JOURNAL

OF THE

Indian Society of Agricultural Statistics

Vol. X] 1958 [Nos. 1-2

ESTIMATES OF VARIANCE COMPONENTS IN FINITE POPULATION

By Daroga Singh

Indian Council of Agricultural Research, New Delhi

1.1. Introduction

In multi-stage sampling design the sampling variance of the mean is generally a function of different stage variance components and in order to know the sampling error of the mean, it becomes necessary for the sampler to know their values. Further, in order to be able to develop a suitable sampling scheme and to distribute the total sample so that maximum efficiency may be attained we require not merely an estimate of the overall sampling error but also a knowledge of different variance components; Cochran (1939) has shown how in case of multi-stage simple sampling from infinite population, this leads to a problem of analysis of total variation into different stage components. Some other authors [Yates (1949), Sukhatme (1954), Crump (1951, 1946) and Henderson (1953)] have also discussed this problem and have indicated methods how in simple cases the variance components can be estimated. But if the sampling method is not simple, if the sizes of the sampling units are unequal, if population is finite, and if samples to be selected from each selected unit at successive stages are not equal, procedure for estimating the variance components at different stages is not straightforward and the sampler is confronted with practical difficulty in making best use of the available data in planning the future survey. In this paper it has been shown how the method of analysis of variance can be suitably used for studying these variance components. The method of analysis of variance has two-fold advantages, firstly calculation procedure is simple, and secondly most of the workers are very well familiar with this method.

2.1. Two-Stage Sampling Design

Let there be N primary sampling units (*psu*) in the population. The i-th *psu* consists of M_i second stage sampling units (*ssu*'s). Assume n *psu*'s are drawn with replacement, the selection probability for the i-th *psu* in the population being p_i ($i = 1, 2, \ldots, N$). Assume further m_i's ($i = 1, 2, \ldots, n$) *ssu*'s are selected from the *psu* drawn at the i-th draw in the sample by method of simple random sampling. m_i is the number which does not depend on the selected *psu*. Let X_{ij} denote the value of the j-th *s-su* ($j = 1, \cdots, M_i$) of the ith *psu* ($i = 1, 2, \ldots, N$) in the population,

$$\bar{X}_{i.} = \frac{1}{M_i} \sum_{j=1}^{M_i} X_{ij},$$

$$\bar{X}_{..} = \sum_{i=1}^{N} p_i \bar{X}_i,$$

$$\sigma_b^2 = \sum_{i=1}^{N} p_i (\bar{X}_{i.} - \bar{X}_{.})^2,$$

$$\sigma_i^2 = \sum_{j=1}^{M_i} \frac{(X_{ij} - \bar{X}_{i.})^2}{M_i - 1},$$

and

$$\bar{\sigma}_w^2 = \sum_{i=1}^{N} p_i \sigma_i^2.$$

Assume,

x_{ij} = the value of the j-th *ssu* of the i-th *psu* in the sample, $\quad (j = 1, \cdots, m_i)$ $(i = 1, 2, \cdots, n)$,

$$\bar{x}_{i.} = \frac{1}{m_i} \sum_{j}^{m_i} x_{ij} \quad ,$$

$$\bar{x}_{..} = \frac{1}{m_0} \sum_{i}^{n} \sum_{j}^{m_i} x_{ij} \qquad \left(m_0 = \sum_{i}^{n} m_i\right),$$

$$s_b^2 = \sum_i^n \frac{m_i (\bar{x}_{i.} - \bar{x}_{..})^2}{n-1},$$

and

$$s_w^2 = \sum_i^n \sum_j^{m_i} \frac{(x_{ij} - \bar{x}_{i.})^2}{m_0 - n}.$$

Now Table I gives the analysis of variance for m_0 observations in the sample.

