# ROTATION SAMPLING

By Albert Ross Eckler[1]

*Princeton University*

**1. Summary.** This paper shows how to find minimum-variance estimates of the mean $\alpha(t_i)$ of a time-dependent population, assuming that one is restricted to the class of linear unbiased estimates. Each minimum-variance estimate is based on a specified sample pattern (a set of sample values drawn from the population at one or more distinct times). Let the random variable $X_{ij}$ denote the value of element $j$ of the population at time $t_i$. The correlation between $X_{ik}$ and $X_{jk}$ is assumed to be $\rho^{|i-j|}$; the correlation between $X_{ij}$ and $X_{ik}$ is assumed to be zero; the variance of $X_{ij}$ is assumed to be $\sigma^2$ independent of time. Iterative methods are developed; the estimate of the population mean $\alpha(t_{i-1})$ is used in determining the population mean $\alpha(t_i)$.

The paper discusses two important methods of sampling: in one-level rotation sampling, the statistician can add to the sample pattern only sample values that have been drawn from the population at the current time; in two-level (and higher-level) rotation sampling, the statistician can add earlier sample values as well as current ones to the pattern. Schematic sample patterns associated with these two methods are illustrated in (3.1) and (4.1).

The optimum structure of a sample pattern is considered from two viewpoints: the variance of a pattern consisting of $n$ sample values drawn at each time $t_i$ is minimized; the number of sample values drawn at time $t_i$ is minimized while the variance of the minimum-variance estimate is held constant.

Finally, the estimation problem is generalized to include minimum-variance estimates of linear functions of two or more population means at different times.

In order to maintain continuity, this paper presents published results along with new results; the latter are summarized below. The paper clarifies Patterson's fundamental method for finding minimum-variance linear unbiased estimates (Sections 2, 3) and extends his methods to two-level and three-level rotation sampling (Sections 4, 6, 8, 10, 13). The paper compares three methods of rotation sampling on a cost basis (Section 11) and shows how the one-level rotation sampling estimate of greatest practical interest can be derived from the two-level estimate (Sections 5, 14). Finally, the paper extends Cochran's work in determining optimum patterns for the one-level rotation sampling estimate (Section 9).

**2. Introduction to rotation sampling.** In survey sampling, the statistician sometimes must estimate at regular intervals of time a population parameter

Received September 2, 1954.

[1] Now a member of technical staff at Bell Telephone Laboratories, Whippany, New Jersey.

which varies with time. If there exists a relationship between the value of an element in the population at time $t$ and the changed value of the same element at the succeeding time $t + \Delta t$, then it is possible to use the information contained in earlier samples to improve the current estimate of the population parameter. In order to use the earlier sample information, one must carry out the sampling in such a way that the two samples drawn at successive times $t$ and $t + \Delta t$ have some elements in common.

The name "rotation sampling" (suggested by Wilks) refers to the process of eliminating some of the old elements from the sample and adding new elements to the sample each time a new sample is drawn. This method of sampling is also called sampling on successive occasions with partial replacement of units (Patterson, Yates) and sampling for a time series (Hansen, Hurwitz and Madow). Double sampling can be regarded mathematically as rotation sampling involving a present sample and one overlapping earlier sample.

We assume that we have a population $\pi$ and a set of times $(t_1, t_2, \cdots, t_m)$. Each element of the population has a set of $m$ values associated with it, one for each time $t_i$. A sample pattern $P$ consists of a set of sample values $x_{ij}$, where $i$ identifies the time $t_i$ when the element was sampled, and $j$ identifies the population element. The sample pattern $P$ can be visualized as an incomplete matrix with $m$ rows and the number of columns equal to the number of distinct elements with values represented in $P$. More definite sample patterns are discussed later.

We assume that the population $\pi$ has an infinite number of elements (eliminating any correlation between the sample values $x_{ij}$ and $x_{ik}$). Let $X_i$ be the random variable representing the population values at time $t_i$. We assume that $E(X_i) = \alpha(t_i)$ and that $\text{var}(X_i) = \sigma^2$ independent of the time. We specify the rest of the second moments of the joint distribution of $(X_1, X_2, \cdots, X_m)$ by means of the exponential correlation assumption: the correlation $\rho(X_i, X_j)$ is equal to $\rho^{|i-j|}$. This assumption implies that all partial correlation coefficients $\rho_{ij \cdot s}$ are zero if $i < s < j$ or $i > s > j$.

We restrict ourselves to estimating the mean of the population $\pi$ at a given time $t_i$, or more generally to estimating a linear combination of the means at several different times (such as $\alpha = c_1\alpha(t_1) + c_2\alpha(t_2) + c_3\alpha(t_3)$). The difference between two successive means is the linear function of greatest practical interest. We restrict ourselves to unbiased linear estimates $L(\alpha) = \sum w_{ij}x_{ij}$ of the population mean; the summation is taken over the values in $P$. Throughout the paper, we use the term unbiased in a stronger sense than usual. We require not only that $E(L(\alpha)) = \alpha$, but also that $E(\sum_j w_{ij}x_{ij}) = c_i\alpha(t_i)$ for all $i$.

Our goal is to determine the minimum-variance estimate of the population mean in the class of linear unbiased estimates based on the sample values in a specified sample pattern. We denote a minimum-variance estimate by the symbol $M(\alpha)$.

Patterson [5] derives a necessary and sufficient condition for a linear unbiased estimate to be a minimum-variance estimate. In view of the importance of his result in deriving minimum-variance estimates, we state it as a theorem but omit the proof.

THEOREM 1. *Assume that we have a sample pattern $P$ of values drawn at times $t_1, t_2, \cdots, t_m$ from a population $\pi$. Assume that the joint distribution of $(X_1, X_2, \cdots, X_m)$ has finite first and second moments. Let $c_1, c_2, \cdots, c_m$ be $m$ constants (not all zero), and let $L(\alpha)$ be a linear unbiased estimate of*

$$\alpha = c_1\alpha(t_1) + c_2\alpha(t_2) + \cdots + c_m\alpha(t_m).$$

*$L(\alpha)$ is the minimum-variance estimate $M(\alpha)$ if and only if* $\text{cov}(x_{ij}, L(\alpha)) = k_{i\alpha}$ *for all combinations of $i$ and $j$ in $P$.* The theorem does not have to be restricted to an infinite population size, a constant variance $\sigma^2$, or the exponential correlation assumption.

Two corollaries of Theorem 1 are frequently useful. The proofs are quite simple so they are omitted.

COROLLARY 1.1. *Let $\bar{x}_i$ be a linear unbiased estimate of $\alpha(t_i)$ based on the sample pattern $P$, and let the assumptions of Theorem 1 hold. Then* $\text{cov}(\bar{x}_i, M(\alpha)) = k_{i\alpha}$, *for $1 \leqq i \leqq m$.*

COROLLARY 1.2. *Let $M_i$ be the minimum-variance estimate of $\alpha(t_i)$ in the class of linear unbiased estimates based on the sample pattern $P$, and let the assumptions of Theorem 1 hold. Then* $\text{var}(M_i) = \text{cov}(x_{ij}, M_i)$, *for $1 \leqq i \leqq m$.* Corollary 1.2 is useful in calculating the variance of complicated minimum-variance estimates.

Frequently many covariance-conditions must be checked in order to determine whether or not an estimate $L(\alpha)$ is minimum-variance. In order to reduce this number, we derive a simplified form of the unbiased linear estimate that still contains the minimum-variance estimate. If the number of different elements in the sample pattern is finite, it is evident that the pattern can be split up into a certain minimum number of subpatterns, each one of which is rectangular (a complete matrix of $x_{ij}$ values).

