EFFECT OF AUTO-CORRELATIONS ON THE OPTIMUM ALLOCATIONS IN TWO PHASE STRATIFIED SAMPLING—A BAYESIAN APPROACH

IRWIN GUTTMAN* AND CHARLES D. PALIT

(Received Jan. 27, 1973; revised July 2, 1973)

Abstract

In this paper we study the problem of optimal allocation when the model for generation of observations in a particular stratum permits for auto-correlations. The object of the study is assumed to be the estimation of the population mean as precisely as possible.

0. Introduction

This paper is a continuation of the papers of Draper and Guttman [1] and Guttman and Palit [5] and discusses the two phase allocation problem for situations in which auto-correlations exist in within stratum samples. By a "two phase design," we mean a design in which two independent samples are selected at two different points in time from each of the strata of the (statified) population under study. In addition to the assumed independence between phases, we also assume independence between strata. Our model for the generation of the observations, which permits for auto-correlations, is further described in Section 1 below. The results of the first phase samples are used to arrive at an optimal design for the second sample. By optimal, we mean a second phase design which minimizes the future expectation of the two-sample (i.e., first and second phase samples) posterior variance—for further discussion, see Draper and Guttman [1] and Guttman and Palit [5].

1. The within stratum model

Let x_{ij} be the jth observation from the ith stratum, where $i=1, \cdots, H$ and $j=1, \cdots, n_i$. Denoting the stratum mean $\mathrm{E}(x_{ij})$ as μ_i, we

* Work supported in part by the National Research Council of Canada under Grant No. A-8743.

assume that the model for x_{ij} is

$$x_{ij}=\mu_i+e_{ij}\ , \tag{1.1}$$

where the error term e_{ij} is assumed to be generated by a first order auto-regressive process with parameter ρ_i, i.e.,

$$e_{ij}=\rho_i e_{i,j-1}+\varepsilon_{ij} \tag{1.2}$$

where the ε_{ij} are independent and normally distributed with mean zero and variance σ_i^2, for each $i=1,\cdots,H$, and with $-\infty<\rho_i<\infty$. (A discussion of such models is given by Kendall and Stuart ([7], Vol. 3, pp. 416–421).)

1.1. *The within stratum posterior distributions*

The subscript "i" is used to denote the stratum from which the sample is drawn. When we discuss the distribution theory for a single stratum, we will drop the subscript "i." Later, when we discuss more than one stratum, we will replace the subscript "i."

(i) The first phase posteriors

For convenience then, we now drop the subscript "i" and derive the within stratum likelihood function for the first phase, following the approach used by Zellner and Tiao [10], and Tiao and Tan [9].

Consider the transformation given by

$$z_0=x_1\ ,\qquad z_t=x_{t+1}-\rho x_t\ ,\quad t=1,2,\cdots,n-1=v_1\ , \tag{1.3}$$

where n is the number of first phase observations drawn from the stratum.

The expectations of these transformed variables may be written as

$$\mathrm{E}\,(z_t\,|\,\rho,\mu,m_1)=\begin{cases}\tau+m_1\ , & t=0\\ \tau\ , & 1\leqq t\leqq v_1\end{cases} \tag{1.4}$$

where $\tau=(1-\rho)\mu$, and m_1 is a level adjustment parameter introduced to take account of the fact that the process was at an unknown level when we began observing it. Now for $t=1,\cdots,v_1=n-1$, we have that

$$z_t=x_{t+1}-\rho x_t\ . \tag{1.5}$$

Using (1.1), we can rewrite (1.5) as

$$z_t=\mu+e_{t+1}-\rho(\mu+e_t)=\mu(1-\rho)+e_{t+1}-\rho e_t\ . \tag{1.5a}$$

Using the relationships given in (1.2), we substitute $(\rho e_t+\varepsilon_{t+1})$ for e_{t+1} in (1.5a) to obtain

$$z_t=\tau+\varepsilon_{t+1}\ ,\qquad t=1,\cdots,v_1\ . \tag{1.6}$$

Hence, the z_t can be considered as independent normal random variables with mean τ and variance σ^2. Using this result, we can write the likelihood function as

$$l(\mu, \rho, \sigma, m_1 | \boldsymbol{x}) \propto \sigma^{-(v_1+1)} \exp\left\{-\frac{1}{2\sigma^2}\left[(z_0-\tau-m_1)^2+\sum_{t=1}^{v_1}(z_t-\tau)^2\right]\right\} \tag{1.7}$$

where $-\infty<m_1<\infty$, $-\infty<\mu<\infty$, $-\infty<\rho<\infty$, $\sigma>0$, and $\tau=(1-\rho)\mu$. Thus if the prior knowledge about the parameters μ, m_1, σ and ρ can be represented by independent priors $p(\mu)$, $p(m_1)$, $p(\sigma)$, and $p(\rho)$ then the posterior distribution of μ, m_1, σ, and ρ can be written as:

$$p(\mu, \sigma, \rho, m_1 | \boldsymbol{x}) = l_1(m_1 | \rho, \mu, \sigma, \boldsymbol{x}) l_2(\rho, \mu, \sigma | \boldsymbol{x}) p(\sigma) p(\mu) p(m_1) p(\rho) \tag{1.8}$$

where

$$l_1(m_1 | \rho, \mu, \sigma, \boldsymbol{x}) \propto \sigma^{-1} \exp\left\{\frac{-(z_0-\tau-m_1)^2}{2\sigma^2}\right\} \tag{1.8a}$$

and

$$l_2(\rho, \mu, \sigma | \boldsymbol{x}) \propto \sigma^{-v_1} \exp\left\{\frac{-\sum_{t=1}^{v_1}(z_t-\tau)^2}{2\sigma^2}\right\} . \tag{1.8b}$$

Expressing our ignorance of m_1 by taking as the prior for m_1

$$p(m_1) \propto \text{constant} ,$$

and integrating (1.8) over m_1, we arrive at the posterior distribution for μ, σ, and ρ which can be written as

$$p(\mu, \sigma, \rho | \boldsymbol{x}) = l_2(\mu, \sigma, \rho | \boldsymbol{x}) p(\mu) p(\sigma) p(\rho) . \tag{1.9}$$

If in addition we choose the priors for μ, σ, and ρ to be locally uniform in μ, $\log\sigma$ and ρ, except* that $p(\rho)$ is equal to zero in the interval $1\pm\delta$ for $\delta>0$, then (1.9) above reduces to

$$p(\mu, \sigma, \rho | \boldsymbol{x}) \propto \sigma^{-(v_1+1)} \exp\left\{\frac{-1}{2\sigma^2}\sum_{t=1}^{v_1}(z_t-\tau)^2\right\} , \tag{1.10}$$

from which we obtain the joint posterior distribution of μ and ρ as

$$p(\mu, \rho | \boldsymbol{x}) \propto \left\{\sum_{t=1}^{v_1}(z_t-\bar{z})^2\right\}^{-v_1/2} \cdot \left\{1+\frac{v_1(v_1-1)(1-\rho)^2[\bar{z}/(1-\rho)-\mu]^2 \Big/ \sum_{t=1}^{v_1}(z_t-\bar{z})^2}{(v_1-1)}\right\}^{-v_1/2} \tag{1.11}$$

* This is equivalent to assuming that our prior knowledge is such that ρ is known not to lie in the interval $(1-\delta, 1+\delta)$, for some known δ.

where $\bar{z}=\sum_{t=1}^{v_1} z_t/v_1$. Integrating (1.11) with respect to μ yields the marginal posterior distribution of ρ as

$$p(\rho\,|\,\boldsymbol{x})\propto\left\{\sum_{t=1}^{v_1}(z_t-\bar{z})^2\right\}^{-(v_1-1)/2}\frac{1}{|1-\rho|}\,, \tag{1.12}$$

where $\rho\notin(1-\delta, 1+\delta)$. The reader should bear in mind that the z's are functions of ρ, besides being functions of the appropriate x's.

