SOME REMARKS ON SYSTEMATIC SAMPLING[1]

By Werner Gautschi[2]

University of California, Berkeley

1. Introduction and summary. Consider a finite population consisting of N elements $y_1, y_2, \cdots, y_N$. Throughout the paper we will assume that $N = nk$. A systematic sample of n elements is drawn by choosing one element at random from the first k elements $y_1, \cdots, y_k$, and then selecting every kth element thereafter. Let $y_{ij} = y_{i+(j-1)k} (i = 1, \cdots, k; j = 1, \cdots, n)$; obviously systematic sampling is equivalent to selecting *one* of the k "clusters"

$$C_i = \{y_{ij}; j = 1, \cdots, n\}$$

at random. From this it follows that the sample mean $\bar{y}_i = 1/n \sum_{j=1}^{n} y_{ij}$ is an unbiased estimate for the population mean $\bar{y} = 1/N \sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij}$ and that $\text{Var}\, \bar{y}_i = 1/k \sum_{i=1}^{k} (\bar{y}_i - \bar{y})^2$. We will denote this variance by $V_{sy}^{(1)}$ indicating by the superscript that only one cluster is selected at random. $V_{sy}^{(1)}$ can be written as

$$V_{sy}^{(1)} = S^2 - \frac{1}{k} \sum_{i=1}^{k} S_i^2, \quad \text{where} \quad S^2 = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y})^2, \tag{1}$$
$$S_i^2 = \frac{1}{n} \sum_{j=1}^{n} (y_{ij} - \bar{y}_i)^2.$$

It is natural to compare systematic sampling with stratified random sampling, where *one* element is chosen independently in each of the n strata $\{y_1, \cdots, y_k\}$, $\{y_{k+1}, \cdots, y_{2k}\}, \cdots$, and with simple random sampling using sample size n. The corresponding variances of the sample mean will be denoted by $V_{st}^{(1)}$ $V_{ran}^{(n)}$ respectively.

We consider now the following generalization of systematic sampling which appears to have been suggested by J. Tukey (see [3], p. 96, [4], [5]). Instead of choosing at first only one element at random we select a simple random sample of size s (without replacement) from the first k elements and then every kth element following those selected. In this way we obtain a sample of ns elements and, if $i_1, i_2, \cdots, i_s$ are the serial numbers of the elements first chosen, the sample mean $1/s(\bar{y}_{i_1} + \cdots + \bar{y}_{i_s})$ can be used as an estimate for the population mean. This sampling procedure is clearly equivalent to drawing a simple random sample of size s from the k clusters $C_i (i = 1, \cdots, k)$. It therefore follows (see, for example, [2], Chapter 2.3 to 2.4) that the sample mean is an unbiased estimate for the population mean and that its variance, which we denote by $V_{sy}^{(s)}$, is given by[3]

Received June 12, 1956.

[1] This investigation was supported (in part) by a research grant (RG-3666) from the Institutes of Health, Public Health Service.

[2] Now at Ohio State University

[3] This formula is not new, but appeared already in [6] and, more recently, in [5].

$$V_{sy}^{(s)} = \frac{k-s}{ks}\frac{1}{k-1}\sum_{i=1}^{k}(\bar{y}_i - \bar{y})^2 = \frac{1}{s}\frac{k-s}{k-1}V_{sy}^{(1)}. \tag{2}$$

Again, it is natural to compare this sampling procedure with stratified random sampling, where a simple random sample of size s is drawn independently in each of the n strata $\{y_1, \cdots, y_k\}$, $\{y_{k+1}, \cdots, y_{2k}\}$, $\cdots$ or with simple random sampling employing sample size ns. We denote the corresponding variances of the sample mean (which in both cases is an unbiased estimate for the population mean) by $V_{st}^{(s)}$, $V_{ran}^{(ns)}$ respectively. From well-known variance formulae (see, for example, [2], Chapters 2.4 and 5.3) it follows that

$$\begin{aligned} V_{st}^{(s)} &= \frac{1}{s}\frac{k-s}{k-1}V_{st}^{(1)}, \\ V_{ran}^{(ns)} &= \frac{N-ns}{s(N-n)}V_{ran}^{(n)} = \frac{1}{s}\frac{k-s}{k-1}V_{ran}^{(n)}. \end{aligned} \tag{3}$$

