling without replacement at least at the first-stage. The following generalised rule, due to Durbin (1953, p. 264), for sampling with varying probabilities and without replacement at least at the first-stage; is obvious from (8.4) :

"The estimate of variance in multi-stage sampling is the sum of two-parts. The first part is equal to the estimate of variance calculated on the assumption that the first-stage values have been measured without error. The second part is calculated as if the first-stage units selected were fixed strata, the contribution from each first-stage units being multiplied by the probability of that unit's inclusion in the sample."

Yates's rule follows as a special case of Durbin's rule if we take $\pi_i = n/N$ for $i=1, 2,\ldots,N$.

REFERENCES

1. Des Raj (1956); A note on the determination of optimum probabilities in sampling without replacement. Sankhya Vol. 17, 197–200.

2. Durbin, J. (1953), Some recent results in sampling theory when the units are selected with unequal probabilities. Jour. Roy Statist. Soe. B Vol. 15, 262–269

3. Ecimovic, J. P. (1956); Three stage sampling with varying probabilities of selection. Jour. Ind. Soc. Agric. Statist. Vol. 7, 14-44.

4. Hansen, M. H. and Hurvitz, W. N. (1943); On the theory of sampling from finite populations. Ann. Math. Statist. Vol.15, 333–362.

5. ———— (1949); On the determination of optimum probabilities. Ibid Vol. 20, 426–432.

6. Gupta, M. G. (1964); Some unpublished results on varying probabilities. Department of Statistics, University of Rajasthan.

7. Hartley, H. O. and Rao, J. N. K. (1962); Sampling with unequal probabilities and without replacement. Ann. Math. Statist. Vol. 33, 350–374.

8. Horvitz, D. G. and Thompson, D. J. (1952); A generalization of sampling without replacement from a finite population. Jour. Amer. Statist. Assoc. Vol. 47, 663-685 (Presented to the Institute of Mathematical Statistics, March 1951).

9. Koop, J. C. (1961); Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities. Memeo series No. 296 Institute of Statistics, North Carolina, U.S A. (Based on Ph.D. thesis 1957).

*10. Madow, William G. (1949); On the theory of systematic sampling 11. Ann. Math. Statist. Vol. 20, 333–354.

*11. Midzuno, Hiroshi (1950); An outline of the theory of sampling systems. Annals of the Institute of the Statistical Mathematics, Vol. 1, 149–56.

*12. ———— (1952); On the sampling system with probability proportional to sum of sizes. Ibid Vol. 3, 99–107.

13. Narain, R. D. (1951); On sampling with replacement with varying probabilities. Jour. Ind. Soc. Agric. Statist. Vol. 3, 169–174. (Presented to the Indian Society of Agricultural Statistics, Nov , 1950).

14. Prabhu Ajgaonkar, S. G. and Tikkiwal, B. D. (1961); Horvitz and Thompson's T-class estimators. Ann. Math. Statist. Vol. 32, 923 (Abstract).

15. Prabhu Ajgaonkar, S. G. (1962); Some aspects of the class of linear estimators with special reference to successive sampling. Ph.D. thesis, Karnatak University, Dharwar, India.

16. Rao, J. N. K.; Hartely, H.O. and Cochran, W. G. (1962); On a simple procedure of unequal probablity sampling without replacement. Jour. Roy. Statist. Soc. B. Vol. 24, 482-491.

17. Sen, A. R. (1953); On the estimate of the variance in sampling with varying probability. Jour. Ind. Soc. Agric. Statist. Vol. 5, 119–127.

*18. ———— (1953); Recent advances in sampling with varying probabilities. Cal. Statist. Assoc Bull. Vol. 5, 1–15.

*19. Singh, D. (1954); On efficiency of sampling with varying probabilities without replacement. Jour. Ind. Soc. Agric. Statist. Vol. 6, 48–57.

20. Sukhatme, P. V. and Narain, R. D. (1952); Sampling with replacement. Jour. Ind. Soc. Agric. Statist. Vol. 4, 42–49.

21. Sukhatme, P. V. (1953); Sampling theory of surveys with applications. New Delhi. : Indian Society of Agricultural Statistics, Ames, Iowa : Iowa State College Press.

22. Tikkiwal, B. D. (1960); On the theory of classical regression and double sampling estimation. Jour. Roy. Statist. Soc. B, Vol. 22, 131–138.

23. Yates, F. (1949); Sampling methods for census and surveys. London : Griffin.

24. Yates, F. and Grundy P. M. (1953); Selection without replecement from within strata with probability proportional to size. Jour. Roy. Statist. Soc. B, Vol. 15, 253-261.

*Not directly referred in the paper.

Paper received : October, 1964.

OSKAR RYSZARD LANGE : 1904-1965

Prof. Oskar Lange one of the eminet economists and statisticians of the contemporary world, Vice-President of the Council of State of Poland, Chairman of Indo-Polish Friendship Society is no more. He died on 2nd Oct., 65. In his death the world in general and India and Poland in particular have suffered a great loss. Prof. Lange was a personal friend of our late Prime Minister Jawahar Lal Nehru. He had left an ever lasting and sincere imprint on the planners of our country with whom he had worked in 1956 for six months in the drafting of the Second Five Year Plan.

