THE COEFFICIENT OF VARIATION
IN INTERVAL ESTIMATION*

By PILAR L. FERRER**

1. Introduction. In allocation problems, where the optimum sample size n is to be determined, the experimenter often wants to know how much he could deviate from the optimum solution without sacrificing the desired precision, nor increasing the cost, nor decreasing the profit angle substantially.

This study attempts specifically to use the coefficient of variation in obtaining confidence intervals for sample sizes n_h in a single-variable problem, and when there are many other characteristics of interest.

In multipurpose surveys, since the characteristics behave differently, i.e., they have different population moments, the allocation of the sample size n to the strata would likewise differ according to the characteristic upon which the allocation was based. There is then need to get some compromise allocation such that the precision set would still be met.

2. The Coefficient of Variation and the Rel-Variance as Estimators. The coefficient of variation, (cv), of any sample estimate is defined as the standard error of the estimate divided by the value being estimated; while the rel-variance, RV, is equal to the square of the coefficient of variation. Thus,

$$cv = S/\bar{X},$$

$$RV = S^2/\bar{X}^2, \text{ where } S^2 = \frac{\sum_{i=1}^{n} (X_1 - \bar{X})^2}{(n-1)}$$

---

* Excerpts from a M.A. Thesis
** Instructor, U.P. Statistical Center

Although the coefficient of variation is neither an unbiased estimate nor an efficient statistic for non-normal distributions, it still remains a popular estimate (Norris). Its main utility lies in the fact that in a sample from a skewed population, the mean and standard deviation tend to change together, thus off-setting and compensating somewhat for any over estimation or under estimation that exist in the sample mean and standard deviation (Gutierrez 1965).

The rel-variance is not an unbiased estimate either, but this estimate $v^2 = s^2 / x^2$ is a consistent estimate of the population rel-variance of the mean ($V^2 = S^2 / X^2$) and for reasonably large samples the bias will be trivial. (Hansen, Hurwitz, and Madow 1966).

3. Single Purpose Surveys.

3.1 Simple Random Sampling: Determination of Sample Size n and its Confidence Interval. The equation of the rel-variance of the mean in simple random sampling is

$$RV_{\bar{X}} = (1 - f) \frac{S^2/n}{\bar{X}^2} = (1 - f) \frac{RV_X}{n}, \qquad (1)$$

and if the finite population correction is ignored,

$$RV_{\bar{X}} = \frac{RV_X}{n} = \frac{(cv_x)^2}{n}. \qquad (2)$$

Thus, if it is desired to have a degree of precision such that t times the coefficient of variation of the mean, $cv_{\bar{x}}$, would be equal to d (where d is the relative difference between the estimated mean from a sample and the true mean), the equation becomes

$$t (cv_{\bar{x}}) = d \qquad (3)$$

and

$$(cv_{\bar{x}})^2 = d^2/t^2. \qquad (4)$$

Thus, substituting in equation (2);

$$\frac{d^2}{t^2} = \frac{(cv_x)^2}{n};$$

or

$$n = \frac{t^2(cv_x)^2}{d^2}. \tag{5}$$

Actually, this is analogous to the equation for sample size n when the aimed-at precision is given in terms of the variance S^2 instead of the rel-variance, $(cv)^2$.

Therefore, the variance of n, using equation (5), would be:

$$Var\ (n) = (t^2/d^2)^2\ Var\ (cv_x^2).$$

$$= (t^4/d^4)\ Var\ (RV_X).$$

Inasmuch as the rel-variance of the estimated rel-variance based on a simple random sample of n units drawn with replacement is:

$$RV(RV) \doteq \frac{\beta_2 - 1}{n} + \frac{4V_X^2}{n} - \frac{4\ \mu_3\ /\ \bar{X}^3 V_X^2}{n},$$

where

$$V_X^2 = (N/N\text{-}1)\ \sigma^2\ /\ \bar{X}^2$$

$$\beta_2 = \mu_4\ /\ \sigma^4$$

The variance of the rel-variance, RV, is given by:

$$Var\ (cv_x^2) = [\ (cv_x^2)\ ]^2\ [Rel\text{-}Var\ (cv_x^2)]$$

$$= (cv_x)^4 \left[\frac{\beta_2 - 1}{n} + \frac{4V_X^2}{n} \right.$$

$$\left. - \frac{4\ \mu_3\ /\ X^3 V_X^2}{n} \right]$$

The variance formula is just an approximation since simple random sampling without replacement is involved here. How-

ever, it is believed that this approximation is close to the true value whenever the number of sampling units in the population is large relative to the number in the sample.