Thus the relative magnitudes of the three variances $V_{sy}^{(s)}$, $V_{st}^{(s)}$, $V_{ran}^{(ns)}$ are the same as for $V_{sy}^{(1)}$, $V_{st}^{(1)}$, $V_{ran}^{(n)}$, of which comparisons were made for several types of populations by W. G. Madow and L. H. Madow [6] and W. G. Cochran [1]. Some of the results will be reviewed in Section 3.

The object of this note is to compare systematic sampling with s random starts, as described above, with systematic sampling employing only one random start but using a sample of the same size ns. To make this comparison we obviously have to assume that k is an integral multiple of s, say $k = ls$. The latter procedure then consists in choosing one element at random from the first l elements $\{y_1, \cdots, y_l\}$ and selecting every lth consecutive element. We denote the variances of the sample mean of the two procedures by $V_k^{(s)}$, $V_l^{(1)}$ respectively, indicating by the subscript the size of the initial "counting interval." (In our notation $V_{sy}^{(s)} \equiv V_k^{(s)}$.) We shall show in Section 4 that $V_l^{(1)} = V_k^{(s)}$ in the case of a population "in random order," but $V_l^{(1)} < V_k^{(s)}$ for a population with a linear trend or with a positive correlation between the elements which is a decreasing convex function of their distance apart. Some numerical results on the relative precision of the two procedures will be given in Section 5 for the case of a large population with an exponential correlogram.

2. Acknowledgment. I wish to express my debt to Professor W. Kruskal for having brought the question treated in this note to my attention.

3. Cochran's approach. Extension of Cochran's results to systematic sampling with multiple random starts. Instead of considering a particular single population $\{y_1, y_2, \cdots, y_N\}$ we assume, following Cochran [1], [2], Chapter 8, that the y_i's are drawn from an infinite population having some specified properties. We are then interested in comparing the expected variance $E(V \mid y_1, \cdots, y_N)$ rather than $(V \mid y_1, \cdots, y_N)$ for the sampling procedures under consideration. More specifically, we consider the following three types of populations.

(i) *Population in random order*. The variates y_i are assumed to be uncorrelated and to have the same expectations. The variances may change with i

$$(4)\qquad \begin{aligned} Ey_i &= \mu, \quad E(y_i - \mu)^2 = \sigma_i^2 \qquad (i = 1, \cdots, N); \\ E(y_i - \mu)(y_j - \mu) &= 0 \qquad (i \neq j). \end{aligned}$$

It is not difficult to show ([2], Chapter 8.5) that in this case

$$(5)\qquad EV_{sy}^{(1)} = EV_{st}^{(1)} = EV_{ran}^{(n)} = \frac{N-n}{N}\frac{\sigma^2}{n} = \frac{k-1}{k}\frac{\sigma^2}{n},$$

where $\sigma^2 = \sum_{i=1}^N \sigma_i^2/N$.

(ii) *Population with a linear trend.* We assume that the y_i's are uncorrelated variates whose expectations change linearly in i, more precisely

$$(6)\qquad \begin{aligned} Ey_i &= \alpha + \beta i, \quad \operatorname{Var} y_i = \sigma^2 \qquad (i = 1, 2, \cdots, N), \\ \operatorname{Cov}(y_i, y_j) &= 0 \qquad (i \neq j). \end{aligned}$$

Applying standard linear regression theory (see, for instance, [7], Chapter 14.2) to the sum of squares in (1), it is easily found that

$$(8)\qquad EV_{sy}^{(1)} = \frac{N-n}{Nn}\sigma^2 + \beta^2\frac{k^2-1}{12} = \frac{k-1}{nk}\sigma^2 + \beta^2\frac{k^2-1}{12}.$$

In a similar way we obtain

$$(9)\qquad \begin{aligned} EV_{st}^{(1)} &= \frac{k-1}{nk}\sigma^2 + \beta^2\frac{k^2-1}{12n}, \\ EV_{ran}^{(n)} &= \frac{k-1}{nk}\sigma^2 + \beta^2\frac{(k-1)(nk+1)}{12}. \end{aligned}$$

Thus

$$(10)\qquad EV_{st}^{(1)} \leqq EV_{sy}^{(1)} \leqq EV_{ran}^{(n)},$$

with equality only if $n = 1$.