TABLE I

Analysis of variance

Source of variation	Degree of freedom	Mean square
Between *psu*	$n - 1$	s_b^2
Within *psu*	$m_0 - n$	s_w^2

It may be easily seen that under the sampling plan,

$$E(s_b^2) = \lambda \sigma_b^2 + \bar{\sigma}_w^2 - \lambda \sum_{i=1}^{N} \frac{p_i \sigma_i^2}{M_i}, \tag{1}$$

where

$$\lambda = \frac{m_0 - \sum_i^n \frac{m_i^2}{m_0}}{n-1}$$

and

$$E(s_w^2) = \bar{\sigma}_w^2. \tag{2}$$

E stands for expectation.

Thus s_w^2 is unbiased estimate of $\bar{\sigma}_w^2$.

Now if M_i is large, the third term in (1) may be negligible and

$$E(s_b^2) = \lambda \sigma_b^2 + \bar{\sigma}_w^2. \tag{3}$$

Therefore,

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2}{\lambda}. \tag{4}$$

But if the third term in (1) is not negligible, the estimate of σ_b^2 given by (4) will be an underestimate and the underestimation will depend on the value of $\sum_{i=1}^{N} p_i \sigma_i^2 / M_i$

If $p_i \propto M_i$, say $p_i = M_i/M_0$ $\left(\text{where } M_0 = \sum_{i=1}^{N} M_i\right)$

$$\sigma_i = \bar{\sigma}_w \; (i = 1, 2, \ldots, N),$$

and

$$E(s_b^2) = \lambda \sigma_b^2 + \bar{\sigma}_w^2 \left(1 - \frac{\lambda}{\bar{M}}\right), \qquad \text{where } \bar{M} = \frac{M_0}{N}.$$

Therefore

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2 \left(1 - \frac{\lambda}{\bar{M}}\right)}{\lambda}. \tag{5}$$

If $\lambda/\bar{M}$ is negligible, estimate of σ_b^2 is given by (4). However, if $s_w^2/\bar{M}$ is small compared to $s_b^2 - s_w^2/\lambda$, bias in accepting (4) as an estimate of σ_b^2 will not be serious. Under other conditions estimate of σ_b^2 given by (4) will be an underestimate.

It can also be seen if M_i's are equal (say $M_i = M$),

$$E(s_b^2) = \lambda \sigma_b^2 + \bar{\sigma}_w^2 \left(1 - \frac{\lambda}{M}\right)$$

and

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2 \left(\frac{1 - \lambda}{M}\right)}{\lambda}. \tag{6}$$

If λ/M is negligible, the estimate of σ_b^2 is again given by (4).

Further, if m_i's are equal (say $m_i = m$),

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2}{m} + \frac{s_w^2}{M}. \tag{7}$$

And if the sampling fraction at the second stage is negligible,

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2}{m}, \tag{8}$$

which is usually obtained from the analysis of variance for large population.

In case of simple random sampling ($p_i = 1/N$), if

$$\sigma_i = \sigma_w \qquad (i = 1, 2, \ldots, N).$$

$$\text{est. } \sigma_b^2 = \frac{s_b^2 - s_w^2\left(1 - \dfrac{\lambda}{\bar{M}_h}\right)}{\lambda}, \tag{9}$$

where $\bar{M}_h$ is harmonic mean. As before, if $\lambda/\bar{M}_h$ is negligible, the est. σ_b^2 will be given by (4). But if $\lambda/\bar{M}_h$ is not small, estimate of σ_b^2 given by (4) will be an underestimate and the underestimation will depend on the estimated variance of the mean per *ssu*.

In large sampling enquiry calculation of value of λ will not be an easy task and one may be tempted in using an approximate value of λ given by

$$\bar{m} = \frac{m_0}{n} \tag{10}$$

It can be seen that

$$\bar{m} - \lambda = \frac{\sigma_m^2}{m_0}, \qquad \text{where } \sigma_m^2 = \frac{\sum_i^n (m_i - \bar{m})^2}{n-1}.$$

Now if variation among m's is not large so that σ_m^2 be a small quantity, in that case $\bar{m}$ may be a good approximation for λ. But since

$$\bar{m} - \lambda \geqslant 0 \qquad \text{(equality will be achieved only if } m_i = m, \; i = 1, \ldots, n),$$

in using $\bar{m}$ there is risk of underestimating σ_b^2. However, in actual practice it is not very common to adopt the system of drawing unequal number of *ssu*'s (m_i) from the i-th *psu* in the sample. In an efficient survey design m_i's will be usually equal although it is likely that through extraneous causes the number of *ssu*'s actually observed may be unequal. Thus in cases where m_i's are not very much different $\bar{m}$ will be a good approximate to λ.

