Survey Methodology, June 1989
Vol. 15, No. 1, pp. 47–57
Statistics Canada

Sample Allocation in Multivariate Surveys

JAMES BETHEL[1]

ABSTRACT

The optimum allocation to strata for multipurpose surveys is often solved in practice by establishing linear variance constraints and then using convex programming to minimize the survey cost. Using the Kuhn-Tucker theorem, this paper gives an expression for the resulting optimum allocation in terms of Lagrangian multipliers. Using this representation, the partial derivative of the cost function with respect to the k-th variance constraint is found to be $-2\alpha_k^* g(x^*)/v_k$, where $g(x^*)$ is the cost of the optimum allocation and where α_k^* and v_k are, respectively, the k-th normalized Lagrangian multiplier and the upper bound on the precision of the k-th variable. Finally, a simple computing algorithm is presented and its convergence properties are discussed. The use of these results in sample design is demonstrated with data from a survey of commercial establishments.

KEY WORDS: Multiple objective sample allocation; Nonlinear programming; Stratified sampling.

1. INTRODUCTION

The problem of optimum sample allocation in surveys with multiple study objectives was first discussed by Neyman (1934) in his development of the theory for solving the univariate optimum allocation problem. Since then, many researchers have studied the multivariate problem and several approaches have been suggested, most of which fall into one of two categories. The first involves forming a weighted average of the stratum variances and finding the optimal allocation for the "average variance" which results. Dalenius (1953), Yates (1960), Folks and Antle (1965), Hartley (1965), and Kish (1976) discuss methods related to this approach. The second basic technique is to require that each variance satisfy an inequality constraint and then use convex programming to obtain the least cost allocation which satisfies all the constraints. Dalenius (1957), Yates (1960), Kokan (1963), Hartley (1965), Kokan and Khan (1967), Chatterjee (1968,1972), Huddleston, Claypool, and Hocking (1970), Bethel (1985), and Chromy (1987) all discuss the use of convex programming in relation to the multivariate optimal allocation problem. Each approach has its advantages and disadvantages. The "weighted average" method is computationally simple, intuitively appealing, and can be solved under a fixed cost assumption, but the choice of the weights is arbitrary and the optimality properties are not clear. The "convex programming" approach gives the optimal solution to the defined problem but the resulting cost may not be acceptable so that a further search is usually required for an optimal solution which falls within the budgetary constraints.

In this paper, a closed expression for the optimal allocation subject to linear inequality constraints will be given in terms of Lagrangian multipliers. In this framework, two results easily follow which substantially overcome the disadvantages of the convex programming approach. The first is that scaling the optimal multivariate allocation results in an allocation which is optimal under constraints which are proportionate to the original ones. Thus, if the optimal solution is too costly, it can be scaled down to the allowable budget directly and the effects of this on the precision of sample estimates can be directly determined. The second result is

[1] James Bethel, Westat, 1650 Research Boulevard, Rockville, MD. 20850 USA.

a simple expression for the partial derivatives of the cost of the sampling allocation with respect to the variance constraints. These quantities, called "shadow prices", show the sensitivity of the cost to variance constraints and are useful in assessing the cost effectiveness of the sample design.

The problem of solving the convex optimization still remains. Much has been written on methods for solving programming problems of this type and there are many software packages available for doing so. Some special programming considerations will be discussed here, however, and a simple method will be presented. This algorithm, essentially a steepest descent procedure, is convergent, straightforward to program, and easy to use, since no initial values are required. An example will be presented which demonstrates this algorithm and the other techniques discussed above.

2. THE ALLOCATION MODEL

Consider the case of stratified random sampling with I strata and J variables. Suppose it is required that the j-th variable satisfy

$$\operatorname{Var}(\bar{y}_j) \approx \sum_{i=1}^{I} W_i^2 S_{ij}^2 / n_i \leq v_j^2, \tag{1}$$

where S_{ij}^2, n_i, and W_i^2, are, respectively, the variance of the j-th response variable, the sample allocation, and the proportion of the population that fall in the i-th stratum, and where v_j is an arbitrary, positive constant. In this paper it will be assumed that the finite population correction factors are negligible. In practice, it is expected that the effects of this assumption, which will be discussed in more detail in Section 7, would be limited.

