AMERICAN JOURNAL OF MATHEMATICAL AND MANAGEMENT SCIENCES

DYNAMIC ALLOCATION IN SURVEY SAMPLING

Z. Govindarajulu
University of Kentucky
Lexington, Ky 40506

Michael N. Katehakis
Rutgers University
Newark, N. J. 07102

SYNOPTIC ABSTRACT

Consider the problem of estimating a linear combination of means from populations with different, unknown variances. We study the Bayes version of the problem and derive the dynamic programming optimality equations for the determination of optimal multi stage sequential sample size allocation procedures for several pertinent loss structures. We point out the relation of these equations to the "inventory equations". We develop heuristic procedures for the case of Bernoulli and normal populations and give numerical comparisons.

Key Words and Phrases: Survey Sampling; Dynamic Programming.

1991, VOL. 11, NOS. 3 & 4, 199-221
0196-6324/91/030199-23 $28.00

1. INTRODUCTION.

Consider the problem of estimating a linear combination of means from populations with different, unknown variances. Let Π_i be a population with unknown mean μ_i and variance σ_i^2, $i = 1, \ldots, m$. We want to estimate a function $h(\underline{\mu})$ of the unknown means, where

$$h(\underline{\mu}) = \sum_{i=1}^{m} w_i \mu_i \ . \tag{1}$$

In (1) w_i are known constants, $i = 1, \ldots, m$, and $\underline{\mu}$ denotes the vector $(\mu_1, \ldots, \mu_m)$. In this paper we consider the Bayes formulation of this problem for several pertinent loss structures.

In this formulation we allow batch sampling. The objective is to develop an adaptive rule for the determination of optimal batch sample sizes with respect to some relevant loss structure. In Section 2 we obtain the pertinent dynamic programming optimality equations for the general problem and discuss two cases of special interest. Case 1: Π_i is $\mathbf{B}(\theta_i)$, $i = 1, \ldots, m$ and case 2: Π_i is $\mathrm{N}(\mu_i, \sigma_i^2)$, $i = 1, , m$. We point out the relation of the optimality equations to those of the optimal inventory equations (cf. Whittle (1983), Ross (1970)). In Section 3 we discuss heuristic approximations for the determination of optimal one stage look ahead allocations. In Section 4 we present numerical evaluations of the heuristic procedures using simulation.

This type of estimation problem typically arises in survey sampling and in Monte Carlo simulation. In the context of survey sampling the $\Pi_i's$ can be thought of as different strata of a larger population, while in Monte Carlo simulation they may represent different regions over which measurements are taken (cf. Halton and Zeidman (1971)). The importance of this problem was pointed out in Ghurye and Robbins (1954) where the case of two normal populations was studied with $w_1 = 1$, $w_2 = -1$. They derived a two stage procedure that was shown to be asymptotically consistent, in the sense that the variance of the estimator tends to that of the optimal one stage allocation procedure when the true values of the variances are known, as the sample sizes become large. This work was extended for arbitrary m and w_i in Peierls and Yahav (1986) for an appropriately defined sequential procedure. In this paper we propose and evaluate with simulation an intermediate procedure of batch sampling which is between the two "extreme" cases of the two stage procedure and the completely sequential one. For additional related work in this area see Robbins (1985) , Govindarajulu (1987), Dudewicz and Bishop (1979), and Kuo and Mukhopadhyay (1990).

2. PROBLEM FORMULATION.

We define the K-stage allocation problem as follows. Suppose that population Π_i is specified by a vector of parameters $\underline{\theta}_i = (\theta_{i1}, \ldots, \theta_{ir_i}) \in \Theta_i \subset \Re^{r_i}$, where θ_i follows a prior density $g_{i0}(\theta_{i1}, \ldots, \theta_{ir_i})$ with respect to a σ- finite measure λ_i. Given θ_i, an observation from population Π_i is a random variable X_i with known conditional probability density $f(x|\theta_i)$. The marginal distribution of X_i is

$$f_i(x) = \int_{\Theta_i} f(x|\underline{\theta}_i) g_{i0}(\underline{\theta}_i) \lambda_i(d\underline{\theta}_i) \ . \tag{2}$$

In the sequel we assume that $E(X_i|\underline{\theta}_i) = \mu_i = \theta_{i1}$ and $\sigma^2(X_i|\underline{\theta}_i) = \sigma_i^2 = \theta_{i2}$, $i = 1, \ldots, M$. We will need the following notation. An observed sample of size n_i from population Π_i, $X_{i1} = x_{i1}$, $X_{i2} = x_{i2}, \ldots, X_{in_i} = x_{in_i}$ will be denoted by $d_i(n_i)$, i.e., $d_i(n_i) = (x_{i1}, \ldots, x_{in_i})$. We also let $\underline{n} = (n_1, \ldots, n_m)$ and $\underline{d}(\underline{n}) = (d_1(n_1), \ldots, d_m(n_m))$, $\underline{\theta} = (\underline{\theta}_1, \ldots, \underline{\theta}_m)$. Given $d_i(n_i)$, $\underline{\theta}_i$ follows a posterior density $g_{in_i}(\underline{\theta}_i|d_i(n_i))$ with respect to the σ-finite measure $\lambda_i(\cdot)$ where

$$g_{in_i}(\underline{\theta}_i|d_i(n_i)) = \frac{g_{i0}(\underline{\theta}_i) f(x_{i1}, \ldots, x_{in_i}|\underline{\theta}_i)}{f(x_{i1}, \ldots, x_{in_i})} = \frac{g_{i0}(\underline{\theta}_i) \prod_{j=1}^{n_i} f(x_{ij}|\underline{\theta}_i)}{f(x_{i1}, \ldots, x_{in_i})}. \tag{3}$$

Define $\underline{g}_{\underline{n}} = \underline{g}_{\underline{n}}(\cdot|\underline{d}(\underline{n})) = (g_{1n_1}(\cdot|d_1(n_1)), \ldots, g_{mn_m}(\cdot|d_m(n_m)))$. Given $d_i(n_i)$, the marginal distribution of a future observation from population Π_i is

$$f_i(x|d_i(n_i)) = \int_{\Theta_i} f(x|\underline{\theta}_i) g_{in_i}(\underline{\theta}_i|d_i(n_i)) \lambda_i(d\underline{\theta}_i). \tag{4}$$

We also define $d_i(n_{i0}) = (x_{i1}, \ldots, x_{in_{i0}})$ to denote some initial sample from population Π_i and think of the prior densities $g_{i0}(\theta_{i1}, \ldots, \theta_{ir_i})$ as being determined by this initial sampling. Given $d_i(n_i)$, if we decide to take a sample of size k_i from population Π_i then we will observe the sample $X_{in_i+1} = x_{in_i+1}, X_{in_i+2} = x_{in_i+2}, \ldots, X_{in_i+k_i} = x_{in_i+k_i}$. In what follows it will be convenient to replace k_i by $\nu_i = n_i + k_i$, i.e, ν_i is the total sample from population Π_i, after we have taken an additional sample of size $\nu_i - n_i$ from population Π_i. If we take no new sample then $\nu_i = n_i$. Given $d_i(n_i)$, before we take the additional $\nu_i - n_i$ "future" observations X_{ij} $j = n_i + 1, \ldots, \nu_i$

the $X_{ij}'s$ are random variables with marginal densities $f(x|d_i(n_i))$ given by (4) . We introduce the notation $D_i(\nu_i, d_i(n_i)) = (d_i(n_i), X_{in_i+1}, \ldots, X_{i\nu_i})$ and $\underline{D}(\underline{\nu}, \underline{d}) = (D_i(\nu_i, d_i(n_i)))_{i=1,\ldots,m}$.