Knowing Var (n), it is now possible to obtain confidence intervals for sample size n from the following relation:

$$n \pm t' \sqrt{\text{Var } n} = n \pm t' \left\{ (t'/d)^4 cv_x^4 \left[\frac{\beta_2 - 1}{n} + \frac{4V_X^2}{n} - \frac{4 \mu_3 / \bar{X}^3 V_X^2}{n} \right] \right\}^{1/4} \quad (6)$$

where t' will take the value corresponding to the degree of confidence one wants to achieve.

3.2 *Stratified Random Sampling*: Determination of stratum s ample size n_h and its Confidence Interval.

In simple random sampling, the equation for the rel-variance of the original variate, RV_X is

$$RV_X = \frac{(N - n)}{n(N - 1)} \frac{S^2}{\bar{X}^2} = \frac{(N - n)}{n(N - 1)} RV_X \quad (7)$$

Suppose that the population is divided into several strata, and that independent estimates of RV_X are obtained for these L strata. The sum

$$\sum_{h=1}^{L} RV_h(X) = \sum_{h=1}^{L} \frac{(N_h - n_h)}{n_h(N_h - 1)} RV_h(X) \quad (8)$$

can be taken. If $\sum_{h=1}^{L} RV_h(X)$ is set equal to a certain desired precision, say RV_{oh}, for the hth stratum,

$$\sum_{h=1}^{L} RV_{oh} = \sum_{h=1}^{L} \frac{(N_h - n_h)}{n_h(N_h - 1)} RV_h(X) \quad (9)$$

To estimate n_h for h = 1, ..., L, $\sum_{h=1}^{L} RV_{oh}$ is to be minimized with the constraint

$$n = \sum_{h=1}^{L} n_h$$

imposed. With the use of Lagrange·multiplier,

$$F = \sum_{h=1}^{L} \frac{(N_h - n_h)}{n_h(N_h - 1)} RV_h(X) + \lambda (n_1 + \cdots + n_L - n)$$

is minimized with respect to n_h, $h = 1, \ldots L$.

The results of the minimization of this function as obtained by Sinsioco (1969) are:

$$n_h = \frac{cv_h}{\sum_{h=1}^{L} cv_h} n \tag{10}$$

and

$$n = \frac{(\sum_{h=1}^{L} cv_h)^2}{\sum_{h=1}^{L} RV_{oh} + \frac{\sum_{h=1}^{L} RV_h}{N_h - 1}},$$

or, if the finite population correction is ignored,

$$n = \frac{(\sum_{h=1}^{L} cv_h)^2}{\sum_{h=1}^{L} RV_{oh}}.$$

From equation (10), a variance formula for sample size n_h can be formulated. However, there would arise several different cases dependent upon the assumptions made with regard to the coefficients of variation. As:

$$\text{Var } n_h = \text{var} \left[\frac{cv_h}{\sum_{h=1}^{L} cv_h} n \right]$$

Case I. Let $\left(\dfrac{cv_h}{\sum_{h=1}^{L} cv_h} \right)$ be constant. Therefore,

$$\text{Var } n_h = \left[\frac{cv_h}{\sum_{h=1}^{L} cv_h} \right] \text{Var } (n).$$

Since
$$n = \frac{\left(\sum\limits_{h=1}^{L} cv_h\right)^2}{\sum\limits_{h=1}^{L} RV_{oh}} ,$$

$$\text{Var } n = \text{Var } \left[\frac{\left(\sum\limits_{h=1}^{L} cv_h\right)^2}{\sum\limits_{h=1}^{L} RV_{oh}}\right]$$

$$= \left[\frac{1}{\sum\limits_{h=1}^{L} RV_{oh}}\right]^2 \text{Var } \left[\left(\sum\limits_{h=1}^{L} cv_h\right)^2\right] \quad (11)$$