Let

$$\begin{aligned} x_i &= 1/n_i \text{ if } n_i \geq 1 \\ &= \infty \quad \text{otherwise} \end{aligned}$$

and assume the cost function

$$g(x) = \sum_{i=1}^{I} c_i / x_i, \quad c_i > 0, \, i = 1, 2, \ldots, I. \tag{2}$$

A constant term for fixed costs could be included, but this would not affect the minimization process and is deleted here to simplify the notation. Define the constants

$$a_{ij} = w_i^2 S_{ij}^2 / v_j^2 \tag{3}$$

which will be referred to as "standardized precision units". Notice that $a_{ij} \geq 0$. Using this notation, the optimal allocation problem can be expressed as follows:

$$\begin{aligned} &\text{Minimize} && g(x) \\ &\text{subject to} && a_j' x \leq 1, \qquad j = 1, 2, \ldots, J \\ & && x > 0 \end{aligned} \tag{4}$$

where a_j is the j-th column vector of the matrix $A = \{a_{ij}\}$.

Kokan (1963) discusses this allocation model extensively and shows how it can be adapted to cover many common sample allocation problems, including cluster sampling and double sampling. Kokan and Khan (1967) give further analytical results in this context; Arthanari and Dodge (1981) restate Kokan and Khan's results. In related work, Kish (1976) describes a class of "linear forms" which occur frequently in survey research and to which many of the results developed here will apply.

3. THE OPTIMUM ALLOCATION

The optimum allocation for a single variable is well known. In that case $J = 1$, and the minimum of $g(x)$ subject to $a_1'x \leq 1$ with $x > 0$, denoted by x^*, is given by

$$\begin{aligned} x_i^* &= \sqrt{c_i} / \left(\sqrt{a_{i1}} \sum_{k=1}^{I} \sqrt{c_k a_{k1}} \right) && \text{if } a_{i1} > 0,\ 1 \leq i \leq I \\ &= \infty && \text{otherwise.} \end{aligned} \tag{5}$$

In this section, formula (5) will be extended to the situation where $J > 1$.

The function g in (2) is strictly convex for $x > 0$, and the constraints given by (4) are linear, so that the basic results in convex programming apply here without difficulty. That an optimal solution always exists was demonstrated by Kokan and Khan (1967). As above, denote the optimal solution by x^*. It follows from the Kuhn-Tucker Theorem (1951) that there exist $\lambda_j \geq 0$ such that

$$\nabla g(x^*) + \sum_{j=1}^{J} \lambda_j a_j = 0 \tag{6}$$

(∇ denotes the gradient) and

$$\lambda_j \left(a_j' x^* - 1 \right) = 0 \tag{7}$$

for $j = 1, 2, \ldots, J$. If $x > 0$ satisfies $\Sigma_{j=1}^{J} \lambda_j a_j' x \leq \Sigma_{j=1}^{J} \lambda_j$, then, combining (6) and (7),

$$-x' \nabla g(x^*) = \sum_{j=1}^{J} \lambda_j a_j' x \leq \sum_{j=1}^{J} \lambda_j = \sum_{j=1}^{J} \lambda_j a_j' x^* = -x^{*\prime} \nabla g(x^*). \tag{8}$$

By convexity, $g(x) - g(x^*) \geq (x - x^*)' \nabla g(x^*)$ (for all $x > 0$ with $x^* > 0$). Thus, from (8)

$$g(x) - g(x^*) \geq (x - x^*)' \nabla g(x^*) \geq 0.$$

It follows that x^* is the minimum of $g(x)$ subject to the conditions

$$\sum_{j=1}^{J} \lambda_j a_j' x \leq \sum_{j=1}^{J} \lambda_j \text{ for all } x > 0.$$

Since the minimization of g is unaffected by positive multiplicative constants, x^* also minimizes $g(x)$ subject to the constraints that $\sum_{j=1}^{J} \alpha_j^* a_j' x \leq 1$ and $x > 0$, where $\alpha_j^* = \lambda_j / \sum_{j=1}^{J} \lambda_j$.