Using (1.12) and (1.11), we have that the conditional posterior of μ, given ρ, is such that

$$p(\mu\,|\,\rho, \boldsymbol{x})\propto\left\{1+\frac{v_1(v_1-1)(1-\rho)^2[\bar{z}/(1-\rho)-\mu]^2\Big/\sum_{t=1}^{v_1}(z_t-\bar{z})^2}{(v_1-1)}\right\}^{-v_1/2} \tag{1.13}$$

for $\rho\notin(1-\delta, 1+\delta)$, so that, conditional on ρ,

$$(\mu-\bar{z}/(1-\rho))\sqrt{\frac{v_1(v_1-1)(1-\rho)^2}{\sum(z_t-\bar{z})^2}} \tag{1.13a}$$

is distributed as a Student-t variable with (v_1-1) degrees of freedom.

From the above result it is easy to see that

$$\mathrm{E}\,(\mu\,|\,\rho, \boldsymbol{x})=\frac{\bar{z}}{1-\rho}=\frac{n\bar{x}}{n-1}-\frac{x_1-\rho x_n}{(n-1)(1-\rho)} \tag{1.14}$$

and that

$$\mathrm{V}\,(\mu\,|\,\rho, \boldsymbol{x})=\sum(z_t-\bar{z})^2/(1-\rho)^2v_1(v_1-3)=S^2(\boldsymbol{z}, \rho)/(n-1)(n-4) \tag{1.15}$$

where

$$S^2(\boldsymbol{z}, \rho)=\sum(z_t-\bar{z})^2/(1-\rho)^2\,. \tag{1.15a}$$

(ii) The second phase posteriors

Continuing to suppress the stratum subscript "i," we turn our attention to the distribution theory after the second phase. We shall use the letter y to denote the second phase observations, and we shall use the first phase posterior as the second phase prior.

As in the first phase the likelihood function for the second phase sample is

$$l(\mu, \rho, \sigma, m_2\,|\,\boldsymbol{y})\propto\sigma^{-(v_2+1)}\exp\left\{-\frac{1}{2\sigma^2}\left[(Z_0-\tau-m_2)^2+\sum_{t=1}^{v_2}(Z_t-\tau)^2\right]\right\}\,, \tag{1.16}$$

where

$v_2+1=N=$ the number of second phase observations drawn from the stratum

y_j = the jth observation in the second phase sample in the stratum

m_2 = the second phase level adjustment parameter

(1.16a) $$Z_0=y_1$$

(1.16b) $$Z_t=y_{t+1}-\rho y_t\ , \qquad t=1, 2, \cdots, v_2\ .$$

Again, we assume that the prior information for m_2 can be represented by the non-informative prior

$$p(m_2)=\text{constant}$$

and as with the first phase, integrate out m_2 from the joint posterior distribution of μ, σ, ρ, and m_2 leaving

(1.17) $$p(\mu, \sigma, \rho\,|\,\boldsymbol{y})=l_2(\mu, \rho, \sigma\,|\,\boldsymbol{y})p(\mu, \rho, \sigma)\ .$$

In this second phase however, we use as the prior for μ, σ, and ρ the posterior distribution for these parameters which result from the first phase as given in (1.10), so that (1.17) may be written as

(1.18) $$p(\mu, \sigma, \rho\,|\,\boldsymbol{y}, \boldsymbol{x})=l_2(\mu, \rho, \sigma\,|\,\boldsymbol{y})p(\mu, \rho, \sigma\,|\,\boldsymbol{x})\ ,$$

where of course, the function l_2 has the same functional form as (1.8b), with y's and Z's replacing x's and z's respectively and v_2 replacing v_1. Rewriting (1.18), then, we have

(1.18a) $$p(\mu, \sigma, \rho\,|\,\boldsymbol{y}, \boldsymbol{x})\propto\sigma^{-(v_1+v_2+1)} \cdot \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{t=1}^{v_1}(z_t-\tau)^2+\sum_{t=1}^{v_2}(Z_t-\tau)^2\right]\right\}\ .$$

After some straightforward integration, and performing some algebra, we find that

(1.19) $$p(\mu\,|\,\rho, \boldsymbol{x}, \boldsymbol{y})\propto\left\{\sum(z_t-\bar{z})^2+\sum(Z_t-\bar{Z})^2+\frac{v_1v_2}{v_1+v_2}(\bar{z}-\bar{Z})^2 +(v_1+v_2)(\tau-R)^2\right\}^{-(v_1+v_2)/2}$$

where

(1.19a) $$\bar{z}=\sum_1^{v_1}z_t/v_1\ , \qquad \bar{Z}=\sum_1^{v_2}Z_t/v_2\ , \quad \text{and} \quad R=\frac{v_1\bar{z}+v_2\bar{Z}}{v_1+v_2}\ .$$

Setting

(1.20) $$SS(\rho, z, \boldsymbol{Z})=\frac{\sum(z_t-\bar{z})^2+\sum(Z_t-\bar{Z})^2+(v_1v_2/(v_1+v_2))(\bar{z}-\bar{Z})^2}{(1-\rho)^2}\ ,$$

it is clear, from (1.19) and the fact that $\tau=(1-\rho)\mu$, that

$$\sqrt{(v_1+v_2)(v_1+v_2-1)\left(\mu-\frac{R}{1-\rho}\right)^2\Big/SS(\rho, \boldsymbol{z}, \boldsymbol{Z})} \tag{1.21}$$

has a Student-t distribution with (v_1+v_2-1) degrees of freedom.

Using (1.21) it follows that

$$\mathrm{E}\,(\mu\,|\,\rho, \boldsymbol{x}, \boldsymbol{y})=\frac{R}{1-\rho}=\frac{n\bar{x}+N\bar{y}}{v_1+v_2}-\frac{v_1(x_1-\rho x_n)+v_2(y_1-\rho y_n)}{(1-\rho)(v_1+v_2)} \tag{1.22}$$

and that

$$\mathrm{V}\,(\mu\,|\,\rho, \boldsymbol{x}, \boldsymbol{y})=\frac{SS(\rho z, Z)}{(v_1+v_2)(v_1+v_2-3)}\ . \tag{1.23}$$

We now return to the consideration of the case where H strata are used, and re-introduce the subscript "i." That is, the stratum mean is now denoted by μ_i, the within stratum variance by σ_i^2, the within stratum auto-correlation parameter by ρ_i, and the first and second phase samples from stratum i by $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, respectively. We further let $\boldsymbol{X}=(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_H)$ and $\boldsymbol{Y}=(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_H)$ denote the entire first and second phase samples, respectively.

The overall population mean is again denoted by μ and is defined as $\mu=\sum_{i=1}^{H}\pi_i\mu_i$, where π_i is the known proportion of the population in the ith stratum. Since the stratum means μ_i are independent (in virtue of the assumption that sampling between strata is done independently), the second phase posterior variance of μ is

$$\mathrm{V}\,(\mu\,|\,\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{\rho})=\sum\pi_i^2\,\mathrm{V}\,(\mu_i\,|\,\boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i)=\sum\pi_i^2\frac{SS_i(\rho_i, \boldsymbol{z}_i, \boldsymbol{Z}_i)}{(v_{1i}+v_{2i})(v_{1i}+v_{2i}-3)} \tag{1.24}$$

where $\boldsymbol{\rho}=(\rho_1,\cdots,\rho_H)'$.

