Chapter II.

One Dimensional Systematic Sampling.

Let us first define systematic sampling for one dimensional field. Let it consist of \( n \) sets of \( K \) elements each, out of which we are to take a sample of size \( m \) in order to estimate the population mean. The elements are arranged in a particular order in the field. Let \( x_{ij} \), \( i = 1, 2, \ldots, n; j = 1, 2, \ldots, K \) be the values of the elements arranged as \( x_{11}, x_{12}, \ldots, x_{1K}, x_{21}, x_{22}, \ldots, x_{2K}, \ldots, x_{n1}, x_{n2}, \ldots, x_{nK} \).

Now, a sample of size \( m \) from these can be taken in three ways. Firstly, we can take \( m \) elements at random out of \( nK \) elements. Secondly, we can consider \( n \) strata, the first stratum containing \( x_{11}, x_{12}, \ldots, x_{1K} \), the second stratum \( x_{21}, x_{22}, \ldots, x_{2K} \), and so on, and take a random sample of size \( m \) from each. Thirdly, we can choose \( nK \) elements at random from the first stratum and then take these as a sample and every \( K \)th one starting from each of these chosen elements. The first one we shall denote as a random sample, the second stratified sample and the third systematic sample.

It can easily be proved that in each case the sample mean is the best unbiased linear estimate of the population mean; and in large samples it will tend to be normally distributed. But its variance will be different for different cases of sampling and considering that efficiency is inversely proportional to variance we can compare different sampling techniques.

For comparing the relative efficiency, we shall consider the universe to be homogeneous in the sense that \( E(x_i') = \mu \), and these have got common variance. We shall also assume that correlation between any two elements is a function of the gap between them. Let \( \rho(u) \) be the correlation between two elements separated by distance \( u \), i.e., the correlation obtained by pairing two elements \( u \)-distance apart, which may have certain fluctuations due to the finiteness of the universe.

If this universe be infinite, such an effect will be absent. If, however, the universe be finite, it may be considered to be a sample from an infinite population, as has been done by Cochran (1946) where \( E(x_i') = \mu \), \( \rho(u) = \sigma^2 \) and covariance between two points \( u \)-distance apart is \( \rho(u) \sigma^2 \) for the infinite population. This \( \rho(u) \)
for different values of \( V \) will form a correlogram. As in Cochran(1946) the expectations of the variances of the ground sample mean, as averaged over all finite populations from the infinite population, under random, stratified and systematic sampling will be denoted by \( \sigma^2_x \), \( \sigma^2_{x|t} \) and \( \sigma^2_{x|s} \) respectively.

We can also find the variance of the sample mean under random, stratified and systematic sampling without referring to the infinite population and denote them by \( \sigma^2_{x|t} \), \( \sigma^2_{x|s} \) and \( \sigma^2_{x|s} \) respectively. The method here is well known for the case of random and stratified cases and here the variances come out as:

\[
\sigma^2_{x|t} = \frac{(k-1)}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2, \quad \sigma^2_{x|s} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2.
\]

In the case of systematic sampling, Madow(1944) have found out the expression. But the following method may be used as an alternative that also be helpful in the general concept of systematic sampling propounded by the author in chap. 6. Systematic sampling as defined in this chapter maybe conceived of as forming a number of sets of the elements in the population and taking a random sample out of these sets. We should of course, define a particular \( \pi \) of these sets. Thus in the present case the \( j^{th} \) set \( S_j \) contains the elements \( x_{ij}, x_{ij} + x_{ij} + \ldots \), and we can define the \( \pi \) of the \( j^{th} \) set \( \pi(S_j) = \frac{1}{n} (x_{ij} + x_{ij} + \ldots) \). Then we are to take a random sample of size \( n \) from \( S_1, S_2, \ldots, S_K \) (vide Madow, 1944). Let \( S_{i1}, S_{i2}, \ldots, S_{in} \) be included in the sample. Then the sample mean