The extension of formula (5) to an expression for the optimum multivariate allocation now consists of applying the former to the weighted sum $\Sigma_{j=1}^{J} \alpha_j^* a_j$:

$$x_i^* = \sqrt{c_i} / \left(\sqrt{\sum_{j=1}^{J} \alpha_j^* a_{ij}} \sum_{k=1}^{I} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^* a_{ij}} \right) \quad \text{if} \sum_{j=1}^{J} \alpha_j^* a_{ij} > 0,\ 1 \le i \le I \tag{9}$$

$$= \infty \quad \text{otherwise.}$$

Notice that since x^* minimizes $g(x)$ subject to $a_j'x \le 1$, with $x > 0$ for $1 \le j \le J$, it follows that mx^* minimizes $g(mx)$ subject to the constraints $a_j'(mx) \le m$, with $x > 0$ for $1 \le j \le J$. Thus, as noted earlier, constraints on variances (or CV's) can be scaled by a factor m (or $\sqrt{m}$) if survey costs are too high.

Formula (9), of course, is computationally useful only if the α_j^* are known. However, this formula is useful for deriving the shadow prices and for developing an algorithm for obtaining x^* and the α_j^*.

4. SENSITIVITY OF SURVEY COST TO VARIANCE CONSTRAINTS

In many optimization problems, it is useful to know how the optimal solution behaves when the constraints are perturbed slightly. This can be especially true in survey research, where trade-offs between costs, survey operations and precision requirements are frequently required. In any case, the "shadow prices", given by $\partial g(x^*)/\partial v_k$, are useful in detecting small shifts in the variance constraints which could substantially reduce the overall survey cost.

Combining (2), (3), and (9), it is easily seen that the cost of the optimum allocation is

$$g(x^*) = \left(\sum_{i=1}^{I} \sqrt{c_i \sum_{j=1}^{J} \alpha_j^* a_{ij}} \right)^2 = \left(\sum_{i=1}^{I} \sqrt{c_i \sum_{j=1}^{J} \alpha_j^* W_i^2 S_{ij}^2 / v_j^2} \right)^2. \tag{10}$$

Thus

$$\frac{\partial g(x^*)}{\partial v_k} = 2\left(\sum_{i=1}^{I} \sqrt{c_i \sum_{j=1}^{J} \alpha_j^* W_i^2 S_{ij}^2 / v_j^2} \right) \sum_{i=1}^{I} \frac{-c_i \alpha_k^* W_i^2 S_{ik}^2 / v_k^3}{\sqrt{c_i \sum_{j=1}^{J} \alpha_j^* W_i^2 S_{ij}^2 / v_j^2}} \tag{11}$$

$$= -2 \frac{\alpha_k^*}{v_k} \sqrt{g(x^*)} \sum_{i=1}^{I} \frac{a_{ik}\sqrt{c_i}}{\sqrt{\sum_{j=1}^{J} \alpha_j^* a_{ij}}}$$

$$= -2 \frac{\alpha_k^*}{v_k} g(x^*) \sum_{i=1}^{I} a_{ik}\sqrt{c_i} / \left(\sqrt{\sum_{j=1}^{J} \alpha_{ij}^* a_{ij}} \sum_{k=1}^{I} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^* a_{kj}} \right)$$

$$= -2 \frac{\alpha_k^*}{v_k} g(x^*)\, a_k' x^*.$$

From (7) it follows necessarily that $\alpha_k^* = 0$ whenever $a_k' x^* < 1$, so that

$$\frac{\partial g(x^*)}{\partial v_k} = -2\frac{\alpha_k^*}{v_k} g(x^*). \tag{12}$$

This formula is somewhat more complicated that the usual expression for shadow prices (*e.g.*, see Luenberger 1984), due to the complex relationship between g and v_j.

Now consider increasing v_k by (100π) %, $0 \leq \pi \leq 1$. Denote by $x^* + \Delta x^*$ the resulting perturbation in x^*. By (12),

$$g(x^* + \Delta x^*) - g(x^*) \approx \pi v_k \frac{\partial g(x^*)}{\partial v_k} = -2\, \pi\, \alpha_k^*\, g(x^*). \tag{13}$$

Thus an increase of (100π)% in the k-th variance constraint results in a $(100)(2\pi\alpha_k^*)$% reduction in the overall survey cost.