(iii) *Population with serial correlation.* It is assumed that two elements y_i, y_j are positively correlated with a correlation which depends only on the "distance" $z = |j - i|$ and which decreases as z increases. The mean and variance of all the y_i are supposed to be constant

$$(11)\qquad \begin{aligned} Ey_i &= \mu, \quad E(y_i - \mu)^2 = \sigma^2 \qquad (i = 1, 2, \cdots, N), \\ E(y_i - \mu)(y_{i+z} - \mu) &= \rho_z\sigma^2, \end{aligned}$$

where $\rho_{z_1} \geqq \rho_{z_2} \geqq 0$ for $z_1 < z_2$. For this type of population Cochran [1] obtained the following results relevant to our purpose:

$$(12)\qquad EV_{sy}^{(1)} = \frac{k-1}{N}\sigma^2\left\{1 - \frac{2}{N(k-1)}\sum_{z=1}^{N-1}(N-z)\rho_z + \frac{2k}{n(k-1)}\sum_{z=1}^{n-1}(n-z)\rho_{kz}\right\},$$

(13) $EV_{st}^{(1)} \leqq EV_{ran}^{(n)}$,

(14) $EV_{sy}^{(1)} \leqq EV_{st}^{(1)}$,

(14) applying if, in addition, ρ_z is convex downwards.

In virtue of (2) and (3) all the results (5), (10), (13) and (14) carry over immediately to the more general sampling procedure discussed in Section 1 and, moreover, the relative sizes of the variances $V_{sy}^{(s)}$, $V_{st}^{(s)}$, $V_{ran}^{(ns)}$ remain the same as those of $V_{sy}^{(1)}$, $V_{st}^{(1)}$, $V_{ran}^{(n)}$. Numerical results of the relative precision

$$EV_{st}^{(1)}/EV_{sy}^{(1)}$$

were given by Cochran [1] for populations with a linear and exponential correlogram.

4. Comparison of systematic sampling and systematic sampling with multiple random starts.

(i) *Population in random order.* From (5), replacing k by l and n by ns, we obtain

$$EV_l^{(1)} = \frac{l-1}{lns}\sigma^2 = \frac{l-1}{N}\sigma^2.$$

On the other hand, by (2) and (5), remembering that $k = sl$,

$$EV_k^{(s)} = \frac{1}{s}\frac{k-s}{k-1}\frac{k-1}{k}\frac{\sigma^2}{n} = \frac{l-1}{N}\sigma^2.$$

Thus

(15) $$EV_l^{(1)} = EV_k^{(s)}.$$

(ii) *Population with linear trend.* By (2) and (8)

$$EV_l^{(1)} = \frac{l-1}{N}\sigma^2 + \beta^2\frac{(l-1)(l+1)}{12},$$

$$EV_k^{(s)} = \frac{1}{s}\frac{k-s}{k-1}\left[\frac{k-1}{nk}\sigma^2 + \beta^2\frac{k^2-1}{12}\right] = \frac{l-1}{N}\sigma^2 + \beta^2\cdot\frac{(l-1)(ls+1)}{12}.$$

Hence

(16) $$EV_l^{(1)} \leqq EV_k^{(s)}$$

with equality only if $s = 1$.

Both these results are, of course, to be expected intuitively. The comparison of $V_l^{(1)}$ and $V_k^{(s)}$ is, perhaps, mostly relevant for a population with a convex decreasing correlogram, since in this case $EV_l^{(1)}$ turns out to be the smallest among all the variances $EV_l^{(1)}$, $EV_k^{(s)}$, $EV_{st}^{(s)}$, $EV_{ran}^{(ns)}$.