Now
$$\text{Var } \left[\left(\sum\limits_{h=1}^{L} cv_h\right)^2\right] = \text{Var } \left[\sum\limits_{h=1}^{L} cv_h^2 + 2 \sum\limits_{i<j}^{L} cv_i cv_j\right]$$

$$= \text{Var } \left[\sum\limits_{h=1}^{L} cv_h^2\right] + 4 \text{ Var } \left[\sum\limits_{i<j}^{L} cv_i \, cv_j\right]$$

$$+ \text{ covariance term}$$

$$= \sum\limits_{h=1}^{L} \text{Var } cv_h^2 + 4 \text{ Var } \sum\limits_{h=1}^{L} cv_i \, cv_j \quad (12)$$

$$+ \text{ covariance term}$$

An approximation of the above equation can be obtained·by dropping the covariance term. It is believed that such procedure will not affect much the final equation for Var n_h since the order of the covariance term is $1/n^2$. Inasmuch as it was also assumed that the samples drawn from each stratum are independent, the stratum coefficients of variation are likewise expected to be independent.

Hansen, Hurwitz, and Madow (1966) give the formula for the variance of a product of two random variable U_1, U_2 as

$$\text{Var } (U_1 U_2) \doteq (U_1 U_2)^2 \left[\frac{\text{Var } U_1}{U_1^2} + \frac{\text{Var } U_2}{U_2^2}\right.$$

$$\left. + 2 \frac{\text{Cov}(U_1 U_2)}{U_1 U_2}\right]$$

Let $U_1 = cv_1$ and $U_2 = cv_2$. The third term of the above equation will drop out since the strata coefficients of variation were assumed to be independent.

If the variance for each pair of cv_h, $h = 1, \ldots$ L, are summed, the second term of equation (12) becomes:

$$\text{Var } (cv_1 cv_2) = (cv_1 cv_2)^2 \left[\frac{\text{Var } cv_1}{cv_1^2} + \frac{\text{Var } cv_2}{cv_2^2} \right]$$

$$\sum_{i<j}^{L} \text{Var } (cv_i cv_j) = \sum_{i<j}^{L} \left\{ (cv_i cv_j)^2 \left[\left[\frac{\text{Var } cv_i}{cv_i^2} + \frac{\text{Var } cv_j}{cv_j^2} \right] \right] \right\}$$

Simplifying, the result is:

$$\sum_{i<j}^{L} \text{Var } (cv_i cv_j) = \sum_{i<j}^{L} \left[\text{Var } cv_h \left(\sum_{i<j}^{L} cv_j \right) \right]$$

where, for a particular stratum h, the subscript j will take on all values, 1, ... L, except h.

Therefore equation (12) becomes:

$$\text{Var} \left[\left(\sum_{h=1}^{L} cv_h \right)^2 \right] = \sum_{h=1}^{L} \text{Var } cv_h^2 + \sum_{h=1}^{L} \left[cv_h^2 . \text{RV}(cv_h^2) \left(\sum_{h=1}^{L} cv_j^2 \right) \right]$$

$$\text{Var} \left[\left(\sum_{i<j}^{L} cv_h \right)^2 \right] = \sum_{h=1}^{L} \left\{ \text{Var}(cv_h^2) \left[1 + \frac{1}{cv_h^2} \left(\sum_{j=1}^{L} cv_j^2 \right) \right] \right\}$$

where $j \neq h$.

If the above is substituted in equation (11), the variance of n is:

$$\text{Var } n = \left[\frac{1}{\sum_{h=1}^{L} \text{RV}_{oh}} \right]^2 \sum_{h=1}^{L} \left\{ \text{Var}(cv_h^2) \left[1 + \frac{1}{cv_h^2} \left(\sum_{j=1}^{L} cv_j^2 \right) \right] \right\} \quad (13)$$

The formula for the variance of a stratum sample size is

$$\text{Var } n_h = \left[\frac{cv_h}{\sum\limits_{h=1}^{L} cv_h} \right]^2 \text{Var } n \qquad (14)$$

where Var n is given as in equation (13).