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2
\]

and we have, \( \pi(S_j) = \frac{1}{n} (x_{ij} + x_{ij} + \ldots) \). Thus the sample mean is the unbiased estimate of the population mean. In chapter 6, it will also be proved to be the best unbiased linear estimate:

\[
\bar{x} = \frac{(k-1)}{(k-1)} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2
\]

and

\[
\sigma^2_x = \frac{(k-1)}{(k-1)} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2
\]

where

\[
\sigma^2_{x|s} = \frac{(k-1)}{(k-1)} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2 + \frac{2}{n} \sum_{u=1}^{n-1} \hat{p}(u) \hat{p}(u)\]
$\beta$ is as defined above. This $\beta(u,k)$ is identical with $\beta_{u,k}$ defined by Maddox (1944).

From (2.1) and (2.2), the variances for $y=1$, comes out as:

$$\hat{\sigma}_y^2 = \frac{\hat{\sigma}_y(\kappa-1)}{(n_k-1)}$$

$$\hat{\sigma}_k^2 = \frac{1}{n_k^2} \sum_{l=1}^{n_k} \hat{\sigma}_y^2,$$

and

$$\hat{\sigma}_\nu^2 = \frac{\hat{\sigma}_\nu}{n} \left[ 1 + \frac{1}{n_k} \sum_{u=1}^{n_k} \hat{\rho}(u,k) \right]$$

Hence, $\hat{\sigma}_y^2 : \hat{\sigma}_k^2 : \hat{\sigma}_\nu^2$ is independent of $\nu$, because for a particular $\nu=1$ the variances for $\nu=1$ are multiplied by the same quantity $\frac{(k-1)}{(n_k-1)}$ for the three cases $\hat{\sigma}_y^2$, $\hat{\sigma}_k^2$, and $\hat{\sigma}_\nu^2$; or in other words, the relative efficiencies $\frac{\hat{\sigma}_y^2}{\hat{\sigma}_k^2} : \frac{\hat{\sigma}_y^2}{\hat{\sigma}_\nu^2}$ of the three types of sampling are independent of $\nu$, and hence the relative efficiencies of the expected variances, $\frac{\alpha_y}{\alpha_k}$, $\frac{\alpha_y}{\alpha_\nu}$, $\frac{\alpha_y}{\alpha_\nu}$ also are independent of $\nu$. This result will be illustrated in course of further investigations.

Before finding out $\alpha_y^2$, $\alpha_k^2$ and $\alpha_\nu^2$ we should establish some lemma:

**Lemma 2**: In a plot containing a row of $N$ cells with $X_i$ as the stochastic variate corresponding to the $i^{th}$ cell:

$$E \left\{ \sum_{l=1}^{N} (X_l - \bar{X})^2 \right\} = (N-1) \alpha_y \left\{ 1 - \Phi(N) \right\}$$

$$+ \sum_{l=1}^{N} (x_i - \bar{X})^2$$

where $\bar{X} = \frac{1}{N} \sum_{l=1}^{N} X_l$, $E(X_l) = x_i$, $\bar{X} = \frac{1}{N} \sum_{l=1}^{N} X_i$, and

$$\Phi(N) = \frac{2}{N(N-1)} \sum_{u=1}^{N} (N-u) \Phi(u).$$

**Proof**: Because, $\sum_{l=1}^{N} (X_l - \bar{X})^2 = \frac{1}{N} \sum_{l=1}^{N} (X_l - \bar{X}^2 + \bar{X}^2)$

$$= \frac{1}{N} \sum_{l=1}^{N} (X_l - \bar{X}^2)$$

$$+ \bar{X}^2$$

$$= \frac{1}{N} \sum_{l=1}^{N} E(X_l^2) - (E(X_l))^2 + (E(X_l))^2$$

$$= \frac{1}{N} \sum_{l=1}^{N} E(X_l^2) - E(X_l)^2$$

$$= \frac{1}{N} \sum_{l=1}^{N} E(X_l^2) - \left( \frac{1}{N} \sum_{l=1}^{N} X_l \right)^2$$