In this Bayesian framework,

$$\hat{\theta}_{ik} = \hat{\theta}_{ik}(d_i(n_i)) = E_{g_{in_i}}(\theta_{ik}) = E(\theta_{ik}|d_i(n_i)). \tag{5}$$

is a reasonable (Bayes) estimate for θ_{ik}, given the sample $d_i(n_i)$. The Bayes risk associated with *the square error loss* for $\hat{\theta}_{ik}$ is given by

$$r(d_i(n_i)) = \sigma^2(\theta_{ik}|d_i(n_i)) = E\big((\theta_{ik} - \hat{\theta}_{ik})^2|d_i(n_i)\big). \tag{6}$$

Thus, the Bayes estimate for $h(\underline{\mu}) = h(\theta_{11}, \ldots, \theta_{m1})$ is

$$T(\underline{d}(\underline{n})) = \sum_{i=1}^{m} w_i\hat{\theta}_{i1}(d_i(n_i)) = \sum_{i=1}^{m} w_i E\big(\theta_{i1}|d_i(n_i)\big), \tag{7}$$

with corresponding risk

$$r(\underline{d}(\underline{n})) = \sum_{i=1}^{m} w_i^2 r(d_i(n_i)) = \sum_{i=1}^{m} w_i^2 E\big((\theta_{i1} - \hat{\theta}_{i1})^2|d_i(n_i)\big). \tag{8}$$

Furthermore, we assume that there is a sampling cost associated with population Π_i which is given by the functions $c_i(k_i)$. We will consider the following special cases

$$c_i(k_i) = a_{i0}\delta(k_i), \tag{9}$$

$$c_i(k_i) = k_i, \tag{10}$$

where in (9) $\delta(k) = 0$ for $k = 0$ and $\delta(k) = 1$ for $k > 0$; a_{i0} denotes a "setup" cost associated with population Π_i . Of course (10) is a special case of (9).

We can now state the allocation problem. The aim is to specify an allocation policy to minimize

$$R(\pi) = E_{\underline{g}_{\underline{n}_0}}\Big(\sum_{t=1}^{K}\sum_{i=1}^{m} c_i(\nu_{it}) + Lr(\underline{d}(\underline{\nu}_K)\Big), \tag{11}$$

subject to

$$\sum_{t=1}^{K}\sum_{i=1}^{m} c_i(\nu_{it}) \le c_0. \tag{12}$$

In (11) L is a known constant and an allocation policy is a rule of the form: $\pi = ((\nu_{11}, \ldots, \nu_{m1}), \ldots, (\nu_{1K}, \ldots, (\nu_{mK})) = (\underline{\nu}_0, \ldots, \underline{\nu}_K)$; where ν_{it} is the sample size "up to" which we sample at stage t, from population Π_i given the data $\underline{d}(\underline{\nu}_{t-1})$.

We will refer to problem (11), (12) as problem (P_1) while the problem of minimizing the risk $R(\pi)$ without any constraint will be referred to as problem (P_0).

In parametric models the posterior densities $g_{in_i}(\cdot) = g_{in_i}(\underline{\theta}_i | d_i(n_i))$ are uniquely determined by a set of sufficient statistics and the known priors g_{i0}. In this situation we will assume that the data $d_i(n_i) = (x_{i1}, \ldots, x_{in_i})$ from population Π_i is represented in terms of the sufficient statistics but we keep the notation $d_i(n_i)$. For the cases under consideration we have the following.

CASE 1: We have $X_i|\theta_i \sim \mathbf{B}(\theta_i)$, where θ_i are independent random variables, $\theta_i \sim \mathbf{Beta}(a_i, b_i)$; a_i, b_i being known constants and $X_i \sim \mathbf{B}(a_i/(a_i+b_i))$. The posterior distribution of θ_i given $d_i(n_i)$ is **Beta** with parameters: $a_i + s_{in_i}$, $b_i + f_{in_i}$, where $s_{in_i} = \sum_{j=1}^{n_i} x_{ij}$ and $f_{in_i} = n_i - s_{in_i}$, i.e., $d_i(n_i) = (n_i, s_{in_i})$ is sufficient for θ_i.

CASE 2: We have $X_i|\theta_i \sim \mathbf{N}(\mu_i, \sigma_i^2)$, $\theta_i = (\mu_i, \sigma_i^2)$. Convenient forms of prior distributions for θ_i are obtained if we assume independent priors for μ_i and σ_i^2 of the form: $\mu_i \sim \mathbf{N}(\mu_{i0}, \sigma_{i0}^2)$ and $\sigma_i^{-2} \sim \Gamma(z_{i0}/2, m_{i0}/s)$, where μ_{i0}, σ_{i0}^2, $z_{i0}/2$, m_{i0}/s, are known constants. The posterior distributions of μ_i and σ_i^{-2} given $d_i(n_i)$ can be computed (cf. Box and Tiao (1973)). In particular,

$$\sigma_i^{-2}|d_i(n_i) \sim \Gamma(\frac{z_{in_i}}{2}, \frac{m_{in_i}}{2}), \tag{13}$$

where $z_{in_i} = z_{i0} + \sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{in_i})^2$, $\bar{x}_{in_i} = \sum_{j=1}^{n_i} x_{ij}/n_i$, $m_{in_i} = (m_{i0} + n_i - 1)$. In this case $d_i(n_i) = (n_i, z_{in_i})$ is sufficient for σ_i^2 and $d_i(n_i) = (n_i, \bar{x}_{in_i}, z_{in_i})$ is sufficient for $(\mu_i, \sigma_i^2) = (\theta_{i1}, \theta_{i2})$.

3. OPTIMALITY EQUATIONS.

The basic idea underlying the derivation of the dynamic programming optimality equations is the fact that for any allocation, $R(\pi)$ is the sum of two cost factors. The first is the cost associated with the initial decision and the second is the cost associated with the remaining decisions. This can be seen from the fact that $R(\pi)$ may be computed recursively as follows.

$$\begin{aligned} R(\pi) &= E(\sum_{t=1}^{K}\sum_{i=1}^{m} c_i(\nu_{it}) + Lr(\underline{d}(\underline{\nu}_K))|\underline{d}(\underline{n}_0)) \\ &= E(\cdots(E(\sum_{t=1}^{K}\sum_{i=1}^{m} c_i(\nu_{it}) + Lr(\underline{d}(\underline{\nu}_K))|\underline{d}(\underline{\nu}_K),\ldots,\underline{d}(\underline{n}_0)))) \\ &= E(\sum_{i=1}^{m} c_i(\nu_{i1})|\underline{d}(\underline{n}_0)) + E(E(\sum_{i=1}^{m} c_i(\nu_{i2})|\underline{d}(\underline{\nu}_1),\underline{d}(\underline{n}_0))) + \cdots \\ &\quad + E(\cdots(E(\sum_{i=1}^{m} c_i(\nu_{iK}) + Lr(\underline{d}(\underline{\nu}_K))|\underline{d}(\underline{\nu}_K),\ldots,\underline{d}(\underline{n}_0)))). \quad (14) \end{aligned}$$