5. PROGRAMMING CONSIDERATIONS

This section discusses some technical aspects of solving for x^* and gives a simple algorithm for finding both x^* and the coefficients α_j^* by searching over weighted averages $\sum_{j=1}^{J} \alpha_j a_j$. Define δ_{ij} by

$$\begin{aligned} \delta_{ij} &= 1 \text{ if } i = j \\ &= 0 \text{ if } i \neq j. \end{aligned}$$

For a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_J)'$, define $\tilde{x}(\alpha)$ by

$$\begin{aligned} \tilde{x}_i(\alpha) &= \sqrt{c_i} \Bigg/ \left(\sqrt{\sum_{j=1}^{J} \alpha_j a_{ij}}\ \sum_{k=1}^{I} \sqrt{c_k \sum_{j=1}^{J} \alpha_j a_{kj}} \right) \text{ if } \sum_{j=1}^{J} \alpha_j a_{ij} > 0,\ 1 \leq i \leq I \\ &= \infty \qquad\qquad \text{otherwise.} \end{aligned}$$

Notice that $\tilde{x}(\alpha^*) = x^*$. Now the iterative algorithm for finding x^* is defined as follows:

1. Take $\alpha_j^{(1)} = \delta_{1j}$, $1 \leq j \leq J$.

2. At step $n \geq 2$, find an index k for which

$$(a_k - a_j)'\, \tilde{x}(\alpha^{(n)}) \geq 0,\ 1 \leq j \leq J. \tag{14}$$

This gives the constraint which the current optimum solution violates by the largest margin. If $a_k' \tilde{x}(\alpha^{(n)}) \leq 1$, then terminate the algorithm. Otherwise, find $t^{(n)} \in (0,1)$ for which

$$g(\tilde{x}(t^{(n)}\, \delta_k + (1 - t^{(n)})\alpha^{(n)}) \geq g(\tilde{x}(t\, \delta_k + (1 - t)\, \alpha^{(n)})) \text{ for all } t \in [0, 1]. \tag{15}$$

3. Take $\alpha_j^{(n+1)} = t^{(n)}\,\delta_{kj} + (1 - t^{(n)})\,\alpha_j^{(n)}$.
4. Terminate when $|\alpha_j^{(n+1)} - \alpha_j^{(n)}| < \epsilon$, $1 \le j \le J$, where ϵ is a predetermined convergence criterion.

To verify the convergence of the algorithm, first note that $\tilde{x}(\alpha)$ minimizes $g(x)$ subject to $\sum_{j=1}^{J}\alpha_j a_j' x \le 1$. Thus, since $\sum_{j=1}^{J}\alpha_j a_j' x^* \le \sum_{j=1}^{J}\alpha_j = 1$,

$$0 \le g(\tilde{x}(\alpha^{(n)})) \le g(x^*) \tag{16}$$

for all n. Furthermore, from (15), $g(\tilde{x}(\alpha^{(n)}))$ is nondecreasing, implying the convergence of $g(\tilde{x}(\alpha^{(n)}))$. To see that $\tilde{x}(\alpha^{(n)}) \to x^*$, first define

$$h_{k\alpha}(t) = \sum_{i=1}^{I}\sqrt{c_i \sum_{j=1}^{J}(t\delta_{kj} + (1-t)\alpha_j)a_{ij}} = \sqrt{g(\tilde{x}(t\delta_k + (1-t)\alpha))}. \tag{17}$$

Since $h_{k\alpha}(t)$ is concave (*i.e.*, $-h_{k\alpha}(t)$ is convex),

$$\begin{aligned} h_{k\alpha}(t) - h_{k\alpha}(0) &= t\,h'(0) + O(t^2) \qquad (18)\\ &= -t\sum_{i=1}^{I}\frac{\sum_{j=1}^{J}(\delta_{kj} - \alpha_j)\,a_{kj}\sqrt{c_k}}{2\sqrt{c_i\sum_{j=1}^{J}\alpha_j a_{ij}}} + O(t^2)\\ &= (t/2)\sqrt{g(\tilde{x}(\alpha))}\,(a_k'\tilde{x}(\alpha) - 1) + O(t^2). \end{aligned}$$

By allowing t to tend toward zero, it follows that there exists $t \in (0,1)$ for which

$$\sqrt{g(\tilde{x}(t\delta_k + (1-t)\alpha))} = h_{k\alpha}(t) > h_{k\alpha}(0) = \sqrt{g(\tilde{x}(\alpha))}$$

if and only if $a_k'\tilde{x}(\alpha) > 1$. Thus it follows from (15) that the constraints are satisfied at convergence; combining this with (16) implies that $\lim_{n\to\infty}\tilde{x}(\alpha^{(n)}) = x^*$.