2.2. If the i-th *psu* is selected with probability pi without replacement and m *ssu*'s units are selected from the i-th selected *psu* in the sample,

$$E(s_b^2) = \frac{1}{n} \sum_{i=1}^{N} P_i \frac{M_i - m}{M_i} \sigma_i^2$$

$$+ \frac{m}{2n(n-1)} \sum_{i \neq j=1}^{N} P_{ij} (\bar{X}_i - \bar{X}_j)^2, \tag{11}$$

$$E(s_w^2) = \frac{1}{n} \sum_{i=1}^{N} P_i \sigma_i^2, \tag{12}$$

where P_i is the probability that the i-th unit will be selected in the sample of size $n \left(\sum_{i=1}^{N} P_i = n \right)$ and P_{ij} is the probability that the i-th and j-th *psu*'s will be selected in the sample of size $n \left(\sum_{i \neq j=1}^{N} P_{ij} = n(n-1) \right)$. If $\sigma_i = \bar{\sigma}_w$ $(i = 1, 2, \ldots, N)$, s_w^2 will be unbiased estimate of $\bar{\sigma}_w^2$. And if $1/n \sum_{i=1}^{N} \frac{P_i}{M_i} \sigma_i^2$ is negligible, $\frac{s_b^2 - s_w^2}{m}$ will be an estimate of a quantity

$$\sigma_b'^2 = \frac{1}{2n(n-1)} \sum_{i \neq j=1}^{N} P_{ij} (\bar{X}_i - \bar{X}_j)^2,$$

which corresponds to the mean square between the *psu*'s in the population. If the method of selection of the *psu* be simple random sampling without replacement

$$E(s_b^2) = m\sigma_b''^2 + \frac{1}{N} \sum_{i=1}^{N} \frac{M_i - m}{M_i} \sigma_i^2, \tag{13}$$

$$E(s_w^2) = \frac{1}{N} \sum \sigma_i^2 = \bar{\sigma}_w'^2, \tag{14}$$

where

$$\sigma_b''^2 = \sum_{i=1}^{N} \frac{(\bar{X}_i - \bar{X}_{..}')^2}{N-1} \qquad \left(\bar{X}_{..}' = \frac{1}{N} \sum \bar{X}_i \right).$$

Equations (13) and (14) will lead to similar results as discussed in (4)–(7), although definition of true variances has changed.

3.1. Three-Stage Sampling Design

Let, as before

N = number of *psu* in the population,

M_i = number of *ssu* in the *i*-th *psu* in the population,

and

K_{ij} = number of ultimate stage units (*usu*'s) in the *j*-th *ssu* of *i*-th *psu*, in the population.

Let small letters (n, m_i, and k_{ij}) denote the corresponding number in the sample.

Let X_{ijl} ($i = 1, \cdots, N$; $j = 1, 2, \ldots, M_i$; $l = 1, 2, \ldots, K_{ij}$) be the value of l-th *usu* of the j-th *ssu* in the i-th *psu* in the population. x_{ijl}'s will denote the corresponding values in the sample.