Thus, the dynamic programming equations are

$$v(K+1,c,\underline{d}) = Lr(\underline{d}) \;, \quad (15)$$

$$v(t,c,\underline{d}) = \min_{\underline{\nu}\in A(c,\underline{d})}\{U(t,c,\underline{d},\underline{\nu})\}, \quad (16)$$

for $c \in \mathcal{C}_t$, $\underline{d} \in \mathcal{D}_t$, $t = K+1,\ldots,1$, where,

$$\begin{aligned} U(t,\underline{d},\underline{n}) &= \sum_{i=1}^{m} c_i(\nu_i - n_i) \\ &\quad + E\big(v(t+1, c - \sum_{i=1}^{m} c_i(\nu_i - n_i), D(\underline{\nu},\underline{d})|\underline{d}(\underline{\nu}_{t-1}) = \underline{d}(\underline{n})\big)\big). \end{aligned} \quad (17)$$

and we use the notation: $\nu_i = \nu_{it}$, $n_i = \nu_{it-1}$. $\mathcal{C}_t$, $\mathcal{D}_t$, $t = K+1, K,\ldots,,1$ are the sets of all possible values of c , $\underline{d}$, at stage t, and c is the remaining cost that can be allocated in stages $t, t+1,\ldots,K$. The sets of allowable actions $A(c,\underline{d})$, are determined by the current data and the constraints as follows. $A(c,\underline{d}) = \{(k_1,\ldots,k_m) \; : \; k_i \geq n_i \; , \; C(c,\underline{\nu},\underline{n}) \geq 0 \; , \; k_i \text{ integers}\}$.

Remark 1. In the case of zero set up costs and linear sampling costs, i.e., (9) with $a_{i0} = 0$, there is no need to keep track of the "remaining cost" since this information is now contained in the data $\underline{d}(\underline{n})$ in terms of the sample size vector $\underline{n}$. Indeed, given that $\nu_{it-1} = n_i$, the remaining cost at stage t is equal to $c_0 - \sum_{i=1}^{m} c_i n_i$. Hence the following simplification to the functional equations is possible:

$$v(K+1,\underline{d}) = Lr(\underline{d}) \;, \quad (18)$$

$$v(t,\underline{d}) = \min_{\underline{\nu}\in A(c,\underline{d})}\{U(t,\underline{d},\underline{\nu})\}, \quad (19)$$

for $c \in \mathcal{C}_t,\ \underline{d} \in \mathcal{D}_t, t = K+1, \ldots, 1$, where

$$U(t, \underline{d}, \underline{n}) = \sum_{i=1}^{m} c_i(\nu_i - n_i) + E_{\underline{g}_\nu}(v(t+1, D(\underline{\nu}, \underline{d})|\underline{d}(\underline{\nu}_{t-1}) = \underline{d}(\underline{n})) \ . \qquad (20)$$

and $A(\underline{d}) = \{(k_1, \ldots, k_m) :\ k_i \geq n_i,\ C(c, \underline{k}, \underline{n}) \geq 0,\ k_i \geq n_i,\ \ k_i \text{ integers}\}$.

For case 1 the functional equations take the form:

$$v(K+1, c, \underline{d}) \ = \ L\sum_{i=1}^{m} w_i^2 \frac{(a_i + s_i)(b_i + n_i - s_i)}{(a_i + b_i + s_i)^2(a_i + b_i + n_i + 1)}, \qquad (21)$$

$$v(t, \underline{d}) \ = \ \min_{\underline{\nu} \in A(c,\underline{d})} \{U(t, \underline{d}, \underline{\nu})\}, \qquad (22)$$

for $c \in \mathcal{C}_t,\ \underline{d} \in \mathcal{D}_t, t = K+1, \ldots, 1$, where,

$$U(t, c, \underline{d}, \underline{\nu}) \ = \ \sum_{i=1}^{m} c_i(\nu_i - n_i) + \sum_{i=1}^{m} \sum_{y_i=0}^{\nu_i - n_i} p(y_i|\nu_i, n_i) \cdot [$$

$$v(t+1, c - \sum_{i=1}^{m} c_i(\nu_i - n_i, G(\underline{y}, \underline{\nu}, \underline{d})))] \qquad (23)$$

and $\underline{d} = ((s_1, n_1), \ldots, (s_m, n_m))$, $G(\underline{y}, \underline{\nu}, \underline{d}) = ((s_1 + y_1, \nu_1), \ldots, (s_m + y_m, \nu_m))$,

$$p(y_i|\nu_i, n_i) = \binom{\nu_i - n_i}{y_i} \frac{(a_i + s_i)^{y_i}(b_i + n_i - s_i)^{\nu_i - n_i - y_i}}{(a_i + b_i + n_i + 1)^{\nu_i - n_i}}, \qquad (24)$$

$y_i = 0, 1, \ldots, \nu_i - n_i$.

In (21), (24) we have used the fact that given the data, $d_i = (s_i, n_i)$ from population Π_i, $\sum_{j=1}^{\nu_i - n_i} X_{ij}$ follows a binomial distribution with parameters $\nu_i - n_i$, $(a_i + s_i)/(a_i + b_i + n_i)$. Note also that $\mathcal{C}_t = [0, c_0]$ and $\mathcal{D}_t = \{(k_1, \ldots, k_m) \ :\ \sum_{i=1}^{m} c_i(k_i) \leq c_0\ ,\ \ k_i \geq 0, \text{ integers}\}$.

Equations (21), (22) can be solved numerically for moderate values of m and c_0. Analogous equations can be derived for case 2. However we will do this in the next section where we derive dynamic programming equations using approximations for the posteriors.

We end this section by pointing out the relation between equations (15), (16) and the optimality equations for the "inventory problem" (cf. Ross (1970) p. 169). First, consider problem (P_0). It decomposes into m independent subproblems associated with each population. For a typical subproblem the optimality equations can be written as follows.

$$
\begin{aligned}
v(K+1,c,d) &= Lr(d), && (25)\\
v(t,c,d) &= \min_{\nu\in A(d)}\{c_i(\nu-n)+E_{g_\nu}(v(t+1,D(\nu,d))|d(\nu_{t-1})=d)\}\\
&= \min_{\nu\geq n}\{H_t(\nu,n)\}\ ,t=K,K-1,\ldots,1, && (26)
\end{aligned}
$$

where $\nu_0 = n_0$ is given. For notational simplicity we have dropped the subscript i since we are dealing with a fixed subproblem. Recall that a function $f(x)$ is called **α-convex** in x if $\alpha+f(x+u)-f(x)\geq u\,\mathrm{d}f(x)/\mathrm{d}x$ for all u and x. Under the appropriate assumptions that insure the a_0 - convexity of $H_t(\nu,n)$ in ν (cf. Ross(1970) p. 173) and if we drop the restriction that ν must be an integer, the following proposition holds.