Oskar Ryszard Lange was born on July 27, 1904 in Tomoszow Mozowiechi, Poland. At a very early age, he became interested in the application of Mathematics to Economics and Social Sciences and at the age of ninteen he read a paper "An Essay in the Theory of the Limits of Production" at Jagellonian University. From this original attempt to apply mathematics to problems of prodution under capitalism and socialism stemmed the broad stream of Oskar Lange's economic studies in the decades to come. His study 'Statistical Investigation of Business Cycles' at Jagellonian University, Cracow, Poland (1931) qualified him for the post of lecturer in Statistics.

Lange in U.S.A. and U.S.S.R. and then back to Poland

Through out the interwar period Oskar Lange had close contacts with the socialistic youth movement. These political activities together with the temper of his writing were a great handicap to his remaining at Jagellonian University. A fellowship from Rockfeller Foundation gave him the opportunity to spend two years in the U.S.A. and in England. There his two essays on the economic theory of socialist economy made him internationally known and secured his place in the history of the political economy of socialism.

Oskar Lange was appointed assistant professor of economics and statistics at the University of Chicago in 1939 and he remained there till 1945 becoming associate professor and then full professor.

At intervals he also lectured at Columbia and Stanford Universities In U.S.A. Prof. Lange published a number of theoretical papers. His most important publications of this period were published in 1944 as a book titled 'Price Flexibility and Employment'.

At the invitation of the Union of Polish Patriots in the U.S.S.R. he travelled to the Soviet Union in the Spring of 1944. There he conferred about the future of Poland with Joseph Stalin, Premier of U.S.S.R.

A Leading Political Figure

In the summer of 1945 Lange returned to Poland; shortly afterwards the government appointed him ambassador to the U.S A. and delegate to the United Nations. In the United Nations he asked the members of U.N. Forum to break diplomatic relations with the Government of General Franco. He also urged co-existence, general disarmament, banning of atomic weapons and broad economic collaborations between all countries. In particular he criticized the use of economic aid as a means of pressure and intervention in the internal affairs of other countries. He was soon recognized as a leading political figure.

At the National Congress of Polish Socialist Party in 1947 Lange was elected to the Central Executive Committee. In 1957 he became Deputy Chairman to the Council of State.

In the international field he had been one of the most active advocates of peaceful coexistence and of international economic co-operation. In 1957-59 he was Chairman of the U.N. Economic Commission for Europe. In 1961-62 he sat on the U.N. Committee of experts for the study of the economic and social consequences of disarmament.

Lange's Contribution to Economics and Statistics

In 1956, Lange received the chair of Political Economy at the University of Warsaw. There, almost each year he started a new course of lectures on a different subject and out of them had come four books. His Lectures on Econometrics (1956-57) were published under the title "Introduction to Econometrics". Besides Polish, the book is available in English, Serbo-Croation, Italian, Japanese and Russian also. In Polish and in English this book has run into two editions. His book on programming viz. Optimum Decision-making" appeared in 1964. 'Theory of Reproduction and Accumulation' which was published in 1961 gives a mathematical treatment of the Marxist theory of reproduction and accumulation. A Russian translation of this book appeared in 1963. In recent years Prof. Lange had been lecturing on Cybernetics and its application to Economics. 'A general theory of System Behaviour' (1963) (published both in English and in Polish) and the recent book 'Introduction to Economic Cybernetics' are valuable contributions in this field. About this time (1957) his book on 'Political Economy' in three volumes was also written. We have already referred to Prof. Lange's 'Statistical Investigation of Business Cycles'. His doctoral thesis was on Business Cycles in Poland' (1924-27) Prof. Lange and contributed papers in disciplines like Sociology and History of Law.

However, even this was not whole of Prof. Lange's scientific output. Among other publications, he had produced over a dozen importand shorter papers, a two-part mimeographed course in the theory of economic growth and a number of public lectures delivered abroad. Besides all these, Prof. Lange contributed a new shape of socialism, his advocacy of socialist democratization being closely geared with his aspiration to put the planning and management of the socialist economy on scientific grounds.

An important sector of his interest was the problem of the countries of the so-called Third World. At the invitation of the Governments of India and Ceylon he participated in the drafting of economic plans for these countries. "Essays on Economic Panning" was the outcome of his visit to India. For the same end he was also invited later by the Planning Commission of the United Arab Republic and of Iraq. In addition to his practical consulations Prof. Lange also worked on a theoretical generalization of the problem of economic unnderdevelopment.

Prof. Oskar Lange was a member of the Polish Academy of Sciences, since 1952. He was also Fellow of the Econometric Society, since 1959, Fellow of the Institute of Social Studies in the Hague, since 1962, Honorary Member of the Royal Statistic Society, since 1964, Member of the International Statistical Institute, since 1955 and Doctor Honoris Causa of the University of Dijon in France in 1962. On May 9, 1964 he was made doctor honoris causa of his almamater Jagellonian University.

Acknowledgement

This obituary note is based on the material supplied by the Polish Embassy, New Delhi.

P. Krishnan
&
U. S. Madan
Department af Statistics
University of Rajasthan.