Now the analysis of variance table for the sample is given as below:

Table II

Analysis of variance

Source of variation	d.f.	sum of square	Mean square
Between *psu* ..	$n-1$	$\sum_i^n k_i (\bar{x}_{i..}-\bar{x}_{...})^2$	s_1^2
Between *ssu* ..	m_0-n	$\sum_i^n \sum_j^{m_i} k_{ij} (\bar{x}_{ij}-\bar{x}_{i..})^2$	s_2^2
Between *usu* ..	k_0-m_0	$\sum_i^n \sum_j^{m_i} \sum_l^{k_{ij}} (x_{ijl}-\bar{x}_{ij.})^2$	s_3^2

$$\left[k_i = \sum_j^{m_i} k_{ij},\ k_0 = \sum_i^n k_i \text{ and } m_0 = \sum_i^n = m_{i.}\right].$$

If the selection procedure of sampling units at successive stages be simple random,

$$E(s_1^2) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} \sum_{j=1}^{M_i} \left(1 - \frac{\lambda_2}{k_{ij}}\right) \sigma^2_{ij(3)}$$

$$+ \frac{1}{N} \sum_{i=1}^{N} \left(\lambda_2 - \frac{\lambda_1}{M_i}\right) \sigma^2_{i(2)} + \lambda_1 \sigma_1^2, \tag{15}$$

$$E(s_2^2) = \frac{\lambda}{N} \sum_{i=1}^{N} \sigma_i{}^2{}_{(2)} + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} \sum_{j=1}^{M_i} \left(1 - \frac{\lambda}{k_{ij}}\right) \sigma_{ij}{}^2{}_{(3)}, \tag{16}$$

and

$$E(s_3^2) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} \sum_{j=1}^{M_i} \sigma_{ij}{}^2{}_{(3)}, \tag{17}$$

where

$$\lambda = \frac{k_0 - \sum_i^n \sum_j^{m_i} \frac{k_{ij}^2}{k_i}}{m_0 - n},$$

$$\lambda_1 = \frac{k_0 - \sum_i^n \frac{k_i^2}{k_0}}{n - 1},$$

$$\lambda_2 = \frac{\sum_i^n \sum_j^{m_i} \frac{k_{ij}^2}{k_i} - \sum_i^n \sum_j^{m_i} \frac{k_{ij}^2}{k_0}}{n - 1},$$

$$\sigma^2_{(1)} = \sum_{i=1}^{N_i} \frac{(\bar{X}_{i..} - \bar{X}_{...})^2}{N - 1},$$

$$\sigma^2_{i(2)} = \sum_{j=1}^{M_i} \frac{(\bar{X}_{ij} - \bar{X}_{i..})^2}{M_i - 1},$$

$$\sigma^2_{ij(3)} = \sum_{l=1}^{K_{ij}} \frac{(X_{ijl} - \bar{X}_{ij})^2}{K_{ij} - 1},$$

$$\bar{X}_{ij.} = \frac{1}{K_{ij}} \sum_{l=1}^{K_{ij}} X_{ijl}, \ \bar{X}_{i..} = \frac{1}{M_i} \sum_{j=1}^{M_i} \bar{X}_{ij.},$$

and

$$\bar{X}_{...} = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_{i..} .$$

If

$$\sigma^2_{ij(3)} = \sigma^2_{(3)} \qquad (j = 1, 2, \ldots, M_i; \ \ i = 1, 2, \ldots, N)$$

and

$$\sigma^2_{i(2)} = \sigma^2_{(2)} \qquad (i = 1, 2, \ldots, N),$$

and sampling fraction at each stage is small,

$$E(s_1^2) = \sigma^2_{(3)} + \lambda_2 \sigma^2_{(2)} + \lambda_1 \sigma^2_{(1)}, \tag{18}$$

$$E(s_2^2) = \sigma^2_{(3)} + \lambda \sigma^2_{(2)}, \tag{19}$$

and

$$E(s_3^2) = \sigma^2_{(3)}. \tag{20}$$

From the equations (18)–(20) we may easily estimate the value of $\sigma_{(1)}$, $\sigma_{(2)}$, and $\sigma_{(3)}$.