THEOREM 2. *Assume that the conditions of Theorem 1 hold. Assume that the sample pattern $P$ has been broken up into a finite number of rectangular subpatterns; let us consider subpattern $P_i$ which forms a complete matrix of $x_{ij}$ values with $r$ rows and $c$ columns. Then the $c$ weights $w_{ij}$ associated with the values of $x_{ij}$ in any one of the $r$ rows are all equal: $w_{11} = w_{12} = \cdots = w_{1c}$, $w_{21} = w_{22} = \cdots = w_{2c}$, $\cdots$, $w_{r1} = w_{r2} = \cdots = w_{rc}$.*

PROOF. According to Theorem 1, the covariance-condition must hold for this subpattern. Consider the identity $\text{cov}(x_{11}, M(\alpha)) - \text{cov}(x_{12}, M(\alpha)) = 0$. Expanding this identity, we obtain an expression of the form

$$a_1(w_{11} - w_{12}) + a_2(w_{21} - w_{22}) + \cdots + a_r(w_{r1} - w_{r2}) = 0.$$

The coefficients $a_i$ depend only on the correlation model and the population size. In order that this expression be identically zero, $w_{i1}$ must be equal to $w_{i2}$, for $1 \leqq i \leqq r$. The theorem is proved by iterating this argument $c - 2$ times. As a consequence of Theorem 2, we can express the minimum-variance estimate $M(\alpha)$ as a linear combination of *means* of sample values; each mean value is formed from the sample values in a row of a rectangular subpattern. In the rest of this paper, we regard the mean value estimate as the canonical form of the linear unbiased estimate.

**3. One-level rotation sampling.** In this section we summarize for future reference a basic result from Patterson [5]: the determination of $M'_i$, the minimum-variance linear unbiased estimate of $\alpha(t_i)$ based on a sample pattern of the type

$$
(3.1) \qquad \begin{array}{ll}
\begin{array}{|l}
xxxxxxx \qquad\qquad \\
\quad xxxxxxx \\
\qquad\quad \cdot \\
\\
\qquad\qquad xxxxxxx \\
\hline
\end{array} & \begin{array}{l} t_1 \\ t_2 \\ \cdot \\ \cdot \\ \cdot \\ t_{i-1} \end{array} \\
\qquad\qquad\quad \underbrace{xx}_{3}\underbrace{xxx}_{2}\underbrace{xx}_{1} & t_i \, .
\end{array}
$$

More precisely, the sample pattern is assumed to have $n$ sample values in it at each time $t_k$, $1 \leqq k \leqq i$. $(1 - \mu)n$ of the elements in the sample at time $t_{k-1}$ are retained in the sample drawn at time $t_k$, and the remaining $\mu n$ elements are replaced with the same number of new ones. The lines indicate how the sample pattern is built up; at time $t_k$ the $k$th row of sample values is added to the $(k - 1)$ rows of earlier values. Since each enlargement of the pattern consists of a set of sample values associated with a single time, we call this one-level rotation sampling on the above pattern.

Patterson shows that the minimum-variance estimate $M'_i$ of $\alpha(t_i)$ based on pattern (3.1) can be written in the iterative form

$$
(3.2) \quad M'_i = A_i\bar{x}_{i,1} + (1 - A_i)\bar{x}_{i,2} - (B_i + C_i)\bar{x}_{i-1,2} + C_i\bar{x}_{i-1,3} + B_iM'_{i-1}
$$

where $A_i$, $B_i$ and $C_i$ are unknown coefficients to be determined, and $M'_{i-1}$ is the minimum-variance linear unbiased estimate of $\alpha(t_{i-1})$ based on the pattern above the line in (3.1). The first subscript of a sample average denotes the time that the sample values were drawn, and the second subscript identifies the elements represented in the sample average (see bottom of (3.1)). The iteration reflects the way in which the sample pattern is built up row by row.

Patterson shows that the iterative estimate is minimum-variance by means of the sufficiency of the covariance-condition (Theorem 1); the four possible covariance-conditions (determined by Theorem 2) can be reduced to three independent equations which can be solved for $A_i$, $B_i$ and $C_i$ in terms of earlier coefficients:

$$
(3.3) \qquad \begin{gathered}
B_i = \rho(1 - A_i), \qquad C_i = 0, \\
1 - A_i = \frac{1 - \mu}{1 - (2\mu - 1)\rho^2 - (1 - \mu)\rho^2(1 - A_{i-1})}.
\end{gathered}
$$

Given $A_1$ (equal to $\mu$), any minimum-variance estimate $M'_i$ can be determined by a repeated application of equations (3.3). The variance of $M'_i$ is determined with the aid of Corollary 1.2,

$$
(3.4) \qquad \operatorname{var}(M'_i) = \operatorname{cov}(M'_i, \bar{x}_{i,1}) = \frac{\sigma^2}{\mu n} A_i \, .
$$

It is easy to show that the sequence of $A_i$ converges. $\text{Var}(M'_{i-1})$ is greater than or equal to $\text{var}(M'_i)$ because the sample pattern associated with $M'_{i-1}$ is contained in the pattern associated with $M'_i$. Since the sequence of variances is bounded from below by zero, we conclude that the sequence of variances (and therefore the sequence of $A_i$) converges. The limiting value is

$$\text{(3.5)} \quad \lim_{i\to\infty} A_i = A = \frac{-(1-\rho^2) + \sqrt{(1-\rho^2)(1-\rho^2(1-4\mu(1-\mu)))}}{2(1-\mu)\rho^2}.$$

Most of the above formulas simplify considerably when $\mu$ is equal to one-half, which will subsequently be shown to be the case of greatest practical interest. For example,

$$\text{(3.6)} \qquad A = \frac{-(1-\rho^2) + \sqrt{1-\rho^2}}{\rho^2}.$$

We tabulate $A_i$ for several values of $\rho$ in Table 3.1.

TABLE 3.1

*Values of the weights $A_i$ as a function of $\rho$ for a replacement rate $\mu$ of one-half (one-level sampling)*

| | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ | $A_6$ | $A_7$ | $A_8$ | $A_9$ | $A_{10}$ | $A$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | .5000 | | | | | | | | | | .5000 |
| .1 | .5000 | .4987 | | | | | | | | | .4987 |
| .2 | .5000 | .4949 | | | | | | | | | .4949 |
| .3 | .5000 | .4885 | | | | | | | | | .4885 |
| .4 | .5000 | .4791 | .4783 | | | | | | | | .4781 |
| .5 | .5000 | .4667 | .4643 | | | | | | | | .4641 |
| .6 | .5000 | .4505 | .4451 | | | | | | | | .4444 |
| .7 | .5000 | .4302 | .4189 | .4169 | | | | | | | .4166 |
| .75 | .5000 | .4182 | .4022 | .3989 | | | | | | | .3980 |
| .8 | .5000 | .4048 | .3824 | .3768 | | | | | | | .3750 |
| .85 | .5000 | .3898 | .3586 | .3492 | .3463 | | | | | | .3450 |
| .9 | .5000 | .3730 | .3298 | .3137 | .3075 | .3051 | .3042 | | | | .3036 |
| .95 | .5000 | .3543 | .2944 | .2664 | .2526 | .2456 | .2420 | .2401 | .2391 | | .2380 |
| 1.00 | .5000 | .3333 | .2500 | .2000 | .1667 | .1427 | .1250 | .1111 | .1000 | .0909 | .0000 |

Variance of the minimum-variance estimate $= \text{var}(M') = \dfrac{2\sigma^2}{n} A$

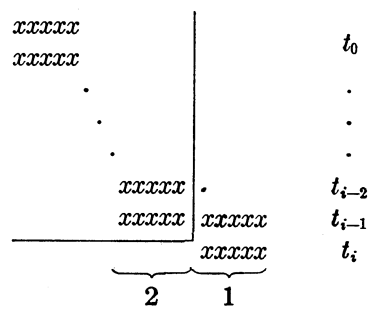
If we examine Table 3.1, we find that the reduction in variance achieved by rotation sampling is quite small for most values of $\rho$ and $A_i$; the variance of the estimate $M'_i$ is not reduced by one-half until the correlation reaches the very high value of .95. It seems quite likely that rotation sampling will be of most value when (a) the correlation is high, and (b) it is so difficult to draw a sample that the sample size must be kept as small as possible. If it costs no more to carry out rotation sampling than independent random sampling, then even a modest reduction of five to ten per cent in variance will be worthwhile.

One-level rotation sampling has been investigated by several authors. The most important work was done by Patterson [5]; Yates [6] presents many results from Patterson's paper without giving any derivations. Jessen [4] considered the special case in which $\mu$ was equal to one-half and the sample values were restricted to times $t_1$ and $t_2$.