Proposition 1. *For problem (P_0), under the assumption of a_{i0} - convexity of the functions $H_{it}(\nu,n)$ in ν, and if we allow non-integer values for the sample sizes ν_{it}, then an optimal allocation policy has the following form. There exist numbers $q_{it}=q_{it}(\nu_{it-1},\underline{d}_{t-1})$, $Q_{it}=Q_{it}(\nu_{it-1},\underline{d}_{t-1})$, $t=1,\ldots,K$ with $q_{it}\leq Q_{it}$ such that the optimal sample (up to) sizes at stage t are given by*

$$
\nu_{it}^0=\nu_{it}^0(\nu_{it-1},\underline{d}_{t-1})=\begin{cases}Q_{it}, & \text{if } \nu_{it-1}\leq q_{it};\\ \nu_{it-1}, & \text{if } \nu_{it-1}>q_{it},\end{cases} \qquad (27)
$$

where Q_{it} is the point at which the function

$$
\tilde{H}_{it}(\nu,n) = H_{it}(\nu,n)-a_{i0}\delta(\nu-n)
$$

attains its minimum value and

$$
q_{it}=\sup\{\nu\leq Q_{it}\ :\ \tilde{H}_{it}(\nu,n)\geq\tilde{H}_{it}(Q_{it},n)+a_{i0}\}.
$$

For the case with no set up cost the following simplification to Proposition 1 is possible.

Proposition 2. *For problem (P_0), under the assumptions that insure the convexity of $H_t(\nu,n)$ in ν, there exist numbers $Q_{it}=Q_{it}(\nu_{it-1},\underline{d}_{t-1})$, $t=1,\ldots,K$, with $\min_{\nu\geq 0}\{H_t(\nu,n)\}=H_t(Q_{it},n)$ such that the optimal (up to) sample sizes at stage t are given by*

$$
\nu_{it}^0=\nu_{it}^0(\nu_{it-1},\underline{d}_{t-1})=\begin{cases}Q_{it}, & \text{if } \nu_{it-1}\leq Q_{it};\\ \nu_{it-1}, & \text{if } \nu_{it-1}>Q_{it}.\end{cases} \qquad (28)
$$

The problem of computing the numbers q_{it}, Q_{it} remains; see Whittle (1983) and references given there. Problem (P_1) is related to the multi-product inventory problem.

4. APPROXIMATIONS ONE-STAGE LOOK AHEAD RULES.

In this section we use approximations for the posteriors to derive asymptotically optimal one stage rules. We suppose that there is no set up cost. The approximations are valid for large sample sizes as is often the case in survey sampling. We start with the following proposition (cf. Lindley (1970) p.132).

Proposition 3. *Suppose that an unknown random vector of parameters* $\underline{\theta} = (\theta_1, \ldots, \theta_r)$ *with prior density function* $g(\theta_1, \ldots, \theta_r)$ *is such that given* $\underline{\theta}$ *, an observable random variable* X *has conditional density* $f(x|\underline{\theta})$ *of known form. If a sample* $\underline{x}$ *of size* n *is taken then, provided that* $g(\underline{\theta}) \neq 0$ *,* $\forall \underline{\theta}$ *, the joint posterior of* $\underline{\theta}$ *is approximately multivariate normal* $\mathbf{N}(\hat{\underline{\theta}}, \Sigma)$ *, where* $\hat{\underline{\theta}} = \hat{\underline{\theta}}(x)$ *is the m.l.e. of* $\underline{\theta}$ *given the sample,*

$$\Sigma^{-1} = -\left(\frac{\partial^2 l(\underline{x}|\underline{\theta})}{\partial\theta_i \partial\theta_j}\right)_{\underline{\theta}=\hat{\underline{\theta}}} \tag{29}$$

and $l(\underline{x}|\underline{\theta})$ *is the log likelihood, i.e.,* $l(\underline{x}|\underline{\theta}) = \log L(\underline{x}|\underline{\theta}) = \sum_{j=1}^{n} \log f(x_j)$. Using improper (or "uniform") priors (c.f., Box and Tiao (1973)) one obtains the following corollaries of the above proposition.

Corollary 1. *For case 1, where* $X|\theta \sim \mathbf{B}(\theta)$ *, the posterior distribution of* θ *given* $d(n) = (s_n, n)$*, is approximately* $\mathbf{N}(s_n/n, s_n(n - s_n)/n^3)$*, where* $s_n = \sum_{j=1}^{n} x_j$.

Corollary 2. *For case 2 , where* $X|\underline{\theta} \sim \mathbf{N}(\mu, \sigma^2)$*, the posterior distribution of* $\underline{\theta} = (\theta_1, \theta_2)$ $(\theta_1 = \mu, \theta_2 = \sigma^2)$ *given* $d(n) = (\bar{x}_n, z_n, n)$*, where* $\bar{x}_n = \sum_{j=1}^{n} x_j/n$, $z_n = \sum_{j=1}^{n} (x_j - \bar{x}_n)^2$ *is such that* μ , σ^2 *are approximately independent and normally distributed r.v.s with* $\mu \sim \mathbf{N}(\bar{x}_n, z_n/n^2) = \mathbf{N}(\bar{x}_n, \hat{\sigma}_n^2/n)$*, and* $\sigma^2 \sim \mathbf{N}(z_n/n, 2z_n^2/n^3) = \mathbf{N}(\hat{\sigma}_n^2, 2(\hat{\sigma}_n^2)^2/n)$.

In the following lemmas we use the above corollaries together with the fact that the pertinent sequences of random variables can be shown to be uniformly integrable and therefore convergence in distribution implies convergence of moments (see Serfling (1980) p. 14). Also, we write the "=" sign with the understanding that the expectations are computed with respect to the asymptotically normal posteriors given by Proposition 3.

Lemma 1. *Under the assumptions of Proposition 3, for a typical population as in case 1 we have:*

$$E\Big(\frac{S_\nu}{\nu}|(s_n,n)\Big) = \frac{s_n}{n} \tag{30}$$

$$E\Big(\frac{S_\nu}{\nu}(1-\frac{S_\nu}{\nu})|(s_n,n)\Big) = \frac{s_n}{n}(1-\frac{s_n}{n})+o(n^{-2}), \tag{31}$$

where $S_\nu = \sum_{j=1}^{\nu} x_j$ and $\nu \geq n$.

Proof. First notice that for $\nu \geq n$,

$$S_\nu = s_n + \sum_{j=1}^{\nu-n} X_{n+j} = s_n + Y \tag{32}$$

where $X_{n+j}|\big(\theta,\ (s_n,n)\big)$ are i.i.d $\mathbf{B}(\theta)$ and Y is a binomial random variable with parameters $(\nu-n)$ and θ. Hence,

$$E\Big(\frac{S_\nu}{\nu}|\theta,(s_n,n)\Big) = E\Big(\frac{s_n+Y}{\nu}|\theta,(s_n,n)\Big) = \frac{s_n+(\nu-n)\theta}{\nu} \tag{33}$$

where by Corollary 1, $\theta \sim \mathbf{N}(s_n/n, s_n(n-s_n)/n^3)$, the result follows. For (31) we have

$$\begin{aligned} E\Big((\frac{S_\nu}{\nu})^2|\theta,(s_n,n)\Big) &= E\Big((\frac{s_n+Y}{\nu})^2|\theta,(s_n,n)\Big) \\ &= \frac{1}{\nu^2}E\big(s_n^2+2s_n^2Y+Y^2|\theta,(s_n,n)\big) \\ &= \frac{1}{\nu^2}(s_n^2+2s_n^2(\nu-n)\theta+(\nu-n)\theta(1-\theta)(\nu-n)^2\theta^2) \end{aligned} \tag{34}$$

the result follows again by Corollary 1.