If in particular case, the same number of samples is selected from successive stage of sampling units, *i.e.*,

$$m_i = m \ (i = 1, \cdots, n),$$

and

$$k_{ij} = k \ (j = 1, \ldots, m \ ; \ \ i = 1, \ldots, n),$$

then

$$\lambda_1 = mk \qquad \text{and} \qquad \lambda = \lambda_2 = k.$$

Hence

$$E(s_1^2) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} \sum_{j=1}^{M_i} \left(1 - \frac{k}{K_{ij}}\right) \sigma^2_{ij(3)}$$

$$+\frac{k}{N}\sum_{i=1}^{N}\left(1-\frac{m}{M_i}\right)\sigma^2_{i(2)}$$

$$+\,mk\,\sigma^2_{(1)}, \tag{21}$$

and

$$E\,(s_2^2)=\frac{k}{N}\sum_{i=1}^{N}\sigma^2_{i(2)}+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M_i}\sum_{j}^{M_i}\left(1-\frac{k}{K_{ij}}\right)\sigma^2_{ij(3)}. \tag{22}$$

Again if

$$\sigma_{i(2)}=\sigma_{(2)} \qquad (i=1,2,\ldots,N),$$

$$\sigma_{ij(3)}=\sigma_{(3)} \qquad (j=1,2,\ldots,M_i;\;\; i=1,2,\ldots,N),$$

and sampling fraction is small, we get the similar set of equations as given by (18)–(20).

It can be seen from (21) and (22) that if the sampling fraction at the second stage is small whatever be the sampling fraction at the third stage it is possible to estimate the value of $\sigma_{(1)}{}^2$ and it will be

$$\frac{s_1{}^2-s_2{}^2}{mk}. \tag{23}$$

Similarly, whatever be the sampling fraction at the second stage, if the sampling fraction at the third stage is small the value of $\sigma_{(2)}{}^2$ may be estimated by

$$\frac{s_2{}^2-s_3{}^2}{k} \tag{24}$$

In particular case, if $M_i=M$ $(i=1,\ldots,N)$, $K_{ij}=K\,(i=1,\ldots,M_i;$ $i=1,\ldots,N)$, $m_i=m\,(i=1,\ \ldots,\ n)$ and $k_{ij}=k\,(j=1,\ \ldots,\ m_i;$ $i=1,\ldots,n)$, we get

$$E\,(s_1{}^2)=\frac{K-k}{K}\,\sigma^2_{(3)}+k\,\frac{M-m}{M}\,\sigma^2_{(2)}+mk\sigma^2_{(1)}, \tag{25}$$

$$E\,(s_2{}^2)=k\sigma^2_{(2)}+\frac{K-k}{K}\,\sigma^2_{(3)}, \tag{26}$$

and

$$E\,(s_3{}^2)=\sigma^2_{(3)}. \tag{27}$$

Therefore

$$\text{est. } \sigma_{(2)}{}^2 = \frac{s_2{}^2 - s_3{}^2}{k} + \frac{s_3{}^2}{K}, \tag{28}$$

and

$$\text{est. } \sigma^2{}_{(1)} = \frac{s_1{}^2 - s_2{}^2}{mk} + \frac{s_2 - s_3{}^2}{kM} + \frac{s_3{}^2}{MK}. \tag{29}$$

If M and K are large, we get the same expression for the estimates of $\sigma^2{}_{(1)}$ and $\sigma^2{}_{(2)}$ as it is obtained in the analysis of variance for large population.

3.2. Now consider the system such that $p_i \left(\sum_{i=1}^{N} p_i = 1 \right)$ = the probability of selecting the i-th *psu* $(i = 1, 2, \ldots, N)$, selection being with replacement.