**4. Two-level rotation sampling.** In the preceding section, we assumed that the sample pattern (3.1) was increased at time $t_k$ by adding sample values of the form $x_{kj}$. Relaxing this assumption, we now permit sample values of the forms $x_{(k-1)j}$ and $x_{kj}$ to be added to the sample pattern at time $t_k$. Clearly, this is possible only if we have records of the earlier values of the elements in the population. We call this two-level rotation sampling to emphasize that both present values and immediately preceding values can be added to the pattern; the generalization to three-level or multi-level sampling is obvious.

Since it is frequently cheaper in a sampling survey to obtain the sample values $x_{(k-1)j}$ and $x_{kj}$ simultaneously instead of at two separate times, we assume a sample pattern of the type

(4.1)




The lines indicate how the sample pattern is built up; at time $t_k$ a new set of $n$ elements is drawn from the population and the associated sample values for the times $t_k$ and $t_{k-1}$ are recorded. In rotation sampling, it is not sufficient to specify a sample pattern; the method by which the pattern is built up in time determines the most suitable iterative form of the minimum-variance estimate.

We now show by means of the covariance-condition (Theorem 1) that the minimum-variance estimate of $\alpha(t_i)$ based on the pattern (4.1) can be written in the iterative form

$$M_i'' = \bar{x}_{i,1} - a_i \bar{x}_{i-1,1} + a_i M_{i-1}'' \tag{4.2}$$

where $a_i$ is to be determined, $M_{i-1}''$ is the minimum-variance linear unbiased estimate based on the sample pattern above the line, and the subscripts of the sample averages are defined as in the preceding section.

Using Theorem 2, we conclude that the only possible covariance-condition is

$$\begin{cases} \operatorname{cov}(\bar{x}_{i-1,1}, M_i'') = \dfrac{\sigma^2}{n}(\rho - a_i), \\ \operatorname{cov}(\bar{x}_{i-1,2}, M_i'') = a_i \operatorname{cov}(\bar{x}_{i-1,2}, M_{i-1}'') = \dfrac{\sigma^2}{n} a_i(1 - a_{i-1}\rho). \end{cases}$$

In other words, we have the solution

$$a_i = \frac{\rho}{2 - a_{i-1}\rho}, \qquad a_1 = 0, \tag{4.3}$$

$$\text{var}\,(M_i'') = \frac{\sigma^2}{n}(1 - a_i\rho). \tag{4.4}$$

It can be shown by an argument similar to the one in the preceding section that the sequence of $a_i$ converges. The limiting value is

$$\lim_{i\to\infty} a_i = a = \frac{1 - \sqrt{1 - \rho^2}}{\rho}, \tag{4.5}$$

$$\lim_{i\to\infty} \text{var}\,(M_i') = \frac{\sigma^2}{n}\sqrt{1 - \rho^2}. \tag{4.6}$$

We tabulate $a_i$ for several values of $\rho$ in Table 4.1, and list the first five terms below.

$$a_1 = 0, \qquad a_4 = \frac{\rho(4 - \rho^2)}{4(2 - \rho^2)},$$

$$a_2 = \tfrac{1}{2}\rho,$$

$$a_3 = \frac{2\rho}{4 - \rho^2}, \qquad a_5 = \frac{4\rho(2 - \rho^2)}{16 - 12\rho^2 + \rho^4}.$$

TABLE 4.1

*Values of the weights $a_i$ as a function of $\rho$*

*(two-level sampling)*

| $\rho$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ | $a_8$ | $a_9$ | $a_{10}$ | $a_{11}$ | $a$ | $1 - a\rho$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | .0000 | | | | | | | | | | | .0000 | 1.0000 |
| .1 | .0000 | .0500 | .0501 | | | | | | | | | .0501 | .9950 |
| .2 | .0000 | .1000 | .1010 | | | | | | | | | .1010 | .9798 |
| .3 | .0000 | .1500 | .1535 | | | | | | | | | .1535 | .9539 |
| .4 | .0000 | .2000 | .2083 | | | | | | | | | .2087 | .9165 |
| .5 | .0000 | .2500 | .2667 | .2679 | | | | | | | | .2679 | .8660 |
| .6 | .0000 | .3000 | .3297 | .3329 | | | | | | | | .3333 | .8000 |
| .7 | .0000 | .3500 | .3989 | .4068 | .4081 | | | | | | | .4084 | .7141 |
| .75 | .0000 | .3750 | .4364 | .4484 | .4508 | | | | | | | .4513 | .6614 |
| .8 | .0000 | .4000 | .4762 | .4941 | .4985 | .4996 | | | | | | .5000 | .6000 |
| .85 | .0000 | .4250 | .5187 | .5452 | .5532 | .5556 | .5564 | | | | | .5567 | .5268 |
| .9 | .0000 | .4500 | .5643 | .6032 | .6176 | .6232 | .6254 | .6262 | | | | .6268 | .4359 |
| .95 | .0000 | .4750 | .6133 | .6703 | .6969 | .7100 | .7167 | .7202 | .7220 | .7229 | .7234 | .7239 | .3123 |
| 1.00 | .0000 | .5000 | .6667 | .7500 | .8000 | .8333 | .8573 | .8750 | .8889 | .9000 | .9091 | 1.0000 | .0000 |

Variance of the minimum-variance estimate = $\text{var}(M'') = \frac{\sigma^2}{n}(1 - a\rho)$

The two-level rotation sampling problem was first solved by Bershad [1]. Using straightforward minimization methods, he determined the unknown coefficients in the general (non-iterative) linear unbiased estimate of $\alpha(t_i)$. The key to his solution is a method of evaluating certain types of large determinants by means of continued fractions. A similar two-level sampling problem is discussed in [3].

**5. Relationship between one-level and two-level rotation sampling.** We preface this section with a lemma which is a generalization of the well-known method used to find the minimum-variance linear combination of two uncorrelated estimates of the same parameter. The proof is simple and therefore is omitted.

LEMMA 1. *We assume that we have a finite sample pattern $P$ which can be partitioned into two uncorrelated subpatterns $P_a$ and $P_b$, the first consisting of sample values all drawn at time $t_k$ or earlier, and the second consisting of sample values all drawn at time $t_k$ or later* ($1 \leqq k \leqq i$). *Let $M$, $M_a$, and $M_b$ denote the minimum-variance unbiased linear estimates of $\alpha(t_k)$ based on the patterns $P$, $P_a$ and $P_b$ respectively. Then* $M = (M_a \operatorname{var}(M_b) + M_b \operatorname{var}(M_a))/(\operatorname{var}(M_a) + \operatorname{var}(M_b))$, *and* $\operatorname{var}(M) = \operatorname{var}(M_a)\operatorname{var}(M_b)/(\operatorname{var}(M_a) + \operatorname{var}(M_b))$.

Using this lemma, we can easily show how the minimum-variance linear unbiased estimate $M_i'$ based on the sample pattern (3.1) with $\mu$ equal to one-half can be determined if we know the minimum-variance linear unbiased estimate $M_i''$ based on the two-level sample pattern (4.1). Consider the following partition of the one-level pattern with $\mu$ equal to one-half:

| | | |
|---|---|---|
| xxxxxxxx | | $t_1$ |
| . | | . |
| . | | . |
| . | | . |
| xxxxxxxx | | $t_{i-1}$ |
| xxxx | xxxx | $t_i$ . |

To the left of the vertical line, the pattern is a two-level sampling pattern with a sample size of $n/2$ instead of $n$ (the sample values drawn at time $t_0$ in the two-level pattern (4.1) have a weight of zero in the minimum-variance estimate and can be ignored). Therefore the minimum-variance estimate of $\alpha(t_i)$ based on the sample values to the left of the line is $M_i''$. The minimum-variance estimate of $\alpha(t_i)$ based on the $n/2$ sample values to the right of the line is the mean of these values. Applying Lemma 1, we find that

$$M_i' = \frac{1}{2 - a_i\rho} M_i'' + \frac{1 - a_i\rho}{2 - a_i\rho}\bar{x}, \qquad \operatorname{var}(M_i') = \frac{2\sigma^2}{n}\cdot\frac{1 - a_i\rho}{2 - a_i\rho}.$$

We compare the second equation with equation (3.4) and conclude that

$$A_i = \frac{1 - a_i\rho}{2 - a_i\rho} \quad \text{or} \quad a_i = \frac{1 - 2A_i}{\rho(1 - A_i)}. \tag{5.1}$$

We have solved the one-level problem by means of the two-level problem by expressing $A_i$ as a function of $a_i$. The identity (5.1) can also be proved by induction.