$$= \frac{1}{N} \sum_{l=1}^{N} E(X_l^2) - \left( \frac{1}{N} \sum_{l=1}^{N} X_l \right)^2$$

$$= \sum_{l=1}^{N} (x_i - \bar{X})^2.$$
To find the expected variance $\sigma^2_{\bar{x}}$ of $\bar{x}$ under systematic sampling, let $x_{1}, x_{2}, \ldots, x_{K}$ be the sample from the first stratum. Then the samples from other strata are automatically fixed up. Let $\bar{x}_{s}$ be the sample mean and $\bar{x}$ the population mean. Then $\bar{x}_{s} = \frac{1}{n} \sum_{i=1}^{n} x_{i}$.

For stratified sampling, from (2.1):

$$E \left[ \left( \bar{x}_{s} - \bar{x} \right)^2 \right] = \frac{(K-1)}{K} \frac{1}{n \cdot \sum_{i=1}^{K} \sigma^2_{x_i}}.$$

But, $E \left[ \bar{x}_{s}^2 \right] = \frac{(K-1)}{K} \frac{1}{n} \sum_{i=1}^{K} \frac{(K-1)}{(K-i)} \frac{1}{n} \sum_{u=1}^{K} x_{u, i}$.

Let $\delta = \frac{n}{n-1} \sum_{u=1}^{K} \frac{(K-1)}{(K-u)} \frac{1}{n} \sum_{u=1}^{K} x_{u, i}$.

To find the expected variance $\sigma^2_{\bar{x}}$ of $\bar{x}$ under systematic sampling, let $x_{11}, x_{12}, \ldots, x_{1n}$ be the sample from the first stratum. Then the samples from other strata are automatically fixed up. Let $\bar{x}_{s}$ be the sample mean and $\bar{x}$ the population mean. Then $\bar{x}_{s} = \frac{1}{n} \sum_{i=1}^{n} x_{i}$.

Let $\frac{1}{n} \sum_{u=1}^{K} x_{u, j} = \bar{x}_{s}$.
Now, we shall take the average of (2.8) over \( K_c_j \) systematic samples. Let \( x_{ij} \) be a random point from the first stratum and \( \bar{x}_{j} \) the mean of the corresponding systematic sample, then for summation of (2.8) over \( K_c_j \) systematic samples any \( \bar{x}_{j} \) will occur \( K_{c_j-1} \) times, and \( \bar{x}_{j} \) and \( \bar{x}_{j+1} \) will occur together \( K_{c_j-1} \) times; in the summation of expression (2.8) over \( K_c_j \) samples \((\bar{x}_{j} - \bar{x})\) will occur \( K_{c_j-1} \) times and \((\bar{x}_{j} - \bar{x})(\bar{x}_{j+1} - \bar{x})\) occur \( K_{c_j-1} \). But

\[
\sum_{j=1}^{K_c_j} (\bar{x}_{j} - \bar{x})(\bar{x}_{j} - \bar{x}) = 0,
\]

\[
\sum_{j=1}^{K_c_j} (\bar{x}_{j} - \bar{x}) = -(\bar{x}_{j} - \bar{x})^2,
\]

so in the summation, \((\bar{x}_{j} - \bar{x})^2\) occurs \( K_{c_j-1} \) times. Hence, the average of (2.8) is:

\[
\bar{\bar{w}} = \frac{1}{K_c_j} \sum_{j=1}^{K_c_j} (\bar{x}_{j} - \bar{x}) = \frac{(K_c_j - 1)}{K_c_j} \bar{\bar{w}}_{\text{within}} - \frac{(K_c_j - 1)}{K_c_j} \bar{\bar{w}}_{\text{between}}.
\]

which is the same expression as has otherwise been obtained in (2.9).