Lemma 2. *Under the assumptions of Proposition 3, for a typical population as in case 2 we have*

$$E\Big(\frac{Z_\nu}{\nu}|(z_n,n)\Big) = \frac{z_n}{n} \tag{35}$$

Proof. First notice that for $\nu \geq n$,

$$Z_\nu = z_n + \sum_{j=0}^{\nu-n-1} U_{nj} \tag{36}$$

where

$$U_{nj} = \frac{n+j}{n+j+1}(X_{n+j+1} - \bar{X}_{n+j})^2 ,$$

$$\bar{X}_{n+j} = \bar{X}_n + \frac{X_{n+1} + \cdots + X_{n+j} - j\bar{X}_n}{n+j} ,$$

and $X_{n+j}|(\mu, \sigma^2, (\bar{x}_n, z_n, n))$ are i.i.d. $\mathbf{N}(\mu, \sigma^2)$.

Let $Y_j = X_j - \mu$. Then, $X_{n+j+1} - \bar{X}_{n+j} = Y_{n+j+1} - \bar{Y}_{n+j}$ and

$$Y_{y+j+1} - \bar{Y}_{n+j} = Y_{n+j+1} - \frac{n\bar{Y}_n + Y_{n+1} + \cdots + Y_{n+j}}{n+j}, \tag{37}$$

where $Y_{n+j}|(\mu, \sigma^2, (\bar{x}_n, z_n, n)) \sim \mathbf{N}(0, \sigma^2)$. It follows that

$$(X_{n+j+1} - \bar{X}_{n+j})|(\mu, \sigma^2, \bar{x}_n, z_n)) \sim \mathbf{N}(\frac{n}{n+j}\bar{Y}_n, \sigma^2(1 + \frac{j}{(n+j)^2})), \tag{38}$$

for $j = 0, \ldots, \nu - n - 1$. Hence,

$$\begin{aligned} E(U_{nj}|(\mu, \sigma^2, \bar{x}_n, z_n)) &= \frac{\mathrm{n+j}}{\mathrm{n+j+1}} E((X_{n+j+1} - \bar{X}_{n+j})^2 |(\mu, \sigma^2, \bar{x}_n, z_n)) \\ &= \frac{\mathrm{n+j}}{\mathrm{n+j+1}}(\sigma^2(1 + \frac{\mathrm{j}}{(\mathrm{n+j})^2} + (\frac{\mathrm{n}}{\mathrm{n+j}})^2(\bar{x}_n - \mu)^2)) \end{aligned} \tag{39}$$

Now it is easy to complete the proof of (35) since by Corollary 2 we have that $E(\sigma^2|(z_n, n)) = z_n/n$ and $E((\bar{x}_n - \mu)^2|(\bar{x}_n, z_n, n)) = z_n/n^2$.

We can now prove the following.

Proposition 4. *Large sample size one stage look ahead allocation rules, given current data $d(n_i)$ from population Π_i, can be determined by computing the values of ν_i , $i = 1, \ldots, m$, that solve the following optimization problem.*

Minimize

$$\sum_{i=1}^{m} c_i\nu_i + \frac{p_i^2}{\nu_i}, \tag{40}$$

subject to

$$\nu_i \geq n_i \tag{41}$$

where p_i^2 is given by (42) below for case 1 and by (43) for case 2.

$$p_i^2 = Lw_i^2 \frac{s_{in_i}}{n_i}(1 - \frac{s_{in_i}}{n_i}), \tag{42}$$

$$p_i^2 = Lw_i^2 \frac{z_{n_i}}{n_i}, \tag{43}$$

Proof. To compute one stage allocation rules we use the dynamic programming optimality equations with $K = 1$. For case 1 we have

$$v(2,\underline{d}) = L\sum_{i=1}^{m} w_i^2 E((\theta_{\,i} - \hat{\theta}_{\,i})^2|d_i) = L\sum_{i=1}^{m} w_i^2 \frac{s_{in_i}(n_i - s_{in_i})}{n_i^3}, \tag{44}$$

where $d_i = (s_{in_i}, n_i)$ and the last equation follows from Lemma 1. Now

$$\begin{aligned} v(1,\underline{d}) &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i(\nu_i - n_i) + E_{g_{in_i}}((v(2,\underline{D}(\underline{\nu},\underline{n}))|\underline{d})\} \\ &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i(\nu_i - n_i) + E_{g_{in_i}}(L\sum_{i=1}^{m} w_i^2 \frac{S_{i\nu_i}(n_i - S_{i\nu_i})}{\nu_i^3}|d_i)\} \\ &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i\nu_i + \sum_{i=1}^{m}\frac{p_i^2}{\nu_i} - \sum_{i=1}^{m} c_i n_i + o(n_i^{-2})\}, \end{aligned} \tag{45}$$

where the last equality follows from Lemma 1 and ((44)).

Similarly for case 2 we have.

$$\begin{aligned} v(2,\underline{d}) &= L\sum_{i=1}^{m} w_i^2 r_i(d_i) = L\sum_{i=1}^{m} w_i^2 E_{g_{in_i}}((\theta_{\,i1} - \hat{\theta}_{\,i1})^2|d_i) \\ &= L\sum_{i=1}^{m} w_i^2 \frac{z_{in_i}}{n_i^2}, \end{aligned} \tag{46}$$

where $d_i = (z_{in_i}, n_i)$, $\hat{\theta}_{i1} = \bar{x}_{n_i} = E(\theta_{i1}|d_i)$, and the last equation follows from Lemma 2. Now

$$\begin{aligned} v(1,\underline{d}) &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i(\nu_i - n_i) + E_{g_{in_i}}((v(2,\underline{D}(\underline{\nu},\underline{n}))|\underline{d})\} \\ &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i(\nu_i - n_i) + E_{g_{in_i}}(L\sum_{i=1}^{m} w_i^2 \frac{Z_{i\nu_i}}{\nu_i^2}|d_i)\} \\ &= \min_{\nu_i \geq n_i}\{\sum_{i=1}^{m} c_i\nu_i + \sum_{i=1}^{m}\frac{p_i^2}{\nu_i} - \sum_{i=1}^{m} c_i n_i\}, \end{aligned} \tag{47}$$

where the last equality follows from Lemma 2 and (46).

A simple algorithm for computing optimal solutions to problem (40), (41) when we drop the constraint that the ν_i's take only integer values is given by Proposition 5 and subsequent remarks below. Consider the problem:

minimize

$$f(x_1,\ldots,x_2)=\sum_{i=1}^{m} x_i+\frac{p_i^2}{x_i}, \tag{48}$$

subject to

$$\sum_{i=1}^{m} x_i \;\le\; c_0, \tag{49}$$

$$x_i \;\ge\; n_i,\ i=1,\ldots,m. \tag{50}$$

where p_i^2 , $c_0, n_i,$ are fixed constants.

Given a subset $\mathcal{G}$ of $\{1,\ldots,m\}$ we define $n(\mathcal{G})=\sum_{i\in\mathcal{G}} n_i$, $p(\mathcal{G})=\sum_{i\in\mathcal{G}} p_i$. We have the following.