**6. Three-level rotation sampling.** We now assume that sample values of the form $x_{(k-2)j}$, $x_{(k-1)j}$ and $x_{kj}$ can be added to the sample pattern at time $t_k$. In analogy with the two-level pattern, we assume a sample pattern of the type

$$(6.1)\qquad \begin{array}{llll|ll}
xxxxx & & & & & t_{-1} \\
xxxxx & \cdot & & & & \cdot \\
xxxxx & & \cdot & & & \cdot \\
 & & \cdot & xxxxx & & \cdot \\
 & & & xxxxxxxxxx & & t_{i-3} \\
 & & & xxxxxxxxxx & xxxxx & t_{i-2} \\
 & & & \quad\;\; xxxxx & xxxxx & t_{i-1} \\
\hline
 & & & & xxxxx & t_i\,. \\
 & & 3 & 2 & 1 &
\end{array}$$

The lines indicate how the sampling pattern is built up by columns; at time $t_k$ a new set of $n$ elements is drawn from the population and the last three sample values are recorded. We let $M_i'''$ denote the minimum-variance estimate of $\alpha(t_i)$ based on pattern (6.1).

The problem of finding an iterative estimate which contains the minimum-variance estimate is twofold: (a) we must find an iterative estimate such that the number of unknown coefficients is equal to the number of independent co-variance-conditions, and (b) we must ascertain that the set of simultaneous equations has a non-trivial solution (the unknown coefficients must not all be zero). Condition (a) becomes non-trivial for the first time in three-level sampling. It can be shown that if we restrict ourselves to an iterative estimate involving sample averages and $M_{i-1}'''$, we always have one too many independent covari-ance-conditions.

Consider the estimate

$$(6.2)\qquad \begin{aligned} M_i''' &= \bar{x}_{i,1} - a_i\bar{x}_{i-1,1} + a_iM_{i-1}''' - b_i\bar{x}_{i-2,1} - (c_i + f_i)\bar{x}_{i-2,2} \\ &\quad + (b_i + c_i)\bar{x}_{i-2,3} - d_i\bar{x}_{i-3,2} - e_i\bar{x}_{i-3,3} + (d_i + e_i)\bar{x}_{i-3,4} + f_iM_{i-2}'''\,. \end{aligned}$$

This estimate contains both $M_{i-1}'''$ and $M_{i-2}'''$; a little consideration shows that condition (a) is now satisfied. In order to simplify the algebra in evaluating the unknown coefficients, we express equation (6.2) in an equivalent form by sub-stituting the term $a_i\bar{x}_{i-1,2}$ for $a_iM_{i-1}'''$.

According to Theorem 2, there are six independent covariance-conditions which can be expressed in terms of the six unknown coefficients. If we can find a non-trivial solution to these equations, we conclude (Theorem 1) that we have found the minimum-variance estimate of $\alpha(t_i)$ based on pattern (6.1). The co-variance-conditions are

$$(6.3)\qquad \begin{cases} \operatorname{cov}(\bar{x}_{i-4,5}, M_i''') = f_i \operatorname{cov}(\bar{x}_{i-4,5}, M_{i-2}''') \\[2ex] \operatorname{cov}(\bar{x}_{i-4,4}, M_i''') = f_i \operatorname{cov}(\bar{x}_{i-4,4}, M_{i-2}''') + \dfrac{\sigma^2}{n}(d_i + e_i)\rho \end{cases}$$

$$\begin{cases} \text{cov}(\bar{x}_{i-3,4}, M_i''') = f_i \, \text{cov}(\bar{x}_{i-3,4}, M_{i-2}''') + \frac{\sigma^2}{n}(d_i + e_i) & (6.4) \\ \text{cov}(\bar{x}_{i-3,3}, M_i''') = f_i \, \text{cov}(\bar{x}_{i-3,3}, M_{i-2}''') + \frac{\sigma^2}{n}[-e_i + (b_i + c_i)\rho] \\ \text{cov}(\bar{x}_{i-3,2}, M_i''') = \frac{\sigma^2}{n}[-d_i - (c_i + f_i)\rho + \rho^2 a_i] & (6.5) \end{cases}$$

$$\begin{cases} \text{cov}(\bar{x}_{i-2,3}, M_i''') = f_i \, \text{var}(M_{i-2}''') + \frac{\sigma^2}{n}[b_i + c_i - \rho e_i] & (6.6) \\ \text{cov}(\bar{x}_{i-2,2}, M_i''') = \frac{\sigma^2}{n}[-\rho d_i + \rho a_i - (c_i + f_i)] \\ \text{cov}(\bar{x}_{i-2,1}, M_i''') = \frac{\sigma^2}{n}[-b_i - \rho a_i + \rho^2] & (6.7) \end{cases}$$

$$(6.8) \quad \begin{cases} \text{cov}(\bar{x}_{i-1,2}, M_i''') = \frac{\sigma^2}{n}[-\rho^2 d_i - \rho(c_i + f_i) + a_i] \\ \text{cov}(\bar{x}_{i-1,1}, M_i''') = \frac{\sigma^2}{n}[\rho - a_i - b_i\rho] \end{cases}$$

If we compare equation (6.7) with (6.8), we find that by setting $a_i$ equal to $\rho/2$ we can reduce these two equations to one independent equation. Furthermore, we conclude from equation (6.3) that $d_i$ is equal to $-e_i$. If we make these substitutions, we have left four equations in four unknowns which can be solved by straightforward algebra. The only term which needs special attention is the cov $(\bar{x}_{i-3,3}, M_{i-2}''')$ in equation (6.5). Starting with the equation

$$\rho \, \text{cov}(\bar{x}_{i-3,3}, \quad M_{i-2}''' - \bar{x}_{i-2,3}) = \text{cov}(\bar{x}_{i-2,3}, \quad M_{i-2}''' - \bar{x}_{i-2,3})$$

we conclude that

$$\text{cov}(\bar{x}_{i-3,3}, M_{i-2}''') = \frac{1}{\rho}\text{var}(M_{i-2}''') + \frac{\sigma^2}{n}\left(\rho - \frac{1}{\rho}\right) = \frac{\sigma^2}{n}\left(\frac{\rho}{2} - b_{i-2}\rho\right).$$

The solution to the four equations is

$$(6.9) \quad \begin{aligned} b_i &= \frac{\rho^2[(3 + \rho^2) - 2b_{i-2}(1 - \rho^2)]}{2[(9 - \rho^2) - 2b_{i-2}(3 + \rho^2)]} \qquad \text{for } i \geqq 4, \\ f_i &= \frac{\rho^2 + 2b_i}{3 - 2b_{i-2}}, \\ c_i &= \frac{b_i(1 + \rho^2) - f_i}{1 - \rho^2}, \\ d_i &= -(b_i + c_i)\rho. \end{aligned}$$

As indicated by this solution, we regard $b_i$ as the fundamental variable; it is the only unknown coefficient to appear in the variance of $M_i'''$.

$$\text{var}(M_i''') = \frac{\sigma^2}{n}\left[1 - \frac{\rho^2}{2} - b_i \rho^2\right]. \tag{6.10}$$

The values $b_1$, $b_2$ and $b_3$ must be determined by an independent method. It is quite easy to evaluate these coefficients by a straightforward minimization of var $(M_1''')$ through var $(M_3''')$. The first five values are

$$b_1 = 0, \qquad b_4 = \frac{\rho^2(3 + \rho^2)}{2(9 - \rho^2)},$$
$$b_2 = 0,$$
$$b_3 = \frac{\rho^2}{6}, \qquad b_5 = \frac{\rho^2(9 + 2\rho^2 + \rho^4)}{2(9 + \rho^2)(3 - \rho^2)}.$$

The $b_i$ converge to the limiting value

$$\lim_{i \to \infty} b_i = b = \frac{(3 - \rho^2) - \sqrt{(1 - \rho^2)(9 - \rho^2)}}{4}$$

and the limiting variance is

$$\text{var}(M''') = \frac{\sigma^2}{n}\left[\frac{(1 - \rho^2)(4 - \rho^2)}{4} + \frac{\rho^2}{4}\sqrt{(1 - \rho^2)(9 - \rho^2)}\right]. \tag{6.11}$$

Unless $\rho$ is close to unity, var (M''') is not much less than $\sigma^2/n$.