The average may be denoted by \( \bar{\bar{w}}_{\text{s}} \) as in (2.9).

Now, apart from the expression in (2.8), \( \bar{\bar{w}}_{\text{s}} \) may also be written as:

\[
\bar{\bar{w}}_{\text{within}} = \frac{(K_c_j - 1)}{K_c_j} \bar{\bar{w}}_{\text{within}} - \frac{(K_c_j - 1)}{K_c_j} \bar{\bar{w}}_{\text{between}}.
\]

It is to note here that the within and between sum of squares are here for \( \lambda = 1 \).

Again putting \( n = n \) and replacing \( p(u) \) by \( p(u|x) \) in (2.5),

\[
E(\text{s.s. within systematic sample}) = K(n-1) \omega^2 \left\{ 1 - \frac{1}{n} \sum_{u=1}^{n} p(u|x) \right\},
\]

and \( E \left\{ \sum_{j=1}^{n} \sum_{j=1}^{K_c_j} (x_{ij} - \bar{x})^2 \right\} \) has already been obtained in deriving (6).
From (8) and (9) it is evident that the relative efficiencies of random, stratified and systematic samples are independent of \( \gamma \), a result that has already been proved.

From (8) and (9) it follows that \( \sigma_f^2 > \sigma_s^2 \), i.e., stratified sampling is more efficient than random sampling, if

\[ \phi(k) > \phi(n_k) \]

For \( \rho(u) \) being a monotonous function of \( u \), Cochran (1946) has proved the following result:

\[
\phi(k) - \phi(k+1) = 2 \sum_{u=1}^{k} \frac{(k-u)}{k(k+1)} \rho(u) - 2 \sum_{u=1}^{k} \frac{(k+1-u)}{k(k+1)} \rho(u)
\]

\[ = 2 \sum_{u=1}^{k} \rho(u) \left( \frac{1}{k} \left( k - \frac{k-u}{k+1} \right) \right)
\]

\[ = \frac{2}{k(k+1)} \sum_{u=1}^{k} \rho(u) (k+1 - 2u)
\]

\[ = -S_k \Delta \rho(u) - S_k \Delta \rho(u) - \cdots - S_k \Delta \rho(u) + S_k \rho(k)
\]

where, \( \Delta \rho(u) = \rho(u+1) - \rho(u) \) and \( S_k = \sum_{i=1}^{k} \frac{2(k+1-i)}{k(k+1)} = \frac{2(k+1-k)}{k(k+1)} \).

\[ \therefore \ S_k = 0 \]

and \( \phi(k) - \phi(k+1) = -S_k \Delta \rho(u) \).

But \( S_k > 0 \), because \( 1 \leq u \leq k \).

Theorem. Under \( \Delta \rho(u) \leq 0 \), \( \sigma_f^2 > \sigma_s^2 \) (Cochran, 1946), irrespective of whether the space-correlation \( \rho(u) \) is positive or negative. But for space-correlation being zero or constant \( \sigma_f^2 = \sigma_s^2 \), i.e., stratified sampling is as efficient as random sampling. Many other results giving the set of sufficient conditions for which \( \sigma_f^2 > \sigma_s^2 \) can be arrived at by considering \( \phi(k) \) and \( \phi(n_k) \).
For comparing stratified sampling with systematic, we shall consider
the case for \( \frac{1}{2} = 1 \).

From (2-1): 
\[
\hat{\sigma}_{st}^2 = \frac{1}{K} \sum_{j=1}^{K} \left( \frac{1}{n_j} \sum_{i,j} (x_{ij} - \bar{x})^2 \right) = \frac{1}{K} \sum_{j=1}^{K} \left( \frac{1}{n_j} \sum_{i,j} (x_{ij} - \bar{x}_j)^2 \right)
\]

where \( \bar{x}_j = \frac{1}{n_j} \sum_{i,j} x_{ij} \).