Proposition 5. *The optimal solution* $\underline{x}^0=(x_1^0,\ldots,x_m^0)$ *of the above minimization problem can be determined by the following procedure.*

1. *Let* $\mathcal{M}=\{1,\ldots,m\}$. *Define the variables* y_i *by*

$$y_i=\begin{cases} p_i & \text{if } \sum_{j=1}^{m} p_j\le c_0;\\ \dfrac{p_i c_0}{p(\mathcal{M})} & \text{if } \sum_{j=1}^{m} p_j> c_0.\end{cases} \tag{51}$$

2. *Let* $\mathcal{G}=\{i: y_i\le n_i\}$. *If* $\mathcal{G}=\emptyset$, *set* $x_j^0=y_j$, $\forall\, j\in\mathcal{M}$, *stop.*
3. *If* $\mathcal{G}\neq$ *then,*
   (a) *set* $x_j^0=n_j^0,\ \forall\, j\in\mathcal{G}$,
   (b) *compute* $n(\mathcal{G})$, $p(\mathcal{G})$, *and*
   (c) *set* $y_j=(c_0-n(\mathcal{G}))p_i/p(\mathcal{M}-\mathcal{G})$, *for* $j\in\mathcal{M}-\mathcal{G}$.
4. *Set* $\mathcal{M}=\mathcal{M}-\mathcal{G}$, $\mathcal{G}=\{i\in\mathcal{M}\ :\ y_i\le n_i\}$, *go to step 2 .*

Proof. We outline the proof for the case in which $m=2$. Consider the Lagrangian

$$L(x_1,x_2,u)=\sum_{i=1}^{2} x_i+\frac{p_i^2}{x_i}+u(\sum_{i=1}^{2} x_i-c_0)=\sum_{i=1}^{2} L_i(x_i,u)-uc_0, \tag{52}$$

where

$$L_i(x_i,u)=x_i+\frac{p_i^2}{x_i}+ux_i \tag{53}$$

Let $x_i^1(u) = p_i/\sqrt{1+u}$. It is easy to see that for fixed u, we have:

$$\min_{x_i \geq n_i} L_i(x_i, u) = L_i(x_i^0(u), u) \tag{54}$$

where

$$x_i^0 = \begin{cases} x_i^1(u) & \text{if } x_i^1(u) \geq n_i; \\ n_i & \text{if } x_i^1(u) < n_i. \end{cases} \tag{55}$$

Now from the complementary slackness conditions (c.f. Lasdon (1970), p. 396) we have that the solution to the optimization problem of eqs. (48) to (50) is given by $x_i^0(u^0)$ where u^0 is the solution of

$$u^0 \cdot (x_i^0(u^0) - c_0) = 0. \tag{56}$$

The proof can be completed by considering various cases. The proof for the general case is easy to complete by induction on m.

Remark 2. It is easy to show that if we replace $f(\cdot)$ in (48) by $f(\underline{x}) = \sum_{i=1}^m p_i^2/x_i$, subject to (49), (50), then the only modification that is needed in Proposition 5 , is to replace the y_i's in step 1 above by $y_i = p_i c_0/p(\mathcal{M})$.

The following variation of problem (48)-(50)

minimize

$$f(x_1, \ldots, x_m) = \sum_{i=1}^m c_i x_i + L\frac{q_i^2}{x_i}, \tag{57}$$

subject to

$$\sum_{i=1}^m c_i x_i \leq c_0, \tag{58}$$

$$c_i x_i \geq n_i, \; i = 1, \ldots, m, \tag{59}$$

can be put in the form of the original problem if we make the transformation $z_i = c_i x_i$, $p_i^2 = Lq_i^2 c_i y$.

5. SIMULATIONS.

The effectiveness of the one - stage look ahead rule was evaluated with a series of Monte Carlo simulations, for cases 1 and 2. In each case we used two populations. Furthermore, we imposed the additional restriction that every time a sample was taken from a population its size could not exceed a maximum "batch" size value m. We call this procedure the batch sequential procedure. Its performance was always found to be very good, i.e., it outperformed (in terms achieving lower variance) the simple allocation policy that chooses sample sizes by distributing uniformly the available sampling "budget" and for the case of normal populations, it achieved sample sizes close to those attainable by the optimal allocation for known variances.

For the case of Bernoulli populations, simulations were done for many possible choices for the parameters. In the appendix we present typical simulation results for Π_1 taken to be $\mathbf{B}(0.5)$ with sampling cost $c_1 = 1$, $w_1 = 1$ and Π_2 taken to be $\mathbf{B}(\theta_2)$ for $\theta_2 = 0.1, 0.5, 0.9$, $c_2 = 0.5, 1, 10$, $w_2 = 1, 2, 10$, $m = 10, 40$ and $c_0 = 200$. We performed 100 independent simulation runs. The results are summarized in Tables 1 to 3, where we denote by $\bar{\mathrm{N}}_i^u$, $i = 1, 2$, the sample sizes prescribed by the uniform allocation of cost procedure, by $\bar{\mathrm{N}}_i$ the average total sample sizes (over the 100 simulation runs) prescribed by the proposed batch sequential allocation procedure, by N_i^l , N_i^m the minimum and maximum sample sizes (over the 100 simulation runs) prescribed by the batch sequential allocation procedure, by b the average number of times (over the 100 simulation runs) that a sample from either population was taken by the batch sequential allocation procedure. Finally, the variable Dsu is the ratio of the estimated values of standard deviation of the estimator for the proposed procedure to the sum of the estimated values of the proposed procedure and that of the uniform allocation of budget procedure.

For example, Table 1 summarizes the results obtained when the total budget is $c_0 = 200$, the population sampling costs are respectively: $c_1 = 1$, $c_2 = 0.5$, the parameter of the first population is equal to $\theta_1 = 0.5$ and the first data row corresponds to a "batch size" of $m = 10$, parameter of the second population equal to $\theta_2 = 0.1$, etc. In Table 2 the only difference is that we have taken $c_2 = 1$; similarly, in Table 3 we have taken $c_2 = 10$.

For the case of Normal populations, simulations were done for many possible choices for the parameters. In the appendix we present typical simulation results for Π_1 taken to be $\mathrm{N}(0, 1)$ with sampling cost $c_1 = 1$, $w_1 =$ and Π_2 taken to be $\mathrm{N}(\mu_2, \sigma_2^2)$, for $\mu_2 = 0, 2, 10$, $\sigma_2^2 = 1, 2, 10$, $c_2 = 0.5, 1, 10$, $w_2 = 1, 2, 10$, $m = 20$, and $c_0 = 200$. We also performed 100 independent simulation runs. The results are summarized in Tables 4, 5, 6, where we use similar notation with that of Tables 1 to 3. In addition we include two columns for the corresponding optimal sample sizes N_1, N_2, for the case of known variances. The variable Dsu is the ratio of the estimated values of standard deviation of the estimator for the proposed procedure to the sum of the estimated values of the proposed procedure and that of the uniform allocation of budget procedure.

Table 4 summarizes the simulation results obtained when the total budget is $c_0 = 200$, the population sampling costs are respectively: $c_1 = 1$, $c_2 = 0.5$, the parameters of the first population are equal to $\mu_1 = 0$, $\sigma_1^2 = 1$, the first data row corresponds to a "batch size" of $m = 10$, and the parameters of the second population are equal to $\mu_2 = 0$, $\sigma_2^2 = 1$. In Table 5 we have taken $c_2 = 1$, and in Table 6 we have taken $c_2 = 10$.

ACKNOWLEDGMENT

The authors wish to thank Dr. Bruce Schmeiser for a useful discussion. This research was supported in part by a Rutgers University, Graduate School of Management, Research Grant.