In carrying out the algorithm, Step 2 requires a search for $t^{(n)}$. Define $h_{k\alpha}(t)$ as in (17). It is clear from the preceding discussion that $a_k'\tilde{x}(t\delta_k + (1-t)\alpha^{(n)}) = 1$ when $h(t)$ (and hence g) is at a maximum. Furthermore, since $h_{k\alpha}(t)$ is strictly concave, $h_{k\alpha}'(t)$ is nonincreasing in t and thus the point where $h_{k\alpha}'(t) = 0$ is unique. It follows that a binary search for the point where $h_{k\alpha}(t)$ is maximized can be implemented by simply checking to see whether $a_k'\tilde{x}(t\delta_k + (1-t)\alpha^{(n)}) = 1$, providing a rapid means of obtaining a close approximation for $t^{(n)}$.

As described above, the algorithm takes a_1 as the initial value. This is completely arbitrary, since any of the a_j, $1 \le j \le J$, would do. In practice, the constraint for which the optimum allocation (*i.e.*, formula (5)) yields the highest cost is generally a good choice for the starting value.

Notice that Step 2 of the algorithm will require IJ calculations in formula (14) and a 10-step (say) binary search of $3I + J + 1$ calculations each in formula (15), while J calculations must be carried out in Step 3. Thus each iteration of the algorithm is $O(IJ)$. From (18), at the n-th interation,

$$h'_{k\alpha}(0) \approx \frac{1}{2} h_{k\alpha}(0) \ (a'_k\tilde{x}(\alpha^{(n)}) - 1)$$

so that $a'_k\tilde{x}(\alpha^{(n)})$ is approximately proportionate to $h'_{k\alpha}$ (0) (up to an additive constant). Heuristically, $h'_{k\alpha}(0)$ is the "slope" of h in the direction of a_k, suggesting that the algorithm is essentially a steepest descent (or ascent, in this case) procedure. This, in turn, suggests a linear rate of convergence (see, for example, Forsyth 1968).

In the author's experience (see Bethel 1985), the algorithm converges quickly for most moderately sized problems. For example, sample allocation problems with 20-30 strata and 5-10 constraints were solved in 3-5 seconds using the algorithm (on a Compaq 38620 with a 30387 math co-processor) versus 6-8 seconds using a sequential unconstrained minimization technique (SUMT) implementing a penalized steepest descent algorithm. Run times vary considerably depending on the magnitude of the problem, the number of active constraints, and, obviously, machine characteristics. The author's computing experience (with problems of 20-30 strata and 5-10 constraints) includes the Macintosh SE (30 seconds to 2 or 3 minutes), Leading Edge Model D (1 to 5 minutes), Zilog System 8000 (5 to 60 seconds), and the Compaq mentioned above (5 to 10 seconds). However, the run times are generally insignificant in comparison with the labor involved in creating files and other preparatory tasks. In particular, it may take several hours to find an acceptable starting value for the SUMT algorithm. Thus a strong feature of the algorithm described in Steps 1-4 above is that it requires no external initial values. Moreover, it is relatively easy to program, requiring only 40 or 50 lines of code.

An even simpler algorithm is given by Chromy (1987). It can be adapted to our notation and general approach as follows: Set $\alpha_j^{(1)} \equiv 1/J$, and, for $n \geq 2$, let

$$\alpha_j^{(n)} = \alpha_j^{(n-1)}(a'_j\tilde{x}(\alpha^{(n-1)}))^2 / \sum_{j=1}^{J} \alpha_j^{(n-1)}(a'_j\tilde{x}(\alpha^{(n-1)}))^2 \quad 1 \leq j \leq J. \tag{19}$$

Like the algorithm described in steps 1-4 above, (19) requires no external initial values; (19), however, requires even less programming effort and, based on several comparisons, it appears to converge considerably more quickly. Unfortunately, there is apparently no formal proof of convergence, although considerable practical experience (see Chromy 1987 for a more detailed discussion) suggests that it has good convergence properties.

6. EXAMPLE

Tables 1-3 present an example drawn from a survey of commercial establishments. (Only the strata for educational institutions are shown here.) Four of the primary variables of interest are given: area of enclosed floorspace, age of building, number of full-time employees, and percent of buildings heated by oil. Table 1 gives the stratum level variance information. Here the standardized precision units are computed as

$$a_{ij} = \frac{W_i^2 S_{ij}^2}{\bar{Y}_j^2 v_j^2}$$

Table 1.
Allocation Example: Survey of Educational Institutions.