(iii) ρ assumed known

Now suppose we have observed $\boldsymbol{X}$, but have not as yet observed $\boldsymbol{Y}=(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_H)$, where $\boldsymbol{y}_i$ is $(N_i\times 1)$. Then (1.24) is a random variable, and so to arrive at an optimum choice of N_i, we perform a preposterior analysis (see Raiffa and Schlaifer [8]) and minimize

$$E_F[\mathrm{V}\,(\mu\,|\,\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{\rho})\,|\,\boldsymbol{X}, \boldsymbol{\rho}]=\sum\pi_i^2\frac{E_F[SS_i(\rho_i, \boldsymbol{Z}_i, \boldsymbol{z}_i)\,|\,\boldsymbol{z}_i, \rho_i]}{(v_{1i}+v_{2i})(v_{1i}+v_{2i}-3)} \tag{1.25}$$

where the operation E_F is expectation with respect to the future or predictive distribution of $\boldsymbol{Y}$, given $\boldsymbol{X}$, $\boldsymbol{\rho}$ (or, equivalently, of $\boldsymbol{Z}$, given $\boldsymbol{z}$, $\boldsymbol{\rho}$ etc.). The minimization of (1.25) over N_i is subject to a cost restriction. Suppose C (in some monetory units) was available for the entire two-phase investigation, and that it was agreed to expend αC for the first phase, that is, denoting the cost per observation in stratum i by c_i, we have that

$$\sum_{i=1}^{H} c_i n_i = \alpha C \, . \tag{1.26}$$

Hence, the N_i are subject to the restriction

$$\sum c_i N_i = (1-\alpha)C = C' \, . \tag{1.27}$$

Now using known results of predictive distributions (see for example, Section 11.7 of Raiffa and Schlaifer [8]), it is easy to see that

$$E_F[\mathrm{V}\,(\tau_i \,|\, \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i) \,|\, \boldsymbol{x}_i, \rho_i] = \frac{\sum (z_{it}-\bar{z}_i)^2}{(v_{1i}+v_{2i})(v_{1i}-3)} \tag{1.28}$$

which, since $\mu_i = \tau_i/(1-\rho_i)$, yields

$$E_F[\mathrm{V}\,(\mu_i \,|\, \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i) \,|\, \boldsymbol{x}_i, \rho_i] = \frac{\sum (z_{it}-\bar{z}_i)^2/(1-\rho_i)^2}{(v_{1i}+v_{2i})(v_{1i}-3)} \, . \tag{1.28a}$$

Thus, on substituting (1.28a) in (1.24) we find that the predicted variance of μ, given the first phase sample, when $\boldsymbol{\rho}$ is known, is

$$\sum_{i=1}^{H} \frac{\pi_i^2 S_i^2(\boldsymbol{z}_i, \rho_i)}{(v_{1i}+v_{2i})(v_{1i}-3)} \tag{1.29}$$

where, as in (1.15a), we have let

$$S_i^2(\boldsymbol{z}_i, \rho_i) = \sum (z_{it}-\bar{z}_i)^2/(1-\rho_i)^2 \, . \tag{1.29a}$$

Minimizing (1.29) subject to (1.27) yields the solution

$$N_i = \left[\frac{C - 2\sum_{1}^{H} c_i}{c_i}\right] \frac{\pi_i u_i \sqrt{c_i}}{\sum \pi_i u_i \sqrt{c_i}} - (n_i - 2) \tag{1.30}$$

where

$$u_i^2 = S_i^2(\boldsymbol{z}_i, \rho_i)/(v_{1i}-3) \, . \tag{1.30a}$$

The solution (1.30) can yield negative values for N_i. As in Draper and Guttman [1] and Guttman and Palit [5], we interpret such an event to mean that (for the given budget) the ith stratum was over-sampled in the first phase and no further observations should be taken from that stratum. Consequently, for the strata that have $N_i < 0$, we set the N_i equal to zero, and recompute the allocations for the *remaining* strata by minimizing (1.29) over the corresponding N_j's.

This does not yet complete the procedure, for the restriction on the N_i for the model (1.1)–(1.2) considered here, must be extended so that if we do not select at least two observations from a stratum in the second phase, we do not select any. The reason for this is as follows:

Suppose for the ith stratum, $N_i=1$, then for the ith stratum the second phase posterior for μ_i, ρ_i and σ_i is exactly the same as the first phase posterior. This is easily seen from the following. We recall from (1.18) that

$$p(\mu_i, \rho_i, \sigma_i | \boldsymbol{y}_i, \boldsymbol{x}_i) = l_2(\mu_i, \rho_i, \sigma_i | \boldsymbol{y}_i) p(\mu_i, \rho_i, \sigma_i | \boldsymbol{x}_i)$$

where from (1.8b)

$$l_2(\mu_i, \rho_i, \sigma_i | \boldsymbol{y}_i) \propto \sigma_i^{-(N_i-1)} \exp\left\{-\frac{1}{2\sigma_i^2}\left[\sum_{t=1}^{N_i-1} (Z_{it}-\tau_i)^2\right]\right\} .$$

It is easy to see that when $N_i=1$, then

$$l_2(\mu_i, \rho_i, \sigma_i | \boldsymbol{y}_i) \propto \text{constant}$$

and hence for the case $N_i=1$,

$$p(\mu_i, \rho_i, \sigma_i | \boldsymbol{y}_i, \boldsymbol{x}_i) = p(\mu_i, \rho_i, \sigma_i | \boldsymbol{x}_i) .$$

Since the selection of only one observation from a stratum in the second phase adds nothing to the precision of our estimator, it would be pointless to select it. For this reason we adopt the rule that if N_i is such that

$$0 < N_i < 1.5$$

then we set N_i equal to zero and recompute the allocations for the other strata with the original budget, and without considering the stratum for which we have set $N_i=0$. If more than one stratum has an N_i such that $0<N_i<1.5$, then we set equal to zero only the *lowest* N_i and drop only that stratum from the system for which allocations are recomputed. If one or more strata still have their N_i such that $0<N_i<1.5$, we again remove the stratum with the lowest N_i and recompute. We continue to do this until all the N_i are acceptable. Finally, we round up or down to arrive at the final integer allocation.

(iv) ρ unknown

In the previous sub-section we obtained an optimal allocation (formula (1.30)) for selecting within stratum sample sizes for the second phase, given that in each stratum we are dealing with a first order auto-regressive process whose parameter ρ_i is known, and in fact, known not to lie in some neighbourhood of one. We now suppose that the ρ_i are unknown, but that it is known that ρ_i does not lie in a small interval about 1. We will see that the optimal allocation arrived at by a preposterior analysis involves the computation of H, N_i-dimensional integrals, where N_i is the second phase sample size for the ith stratum, and $i=1,\cdots,H$. Even with the aid of high speed computer, this is a cumbersome procedure (the N_i are to be determined, so for each trial set, H multiple integrals of dimension $(N_1,\cdots,N_H)$ must be evaluated etc.), so that after discussing the preposterior analysis in this subsection, we will propose some alternative procedures in Section 2.

But let us first discuss and illustrate the problems mentioned in the last paragraph. From (1.18a), it is easy to see that

$$(1.31)\quad p(\mu_i,\rho_i|\boldsymbol{x}_i,\boldsymbol{y}_i)\propto\left\{\left[\sum_{t=1}^{v_{1i}}(z_{ti}-\bar{z}_i)^2+v_{1i}(\bar{z}_i-\mu_i(1-\rho_i))^2+\sum_{t=1}^{v_{2i}}(Z_{ti}-\bar{Z}_i)^2\right.\right.$$
$$\left.\left.+v_{2i}(\bar{Z}_i-\mu_i(1-\rho_i))^2\right]\right\}^{-(v_{1i}+v_{2i})/2}.$$