Appendix

Table 1: Binomial populations, $c_1 = 1$, $c_2 = .5$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation.								
m,	θ_2	$\bar{N}_1^u$	$\bar{N}_2^u$	$\bar{N}_1$	$\bar{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	DSU
10	0.1	100	200	140	122	130	152	96	140	14.3	1	0.46
10	0.5	100	200	117	166	116	119	162	168	17.0	1	0.42
10	0.9	100	200	140	120	131	158	84	138	14.4	1	0.33
40	0.1	100	200	139	122	131	158	84	138	5.0	1	0.35
40	0.5	100	200	117	166	115	119	162	170	5.0	1	0.37
40	0.9	100	200	140	120	130	162	74	140	5.0	1	0.31
10	0.1	100	200	110	180	97	129	142	206	18.4	2	0.39
10	0.5	100	200	82	235	80	84	232	240	23.9	2	0.41
10	0.9	99	199	110	180	99	125	150	202	18.3	2	0.28
40	0.1	99	199	109	182	99	124	152	202	5.3	2	0.34
40	0.5	100	200	82	235	81	85	230	238	6.0	2	0.39
40	0.9	100	200	110	181	98	128	144	204	5.3	2	0.27
10	0.1	100	200	41	317	33	71	258	334	31.9	10	0.41
10	0.5	100	200	24	351	19	27	346	362	35.1	10	0.45
10	0.9	100	200	41	317	33	53	294	334	31.9	10	0.27
40	0.1	100	200	40	319	34	62	276	332	8.5	10	0.27
40	0.5	100	200	24	352	24	25	350	352	9.0	10	0.41
40	0.9	99	199	41	318	35	61	278	330	8.4	10	0.25

Table 2: Binomial populations, $c_1 = 1$, $c_2 = 1$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation.								
m,	θ_2	$\bar{N}_1^u$	$\bar{N}_2^u$	$\bar{N}_1$	$\bar{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	DSU
10	0.1	100	100	124	76	10	154	46	89	12.6	1	0.34
10	0.5	99	99	100	100	98	103	97	101	10.2	1	0.37
10	0.9	100	100	124	76	111	153	47	89	12.7	1	0.30
40	0.1	99	99	123	77	110	153	47	90	5.0	1	0.33
40	0.5	100	100	100	100	95	104	96	105	5.0	1	0.37
40	0.9	99	99	122	77	112	143	57	88	5.0	1	0.31
10	0.1	100	100	91	109	76	108	92	124	11.5	2	0.27
10	0.5	100	100	66	134	63	68	132	137	13.9	2	0.39
10	0.9	100	100	90	110	78	108	92	122	11.5	2	0.26
40	0.1	100	100	89	111	79	114	86	121	5.0	2	0.33
40	0.5	100	100	66	134	62	72	128	138	5.0	2	0.40
40	0.9	100	100	90	110	78	114	86	122	5.0	2	0.26
10	0.1	100	100	31	169	25	49	151	175	17.1	10	0.31
10	0.5	100	100	18	182	12	24	176	188	18.1	10	0.40
10	0.9	100	100	31	169	23	52	148	177	17.1	10	0.27
40	0.1	100	100	30	170	23	48	152	177	5.1	10	0.28
40	0.5	100	100	18	182	18	20	180	182	5.0	10	0.40
40	0.9	100	100	29	171	22	43	157	178	5.1	10	0.24

Table 3: Binomial populations, $c_1 = 1$, $c_2 = 10$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation.								
m,	θ_2	$\bar{N}_1^u$	$\bar{N}_2^u$	$\bar{N}_1$	$\bar{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	DSU
10	0.1	96	9	62	13	50	72	12	15	6.3	1	0.34
10	0.5	98	9	48	15	40	55	14	16	5.1	1	0.38
10	0.9	95	9	61	14	50	72	12	15	6.2	1	0.28
40	0.1	100	10	50	15	50	50	15	15	5.0	1	0.32
40	0.5	99	9	48	15	45	50	15	15	5.0	1	0.36
40	0.9	100	10	50	15	50	50	15	15	5.0	1	0.30
10	0.1	99	9	37	16	27	48	15	17	5.0	2	0.33
10	0.5	97	9	27	17	25	30	17	17	5.0	2	0.36
10	0.9	99	9	37	16	27	48	15	17	5.0	2	0.27
40	0.1	100	10	30	17	30	30	17	17	5.0	2	0.29
40	0.5	98	9	27	17	27	30	17	17	5.0	2	0.38
40	0.9	100	10	30	17	28	30	17	17	5.0	2	0.30
10	0.1	98	9	9	19	6	12	18	19	5.0	10	0.37
10	0.5	98	9	6	19	6	7	19	19	5.0	10	0.39
10	0.9	99	9	9	19	6	12	18	19	5.0	10	0.31
40	0.1	98	9	8	19	6	10	19	19	5.0	10	0.38
40	0.5	98	9	6	19	6	7	19	19	5.0	10	0.21
40	0.9	99	9	9	19	6	10	19	19	5.0	10	0.26

Table 4: Normal populations, $c_1 = 1$, $c_2 = .5$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation, m=20								Known σ_i^2		
μ_2,	σ_2^2	$\tilde{N}_1^u$	$\tilde{N}_2^u$	$\tilde{N}_1$	$\tilde{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	N_1	N_2	DSU
0	1	100	200	117	165	107	127	146	186	8.8	1	117	164	0.50
0	2	100	200	101	198	86	112	176	228	10.3	1	100	200	0.50
0	10	100	200	63	275	47	77	246	306	14.2	1	61	276	0.46
2	1	100	200	117	166	107	127	146	184	8.9	1	117	164	0.49
2	2	100	200	100	200	88	115	170	224	10.4	1	100	200	0.50
2	10	100	200	63	274	50	77	246	300	14.1	1	61	276	0.46
10	1	100	200	104	192	97	112	174	206	10.4	1	117	164	0.50
10	2	100	200	93	214	83	102	196	234	11.4	1	100	200	0.49
10	10	100	200	60	281	46	74	252	308	14.5	1	61	276	0.46
0	1	100	200	85	231	68	98	204	264	12.0	2	82	234	0.49
0	2	100	200	68	264	56	81	238	288	13.7	2	66	266	0.47
0	10	100	200	38	323	28	48	304	344	16.6	2	36	326	0.42
2	1	100	200	81	237	69	97	206	262	12.3	2	82	234	0.49
2	2	100	200	67	266	56	79	242	288	13.8	2	66	266	0.47
2	10	100	200	38	323	28	52	296	344	16.6	2	36	326	0.41
10	1	100	200	71	257	60	81	236	280	13.5	2	82	234	0.46
10	2	100	200	60	279	48	72	256	304	14.5	2	66	266	0.44
10	10	100	200	35	330	25	45	310	350	17.0	2	36	326	0.41
0	1	100	200	26	348	21	41	318	358	17.8	10	24	350	0.38
0	2	100	200	22	357	21	31	338	358	18.0	10	18	362	0.38
0	10	100	200	21	358	21	21	358	358	18.0	10	8	382	0.37
2	1	100	200	24	352	21	34	332	358	17.9	10	24	350	0.38
2	2	100	200	21	357	21	27	346	358	18.0	10	18	362	0.38
2	10	100	200	21	358	21	21	358	358	18.0	10	8	382	0.36
10	1	100	200	21	357	21	26	348	358	18.0	10	24	350	0.33
10	2	100	200	21	358	21	21	358	358	18.0	10	18	362	0.35
10	10	100	200	21	358	21	21	358	358	18.0	10	8	382	0.36