Again, 
\[
\hat{\sigma}_{st}^2 = \frac{1}{mnK} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2 - \frac{1}{mnK} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 - \frac{1}{mnK} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

\[
= \frac{2mnK - 2mnK}{mnK} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

Because, the sum of the product term vanishes and
\[
\hat{\sigma}_{st}^2 = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

\[
= \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

Hence, from (2-11):
\[
\hat{\sigma}_{st}^2 = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

\[
= \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

\[
= \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2
\]

Hence, from (2-12) and (2-11):
\[
\hat{\sigma}_{st}^2 - \hat{\sigma}_{st}^2 = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{K} \left( (x_{ij} - \bar{x}_j)^2 - \frac{1}{mnK} \sum_{i,j} (x_{ij} - \bar{x}_j)^2 \right)
\]

Plugging \( S_{st}(i,j) \) = \( \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 \) and \( S_{st}(i,j) \) = \( \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 \),
\[
S_{st}(i,j) = \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x})^2 - (x_{ij} - \bar{x}_j)^2 + (\bar{x}_j - \bar{x})^2
\]

\[
 = \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 + (\bar{x}_j - \bar{x})^2
\]

\[
= \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j)^2 + (\bar{x}_j - \bar{x})^2
\]

\[
= 2 \cdot \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \bar{x}_j) (x_{ij} - \bar{x}_j)
\]
and \( s_{at}(\xi, \tau) = \frac{1}{2K} \left\{ \sum_{j=1}^{K} \left( (x_{ij} - \overline{x_i})^2 + (x_{ij} - \overline{x_j}) - (x_{ij} - \overline{x_{ij}}) \right) \right\} \)

\[ = \frac{1}{K^2} \sum_{j,j'=1}^{K} \left( (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \]

\[ = \overline{y}^2 + \overline{y}_p^2 + (\overline{y}_i - \overline{y}_j)^2 \]

\[ s_{at}(\xi, \tau) = \frac{1}{K^2} \sum_{j=1}^{K} \left( (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \]

whence it follows that \( \sigma_a^2 > \sigma_b^2 \), i.e., systematic sampling is more efficient than stratified sampling if the average sum of product between the associated elements of different pairs of strata be negative; by associated elements we mean the elements that should come together in systematic sampling (vide, chap. 6).

Now referring (112) to the infinite population:

\[ \sigma_a^2 - \sigma_b^2 = \frac{1}{K^2} \sum_{j=1}^{K} \left( \sum_{i=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]

\[ = \frac{1}{K^2} \left\{ \sum_{i=1}^{K} \left( \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 - \frac{1}{K} \sum_{j=1}^{K} (x_{ij} - \overline{x_i})^2 + (\overline{x_i} - \overline{x_j}) \right) \right\} \]
Hence, the following theorem due to Coohran (1946):

**Theorem (2.3)** For all infinite populations in which the correlogram is concave upwards, i.e., $\gamma^2 p(u) > 0$, systematic sampling is more efficient than stratified sampling. The restrictions in this theorem should, of course, be noted.

From (2.14), another important result follows:

From (2.14), another important result follows:

Theorem (2.3) For all infinite populations in which the correlogram is concave upwards, i.e., $\gamma^2 p(u) > 0$, systematic sampling is more efficient than stratified sampling. The restrictions in this theorem should, of course, be noted.

From (2.14), another important result follows:

Theorem (2.3) For all infinite populations in which the correlogram is concave upwards, i.e., $\gamma^2 p(u) > 0$, systematic sampling is more efficient than stratified sampling. The restrictions in this theorem should, of course, be noted.
From (13), many other results can be obtained for different special types of oscillating and other correlograms. But no general result can be suggested.

From theorems stated above, we can compare systematic with random samples. But for populations not satisfying the conditions of these theorems, it is better to use equations (16) and (17) and (18), when the strata have got different means and variances, and their modified forms with the help of Lemma (1) when these are different for different strata.

Examples illustrating the results in this chapter can be had from Maddows (1944) and Cochran (1946).