Now, the expression in the square brackets in (1.31) may be written as

$$(1.32)\qquad A_i\left(\rho_i-\frac{B_i}{A_i}\right)^2+C_i-\frac{B_i^2}{A_i}$$

where

$$(1.32a)\quad A_i=S^2_{x_{i,t}}+S^2_{y_{i,t}}+v_{1i}(\mu_i-\bar{x}_{i,t})^2+v_{2i}(\mu_i-\bar{y}_{i,t})^2$$

$$(1.32b)\quad B_i=CS_{x_i}+CS_{y_i}+v_{1i}(\mu_i-\bar{x}_{i,t+1})(\mu_i-\bar{x}_{i,t})+v_{2i}(\mu_i-\bar{y}_{i,t+1})(\mu_i-\bar{y}_{i,t})$$

$$(1.32c)\quad C_i=S^2_{x_{i,t+1}}+S^2_{y_{i,t+1}}+v_{1i}(\mu_i-\bar{x}_{i,t+1})^2+v_{2i}(\mu_i-\bar{y}_{i,t+1})^2$$

with

$$(1.32d)\qquad S^2_{y_{i,t}}=\sum_{t=1}^{v_{2i}}(y_{it}-\bar{y}_{i,t})^2\,,\qquad \bar{y}_{i,t}=\frac{1}{v_{2i}}\sum_{t=1}^{v_{2i}}y_{it}$$

$$(1.32e)\qquad S^2_{y_{i,t+1}}=\sum_{t=1}^{v_{2i}}(y_{i,t+1}-\bar{y}_{i,t+1})^2\,,\qquad \bar{y}_{i,t+1}=\frac{1}{v_{2i}}\sum_{t=1}^{v_{2i}}y_{i,t+1}$$

and the obvious similar definitions for $S^2_{x_{i,t}}$, $S^2_{x_{i,t+1}}$, $\bar{x}_{i,t}$ and $\bar{x}_{i,t+1}$, with

$$(1.32f)\qquad v_{1i}=n_i-1\,,\qquad v_{2i}=N_i-1$$

and where

$$(1.32g)\qquad \begin{aligned}CS_{x_i}&=\sum_{t=1}^{v_{1i}}(x_{it}-\bar{x}_{i,t})(x_{i,t+1}-\bar{x}_{i,t+1})\,,\\ CS_{y_i}&=\sum_{t=1}^{v_{2i}}(y_{it}-\bar{y}_{i,t})(y_{i,t+1}-\bar{y}_{i,t+1})\,.\end{aligned}$$

Integrating (1.31) over the range of ρ_i, and assuming that $(1-\delta_i, 1+\delta_i)$

is a very small interval, (that is, assuming that ρ_i does "not equal 1"), we find that

$$p(\mu_i\,|\,\boldsymbol{x}_i,\boldsymbol{y}_i)\propto A_i^{-1/2}[C_i-B_i^2/A_i]^{-(v_{1i}+v_{2i}-1)/2}\,. \tag{1.33}$$

Recalling the definitions of A_i, B_i and C_i given in (1.32a, b and c), this implies that the two phase posterior variance of μ_i is a function of CS_{y_i}, $CS^2_{y_{i,t}}$, $S^2_{y_{i,t+1}}$, $\bar{y}_{i,t}$ and $\bar{y}_{i,t+1}$, that is, once $\boldsymbol{x}_i$ is observed, and $\boldsymbol{y}_i$ is as yet unobserved the posterior variance of μ_i is a random variable. To find its future distribution, we need a closed form for the joint distribution of the five quantities mentioned above, given μ_i, ρ_i and σ_i^2. In principle, however,

$$E_F[\mathrm{Var}\,(\mu_i\,|\,\boldsymbol{x}_i,\boldsymbol{y}_i)\,|\,\boldsymbol{x}_i] \tag{1.34}$$

can be found directly by integrating the product of the second phase posterior variance and the joint future distribution of $\boldsymbol{y}_i$ over the range of $\boldsymbol{y}_i$, i.e. by evaluating

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathrm{Var}\,(\mu_i\,|\,\boldsymbol{x}_i,\boldsymbol{y}_i)p(\boldsymbol{y}_i\,|\,\boldsymbol{x}_i)d\boldsymbol{y}_i\ , \tag{1.35}$$

an N_i-dimensional integral, where

$$p(\boldsymbol{y}_i\,|\,\boldsymbol{x}_i)=\int\int\int l_2(\mu_i,\rho_i,\sigma_i\,|\,\boldsymbol{y}_i)p(\mu_i,\rho_i,\sigma_i\,|\,\boldsymbol{x}_i)d\mu_i d\sigma_i d\rho_i\ . \tag{1.36}$$

The difficulties involved here are many—for one, note that the N_i are as yet unknown—a trial set would have to be used, (1.36) and (1.35) computed, and then, using these results, the corresponding value of

$$\sum_{i=1}^{H}E_F[\mathrm{Var}\,(\mu_i\,|\,\boldsymbol{x}_i,\boldsymbol{y}_i)\,|\,\boldsymbol{x}_i] \tag{1.37}$$

computed. A search procedure would then be instituted—another set of N_i would be closen appropriately (eg, a Gauss-Newton procedure could be employed) and the set $\{N_i\}$ that minimizes (1.37) could be found by an appropriate iterative procedure. But this is indeed combersome and time consuming, and so in the next section we propose three alternative procedures which are of greater practical value in determining an "optimal" set of N_i. The third of these procedures has been utilized in Guttman and Palit [4].

2. Alternatives and approximations

The difficulties encountered in the evaluation of $E_F[\mathrm{Var}\,(\mu_i\,|\,\boldsymbol{x}_i,\boldsymbol{y}_i)\,|\,\boldsymbol{x}_i]$ and hence in the evaluation of the future expectation of the two phase posterior variance of the population mean μ, which is given by

$$\sum_{i=1}^{H} \pi_i^2 E_F[\mathrm{Var}\,(\mu_i | \boldsymbol{x}_i, \boldsymbol{y}_i) | \boldsymbol{x}_i] , \tag{2.1}$$

lead to the consideration of alternative and approximate procedures for optimizing designs. Three alternative procedures are considered in this section.

Two of these procedures consist of substituting estimates for the ρ_i, appearing in $E_F[\mathrm{Var}\,(\mu_i | \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i) | \boldsymbol{x}_i, \rho_i]$. One of the estimators for ρ_i, is called $\hat{\rho}_i$, and is of a familiar form (see for example Jenkins and Watts [6], p. 180), viz.

$$\hat{\rho}_i = \frac{CS_{x_i}}{S^2_{x_{it}}} \tag{2.2}$$

where CS_{x_i} and $S^2_{x_{it}}$ are defined in (1.32). Another estimator for ρ_i which we shall discuss is the first phase posterior expectation for ρ_i, given $\boldsymbol{x}_i$, that is

$$\hat{\hat{\rho}}_i = \mathrm{E}\,(\rho_i | \boldsymbol{x}_i) . \tag{2.3}$$

The third procedure attempts to make use of more of the information about ρ_i which is contained in the sample $\boldsymbol{x}_i$. In this procedure, the future expectation of the posterior variance of μ_i, given $(\boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i)$, that is the future expectation of

$$\mathrm{V}\,(\mu_i | \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i) = \int_{-\infty}^{\infty} [\mu_i - \mathrm{E}\,(\mu_i | \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i)]^2 p(\mu_i | \boldsymbol{x}_i, \boldsymbol{y}_i, \rho_i) d\mu_i , \tag{2.4}$$

is averaged over the posterior distribution of ρ_i after the first phase, i.e., over $p(\rho_i | \boldsymbol{x}_i)$. A similar process was used in Sections 3 and 4 of Guttman and Palit [4].

Since this last alternative makes "greater use" of the sample information about ρ_i, we might expect the results to be superior to the first two procedures. An empirical investigation reveals that this is generally the case. Before discussing the results of this investigation, we present the computational algorithms associated with each of the three procedures mentioned above. The restrictions regarding admissible values for the N_i which apply to (1.30), and discussed below (1.30), continue to apply here.

2.1. *The computational algorithms*

(i) Substitution of $\hat{\rho}_i$ for ρ_i

The procedure in which $\hat{\rho}_i$ is substituted for ρ_i in (1.30) produces the algorithm

(2.5) $$N_i(\hat{\rho}_i)=\frac{\left[C-2\sum_1^H c_i\right]}{c_i}\frac{\pi_i\hat{u}_i\sqrt{c_i}}{\sum \pi_i\hat{u}_i\sqrt{c_i}}-(n_i-2)\ ,$$

where

(2.5a) $$\hat{u}_i^2=S_i^2(\boldsymbol{z}_i,\ \hat{\rho}_i)/(v_{1i}-3)$$

and where $S_i^2(\boldsymbol{z}_i,\ \rho_i)$ is as defined in (1.29a).

We have found (2.5) by minimizing $\sum_{i=1}^H (\pi_i^2 S_i^2(\boldsymbol{z}_i,\ \hat{\rho}_i))/((v_{1i}-3)(v_{1i}-1+N_i))$ (see (1.29)-(1.29a)).