	Stratum Standard Deviation					
	Weight	Floorspace	Age	Employees	Pct. Oil Heating
Stratum					
1	.5158	22,319.11	43.71	25.72	48.15
2	.2632	24,056.21	16.68	27.09	36.79
3	.1184	54,201.75	24.70	17.11	48.04
4	.0711	155,514.21	16.01	59.46	38.07
5	.0184	125,239.21	14.74	51.27	48.80
6	.0132	355,392.69	20.90	212.13	57.74
Mean:		54,641.85	43.03	45.23	67.58
v_k:		.06	.06	.06	.06

	Standardized Precision Units			
	Floorspace	Age	Employees	Pct. Oil Heating
Stratum				
1	12.33	76.24	23.90	37.52
2	3.73	2.89	6.93	5.70
3	3.83	1.28	.56	1.96
4	11.36	.19	2.44	.45
5	7.37	.01	.12	.05
6	2.03	.01	1.06	.04
Required Sample Size:	222	149	127	121

where $v_j = .06$ for all variables (so that the half-width of a 90% confidence interval will be approximately 10% of the mean). Also given are the sample sizes required for Neyman allocation for each of the variables taken individually. Survey costs are assumed to be constant across strata.

Table 2 gives the first-pass solution, which requires a sample of 241 units. The normalized Lagrangian coefficients and the achieved precision levels are given, from which it is apparent that floorspace and building age are dominating the solution while the other variables are not "active". Here the starting value $\alpha^{(1)} = (1,0,0,0)$ was used; because the third and fourth constraints were always satisfied, there was only one iteration with a 9-step binary search for $t^{(1)}$. (The successive estimates for the optimal t were 1/2, 1/4, 3/8, 5/16, 11/32, 21/64, 43/128, 85/256, and 171/512.) Also given in Table 2 are the 10% shadow prices: 10% increases in the first (or second) constraints would result in a sample size reduction of approximately 32 (or 16) units. Since the third and fourth constraints are not active in the solution, changing their CV requirements would have no effect on the allocation or the sampling costs.

Table 3 gives a second pass solution under the requirement that the total sample size is no larger than 200. The optimal solutions are thus scaled by 241/200 (so that the optimal allocation goes down by 200/241) and the resulting CV's are scaled by $\sqrt{241/200}$. The new 10% shadow prices are −27 and −13 for the first and second constraints, reflecting the decrease in the overall survey cost. Notice that there is approximately a 10% increase in the CV's (from the original ones in Table 1), so that the sample reduction of 48 predicted by the shadow prices in Table 2 compares favorably with the actual 41 unit reduction. (The shadow price predictions will always be somewhat optimistic due to the linear approximation.)

Table 2.

Allocation Example: First Pass Optimum Solution.

Stratum	$\Sigma\ \alpha_j^*\ a_{ji}$	x_i^*	Optimum Allocation
1	33.6749	.0111	90
2	3.4495	.0347	29
3	2.9783	.0373	27
4	7.6294	.0233	43
5	4.9119	.0291	34
6	1.3554	.0553	18
Total:			241

	Floorspace	Age	Employees	Pct. Oil Heating
Lagrangian Multiplier (Normalized):	.6660	.3340	.0000	.0000
Achieved Precision:	.0600	.0600	.0481	.0502
10% Shadow Prices:	-32	-16	0	0

Table 3.

Allocation Example: Optimum Solution for Sample Size Limited to 200.

Stratum	$\Sigma\ \alpha_j^*\ a_{ji}$	x_i^*	Optimum Allocation
1	33.6749	.0134	75
2	3.4495	.0418	24
3	2.9783	.0449	22
4	7.6294	.0281	36
5	4.9119	.0351	29
6	1.3554	.0666	15
Total:			201

	Floorspace	Age	Employees	Pct. Oil Heating
Lagrangian Multiplier (Normalized):	.6660	.3340	.0000	.0000
Achieved Precision:	.0657	.0658	.0528	.0551
10% Shadow Prices:	-27	-13	0	0

7. DISCUSSION

In this paper we have given a formal representation for the optimal sample allocation for a multipurpose survey with linear variance constraints, and derived expressions for the partial derivatives of the cost function with respect to the precision constraints. The latter result, in particular, provides approximations that are useful in survey planning, permitting a great deal of exploratory work without exact computer calculations.