(iii) *Population with serial correlation.* From (12) and (2),

$$EV_l^{(1)} = \frac{l-1}{N}\sigma^2\left\{1 - \left[\frac{2}{N(l-1)}\sum_{z=1}^{N-1}(N-z)\rho_z\right.\right.$$

$$- \frac{2l}{ns(l-1)} \sum_{z=1}^{ns-1} (ns - z)\rho_{lz} \Big]\Big\} \tag{17}$$

$$= \frac{l-1}{N} \sigma^2 \{1 - L_1\},$$

$$EV_k^{(s)} = \frac{l-1}{N} \sigma^2 \Big\{1 - \Big[\frac{2}{N(k-1)} \sum_{z=1}^{N-1} (N - z)\rho_z$$

$$- \frac{2k}{n(k-1)} \sum_{z=1}^{n-1} (n - z)\rho_{kz} \Big]\Big\} \tag{18}$$

$$= \frac{l-1}{N} \sigma^2 \{1 - L_2\}.$$

It is easy to check that both L_1 and L_2 are linear forms in the ρ_z's in each of which the sum of coefficients is equal to 1. Hence, in order to show that $EV_l^{(1)} \leqq EV_k^{(s)}$, it is enough to prove that

$$L = L_1 - L_2 \geqq 0, \tag{19}$$

L being a linear form of the ρ_z's whose sum of coefficients is zero. If in addition to the monotonicity the ρ_z are assumed to be convex, the following lemma, which is analogous to the lemma proved in [1], is applicable to forms of this type.

LEMMA. *Let S be the set of $\rho = \{\rho_1, \rho_2, \cdots, \rho_m\}$ for which*

$$\rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_m \geqq 0 \tag{20}$$

and

$$\Delta^2 \rho_{\mu-1} = \rho_{\mu+1} - 2\rho_\mu + \rho_{\mu-1} \geqq 0 \qquad (\mu = 2, 3, \cdots, m-1). \tag{21}$$

Let $\alpha_1, \cdots, \alpha_m$ be constants such that $\sum_{\mu=1}^m \alpha_\mu = 0$ and put $A_i = \sum_{\mu=1}^i \alpha_\mu$. Then

$$L = \sum_{\mu=1}^m \alpha_\mu \rho_\mu \geqq 0 \qquad \text{for all } \rho \varepsilon S$$

if and only if

$$B_j = \sum_{i=1}^j A_i \geqq 0 \qquad \text{for } j = 1, 2, \cdots, m-1. \tag{22}$$

Moreover, if in addition to (20) *and* (21) *strict inequality holds in* (22), *then $L > 0$ unless $\rho_1 = \cdots = \rho_m$.*

PROOF. Writing $\alpha_\mu = A_\mu - A_{\mu-1} (\mu = 1, \cdots, m; A_0 = 0)$ and using the fact that $A_m = 0$, we find

$$L = \sum_{\mu=1}^m A_\mu \rho_\mu - \sum_{\mu=1}^m A_{\mu-1} \rho_\mu = - \sum_{\mu=1}^{m-1} A_\mu \Delta\rho_\mu .$$

Similarly,

$$\sum_{\mu=1}^{m-1} A_\mu \Delta\rho_\mu = - \sum_{\mu=1}^{m-2} B_\mu \Delta^2 \rho_\mu + B_{m-1}(\rho_m - \rho_{m-1}).$$

Thus

$$L = \sum_{\mu=1}^{m-2} B_\mu \Delta^2 \rho_\mu + B_{m-1}(\rho_{m-1} - \rho_m). \tag{23}$$

Since, by hypothesis, the coefficients of all the B_μ are nonnegative, the sufficiency of (22) is clear. On the other hand, if $B_{m-1} < 0$, we could choose the ρ_μ linearly decreasing and obtain $L < 0$. If $B_j < 0$, $1 \leqq j \leqq m - 2$, L could be made negative by taking, for example,

$$\rho_\mu = \begin{cases} j + 2 - \mu, & 1 \leqq \mu < j + 1, \\ 1, & j + 1 \leqq \mu \leqq m. \end{cases}$$

Thus (22) is also a necessary condition. If all the B_j are positive, then $L = 0$ implies $\Delta^2 \rho_\mu = 0(\mu = 1, \cdots, m - 2)$, $\rho_{m-1} = \rho_m$. This in turn implies that $\rho_{m-2} = \rho_{m-1}$, $\rho_{m-3} = \rho_{m-2}$, $\cdots$, $\rho_1 = \rho_2$.