(ii) Substitution of $\hat{\rho}_i$ for ρ_i

The procedure in which $\mathrm{E}(\rho_i|\boldsymbol{x}_i)=\hat{\rho}_i$ is substituted for ρ_i produces the algorithm

(2.6) $$N_i(\hat{\rho}_i)=\frac{\left[C-2\sum_1^H c_i\right]}{c_i}\frac{\pi_i\hat{u}_i\sqrt{c_i}}{\sum \pi_i\hat{u}_i\sqrt{c_i}}-(n_i-2)$$

with

(2.6a) $$\hat{u}_i^2=S_i^2(\boldsymbol{z}_i,\ \hat{\rho}_i)/(v_{1i}-3)\ .$$

We have found (2.6) by minimizing $\sum_{i=1}^H (\pi_i^2 S_i^2(\boldsymbol{z}_i,\ \hat{\rho}_i))/((v_{1i}-3)(v_{1i}-1+N_i))$ (see (1.29)-(1.29a)). Hence for the above procedure, we need $\mathrm{E}(\rho_i|\boldsymbol{x}_i)$. The reader will recall that

(2.7) $$p(\rho_i|\boldsymbol{x}_i)\propto\left\{\sum_{t=1}^{v_1}(z_{it}-\bar{z}_i)^2\right\}^{-(v_{1i}-1)/2}\frac{1}{|1-\rho_i|}\ .$$

Since $z_{it}=x_{i,t+1}-\rho x_{it}$, it is easily seen that

(2.8) $$\sum_{t=1}^{v_1}(z_{it}-\bar{z}_i)^2=S_{x_{it}}^2\left[\rho_i-\left(\sum_{t=1}^{v_1}a_{i,t+1}b_{it}/S_{x_{it}}^2\right)\right]^2+S_{x_{i,t+1}}^2-[(\sum a_{i,t+1}b_{it})^2/S_{x_{it}}^2]$$

where

(2.8a) $$a_{i,t+1}=(x_{i,t+1}-\bar{x}_{i,t+1})\qquad b_{it}=(x_{it}-\bar{x}_{it})$$

with $\bar{x}_{i,t+1}$, $\bar{x}_{it}$, $S_{x_{i,t}}^2$ and $S_{x_{i,t+1}}^2$ as defined in (1.32). This mean that we may rewrite (2.7) as

(2.9) $$p(\rho_i|\boldsymbol{x}_i)\propto|1-\rho_i|^{-1}\left\{1+\frac{L_i(\rho_i-D_i)^2}{v_{1i}-2}\right\}^{-(v_{1i}-1)/2}$$

where

(2.9a) $$D_i=\sum_{t=1}^{v_1}a_{i,t+1}b_{it}/S_{x_{it}}^2\qquad L_i=(v_{1i}-2)(S_{x_{it}}^2)/[S_{x_{i,t+1}}^2-S_{x_{it}}^2D_i^2]\ .$$

Using (2.9) we have that

$$\text{(2.10)}\qquad \mathrm{E}(1-\rho_i|\boldsymbol{x}_i)=K_i\int_{\Omega_i}\frac{1-\rho_i}{|1-\rho_i|}\left\{1+\frac{L_i(\rho_i-D_i)^2}{v_{1i}-2}\right\}^{-(v_{1i}-1)/2}d\rho_i$$

where Ω_i is the union of the two regions $(-\infty, 1-\delta_i)$ and $(1+\delta_i, \infty)$. The constant K_i is the normalizing constant for (2.9) and is such that

$$\text{(2.11)}\qquad K_i^{-1}=\int_{\Omega_i}|1-\rho_i|^{-1}\left\{1+\frac{L_i(\rho_i-d_i)^2}{v_{1i}-2}\right\}^{-(v_{1i}-1)/2}d\rho_i\,.$$

Now we may write

$$\text{(2.12)}\qquad \mathrm{E}(1-\rho_i|\boldsymbol{x}_i)=\frac{K_iK_i'}{\sqrt{L_i}}\{\mathrm{P}[t_i\leqq(1-D_i-\delta_i)\sqrt{L_i}]-\mathrm{P}[t_i\geqq(1-D_i+\delta_i)\sqrt{L_i}]\}$$

where K_i' is such that

$$\text{(2.12a)}\qquad (K_i')^{-1}=\Gamma\left(\frac{v_{1i}-1}{2}\right)\Big/\Gamma\left(\frac{v_{1i}-2}{2}\right)\sqrt{\pi(v_{1i}-2)}\,.$$

Hence, if δ_i small, we may write

$$\text{(2.13)}\qquad \mathrm{E}(1-\rho_i)\cong\frac{K_iK_i'}{\sqrt{L_i}}\{2\,\mathrm{P}[t_i\leqq\sqrt{L_i}(1-D_i)]-1\}$$

so that, for small δ_i,

$$\text{(2.14)}\qquad \hat{\rho}_i=\mathrm{E}(\rho_i|\boldsymbol{x}_i)\cong1-\frac{K_iK_i'}{\sqrt{L_i}}\{2\,\mathrm{P}[t_i\leqq\sqrt{L_i}(1-D_i)]-1\}$$

where $\mathrm{P}(t_i\leqq t_{0i})$ is the cumulative distribution of the Student-t with $(v_{1i}-2)$ degrees of freedom.

(iii) Integrating over the first phase posterior for ρ_i

The third procedure is as follows. Acting as if we know the N_i, we find the form of the posterior of μ_i, given $\boldsymbol{x}_i$, $\boldsymbol{y}_i$ and ρ_i as in (1.21). We then find the variance of this posterior distribution of μ_i, as in (1.23). Hence, we may find the posterior of $\mu=\sum\pi_i\mu_i$, which is given in (1.24). Note that this is a function of $\boldsymbol{X}=(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_H)$, $\rho=(\rho_1,\cdots,\rho_H)'$ and the as yet unobserved $\boldsymbol{Y}=(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_H)$. We proceed to find the future expectation of the posterior variance of μ, and this result has been given in (1.29)–(1.29a), and we repeat it here, that is, we have

$$\text{(2.15)}\qquad E_F[\mathrm{V}(\mu|\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho})|\boldsymbol{x},\boldsymbol{\rho}]=\sum_{i=1}^{H}\frac{\pi_i^2S_i^2(\boldsymbol{z}_t,\rho_i)}{(v_{1i}-3)(v_{1i}-1+N_i)}\,.$$

Finally, instead of proceeding as in (i) and (ii), we minimize

$$(2.16)\qquad \int\cdots\int E_F[\mathrm{V}(\mu|\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho}|\boldsymbol{x},\boldsymbol{\rho})p(\rho|\boldsymbol{x})d\boldsymbol{\rho} = \sum_{i=1}^{H}\frac{\pi_i^2}{(v_{1i}-3)(v_{1i}-1+N_i)}\int_{\Omega_i} S_i^2(\boldsymbol{z}_i,\rho_i)p(\rho_i|\boldsymbol{x}_i)d\rho_i$$

where $S_i^2(\boldsymbol{z}_i,\rho_i)=\sum_{t=1}^{v_{1i}}(z_{it}-\bar{z}_i)^2/(1-\rho_i)^2$ and $p(\rho_i|\boldsymbol{x}_i)$ is given in (2.7). Now from (2.8) and (2.9a), it is easily seen that

$$(2.17)\qquad S_i^2(\boldsymbol{z}_i,\rho_i)=\frac{1}{(1-\rho_i)^2}[S^2_{x_{i,t+1}}-D_i^2S^2_{x_{it}}]\left\{1+\frac{L_i(\rho_i-D_i)^2}{v_{1i}-2}\right\}$$

so that we may rewrite (2.16) as

$$(2.18)\qquad \sum_{i=1}^{H}\frac{\pi_i^2q_i^2}{(v_{1i}-1+N_i)}$$

where

$$(2.18a)\qquad q_i^2=\frac{K_i[S^2_{x_{i,t+1}}-D_i^2S^2_{x_{it}}]}{(v_{1i}-3)}\int_{\Omega_i}\frac{d\rho_i}{|1-\rho_i|^3}\left\{1+\frac{L_i(\rho_i-D_i)^2}{v_{1i}-2}\right\}^{-(v_{1i}-3)/2}.$$

Minimizing (2.18), subject to $\sum c_iN_i=C'=(1-\alpha)C$, we find

$$(2.19)\qquad N_i=\left[\frac{C-2\sum_1^H c_i}{c_i}\right]\frac{\pi_iq_i\sqrt{c_i}}{\sum\pi_iq_i\sqrt{c_i}}-(n_i-2)\,.$$

3. A Monte Carlo study of the procedures

In this section, a Monte Carlo study is made to compare the procedures discussed in the previous sections. This study involves six strata, and for each of the six strata, we generate a different sample using a first order auto-regressive process with known parameters μ_i, ρ_i, σ_i, $i=1,\cdots,6$. The stratified sample so obtained is used as a first phase sample in the allocation procedures of Section 2 to produce the second phase allocations $\boldsymbol{N}(\hat{\rho})$, $\boldsymbol{N}(\hat{\hat{\rho}})$, and N^* of algorithms (2.5), (2.6) and (2.19) respectively.