$p_{ij} \left(\sum_{j}^{M_i} p_{ij} \, (i = 1) \right)$ = the probaility of selecting the j-th *ssu* of i-th *psu* with replacement $(j = 1, \ldots, M_i;\ i = 1, \ldots, N)$. Assume *usu*'s are selected with equal probability without replacement. Then we get

$$E\,(s_1{}^2) = \sum_{i=1}^{N} p_i \sum_{j}^{M_i} p_{ij} \left(1 - \frac{\lambda_2}{K_{ij}}\right) \sigma^2{}_{ij(3)}$$

$$+ \lambda_2 \sum_{i=1}^{N} p_i \sigma^2{}_{i(2)} + \lambda_1 \sigma^2{}_{(1)}{}', \tag{30}$$

$$E\,(s_2{}^2) = \sum_{i=1}^{N} p_i \sum_{j=1}^{M_i} p_{ij} \left(1 - \frac{\lambda}{K_{ij}}\right) \sigma_{ij}{}^2{}_{(3)}$$

$$+ \lambda \sum_{i=1}^{N} p_i \sigma^2{}_{i(2)}, \tag{31}$$

and

$$E\,(s_3{}^2) = \sum_{i=1}^{N} p_i \sum_{j=1}^{M_i} p_{ij} \sigma^2{}_{ij(3)}, \tag{32}$$

where

$$\sigma'^2{}_{i(2)} = \sum_{j=1}^{M_i} p_{ij} \, (\bar{X}_{ij.} - \bar{X}_i'{}_{..})^2,$$

$$\sigma'^2 = \sum_{i=1}^{N} p_i (\bar{X}'_{i..} - \bar{X}'_{...})^2,$$

$$\bar{X}'_{i..} = \sum_{j=1}^{M_i} p_{ij} \bar{X}_{ij.},$$

and

$$\bar{X}'_{...} = \sum_{i=1}^{N} p_i \bar{X}'_{i..}.$$

Thus it is seen that when the system of probability of selection of sampling units is changed, definition of variance components occurring in the expectation of mean square in the analysis of variance is also changed. If the sampling fraction at the third stage is small, the estimates of $\sigma_1'^2$ and $\sigma_2'^2$ $(= \Sigma p_i \sigma_{i(2)}'^2)$ can be easily obtained from the equations (30) and (31). However, if sampling fraction is not small and if

$$\sigma^2_{ij} = \sigma^2_3 \; (j = 1, \ldots, M_i;\; i = 1, \ldots, N),$$

$$\text{est.}\, \sigma'^2_{(2)} = \frac{s_2^2 - s_3^2}{\lambda} + \frac{s_3^2}{\bar{K}_h'}, \tag{33}$$

and

$$\text{est.}\, \sigma'^2_{(1)} = \frac{(s_1^2 - s_3^2) - \dfrac{\lambda_2}{\lambda}(s_2^2 - s_3^2)}{\lambda_1}; \tag{34}$$

where

$$\frac{1}{\bar{K}_h'} = \sum_{i=1}^{N} p_i \sum_{j=1}^{M_i} \frac{p_{ij}}{K_{ij}}.$$

$\bar{K}_h'$ may be defined as 'weighted harmonic mean'.

Thus estimate of $\sigma'^2_{(1)}$ is independent of $\bar{K}_h'$. If the values of λ_2 and λ are of the same order so that $\lambda_2/\lambda \sim 1$ (this will be generally achieved only when the number of sampling units in the sample at successive stage is of the same order),

$$\text{est.}\, \sigma'^2_{(1)} = \frac{s_1^2 - s_2^2}{\lambda_1}.$$

3.3. If the selection procedure at the second stage is also simple without replacement,

$$E(s_1^2) = \sum_{i=1}^{N} \frac{p_i}{M_i} \sum_{j=1}^{M_i} \left(1 - \frac{\lambda_2}{K_{ij}}\right) \sigma_{ij}{}^2{}_{(3)}$$

$$+ \sum_{i=1}^{N} p_i \left(\lambda_2 - \frac{\lambda_1}{M_i}\right) \sigma_i{}^2{}_{(2)}$$

$$+ \lambda_1 \sigma_{(1)}{}''^2, \qquad (35)$$

$$E(s_2^2) = \sum_{i=1}^{N} \frac{p_i}{M_i} \sum_{j=1}^{M_i} \left(1 - \frac{\lambda}{K_{ij}}\right) \sigma_{ij}{}^2{}_{(3)}$$