Throughout the paper, the normalized Lagrangian multipliers, α_j^*, play a key role. In particular, we have noted that whenever the j-th variance constraint is not "active" in the solution to the allocation problem, the j-th Lagrangian $\alpha_j^* = 0$.

The optimization approach discussed in this article yields a continuous solution, which must then be rounded in some way to provide integer stratum sample sizes. Clearly this rounding will cause some deviation from optimality. However, the objective function here is generally considered to be rather insensitive to small deviations from optimality (see Cochran 1977), so that exact integer solutions are probably not cost effective. In fact, it seems likely that round-off error would be insignificant in comparison with the sampling errors in estimates of means and variances that would normally be available for developing an optimized survey design.

Finally the reader will recall that finite population correction factors have been ignored throughout this paper. It is easy to include these in the allocation model by manipulating equations (1) and (3), although that would cause equation (13) to be somewhat imprecise. However, it should be kept in mind that even when the FPC is non-negligible for some of the strata, the overall effect usually is negligible. In any case, the FPC term, $\sum_{i=1}^{I} W_i^2 S_{ij}^2/N_i$, can always be calculated in order to evaluate the situation and, if necessary, it can be added to v_k in formula (13) to obtain exact results.

ACKNOWLEDGEMENTS

The author acknowledges the comments of the referrees and useful discussions with Rick Williams, of Research Triangle Institute, and Patrick McCarthy, of Applied Management Sciences, all of which led to substantial improvements in the article.

REFERENCES

ARTHANARI, T.S., and DODGE, Y., (1981). *Mathematical Programming in Statistics*. New York: John Wiley and Sons.

BETHEL, J.W. (1985). An optimum allocation algorithm for multivariate surveys. *Proceedings of the Survey Research Section, American Statistical Association*, 209-212.

CHATTERJEE, S. (1968). Multivariate stratified surveys. *Journal of the American Statistical Association*, 63, 530-534.

CHATTERJEE, S. (1972). A study of optimum allocation in multivariate stratified surveys. *Skandinavisk Actuarietidskrift*, 55, 73-80.

CHROMY, J.R. (1987). Design optimization with multiple objectives. *Proceedings of the Survey Research Section, American Statistical Association*, 194-199.

DALENIUS, T. (1953). The multivariate sampling problem. *Skandinavisk Actuarietidskrift*, 36, 92-102.

DALENIUS, T. (1957). *Sampling in Sweden*. Stockholm: Almquist and Wicksall.

FOLKS, J.L., and ANTLE, C.E. (1965). Optimum allocation of sampling units when there are R responses of interest. *Journal of the American Statistical Association*, 60, 225-233.

FORSYTH, G.E. (1968). On the asymptotic directions of the s-dimensional optimum gradient method. *Numerische Mathematik*, 11, 57-76.

HARTLEY, H.O. (1965). Multiple purpose optimum allocation in stratified sampling. *Proceedings of the Social Statistics Section, American Statistical Association*, 258-261.

HUDDLESTON, H.F., CLAYPOOL, P.L., and HOCKING, R.R. (1970). Optimum sample allocation to strata using convex programming. *Applied Statistics*, 19, 273-278.

KISH, L. (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society A.*, 139, 80-95.

KOKAN, A.R. (1963). Optimum allocation in multivariate surveys. *Journal of the Royal Statistical Society A.*, 126, 557-565.

KOKAN, A.R., and KHAN, S. (1967). Optimum allocation in multivariate surveys: an analytical solution. *Journal of the Royal Statistical Society B.*, 29, 115-125.

KUHN, H.W., and TUCKER, A.W. (1951). Nonlinear programming. *Proceedings 2nd Berkeley Symposium Mathematical Statistics and Probability.*

LUENBERGER, D.G. (1984). *Linear and Nonlinear Programming.* Reading, Massachusets: Addison-Wesley.

NEYMAN, J. (1934). On the two different aspects of the representative method: the method of representative sampling and the method of purposive sampling. *Journal of the Royal Statistical Society*, 558-625.

YATES, F. (1960). *Sampling Methods for Censuses and Surveys.* London: Charles Griffin and Company.