In addition, we use the sample with the algorithm in (1.30) to compute the second phase allocations, given the true value of the ρ_i, that is, the value of the ρ_i used to produce the first phase sample. We call this allocation set $\boldsymbol{N}(\boldsymbol{\rho})$.

The cost of sampling within each stratum is taken as one unit and the second phase budget is set at 600 units. For each set of 18 population parameters, $(\mu_1,\rho_1,\sigma_1,\cdots,\mu_i,\rho_i,\sigma_i,\cdots,\mu_6,\rho_6,\sigma_6)$, 14 first phase samples of $n_i\equiv m$ observations each are generated and used to produce second phase allocations. In this study, we let m take the values 5(5)25.

The results for $n_i=5$ and 15 are displayed in Tables 3.1 and 3.2, respectively. (The results for $n_i=10$, 20 and 25 are given in Guttman and Palit [5]: copies available on request.)

Now the reader will recall from (1.29), that the future expectation of the posterior variance (given $\boldsymbol{X}$ and $\boldsymbol{\rho}$) is

$$E_F[\mathrm{Var}(\mu|\boldsymbol{X},\boldsymbol{Y},\boldsymbol{\rho},\boldsymbol{N})|\boldsymbol{X},\boldsymbol{\rho}]=\sum_{i=1}^{6}\frac{\pi_i^2 S_i^2(\boldsymbol{z}_i,\rho_i)}{[V_{1i}+N_i-1][v_{1i}-3]}\,. \tag{3.1}$$

To help us evaluate the efficiency of an allocation method, we also tabulate in Tables 3.1–3.2, the value of (3.1) when $\boldsymbol{N}$ of (3.1) is set equal to each of the allocations $\boldsymbol{N}(\hat{\boldsymbol{\rho}})$, $\boldsymbol{N}(\hat{\boldsymbol{\rho}})$, $\boldsymbol{N}^*$ and $\boldsymbol{N}(\boldsymbol{\rho})$. The vector $\boldsymbol{\rho}$ is the vector used in generating the samples $\boldsymbol{X}$. We also tabulate in the tables the measure of efficiency for any allocation set, say $\boldsymbol{N}^{**}$, given by

$$\text{Efficiency of } \boldsymbol{N}^{**}=\frac{E_F[\mathrm{Var}(\mu|\boldsymbol{X},\boldsymbol{Y},\boldsymbol{\rho},\boldsymbol{N}(\boldsymbol{\rho}))|\boldsymbol{X},\boldsymbol{\rho}]}{E_F[\mathrm{Var}(\mu|\boldsymbol{X},\boldsymbol{Y},\boldsymbol{\rho},\boldsymbol{N}^{**})|\boldsymbol{X},\boldsymbol{\rho}]}\,. \tag{3.2}$$

The first six lines of each of the two tables show the parameters used to generate the first phase samples. The second phase allocations are shown in succeeding lines. The legends used for the parameters in the first six lines are, respectively:

c is the cost of taking an observation in the stratum. (We take this to be one always)

π is π_i the proportion of the population in the ith stratum

Var is the variance of ε_{ij} in (1.2)

U is the stratum mean μ_i

ρ is the auto-regressive parameter ρ_i, and

N is the first phase size n_i.

The rest of the table is divided into 14 groups of four lines each. Each of these groups contains the four allocation sets which are produced from the same first phase sample using the four allocation procedures previously discussed, the predicted variance for the allocation as defined in (3.1) and the efficiency ratio defined in (3.2). The legends used are as follows:

(i) The first line in every group is labelled NN. This line contains the allocation set for the ρ_i known procedure in (1.30), and its predicted variance.

(ii) The second line of the group *labelled ENN contains* the allocation set for the procedure in (2.6), in which $\hat{\rho}_i$ is substituted for ρ_i, its associated predicted variance, and its efficiency.

(iii) The third line *labelled CNN contains* the procedure in (2.5), where $\hat{\rho}_i$ is substituted for ρ_i.

(iv) The fourth line *labelled ANN contains* the equivalent results for

Table 3.1 Summary of Monte Carlo results for $n_i=5$

Budget=600 6 strata considered

Stratum No.	(1)	(2)	(3)	(4)	(5)	(6)		
c	1.000	1.000	1.000	1.000	1.000	1.000		
π	0.150	0.200	0.100	0.150	0.250	0.150		
Var	1.000	2.000	4.000	16.000	25.000	36.000		
U	5.000	6.000	7.000	8.000	9.000	10.000		
ρ	0.000	0.100	−0.200	0.500	0.600	0.300		
N	5	5	5	5	5	5		
NN	27.923	27.544	19.068	96.843	369.551	59.072	0.04128	1.00000
ENN	37.899	13.566	20.580	66.471	438.050	23.433	0.04705	0.87748
CNN	44.260	13.418	22.308	78.025	418.535	23.453	0.04637	0.89019
ANN	57.230	52.200	57.927	94.775	262.138	75.731	0.04777	0.86416
NN	15.153	29.381	15.189	210.247	127.293	202.736	0.07591	1.00000
ENN	41.191	63.828	16.984	73.326	250.640	154.032	0.11938	0.63582
CNN	36.751	73.580	7.726	53.292	307.668	120.983	0.15436	0.49176
ANN	33.858	33.077	21.401	160.853	108.849	241.962	0.08029	0.94544
NN	2.100	28.567	18.967	58.329	362.922	129.116	0.09129	1.00000
ENN	1.922	74.287	69.267	25.299	292.506	136.719	0.10869	0.83989
CNN	0.000	79.568	78.516	21.033	294.011	126.873	0.11350	0.80429
ANN	1.723	59.959	40.272	41.294	275.757	180.995	0.10231	0.89228
NN	5.605	48.197	5.628	70.700	364.761	105.108	0.11506	1.00000
ENN	8.365	82.456	5.130	25.265	324.956	153.828	0.13506	0.85190
CNN	7.857	79.021	3.428	21.320	317.270	171.104	0.14208	0.80983
ANN	16.697	85.938	18.530	51.827	272.760	154.248	0.13042	0.88218
NN	14.523	32.956	22.994	180.676	175.648	173.203	0.07703	1.00000
ENN	12.239	29.807	24.319	51.103	401.996	80.535	0.14437	0.53353
CNN	9.445	25.227	19.376	39.419	442.855	63.678	0.17863	0.43120
ANN	32.172	61.252	59.035	138.549	103.111	205.881	0.09067	0.84949
NN	7.808	36.267	27.719	37.325	392.545	98.336	0.09209	1.00000
ENN	12.336	38.813	57.193	44.084	383.111	64.463	0.09718	0.94766
CNN	10.963	33.752	52.266	47.208	411.086	44.726	0.10323	0.89206
ANN	20.187	62.819	86.822	39.683	321.602	68.886	0.10459	0.88045
NN	13.461	44.075	20.695	235.001	177.062	109.706	0.05263	1.00000
ENN	16.758	67.532	41.301	159.344	198.685	116.380	0.05739	0.91707
CNN	14.683	65.027	38.306	153.004	217.244	111.737	0.05812	0.90548
ANN	26.572	78.002	54.161	185.876	120.128	135.260	0.05974	0.88104
NN	11.143	30.680	19.093	70.243	314.617	154.225	0.10547	1.00000
ENN	17.931	105.585	111.415	47.825	216.847	100.397	0.14127	0.74659
CNN	16.034	118.084	118.073	44.398	218.052	85.359	0.14903	0.70773
ANN	27.963	52.841	39.525	64.421	230.547	184.702	0.11630	0.90687
NN	4.725	55.543	9.802	56.736	368.040	105.153	0.06886	1.00000
ENN	10.102	118.282	49.577	107.578	98.493	215.968	0.16469	0.41811
CNN	3.015	116.891	55.401	121.580	75.343	227.769	0.20924	0.32909
ANN	4.864	132.466	33.656	70.012	170.426	188.576	0.10486	0.65666
NN	15.964	38.740	19.083	95.143	197.600	233.469	0.06142	1.00000
ENN	33.961	53.879	18.148	113.590	166.845	213.577	0.06372	0.96389
CNN	33.778	55.835	14.151	130.815	165.351	200.070	0.06502	0.94467
ANN	24.583	61.889	46.017	76.590	120.683	270.239	0.06965	0.88178