THEOREM. *For any population in which*

$$\rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_{N-1} \geqq 0,$$

$$\Delta^2 \rho_{z-1} = \rho_{z+1} - 2\rho_z + \rho_{z-1} \geqq 0 \quad (z = 2, \cdots, N - 2)$$

we have

$$EV_l^{(1)} \leqq EV_k^{(s)} \tag{24}$$

with equality only if $s = 1$ *or* $\rho_1 = \cdots = \rho_{N-1}$.

PROOF. There is nothing to prove if $s = 1$. If $s > 1$ we apply the above lemma (with $m = N - 1$ and L given by (17), (18) and (19)) and show that

$$B_j > 0 \qquad j = 1, 2, \cdots, N - 2. \tag{25}$$

We notice that

$$\frac{N}{2} L_1 = \frac{1}{l - 1} \left[\sum_{z=1}^{N-1} (N - z)\rho_z - l^2 \sum_{z=1}^{ns-1} (ns - z)\rho_{lz} \right]$$

$$\frac{N}{2} L_2 = \frac{1}{ls - 1} \left[\sum_{z=1}^{N-1} (N - z)\rho_z - (ls)^2 \sum_{z=1}^{n-1} (n - z)\rho_{(ls)z} \right].$$

To prove (25) it is enough to show that the sums B_j are positive for the form $NL/2 = NL_1/2 - NL_2/2$. We compute these sums separately for $NL_1/2$, $NL_2/2$ and then take their differences. Put[4]

$$\begin{aligned} j &= \nu k + \sigma l + \lambda = (\nu s + \sigma) l + \lambda, \qquad \text{where } \nu = 0, 1, \cdots, n - 1; \\ \sigma &= 0, 1, \cdots, s - 1; \qquad \lambda = 0, 1, \cdots, l - 1. \end{aligned} \tag{26}$$

[4] We use the Greek letters ν, σ, λ to indicate their range $n - 1$, $s - 1$, $l - 1$, respectively; σ, λ should not be confused with the variance symbol and the parameter to be introduced in Section 5.

By elementary computations the sums $B_j^{(1)}$ for $NL_1/2$ are found to be

$$B_j^{(1)} = \frac{1}{l-1}\{\mathrm{I} - \mathrm{II}\},$$

where

$$\mathrm{I} = \sum_{i=1}^{(\nu s+\sigma)l+\lambda} \frac{i(2N-i-1)}{2}$$

$$= \tfrac{1}{6}[(\nu s+\sigma)l+\lambda][(\nu s+\sigma)l+\lambda+1][3N-(\nu s+\sigma)l-\lambda-2]$$

$$\mathrm{II} = l^2\left[l\sum_{i=1}^{\nu s+\sigma-1}\frac{i(2ns-i-1)}{2} + (\lambda+1)\frac{(\nu s+\sigma)(2ns-\nu s-\sigma-1)}{2}\right]$$

$$= \frac{l^2(\nu s+\sigma)}{6}[l(\nu s+\sigma-1)(3ns-\nu s-\sigma-1) + 3(\lambda+1)(2ns-\nu s-\sigma-1)].$$

Similarly the sums $B_j^{(2)}$ for $NL_2/2$ are obtained as

$$B_j^{(2)} = \frac{1}{ls-1}\{\mathrm{I} - \mathrm{III}\},$$

where

$$\mathrm{III} = (ls)^2\left[ls\sum_{i=1}^{\nu-1}\frac{i(2n-i-1)}{2} + (\sigma l+\lambda+1)\frac{\nu(2n-\nu-1)}{2}\right]$$

$$= \frac{\nu(ls)^2}{6}[ls(\nu-1)(3n-\nu-1) + 3(\sigma l+\lambda+1)(2n-\nu-1)].$$