The method used in solving the three-level rotation sampling problem can be extended without any conceptual difficulty to higher-level sampling. However, the smaller variance obtained by multi-level procedures over single-level procedures (discussed in a later section) is probably not worth the very great increase in algebraic complexity. For example, the four-level rotation sampling problem requires the solution of at least eight simultaneous linear equations with algebraic coefficients.

**7. Truncated patterns.** The one-level and multi-level rotation sampling estimates discussed in the previous sections were based on sample patterns that extended over any number of distinct time-levels. However, there are several practical reasons for truncating these patterns—that is, ignoring all sample values except those associated with the $N$ most recent times $t_i$, $t_{i-1}$, $\cdots$, $t_{i-N+1}$.

Most of the reduction in variance accomplished by rotation sampling is attributable to the sample values that were drawn most recently. For example, suppose we want the variance of the minimum-variance estimate based on a truncated pattern to be no more than ten per cent larger than the variance of the corresponding minimum-variance estimate based on an infinitely long pattern. If we are carrying out one-level sampling, and if $\rho$ is equal to .5, .7, .9 or .95, we should use a pattern with sample values restricted to 2, 3, 4 or 5 time-levels, respectively. These and other variance comparisons can be easily obtained from

Tables 3.1 and 4.1. It seems pointless to continue using older sample values which make virtually no contribution to the estimate.

A more important reason for using truncated patterns is the possibility that the exponential correlation model may describe the behavior of the parent population only locally; it may be adequate for sample values associated with nearby times, but not for sample values farther apart in time. For example, in economic populations with month-to-month correlation $\rho$, the year-to-year correlation often is much larger than $\rho^{12}$. In other words, an underlying cyclic behavior of the population may upset the exponential correlation model unless the pattern length is a small part of the period.

Finally, a truncated pattern is easier to handle computationally than an ever-increasing pattern. After deciding how far back to truncate, we compute a set of coefficients to be multiplied into the sample averages. At each time $t_i$, we use the same set of coefficients but apply them to a different set of sample averages.

**8. Generalization of the sample pattern.** In this section and the next two sections, we consider what happens when we allow more freedom in the choice of a sample pattern. We first describe the modifications necessary when the number of sample values added to the pattern varies with time.

To be specific, we assume a sample pattern of the type shown in (3.1), but with $n_k$ sample values associated with time $t_k$, $1 \leqq k \leqq i$. These $n_k$ values can be divided into two classes: $n_k'$ are associated with popualtion elements represented in the sample pattern at time $t_{k-1}$, and the remaining $n_k''$ are associated with elements entering the pattern for the first time. We assume one-level rotation sampling; the pattern is built up row by row.

The minimum-variance estimate $M_i'$ based on this sample pattern can be found by the same methods as before. We quote Patterson's results; equation (3.3) generalizes to

$$(8.1) \qquad 1 - A_i' = \frac{n_i' n_{i-1}''}{n_i n_{i-1}'' - \rho^2 n_i''(n_{i-1}'' - A_{i-1}' n_i')}$$

and the variance of the estimate is

$$(8.2) \qquad \operatorname{var}(M_i') = \begin{cases} \dfrac{\sigma^2}{n_i''} A_i', & \text{if } n_i'' \neq 0, n_{i-1}'' \neq 0 \\[2ex] \sigma^2 \left[ \dfrac{1 - \rho^2}{n_i} + \dfrac{\rho^2 A_{i-1}'}{n_{i-1}''} \right], & \text{if } n_i'' = 0, n_{i-1}'' \neq 0. \end{cases}$$

Analogously, we assume a sample pattern of the type shown in (4.1). At time $t_k$ we add to this pattern by drawing a new set of $n_k$ elements from the population and recording the associated sample values for times $t_k$ and $t_{k-1}$ (two-level rotation sampling).

The minimum-variance estimate $M_i''$ based on this sample pattern can be found by the same methods as before. We omit details and give the results.

Equation (4.3) generalizes to

$$a_i' = \frac{\rho n_{i-1}}{n_{i-1} + n_i(1 - a_{i-1}' \rho)} \tag{8.3}$$

and the variance of the estimate is

$$\text{var}(M_i'') = \frac{\sigma^2}{n_i}(1 - a_i' \rho). \tag{8.4}$$

**9. Optimum choice of $\mu$ in one-level rotation sampling.** We have shown in Section 3 how to find $M_i'$ for any preassigned value of $\mu$; we now determine that value of $\mu$ which corresponds to the minimum value of the function var $(M_i'(\mu))$. The solution to this problem depends only on the point at which the pattern is truncated.

If we wish to find the optimum $\mu$ for infinitely long patterns, the problem is easy to solve: we differentiate the variable $A/\mu$ of equation (3.5) with respect to $\mu$, set the result equal to zero, and solve for $\mu$. If $\rho$ is less than unity, a little calculation leads to the result $\mu = \frac{1}{2}$; if $\rho$ is equal to unity, it is intuitively clear that the optimum value of $\mu$ is unity.

If we adopt the practical viewpoint of Section 7 and decide to use truncated patterns, the problem of finding a minimum-variance $\mu$ is conceptually simple but computationally tedious. If we have a pattern consisting of sample values associated with the time-levels $t_i$ and $t_{i-1}$ only, we differentiate $A_2/\mu$ with respect to $\mu$, set the result equal to zero, and solve a quadratic equation in $\mu$. The optimum value of $\mu$ is

$$\mu = \frac{1 - \sqrt{1 - \rho^2}}{\rho^2}. \tag{9.1}$$

Similarly, we can determine the optimum $\mu$ for patterns consisting of sample values associated with three or four time-levels. The algebra is laborious; for example, when the pattern is four time-levels long, one must solve a sixth-degree equation in $\mu$ with coefficients that are fourth-degree polynomials in $\rho^2$. Omitting these equations, we tabulate optimum values of $\mu$ for selected values of $\rho$:

| Pattern Length | $\rho$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | .4 | .6 | .8 | .9 | .95 | 1.00 |
| 2 | .500 | .522 | .556 | .625 | .698 | .762 | 1.000 |
| 3 | .500 | — | .517 | .552 | .612 | .670 | 1.000 |
| 4 | .500 | — | — | — | .563 | .623 | 1.000 |
| $\infty$ | .500 | .500 | .500 | .500 | .500 | .500 | 1.000 |

The corresponding variances are given in Table 9.1 at the end of this section.

We now consider the problem of choosing an optimum set of $\mu$ for the one-level rotation sampling pattern when $\mu$ is not restricted to a constant value in time, as it is in the sampling pattern (3.1). This problem has two parts: one must determine the set $(1, \mu_2, \mu_3, \cdots, \mu_i)$ which corresponds to the minimum-variance estimate $M_i^*$, and one must calculate the variance of this estimate.

This problem was first solved by Cochran [2]; we summarize his results for the purpose of comparison. The solution is simplified by the fact that the optimum set $(1, \mu_2, \mu_3, \cdots, \mu_k)$ contains the optimum set $(1, \mu_2, \mu_3, \cdots, \mu_{k-1})$, $2 \leqq k \leqq i$; therefore the solution can be obtained in an iterative form. If we define $G_i$ by the equation $\text{var}(M_i^*) = (\sigma^2/n)G_i$, then the variance can be calculated from the iterative equation

$$\frac{1}{G_i} = 1 + \frac{(1 + \sqrt{1 - \rho^2})^2}{\rho^2 G_{i-1}} \quad \text{and} \quad G_1 = 1 \tag{9.2}$$

and the optimum values of $\mu_k$, $1 \leqq k \leqq i$, are given by

$$\mu_k = 1 + \frac{(1 - \rho^2) - \sqrt{1 - \rho^2}}{\rho^2 G_{k-1}}. \tag{9.3}$$

The limiting values of $G_i$ and $\mu_k$ are

$$\lim_{i \to \infty} G_i = G = \frac{2(\sqrt{1 - \rho^2} - (1 - \rho^2))}{\rho^2}, \tag{9.4}$$

$$\lim_{k \to \infty} \mu_k = \mu = \tfrac{1}{2}. \tag{9.5}$$

We tabulate $\mu_2$, $\mu_3$ and $\mu_4$ for selected values of $\rho$; the corresponding variances are presented in Table 9.1 at the end of this section. When $\rho$ is equal to zero, it does not matter what replacement rate is used.

| $\rho$ | 0 | .4 | .6 | .8 | .9 | .95 | 1.00 |
|---|---|---|---|---|---|---|---|
| $\mu_2$ | .500 | .522 | .556 | .625 | .698 | .762 | 1.000 |
| $\mu_3$ | .500 | .502 | .506 | .531 | .579 | .637 | 1.000 |
| $\mu_4$ | .500 | .500 | .501 | .508 | .533 | .572 | 1.000 |
| $\mu$ | .500 | .500 | .500 | .500 | .500 | .500 | 1.000 |

Since the successive values of $\mu_k$ are different, the optimum patterns cannot be conveniently truncated. The truncated patterns gradually change in form from the optimum pattern with $\mu_k$ given by equation (9.3) to the limiting pattern in which all $\mu_k$ are equal to one-half.