If the assumption in Case I is not made, the variance of n_h can be obtained thus:

$$\text{Var } n_h = \text{Var} \left[\frac{cv_h}{\sum\limits_{h=1}^{L} cv_h} \cdot n \right]$$

where
$$n = \frac{(\sum\limits_{h=1}^{L} cv_h)^2}{\sum\limits_{h=1}^{L} RV_{oh}}$$

$$\text{Var } n_h = \left[\frac{cv_h \cdot n}{\sum\limits_{h=1}^{L} cv_h} \right]^2 \left[\frac{\text{Var } cv_h}{cv_h^2} + \frac{\text{Var } (n/\sum\limits_{h=1}^{L} cv_h)}{(n/\sum\limits_{h=1}^{L} cv_h)^2} \right.$$

$$\left. + 2 \frac{\text{Cov } (cv_h, n/\sum\limits_{h=1}^{L} cv_h)}{\dfrac{cv_h \cdot n}{\sum\limits_{h=1}^{L} cv_h}} \right]$$

Simplifying the numerator of the second term

$$\text{Var } \frac{n}{\sum\limits_{h=1}^{L} cv_h} = \text{Var} \left[\frac{(\sum\limits_{h=1}^{L} cv_h)^2/ K_o}{\sum\limits_{h=1}^{L} cv_h} \right],$$

where $K_o = \sum\limits_{h=1}^{L} RV_{oh}$

$$= \text{Var} \left[\frac{\sum\limits_{h=1}^{L} cv_h}{K_o} \right]$$

PILAR L. FERRER

$$= \frac{1}{K_o^2} \left[\text{Var} \left(\sum_{h=1}^{L} cv_h \right) \right]$$

$$= \frac{1}{K_o^2} \left[\sum_{h=1}^{L} \text{Var } cv_h \right]$$

and likewise the numerator of the third term,

$$\text{Cov} \left(cv_h, \frac{n}{\sum_{h=1}^{L} cv_h} \right) = \text{Cov} \left(cv_h, \frac{\sum_{h=1}^{L} cv_h}{K_o} \right)$$

$$= \frac{1}{K_o} \left[\text{Var } cv_h \right] ,$$

the variance formula for n_h becomes

$$\text{Var } n_h = \frac{(cv_h.n)^2}{\left(\sum_{h=1}^{L} cv_h \right)^2} \left[\text{Rel-Var } (cv_h) + \frac{\sum_{h=1}^{L} \text{Var } cv_h \left(\sum_{h=1}^{L} cv_h \right)^2}{(K_o.n)^2} \right.$$

$$\left. + 2 \frac{\text{Var } cv_h \left(\sum_{h=1}^{L} cv_h \right)}{K_o.cv_h.n} \right]. \qquad (15)$$

Cases II and III employ the above equation in getting confidence intervals for n_h, where different assumption about the coefficients of variation are made.

Case II. Specify the precision by making the supposition that

$$\sum_{h=1}^{L} cv_h = \sum_{h=1}^{L} \text{ (all upper limits of } cv_h)$$

$$= \sum_{h=1}^{L} cv_{hu}$$

where cv_{hu} is obtained from $cv_h + t \sqrt{\text{Var } cv_h}$

Case III. For this case, it is assumed that

$$\sum_{h=1}^{L} cv_h = K \text{ (largest } cv_h)$$

where K = number of strata.

Confidence limits for the stratum sample sizes n_h can then be obtained by using either equations 14 or 15, and taking into account the different assumptions made in the three methods of estimating Var (n_h).

4. Multipurpose Surveys. In sampling practice it is quite unusual if only one population characteristic is being estimated. Without loss of generality, it will be assumed that interest lies on the estimates of the population means of several characteristics. The notations to be used are:

$\bar{x}_{ij}$ — mean of the ith stratum of the jth variable

n_{ij} — sample size for the ith stratum of the jth variable

N_{ij} — number of elements of ith stratum of jth variable

cv_{ij} — coefficient of variation of ith stratum of jth variable

RV_{ij} — rel-variance of the ith stratum of jth variable

U_{ij} — the reciprocal of the sample size n_{ij}

For each characteristic, the sum of the rel-variances over all the L strata is desired to be less than or equal to a specific value, say RV_{oj}, for the jth characteristic.