We have to show that

$$B_j = B_j^{(1)} - B_j^{(2)} = \frac{1}{6}\frac{l}{(l-1)(ls-1)} \cdot \left[(s-1)6\mathrm{I} - (ls-1)\frac{6\mathrm{II}}{l} + (l-1)\frac{6\mathrm{III}}{l}\right] > 0.$$

After some elementary algebra the expression in brackets is found to be a polynomial $f(\sigma)$ in σ of third degree with the following coefficients

$$\begin{aligned}
&\sigma^3: l^2(l-1)\\
&\sigma^2: -3l(l-1)[(n-\nu)ls-(\lambda+1)]\\
(27)\quad &\sigma^1: l\{(l-1)[3s(n-\nu)(sl-2(\lambda+1)) - sl + 3s\lambda + 2s + 1] - 3\lambda(\lambda+1)(s-1)\}\\
&\sigma^0: (s-1)\{\nu ls(l-1)(ls-1) + \lambda(\lambda+1)[3ls(n-\nu) - (\lambda+2)]\}.
\end{aligned}$$

We notice that the second derivative $f''(\sigma)$ vanishes at

$$\sigma^* = (n - \nu)s - \frac{\lambda + 1}{l}$$

which is $\geqq s - 1$ whatever be the values of ν, λ specified by (26). For any of those values $f(\sigma)$ is therefore concave between $\sigma = 0$ and $\sigma = s - 1$ so that it is enough to show $f(0) > 0$, $f(s - 1) > 0$. Now, if $\sigma = 0$ then not both ν, λ can vanish. Hence, $f(0) > 0$ follows immediately from (27). On the other hand, $f(s - 1)/(s - 1)$, after some slight rearranging, can be written as

$$\begin{aligned} \frac{f(s-1)}{s-1} &= 3(n - \nu)sl[(l - 1)(l - 2(\lambda + 1)) + \lambda(\lambda + 1)] \\ &+ l\{l(l - 1)((s - 1)^2 - s) + 3(s - 1)(\lambda + 1)(l - 1 - \lambda)\} \\ &+ \lambda\{3sl(l - 1) - (\lambda + 1)(\lambda + 2)\} + l(l - 1)\{2s + \nu s(ls - 1) + 1\}. \end{aligned} \tag{28}$$

The expression in brackets is a polynomial of second degree in λ with a positive leading coefficient and with roots $\lambda = l - 2$, $\lambda = l - 1$. It is therefore nonnegative for $\lambda = 0, 1, \cdots, l - 1$. It is easily verified that the quantities in the three braces are nonnegative for $l > 1$, $s > 2$ and λ, ν satisfying (26). Furthermore, the last term is positive. It remains to consider (28) for the particular case $s = 2$. We have

$$\begin{aligned} f(1) \geqq 6l[(l - 1)(l - 2(\lambda + 1)) + \lambda(\lambda + 1)] \\ + l\{3(\lambda + 1)(l - 1 - \lambda) - l(l - 1)\} \\ + \lambda\{6l(l - 1) - (\lambda + 1)(\lambda + 2)\} + 5l(l - 1). \end{aligned}$$

The right-hand side is a polynomial $\varphi(\lambda)$ of third degree,

$$\varphi(\lambda) = -\lambda^3 + 3(l - 1)\lambda^2 - (3l^2 - 6l + 2)\lambda + l(l - 1)(5l - 4),$$

whose second derivative $\varphi''(\lambda)$ vanishes at $\lambda = l - 1$. It is easy to verify that $\varphi(\lambda)$ has its relative minimum at $\lambda = l - 1 - \sqrt{3}/3$. Hence $\varphi(\lambda) > 0$ for $\lambda = 0, 1, \cdots, l - 1$ follows from

$$\varphi(l - 2) = \varphi(l - 1) = 2l(2l - 1)(l - 1) > 0.$$

This completes the proof of our theorem.