Table 5: Normal populations, $c_1 = 1$, $c_2 = 1$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation, m=20								Known σ_i^2		
μ_2,	σ_2^2	$\bar{N}_1^u$	$\bar{N}_2^u$	$\bar{N}_1$	$\bar{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	N_1	N_2	DSU
0	1	100	100	100	100	89	111	89	111	6.1	1	100	100	0.50
0	2	100	100	83	117	69	94	106	131	6.4	1	82	117	0.49
0	10	100	100	50	150	37	62	138	163	8.0	1	48	151	0.44
2	1	100	100	99	101	84	116	84	116	6.1	1	100	100	0.50
2	2	100	100	83	117	69	97	103	131	6.4	1	82	117	0.49
2	10	100	100	48	152	32	61	139	168	8.0	1	48	151	0.44
10	1	100	100	80	119	68	89	111	131	7.1	1	100	100	0.47
10	2	100	100	72	128	59	83	117	141	7.4	1	82	117	0.46
10	10	100	100	45	155	34	55	145	166	8.1	1	48	151	0.43
0	1	100	100	67	133	53	81	119	147	7.1	2	66	133	0.47
0	2	100	100	54	146	41	66	134	159	7.6	2	52	147	0.45
0	10	100	100	29	171	21	41	159	179	8.9	2	27	172	0.39
2	1	100	100	66	134	54	[illegible]	122	145	7.2	2	66	133	0.47
2	2	100	100	52	148	42	65	135	158	7.9	2	52	147	0.44
2	10	100	100	28	172	21	41	159	179	9.0	2	27	172	0.40
10	1	100	100	52	148	41	61	139	158	8.4	2	66	133	0.39
10	2	100	100	43	157	30	52	147	170	8.5	2	52	147	0.39
10	10	100	100	25	175	21	33	167	179	9.0	2	27	172	0.38
0	1	100	100	20	180	18	33	167	182	9.6	10	18	181	0.38
0	2	100	100	18	182	18	20	180	182	10.0	10	13	186	0.36
0	10	100	100	18	182	18	18	182	182	10.0	10	6	193	0.37
2	1	100	100	19	181	18	23	177	182	9.7	10	18	181	0.38
2	2	100	100	18	182	18	25	175	182	10.0	10	13	186	0.36
2	10	100	100	18	182	18	18	182	182	10.0	10	6	193	0.36
10	1	100	100	18	182	18	18	181	182	10.0	10	18	181	0.30
10	2	100	100	18	182	18	18	182	182	10.0	10	13	186	0.31
10	10	100	100	18	182	18	18	182	182	10.0	10	6	193	0.35

Table 6: Normal populations, $c_1 = 1$, $c_2 = 10$, $c_0 = 200$.

Param.		Un. Aloc.		Batch Sequential Allocation, m=20								Known σ_i^2		
μ_2,	σ_2^2	$\bar{N}_1^u$	$\bar{N}_2^u$	$\bar{N}_1$	$\bar{N}_2$	N_1^l	N_1^m	N_2^l	N_2^m	b	w_2	N_1	N_2	DSU
0	1	100	10	17	15	31	50	15	17	5.0	1	48	15	0.44
0	2	100	10	39	16	23	50	15	18	5.0	1	36	16	0.43
0	10	100	10	23	18	21	41	16	18	5.0	1	18	18	0.39
2	1	100	10	44	16	31	50	15	17	5.0	1	48	15	0.42
2	2	100	10	36	16	24	50	15	18	5.0	1	36	16	0.41
2	10	100	10	23	18	21	38	16	18	5.0	1	18	18	0.39
10	1	100	10	21	18	21	24	18	18	5.0	1	48	15	0.30
10	2	100	10	21	18	21	22	18	18	5.0	1	36	16	0.30
10	10	100	10	21	18	21	21	18	18	5.0	1	18	18	0.33
0	1	100	10	27	17	21	30	17	18	5.0	2	2	17	0.42
0	2	100	10	23	18	21	30	17	18	5.0	2	20	18	0.42
0	10	100	10	21	18	21	21	18	18	5.0	2	9	19	0.41
2	1	100	10	25	18	21	30	17	18	5.0	2	27	17	0.38
2	2	100	10	22	18	21	30	17	18	5.0	2	20	18	0.40
2	10	100	10	21	18	21	21	18	18	5.0	2	9	19	0.41
10	1	100	10	21	18	21	21	18	18	5.0	2	27	17	0.29
10	2	100	10	21	18	21	21	18	18	5.0	2	20	18	0.29
10	10	100	10	21	18	21	21	18	18	5.0	2	9	19	0.33
0	1	98	9	6	19	6	10	19	19	5.0	10	6	19	0.32
0	2	98	9	6	19	6	8	19	19	5.0	10	4	20	0.35
0	10	98	9	6	19	6	6	19	19	5.0	10	1	20	0.35
2	1	98	9	6	19	6	7	19	19	5.0	10	6	19	0.30
2	2	98	9	6	19	6	6	19	19	5.0	10	4	20	0.33
2	10	98	9	6	19	6	6	19	19	5.0	10	1	20	0.34
10	1	98	9	6	19	6	6	19	19	5.0	10	6	19	0.24
10	2	98	9	6	19	6	6	19	19	5.0	10	4	20	0.24
10	10	98	9	6	19	6	6	19	19	5.0	10	1	20	0.27

REFERENCES

Box, G. E. P. and G. C. Tiao (1973). Bayesian Inference in Statistical Analysis. Addison Wesley, Reading Mass.

Ghurye, S.G. and Robbins H. (1954). Two stage procedures for estimating the difference between means. Biometrika 41 146-152.

Govindarajulu, Z. (1987). The Sequential Statistical Analysis of Hypothesis Testing, Point and Interval Estimation, and Decision Theory. American Series in Mathematical and Management Sciences, Columbus OH.

Dudewicz, E. J. and T. Bishop (1979). The heteroscedastic method. Optimization Methods in Statistics (J. S. Rustagi ed.). Academic Press, New York, 183-203.

Halton, J.H. and E. A. Zeidman (1971). The Evaluation of Multi Dimensional Integrals by the Monte Carlo Sequential Stratification Techni-que. Technical Report No 137. Department of Computer Sci., University of Wisconsin .

Kuo, L. and N. Mukhopadhyay (1990). Multistage Point and Interval Estimation of the Largest Mean of k Normal Populations and the Associated Second Order Properties. Technical Report, 13 pp., Department of Statistics, University of Connecticut.

Lasdon, L. S. (1970). Optimization Theory for Large Systems. MacMillan, New York.

Lindley, D. V. (1980). Introduction to Probability and Statistics. Part 2 Inference. University Press, Cambridge.

Peierls, R. F. and Yahav J. A. (1986). Adaptive allocation in importance sampling. Adaptive Statistical Procedures and Related Topics (J. Van Ryzin ed.). I.M.S. Lecture Notes-Monograph Series, 8, 29-39.

Robbins H. (1985). Selected Papers (Lai, T. L. and Siegmund, D. eds.), Springer Verlag.

Ross, S. M (1970). Introduction to Probability Models with Optimization Applications. Holden Day , San Francisco.

Serfling, R. J. (1980). Approximation Theorems of Mathematical Statistics. Wiley, New York.