Table 9.1 gives the variances of the estimates that have been discussed in this section. For comparison, we include the variance of the minimum-variance estimate based on the sample pattern (3.1) with $\mu$ equal to one-half. The most striking characteristic of this table is the small difference in variance between the two kinds of optimum estimates and the one-half replacement rate estimate. In practical applications of rotation sampling, one might as well use the latter pattern, since it is much simpler to apply and can be conveniently truncated.

TABLE 9.1

*Comparison of the variances of three minimum-variance estimates based on different restrictions on the replacement rates $\mu_i$*

*(one-level sampling)*

| Length of pattern | Restriction on $\mu$-values | $\rho=0$ | .2 | .4 | .6 | .8 | .9 | .95 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | No restriction | 1.000 | .990 | .958 | .900 | .800 | .718 | .656 | .500 |
| | All $\mu_i$ are equal | 1.000 | .990 | .958 | .900 | .800 | .718 | .656 | .500 |
| | All $\mu_i$ equal 1/2 | 1.000 | .990 | .958 | .901 | .810 | .746 | .709 | .667 |
| 3 | No restriction | 1.000 | .990 | .956 | .890 | .762 | .646 | .556 | .333 |
| | All $\mu_i$ are equal | 1.000 | .990 | .956 | .890 | .763 | .649 | .560 | .333 |
| | All $\mu_i$ equal 1/2 | 1.000 | .990 | .956 | .890 | .765 | .660 | .589 | .500 |
| 4 | No restriction | 1.000 | .990 | .956 | .889 | .753 | .622 | .515 | .250 |
| | All $\mu_i$ are equal | 1.000 | .990 | .956 | .889 | .753 | .624 | .518 | .250 |
| | All $\mu_i$ equal 1/2 | 1.000 | .990 | .956 | .889 | .754 | .628 | .533 | .400 |
| $\infty$ | All three methods are equivalent | 1.000 | .990 | .956 | .889 | .750 | .607 | .476 | .000 |

All entries should be multiplied by $\sigma^2/n$ to obtain variances

**10. Minimizing the sample size while holding the variance constant in time.** In this section we consider the problem of minimizing the sample size of the pattern at each time $t_i$ while holding constant in time the variance of $M_i$, the minimum-variance linear unbiased estimate of $\alpha(t_i)$ based on the sample pattern. The solution can be obtained for one-level or two-level rotation sampling.

We consider the two-level sampling problem first because its solution is simpler. We assume a sample pattern of the type illustrated by (4.1); at time $t_k$ we add to the pattern by drawing a new set of $n_k$ elements from the population and recording the associated sample values for times $t_k$ and $t_{k-1}$, $1 \leqq k \leqq i$. We assume that var $(M_i'')$ is equal to $\sigma^2/N$ for all values of $i$. Obviously, $n_1$ is equal to $N$. If var $(M_2'')$ is to be equal to $\sigma^2/N$, then

$$\frac{1}{n_2}(1 - a_2' \rho) = \frac{1}{N}.$$

We substitute equation (8.3) for $a_2'$ and solve for $n_2$.

$$n_2 = N\sqrt{1 - \rho^2}, \qquad a_2' = \frac{1 - \sqrt{1 - \rho^2}}{\rho}.$$

We can similarly evaluate $n_3$, $n_4$, etc.; we find by induction that all succeeding $n_i$ and $a_i'$ are equal to $n_2$ and $a_2'$, respectively. In other words, we should draw a sample of $N$ elements at time $t_1$, and a sample of $N\sqrt{1 - \rho^2}$ elements at all succeeding times. It is rather surprising to find that the minimum value of $n_i$ is attained by time $t_2$; when we considered in Section 4 the inverse problem of minimizing the variance while keeping the sample size constant, we found that the variance approached the limiting value $(\sigma^2/n)\sqrt{1 - \rho^2}$ as $i$ approached infinity.

Patterson [5] has solved the problem of minimizing the sample size while holding the variance constant using a one-level rotation sampling pattern; we summarize his results for the sake of comparison. The problem was first considered by Jessen [4]; he gave an incorrect solution for a pattern consisting of sample values from times $t_1$ and $t_2$ only.

The solution to the problem is carried out by induction. Using the terminology of Section 8, we seek to minimize the sample size $n_k = n_k' + n_k''$. By means of equation (8.1) and the restriction that $A_{k-1}'/n_{k-1}''$ be equal to $1/N$ (the induction hypothesis), we can express $n_k$ as a function of $n_k'$ alone and carry out the minimization with respect to this variable. We find that

$$n_k' = N\left[\frac{-(1 - \rho^2) + \sqrt{1 - \rho^2}}{\rho^2}\right]$$

and that $n_k''$ is equal to $n_k'$ for $k$ greater than or equal to two. In other words, from time $t_2$ onward one should draw a sample of $2n_k'$ elements and use a replacement rate of one-half.

For both one-level and two-level rotation sampling, the behavior of the minimum-sample-size problem can be summarized by the relation

$$\frac{\text{minimum variance at time } t_1}{\text{minimum variance at time } t_\infty} \text{ using the constant sample size method}$$

$$= \frac{\text{minimum sample size at time } t_1}{\text{minimum sample size at time } t_2} \text{ using the constant variance method.}$$

If one wishes to carry out rotation sampling on a truncated pattern, then one is restricted to minimum-variance estimates, since patterns with constant sample sizes and constant replacement rates are the only ones that can be conveniently truncated.

**11. One-level versus multi-level rotation sampling.** In this section we derive a criterion for deciding when to use multi-level sampling instead of one-level sampling. The criterion is given in the form of a graph in which the correlation $\rho$ is plotted against a parameter $k$ which compares the cost of sampling several different values of an element at one time with the cost of sampling only one value at a time.

We assume that it costs $c$ to obtain a single sample value at time $t_i$, and $c(1 + k)$ to obtain both sample values $x_{ij}$ and $x_{(i-1)j}$ at time $t_i$, where $0 \leqq k \leqq 1$. In other words, we allow for the fact that it may cost less to obtain two sample values at a single time than it does to obtain them at two separate times. The three-level sampling cost is assumed to be $c(1 + 2k)$ per element.

Suppose that we have a fixed amount $T$ to spend on our sample at time $t_i$; how many sample values can we draw, using each of the three methods of sampling? If we restrict ourselves to patterns of constant sample size $n$, then for one-level sampling the sample size $n'$ is equal to $T/c$, for two-level sampling $n''$ is equal to $T/c(1 + k)$, and for three-level sampling $n'''$ is equal to $T/c(1 + 2k)$.