Table 3.1 (Continued)

NN	5.249	18.603	17.335	145.275	326.240	87.298	0.15011	1.00000
ENN	5.697	50.969	51.436	78.329	197.813	215.756	0.21167	0.70915
CNN	3.396	53.315	49.908	68.302	188.891	236.187	0.22688	0.66162
ANN	9.341	42.544	58.977	113.710	238.663	136.765	0.17432	0.86108
NN	8.851	42.973	20.509	136.803	281.933	108.931	0.02893	1.00000
ENN	11.254	73.821	44.349	126.374	303.982	40.220	0.03532	0.81913
CNN	9.122	69.392	36.876	122.393	338.610	23.607	0.04302	0.67244
ANN	21.663	84.536	67.233	132.457	231.033	63.077	0.03363	0.86012
NN	10.770	21.831	15.035	143.078	310.306	98.979	0.05234	1.00000
ENN	23.111	44.733	25.357	97.378	239.630	169.791	0.06004	0.87166
CNN	22.326	44.102	21.925	92.048	240.477	179.122	0.06084	0.86025
ANN	27.193	41.596	42.798	122.682	221.373	144.358	0.05971	0.87653
NN	7.913	27.643	12.217	100.951	332.999	118.278	0.13858	1.00000
ENN	5.880	246.509	21.517	41.753	184.576	99.766	0.22717	0.61003
CNN	4.376	255.226	19.727	37.664	188.083	94.924	0.23247	0.59611
ANN	17.729	37.216	42.070	82.225	250.804	169.956	0.15492	0.89451

Table 3.2 Summary of Monte Carlo results for $n_i=15$

Budget=600 6 strata considered

Stratum No.	(1)	(2)	(3)	(4)	(5)	(6)		
c	1.000	1.000	1.000	1.000	1.000	1.000		
π	0.150	0.200	0.100	0.150	0.250	0.150		
Var	1.000	2.000	4.000	16.000	25.000	36.000		
u	5.000	6.000	7.000	8.000	9.000	10.000		
ρ	0.000	0.100	−0.200	0.500	0.600	0.300		
n	15	15	15	15	15	15		
NN	0.000	34.290	10.517	98.875	381.695	74.623	0.04323	1.00000
ENN	0.000	112.460	13.437	145.017	260.343	68.742	0.05065	0.85343
CNN	3.161	100.076	20.167	131.562	259.902	85.133	0.04985	0.86718
ANN	14.038	62.887	42.252	97.096	279.401	104.326	0.04822	0.89646
NN	6.354	17.995	12.670	157.862	279.554	125.564	0.04640	1.00000
ENN	14.802	42.605	39.318	226.844	155.167	121.264	0.05592	0.82985
CNN	17.582	41.753	40.926	199.634	165.710	134.395	0.05397	0.85987
ANN	23.801	40.224	43.238	148.002	182.332	162.403	0.05262	0.88178
NN	0.000	20.601	2.846	196.490	252.211	127.851	0.06409	1.00000
ENN	0.000	8.789	4.592	282.488	208.525	95.606	0.06879	0.93172
CNN	3.723	15.251	8.244	252.335	199.141	121.306	0.06676	0.96011
ANN	12.496	40.169	21.821	179.572	179.855	166.087	0.06973	0.91920
NN	4.498	25.312	3.236	111.895	319.481	135.578	0.07146	1.00000
ENN	7.402	72.745	34.495	52.842	373.215	59.301	0.09129	0.78279
CNN	11.901	68.394	31.967	66.997	337.723	83.018	0.08182	0.87339
ANN	22.276	58.885	24.176	94.363	255.618	144.683	0.07726	0.92496
NN	10.708	18.025	0.000	81.058	384.651	105.558	0.05683	1.00000
ENN	5.360	8.377	0.000	40.414	512.552	33.297	0.07191	0.79020
CNN	11.434	15.231	0.000	49.180	469.450	54.706	0.06265	0.90708
ANN	33.470	41.828	20.302	77.443	290.320	136.637	0.06251	0.90909

Table 3.2 (Continued)

NN	9.480	21.498	0.000	89.528	363.165	116.329	0.06090	1.00000
ENN	23.988	53.196	3.694	132.604	275.402	111.115	0.06636	0.91763
CNN	25.432	49.856	7.216	116.654	278.807	122.036	0.06553	0.92927
ANN	29.721	45.557	18.662	85.418	262.954	157.689	0.06772	0.89927
ENN	0.000	27.021	0.000	100.913	376.088	95.977	0.06174	0.97210
CNN	3.063	31.491	4.753	104.600	335.023	121.071	0.06064	0.98983
ANN	9.225	42.657	18.505	100.674	253.825	175.115	0.06542	0.91740
NN	1.890	21.839	4.441	65.602	393.529	112.699	0.07388	1.00000
ENN	3.163	14.959	0.000	37.136	437.072	107.669	0.07649	0.96596
CNN	6.798	22.485	4.633	42.734	398.293	125.057	0.07517	0.98293
ANN	15.664	46.204	25.700	59.058	297.165	156.209	0.08152	0.90636
NN	0.000	15.040	0.000	126.228	305.510	153.222	0.06408	1.00000
ENN	10.684	5.369	5.128	184.434	259.590	134.796	0.06762	0.94758
CNN	12.633	10.500	9.101	163.643	249.224	154.900	0.06690	0.95776
ANN	14.919	28.131	14.850	120.959	221.387	199.753	0.06971	0.91914
NN	0.000	21.720	1.841	119.989	320.330	136.121	0.08306	1.00000
ENN	0.000	0.000	0.000	184.412	242.110	173.478	0.09399	0.88376
CNN	1.683	7.150	3.353	158.705	253.898	175.211	0.08851	0.93849
ANN	10.754	35.584	20.572	111.374	236.426	185.289	0.09052	0.91763
NN	1.800	25.870	3.895	102.011	363.896	102.529	0.07975	1.00000
ENN	0.000	25.623	5.476	166.269	284.241	118.391	0.08524	0.93559
CNN	0.000	31.216	10.026	154.122	283.132	121.504	0.08487	0.93973
ANN	13.436	54.009	25.723	96.453	267.717	142.661	0.08832	0.90301
NN	3.056	23.106	4.159	146.042	311.340	112.296	0.05534	1.00000
ENN	11.393	49.441	8.029	221.507	228.633	80.996	0.06168	0.89720
CNN	12.912	49.858	12.459	199.514	227.654	97.604	0.06043	0.91579
ANN	17.588	48.188	26.220	138.865	223.683	145.456	0.06100	0.90722
NN	8.486	14.544	0.000	162.363	316.222	98.384	0.06305	1.00000
ENN	0.000	0.000	0.000	54.851	523.458	21.690	0.10412	0.60555
CNN	2.541	3.717	0.000	66.434	490.904	36.404	0.08754	0.72022
ANN	29.876	36.445	17.830	156.838	228.051	130.960	0.06959	0.90599
NN	2.146	21.090	0.000	76.054	366.259	134.450	0.09635	1.00000
ENN	0.000	15.599	7.070	64.553	431.297	81.481	0.10253	0.93974
CNN	4.626	22.687	11.516	73.529	381.121	106.521	0.09797	0.98344
ANN	15.060	42.375	22.731	71.701	292.180	155.953	0.10306	0.93485

the procedure in (2.19) in which the second phase posterior variance is expected over $p(\rho_i|\boldsymbol{x}_i)$, $i=1,\cdots,6$.