$$+ \lambda \sum_{i=1}^{N} p_i \sigma^2{}_{i(2)}, \qquad (36)$$

and

$$E(s_3^2) = \sum_{i=1}^{N} \frac{p_i}{M_i} \sum_{j=1}^{M_i} \sigma^2{}_{ij(3)}, \qquad (37)$$

where $\sigma^2{}_{i(2)}$ and $\sigma^2{}_{ij(3)}$ are given by (15)–(17), and

$$\sigma''^2{}_{(1)} = \sum_{i=1}^{N} p_i (\bar{X}_{i..} - \bar{X}_{...})^2.$$

If $p_i \propto M_i$ and $\sigma^2{}_{ij(3)} = \sigma^2{}_{(3)}$ $(j = 1, \ldots, M_i;\ i = 1, \ldots, N)$ and $\sigma_{i(2)}{}^2 = \sigma_{(2)}{}^2$,

$$E(s_1^2) = \left(1 - \frac{\lambda_2}{\bar{K}_h''}\right) \sigma^2{}_{(3)} + \lambda_2 \sigma^2{}_{(2)} - \frac{\lambda_1}{\bar{M}} \sigma^2{}_{(2)}$$

$$+ \lambda_1 \sigma''^2{}_{(1)}, \qquad (38)$$

$$E(s_2^2) = \lambda \sigma^2{}_{(2)} + \left(1 - \frac{\lambda}{\bar{K}_h''}\right) \sigma^2{}_{(3)}, \qquad (39)$$

and

$$E(s_3^2) = \sigma^2{}_{(3)}$$

where

$$\frac{1}{\bar{K}_h''} = \frac{1}{M_0} \sum_{i=1}^{N} \sum_{j}^{M_i} \frac{1}{K_{ij}}. \qquad (40)$$

From (38)–(40) it may be possible to estimate the values of $\sigma^2_{(3)}$, $\sigma^2_{(2)}$ and $\sigma''^2_{(1)}$.

4.1. It is indicated that with the help of the analysis of variance it should be possible to estimate the variance components in multi-stage sampling although their definition and actual form will depend upon the system of selection probability at different stages. In the present paper, the results have been obtained only up to three-stage sampling design, which we generally come across in this country but the method will, however, be applicable for any stage of sampling although calculation of the expectation of mean squares as obtained in the analysis of variance table will be tedious work for multi-stage sampling. Nothing is known about the precision of these estimates. One way of judging the precision of these estimates will be by calculating their sampling errors. In this connection, mention may be made of the interesting work done by Tukey (1956–57). Although the papers of Tukey give general method of calculating the sampling variance of estimated variance components, its application to the present problem will not be straightforward. Author is studying the subject and the result will be published in subsequent paper.

5.1. SUMMARY

It has been shown how the method of analysis of variance can be suitably used for studying variance components in multi-stage-sampling from finite population. It has also been shown how a fairly good approximation of these estimates can be obtained without undertaking detailed computation.

REFERENCES

1. Cochran, W. G. .. "The use of analysis of variance in enumeration by sampling," *J. Am. Stat. Assoc.*, 1939, **34**, 492.

2. Sukhatme, P. V. .. "Sampling theory of surveys with applications," *I.S.A.S.*, 1954, p. 335.

3. Yates, F. .. *Sampling Methods for Censuses and Surveys*, Charles Griffin & Co., London, 1949.

4. Crump, S. L. .. "Present status of variance component analysis," *Biometrics*, 1951, **7**, 1.

5. ———— .. "Estimation of variance component in analysis of variance," *Ibid.*, 1946, **2**, 7.

6. Henderson, C. R. .. "Estimation of Variance and covariance components," *Biometrics*, 1953, **9**, 226.

7. Tukey, J. W. .. "Variance of variance component—I," *A.M.S.*, 1956, **27**, 722.

8. ———— .. "Variance of variance component—II," *Ibid.*, 1957, **28**, 43.

9. Kitagawa .. "Some contributions to the design of sample surveys," *Sankhya*, **17** (1), 1.