We begin by assuming that we have patterns of infinite length. Equating the variance of the one-level estimate for $\mu$ equal to one-half with the variance of the two-level estimate, we have

$$\frac{2\sigma^2}{n'}\left[\frac{-(1-\rho^2)+\sqrt{1-\rho^2}}{\rho^2}\right] = \frac{\sigma^2}{n''}\sqrt{1-\rho^2}.$$

Substituting in the values of $n'$ and $n''$ and solving for $k$, we find the curve on which var $(M')$ is equal to var $(M'')$:

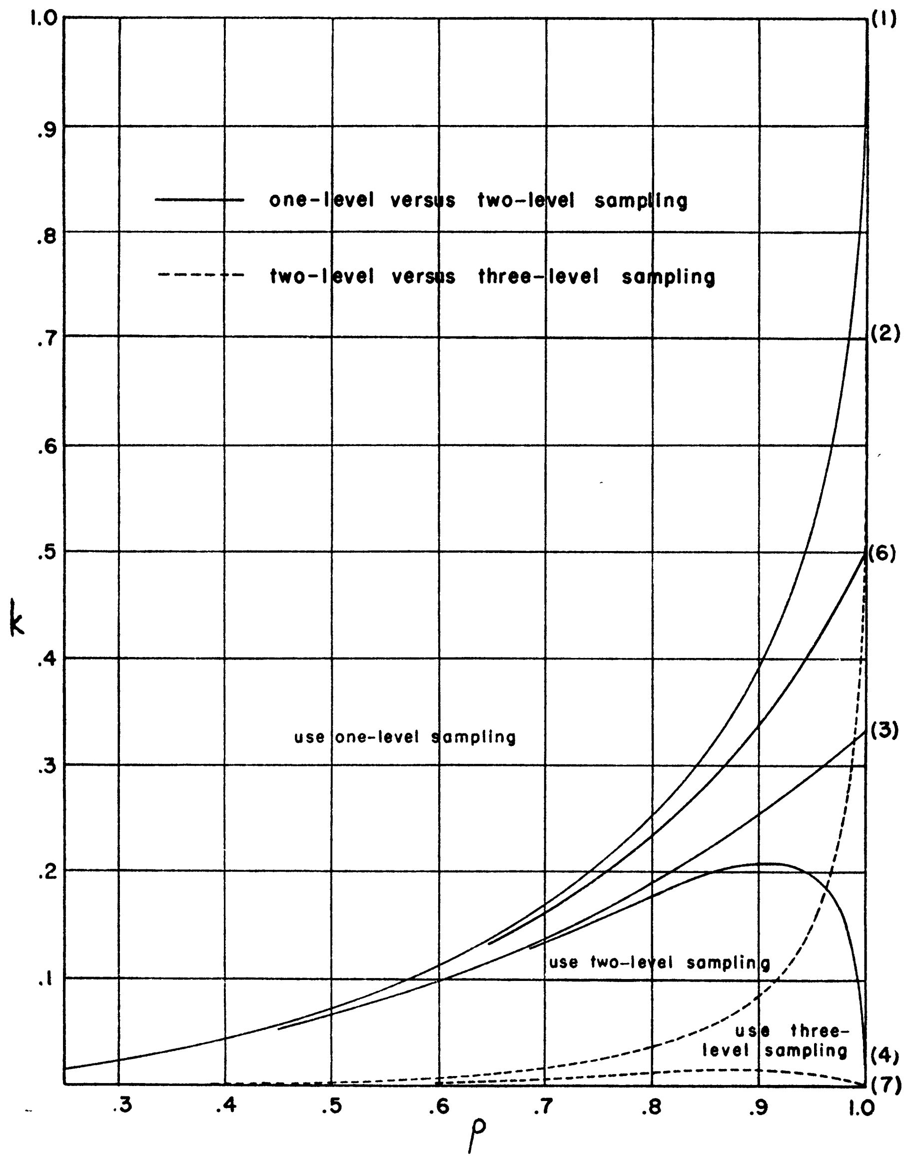
$$k = \left[\frac{\sqrt{1-\rho^2}-1}{\rho}\right]^2.$$

In order to decide when to use a three-level estimate instead of a two-level one, we equate the variance of the two-level estimate with the variance of the three-level estimate, and solve for $k$. Both of these curves are plotted in Graph 11.1; the one-level versus two-level curve is labeled (1) and the two-level versus three-level curve is labeled (2). The two curves partition the $(k, \rho)$ unit square into three areas: in the large area at the upper left, var $(M')$ is less than var $(M'')$ or var $(M''')$; in the central area var $(M'')$ is the minimum; in the small area area at the lower right var $(M''')$ is the minimum.

If we have very short truncated patterns, the preceding analysis no longer holds. In order to see what happens, we solve the equations var $(M_i') =$ var $(M_i'')$ and var $(M_i'') =$ var $(M_i''')$ for $i$ equal to two and three. In comparing the variances of $M_2'$ and $M_2''$, we have two possibilities to consider: we can use the simple one-level pattern in which $\mu$ is equal to one-half, or we can use the optimum one-level pattern in which $\mu$ is equal to $(1 - \sqrt{1-\rho^2})/\rho^2$. We summarize all of these results in the table below and in Graph 11.1.

| Variance of | $\rho =$ | .0 | .2 | .4 | .6 | .8 | .9 | .95 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|
| (1) $M' = M''$ | $k =$ | .000 | .010 | .044 | .111 | .250 | .393 | .524 | 1.000 |
| (2) $M'' = M'''$ | | .000 | .000 | .001 | .008 | .036 | .085 | .151 | .707 |
| (3) $M_2' = M_2''$, $\mu = \frac{1}{2}$ | $k =$ | .000 | .010 | .042 | .099 | .191 | .254 | .291 | .333 |
| (4) $M_2' = M_2''$, $\mu =$ opt. | | .000 | .010 | .042 | .098 | .177 | .207 | .196 | .000 |
| (5) $M_2'' = M_2'''$ | | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| (6) $M_3' = M_3''$, $\mu = \frac{1}{2}$ | $k =$ | .000 | .010 | .044 | .110 | .235 | .340 | .411 | .500 |
| (7) $M_3'' = M_3'''$ | | .000 | .000 | .001 | .005 | .012 | .014 | .011 | .000 |

Graph 11.1 clearly shows that the higher-level sampling patterns are optimum over a very restricted area; it is advantageous to undertake four-level or higher-level rotation sampling only for very low $k$, very high $\rho$, and relatively long patterns.




GRAPH 11.1

*Comparison of one-level rotation sampling with multi-level rotation sampling*

**12. Estimation of a linear function of the population means.** We have discussed at some length the problem of finding minimum-variance linear unbiased estimates $M_i$ of the population mean $\alpha(t_i)$ based on several different sample patterns. In the next two sections, we consider the more general problem of finding $M(\alpha)$, the minimum-variance linear unbiased estimate of a linear function of the population means at several different times:

$$\alpha = c_i\alpha(t_i) + c_j\alpha(t_j) + \cdots + c_k\alpha(t_k),$$

where the letter $c$ denotes a constant.

The problem of finding minimum-variance estimates $M(\alpha)$ contains the problem of improving the minimum-variance estimate $M_i$ of $\alpha(t_i)$ by using observations drawn at times $t_{i+1}$ or later. (Previously, we assumed that these later observations were not available.) The connection between these problems is presented in Theorem 3 below, which can easily be proved using Theorem 1 in both directions. Following Patterson's notation, we denote the minimum-variance linear unbiased estimate of $\alpha(t_i)$ based on a sample pattern containing values drawn at times $t_1, \cdots, t_i, \cdots, t_j$ by the symbol ${}_jM_i$. When $j$ is equal to $i$, we write $M_i$ as before.

THEOREM 3. *Assume the conditions of Theorem* 1. *Assume also that we know the estimates* ${}_mM_i$, $i = 1, 2, \cdots, m$, *based on the sample pattern* $P$. *Then the minimum-variance estimate* $M(\alpha)$ *of* $\alpha = c_1\alpha(t_1) + c_2\alpha(t_2) + \cdots + c_m\alpha(t_m)$ *in the class of linear unbiased estimates based on the sample pattern* $P$ *is*

$$M(\alpha) = c_1({}_mM_1) + c_2({}_mM_2) + \cdots + c_m(M_m).$$

Using this theorem and the results derived in the next section, it is a straightforward but tedious job to write down minimum-variance estimates of any linear function of the population means; therefore we omit specific details or examples. It is equally simple to calculate the variances of these estimates. The linear function of greatest practical interest and most frequently discussed in the literature is the difference between two successive means $\alpha(t_i) - \alpha(t_{i-1})$. Other functions of potential interest are higher-order differences and moving averages.