4. Discussion

Inspection of Tables 3.1 and 3.2 will reveal that though the procedures ENN and CNN often produce results which have greater efficiency than ANN, the results of the procedure labelled ANN are more consistent than the other two procedures. (This holds true for the other cases $n_i=10$, 20 and 25, tabled in Guttman and Palit [5].) Figure 3.1 shows a plot of the efficiencies of the procedures ENN, CNN, and ANN for five first phase sample sizes investigated, that is, the cases given here ($n_i=5$ and 15) and the other cases ($n_i=10$, 20 and 25) given in Guttman and Palit [5].

From Fig. 3.1, we observe that as a rule the distribution of the

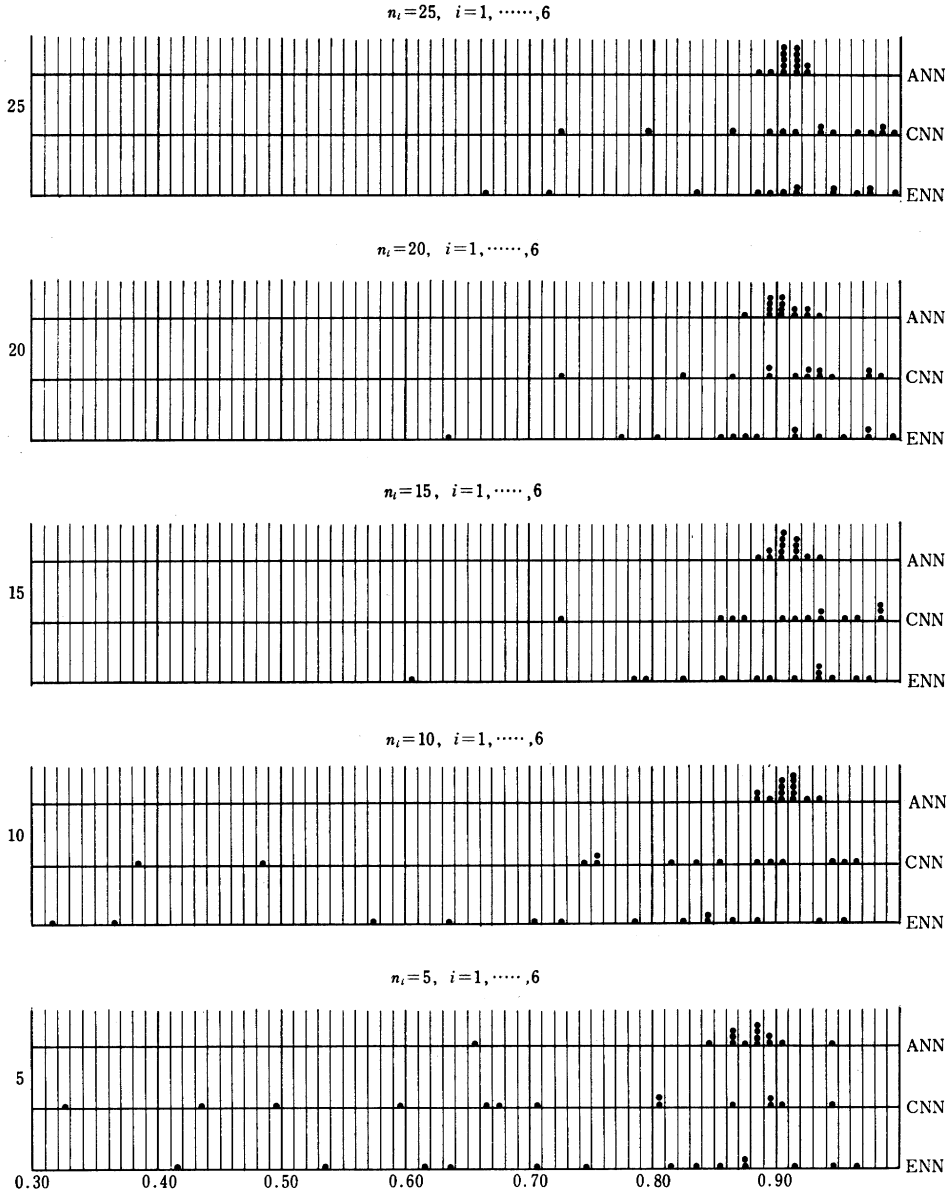

Fig. 3.1 Plot of the empirical distribution of efficiencies of the three allocation procedures for ρ unknown and for first phase samples of 5, 10, 15, 20, and 25.

efficiencies for the ANN procedure is not as spread out as the distributions for the CNN and ENN procedures. In fact the distribution for the ANN procedure when the first phase within stratum sample sizes are $n_i=5$, $i=1,\cdots,6$, appears to be at least as tight as the distribution of efficiencies for the CNN or ENN procedures where the first phase within stratum sample sizes are $n_i=25$, $i=1,\cdots,6$. It follows then that even though the use of the ANN procedure may occasionally lead to slightly lower efficiencies, its generally greater reliability makes it the preferred procedure when ρ_i are unknown.

We observe also that the efficiency of the ANN procedure is surprisingly good, even for small first phase samples. For example the smallest usuable first phase sample in each stratum, i.e., $n_i=5$, $i=1, 2,\cdots,6$ have efficiencies that tend to cluster between 85% and 91%.

UNIVERSITY OF TORONTO
UNIVERSITY OF WISCONSIN

REFERENCES

[1] Draper, N. R. and Guttman, Irwin (1968a). Some Bayesian stratified two-phase sampling results, *Biometrika*, **55**, 131-139.

[2] Draper, N. R. and Guttman, Irwin (1968b). Bayesian stratified two-phase sampling results: *k* characteristics, *Biometrika*, **55**, 587-589.

[3] Guttman, Irwin and Palit, C. D. (1971a). Optimal allocation for two-phase stratified samples when the within stratum observations are correlated—A Bayesian approach I. Known correlation—A sensitivity study, *Technical Report* No. 128, Centre de Recherches Mathématiques, Université de Montréal.

[4] Guttman, Irwin and Palit, C. D. (1971b). Optimal allocation for two-phase stratified samples when the within stratum observations are correlated—A Bayesian approach II. Unknown correlation, *Technical Report* No. 129, Centre de Recherches Mathématiques, Université de Montréal.

[5] Guttman, Irwin and Palit, C. D. (1972). Auto correlations and effect on optimum allocation—A Bayesian approach, *Technical Report* No. 130, Centre de Recherches Mathématiques, Université de Montréal.

[6] Jenkins, G. M. and Watts, D. G. (1968). *Spectral Analysis and Its Applications*, Holden Day, San Francisco, California.

[7] Kendall, M. G. and Stuart, A. (1966). *The Advanced Theory of Statistics*, Vol. 3, Griffin and Co., London.

[8] Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory*, Division of Research, Harvard Business School, Harvard University, Boston.

[9] Tiao, G. C. and Tan, W. Y. (1966). Bayesian analysis of random-effect models in the analysis of variance II. Effect of autocorrelated errors, *Biometrika*, **53**, 477-495.

[10] Zellner, A. and Tiao, G. C. (1964). Bayesian analysis of the regression model with autocorrelated errors, *J. Amer. Statist. Ass.*, **59**, 763-768.