For populations with serial correlation the result (24) is to be expected also on intuitive grounds; in fact, the systematic sample is spread more evenly through the population than the sample with multiple random starts which may contain elements very close together, giving about the same information. Our proof, however, does not make clear why (24) only holds for populations with a *convex* correlogram. That (24) does not generally hold for any monotone decreasing correlogram can readily be seen by trying to apply Cochran's lemma [1] to the linear form (19). It turns out that, for example, the sum of the first l coefficients of $NL/2$ is equal to

$$\frac{-l^2}{2(ls-1)}[(2n-1)s-1] < 0,$$

One might suspect that $EV_l^{(1)} \geqq EV_k^{(s)}$ for all populations with a *concave* decreasing correlogram. However, according to our theorem $EV_l^{(1)} < EV_k^{(s)}$ for the example of a linear correlogram, so that the conjecture is not generally true.

5. Asymptotic results in the case of an exponential correlogram. We assume that $\rho_z = e^{-\lambda z}(z = 1, \cdots, N-1)$ and that both l and n are large. For n, k large Cochran [1] showed that the expression in braces of (12) is approximately equal to $1 - 2/\lambda k + 2/(e^{\lambda k} - 1)$. Since the corresponding expression $1 - L_1$ in (17) is obtained by replacing k by l and n by ns, we find

$$(29) \qquad 1 - L_1 \sim 1 - \frac{2}{\lambda l} + \frac{2}{e^{\lambda l} - 1}.$$

On the other hand, replacing l by $k = ls$, s by 1 in the brace of (17), we obtain $1 - L_2$ of (18). Thus

$$1 - L_2 \sim 1 - \frac{2}{\lambda ls} + \frac{2}{e^{\lambda ls} - 1}.$$

Introducing $\rho = e^{-\lambda l}$, we see that the relative precision of systematic sampling over systematic sampling with multiple random starts

$$RP = \frac{EV_k^{(s)}}{EV_l^{(1)}} \sim \frac{1 + \dfrac{2}{s\log\rho} + \dfrac{2\rho^s}{1-\rho^s}}{1 + \dfrac{2}{\log\rho} + \dfrac{2\rho}{1-\rho}}$$

depends, apart from s, only on the correlation ρ of elements of a distance l apart. Clearly $\lim_{\rho\downarrow 0} RP = 1$; also, expanding numerator and denominator in power series, it is readily seen that $\lim_{\rho\uparrow 1} RP = s$. The numerical values in Table 1 show that the limit as $\rho \downarrow 0$ is approached rather slowly.

6. Concluding remark. When the statistician has a choice between systematic sampling and systematic sampling with multiple random starts, he is more

TABLE 1

Relative precision RP of systematic sampling over systematic sampling with multiple random starts for an exponential correlogram

s	ρ										
	.01	.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	1.34	1.53	1.66	1.80	1.87	1.92	1.95	1.97	1.99	2.00	2.00
5	1.56	1.98	2.34	2.92	3.43	3.88	4.25	4.55	4.76	4.92	4.99
10	1.63	2.13	2.58	3.40	4.26	5.19	6.23	7.32	8.39	9.31	9.85

likely to use the latter procedure because its variance can be estimated from the sample and the estimate is unbiased whatever be the form of the population. On the other hand, as we have seen in Section 5, systematic sampling is considerably more precise in the case of a population with an exponential correlogram. Thus, it may be worth while to try to find an estimate for the variance of systematic sampling which is at least consistent in some sense if the underlying assumption of an exponential correlogram is realized. In view of (17) or (29) this would involve estimating the correlation between the elements as well as σ^2.

REFERENCES

[1] W. G. Cochran, "Relative accuracy of systematic and stratified random samples for a certain class of populations," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 164–177.

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

[3] W. E. Deming, *Some Theory of Sampling*, John Wiley and Sons, 1950.

[4] H. L. Jones, "The application of sampling procedures to business operations," *J. Amer. Stat. Assoc.*, Vol. 50 (1955), pp. 763–774.

[5] H. L. Jones, "Investigating the properties of a sample mean by employing random subsample means," *J. Amer. Stat. Assoc.*, Vol. 51 (1956), pp. 54–83.

[6] W. G. Madow and L. H. Madow, "On the theory of systematic sampling," *Ann. Math. Stat.*, Vol. 15 (1944), pp. 1–24.

[7] A. Mood, *Introduction to the Theory of Statistics*, McGraw-Hill Co., 1950.