**13. Improvement of minimum-variance estimates of the mean.** Patterson [5] has solved the problem of finding improved minimum-variance linear unbiased estimates ${}_jM_i$ of $\alpha(t_i)$ based on one-level rotation sampling of patterns of the type illustrated by (3.1). He shows how to solve for the estimates iteratively. For example:

$$(13.1) \qquad {}_iM'_{i-1} = M'_{i-1} - \rho A_{i-1}M'_i + \rho A_{i-1}\bar{x}_{i,1},$$

$$(13.2) \qquad {}_iM'_{i-2} = {}_{i-1}M'_{i-2} - \rho^2 A_{i-2}(1 - A_{i-1})M'_i + \rho^2 A_{i-2}(1 - A_{i-1})\bar{x}_{i,1}.$$

Patterson also derives variances and covariances based on the improved estimates. However, when $j$ becomes much larger than $i$ in the estimate ${}_jM_i$, these expressions become quite cumbersome; consequently Patterson derives non-minimum-variance linear unbiased estimates of $\alpha(t_i)$ that have a somewhat simpler form.

The problem of finding ${}_jM''_i$ using two-level rotation sampling can be solved rather completely; it is not necessary to resort to non-minimum-variance estimates in order to obtain manageable expressions for the estimates and the variances and covariances between them. In fact, it is possible to obtain variances and covariances of improved estimates based on infinitely long patterns.

We assume that we know $M''_1, M''_2, \cdots, M''_i$, and wish to obtain ${}_iM''_{i-k}$ when $k$ is greater than or equal to one. We can partition the sample pattern illustrated in (4.1) into two statistically independent parts, the left-hand one a two-level sampling pattern running from time $t_0$ through time $t_{i-k}$, and the right-hand one a two-level sampling pattern running from time $t_{i-k}$ through time $t_i$. Each part can be used to obtain a minimum-variance linear unbiased estimate of $\alpha(t_{i-k})$; the estimate based on the right-hand pattern is $M''_k$ (with inverted time), and the estimate based on the left-hand pattern is $M''_{i-k}$. Applying Lemma 1 of Section 5, we conclude that the minimum-variance linear unbiased estimate of $\alpha(t_{i-k})$ based on the entire pattern is

$${}_iM''_{i-k} = \frac{a}{a+b} M''_k + \frac{b}{a+b} M''_{i-k}$$

where $a$ is equal to var $(M''_{i-k})$ and $b$ is equal to var $(M''_k)$. Furthermore, the variance of this estimate is

$$\text{var}({}_iM''_{i-k}) = \frac{\sigma^2}{n} \frac{(1 - a_{i-k}\,\rho)(1 - a_k\,\rho)}{(1 - a_{i-k}\,\rho) + (1 - a_k\,\rho)}.$$

The first three terms can be written down in an equivalent form:

$$\text{var}({}_iM''_{i-1}) = \frac{\sigma^2}{n}\frac{1}{\rho}\, a_i(1 - a_{i-1}\,\rho),$$

$$\text{var}({}_iM''_{i-2}) = \frac{\sigma^2}{n}\frac{2 - \rho^2}{\rho^2}\, a_i\, a_{i-1}(1 - a_{i-2}\,\rho),$$

$$\text{var}({}_iM''_{i-3}) = \frac{\sigma^2}{n}\frac{4 - 3\rho^2}{\rho^3}\, a_i\, a_{i-1}\, a_{i-2}(1 - a_{i-3}\,\rho).$$

If we let $i$ equal $2k$, then the variance becomes

$$\text{var}({}_{2k}M''_k) = \frac{\sigma^2}{n}\frac{1 - a_k\,\rho}{2} = \tfrac{1}{2}\,\text{var}(M''_k)$$

which could have been predicted at once from (4.1). The variance of this estimate is less than that of any other estimate of a single mean based on a pattern of the same length.

The covariance terms are equally easy to derive. Using Corollary 1.1 and the iterative form of the two-level estimate, we obtain

$$\text{cov}({}_iM''_{i-k}, M''_i) = \frac{\sigma^2}{n}\, a_i\, a_{i-1} \cdots a_{i-k+1}(1 - a_{i-k}\,\rho). \tag{13.3}$$

The following covariance-equation contains all the preceding results:

$$(13.4) \quad \operatorname{cov}({}_iM''_{i-k}, {}_iM''_{i-j}) = \frac{\sigma^2}{n} \frac{(1 - a_j\rho)(1 - a_{i-k}\rho)}{(1 - a_j\rho) + (1 - a_{i-j}\rho)} a_{i-j}a_{i-j-1} \cdots a_{i-k+1}.$$

We require that $0 < j \leq k < i$; the second inequality is not a restriction. The derivation of equation (13.4) is straightforward:

$$\operatorname{cov}({}_iM''_{i-k}, {}_iM''_{i-j}) = \operatorname{cov}(M''_{i-k}, {}_iM''_{i-j})$$

$$= \frac{1 - a_j\rho}{(1 - a_j\rho) + (1 - a_{i-j}\rho)} \operatorname{cov}(M''_{i-k}, M''_{i-j}).$$

If we substitute equation (13.3) into the above expression, we obtain (13.4) at once. When $k$ equals $j$, we assume that $a_{i-j} \cdots a_{i-k+1}$ is equal to one.

It is easy to derive limiting results from equation (13.4). For example:

$$\lim_{i \to \infty} \operatorname{cov}({}_{2i}M''_i, {}_{2i}M''_{i-k}) = \frac{\sigma^2}{n} \cdot \frac{(1 - a\rho)a^k}{2}.$$

**14. A further relationship between one-level and two-level rotation sampling.** If we restrict ourselves to a replacement rate $\mu$ of one-half while using one-level rotation sampling, then it is not necessary to use the cumbersome formulas developed by Patterson in [5]; we can write down the estimates ${}_jM'_i$ and the associated variances and covariances in terms of the two-level estimates ${}_jM''_i$. This section extends the results of Section 5, which showed how to write $M'_i$ as a function of $M''_i$. The simplified results derived below can be quite useful in practice, since a replacement rate of one-half was shown to be nearly optimum in Section 9.

Consider the minimum-variance estimate ${}_{i+1}M''_i$, which is based on the sample pattern (4.1) extending from time $t_0$ to time $t_{i+1}$. The estimate ${}_{i+1}M''_i$ is also a minimum-variance linear unbiased estimate based on the sample values extending from time $t_1$ to time $t_i$; the sample values associated with time $t_0$ and time $t_{i+1}$ are assigned zero weights in the minimum-variance estimate and can be disregarded. In other words, ${}_{i+1}M''_i$ is the minimum-variance estimate based on pattern (3.1) with $\mu$ equal to one-half and a sample size of $2n$ instead of $n$; ${}_{i+1}M''_i$ is identically equal to $M'_i$. Using a similar argument, one can easily show that ${}_{i+1}M''_{i-k}$ (based on a sample size of $n$) is identically equal to $M'_{i-k}$ (with $\mu$ equal to ½ and a sample size of $2n$). The variance-covariance formulas are summarized by the general expression

$$\operatorname{cov}({}_iM'_{i-j}, {}_iM'_{i-k}) = 2\operatorname{cov}({}_{i+1}M''_{i-j}, {}_{i+1}M''_{i-k}).$$

**15. Acknowledgments.** The author is indebted to Professor S. S. Wilks for suggesting the subject of this paper, and for his advice and encouragement during the course of the work.

## REFERENCES

[1] M. A. Bershad, "Best linear estimate—sampling for a time series," unpublished memorandum, U. S. Bureau of the Census (1952).

[2] W. G. Cochran, *Sampling Techniques*, John Wiley and Sons, 1953, pp. 287–290.

[3] M. H. Hansen, W. N. Hurwitz, and W. G. Madow, *Sample Survey Methods and Theory*, John Wiley and Sons, 1953, Vol. 1, pp. 490–503, Vol. 2, pp. 268–279.

[4] R. J. Jessen, "Statistical Investigation of a Sample Survey for Obtaining Farm Facts," *Iowa State College of Agriculture and Mechanic Arts Res. Bull. 304* (1942), pp. 54–59.

[5] H. D. Patterson, "Sampling on Successive Occasions with Partial Replacement of Units," *J. Roy. Stat. Soc. Ser. B*, *Vol. 12* (1950), pp. 241–255.

[6] F. Yates, *Sampling Methods for Censuses and Surveys*, Griffin, 1949, pp. 175–182, 233–235, 260–262.