$$RV_{1j} + RV_{2j} + \ldots + RV_{Lj} \leq RV_{oj}; \ j = 1, \ldots p$$

So, all in all, there would be p constraints of the above type and L restrictions of the form

$$O \leq n_i \leq N_i; \qquad i = 1, \ldots L$$

4.1 Linear Programming: Charnes and Lemke's Method.

The linear programming technique is to minimize or maximize an objective function subject to certain restrictions. In this study, since each stratum sample size n_{ij} estimated has a com-

puted confidence interval, there would.arise restrictions of the form

$$n_{ijv} \leqq n_{ij} \leqq n_{iju}; \quad i = 1, \ldots L$$
$$j = 1, \ldots p$$

where n_{ijv} is the lower limit of n_{ij} and n_{jiu} the upper limit.

For this bounded-variable problem as it is known in lite-rature, the interest is not in getting higher precision than that already guaranteed by the stratum sample size n_{iu}, where n_{iu} is the largest upper bound for a particular stratum i among all n_{iju}, j = 1, ... p. Neither is the interest on any decrease in precision by reducing sample size n_i such that it is less than n_{iv}, where n_{iv} is the least lower bound among all n_{ijv}, j = 1, ... p.

If it is assumed that the problem is that of maximization, the bounded-variable problem can be written as follows:

a)   Maximize $f(U) = c_1 U_1 + \ldots + c_L U_L$,

b)   subject to: $a_{11} U_1 + \ldots + a_{L1} U_L \leqq b_1$

$$a_{1p} U_1 + \ldots + a_{Lp} U_L \leqq b_p ;$$

c) $d_{vj} \leqq U_j \leqq d_{uj}$ j = 1, ... p

and

d)   . $U_j \geqq 0$ j = 1, ... p

where d_{vj} is the lower bound of the variable U_j and d_{uj}, the upper bound. The lower bound d_{vj} may all be zero.

By adding the necessary slack variables $(U_{L+1} \ldots U_{L+p})$, $(X_1, \ldots X_L)$, the problem becomes:

a)   Maximize· $f(U) = c_1 U_1 + \ldots + c_L U_L$;

b)   subject to: $a_{11} U_1 + \ldots + a_{L1} U_L + U_{L+1} = b_1$

$$a_{1p} U_1 + \ldots + a_{Lp} U_L + U_{L+p} = b_p ;$$

c) $\qquad U_1 \qquad\qquad + \quad X_1 \quad = d_1$

$$\vdots \qquad\qquad \vdots \qquad \vdots$$

$$U_L + X_L \qquad = d_L$$

and

d) $\qquad\qquad U_i, X_i \geqq 0 \qquad\qquad\qquad i = 1, \ldots L.$

All in all, there are $p + L$ constraints (p structural and L upper-bound constraints.)

5. Application. The data used throughout are the results of a research (1967) conducted by the Department of Varietal Improvement of the International Rice Research Institute. The objective of the experiment was to determine the association between grain yield and agronomic traits in Peta X T(N) Backcrosses. Table 1 shows the coefficients of variation of the 5 different strata and characteristics.

5.1 Single Purpose Survey: Stratified Random Sampling.

Using equations 10, 14, and 15 (section 3.2), and taking into account the different assumptions of each method, stratum sample size n_h, $h = 1, \ldots L$ and its confidence limits were estimated for the first characteristic X_1 in Table 1. The precision set in terms of RV_{oh} were:

$$RV_{01} = RV_{02} = RV_{03} = .05$$
$$RV_{04} = RV_{05} = .0045$$

The results are presented in Tables 2 and 3.

5.2 Multipurpose Survey: Charnes and Lemke's Method.

The objective function in terms of $U_i = 1/n_i$ is:

Maximize $K = c_1 U_1 + c_2 U_2 + \ldots + c_L U_L$

while the structural constraints become

$$\sum_{i=1}^{L} RV_i(X_j)U_i \leq RV_{0j}. \qquad j = 1, \ldots p$$

Recal that $RV(X) = RV(X)/n$, and thus, the above result.

PILAR L. FERRER

TABLE I. STRATA COEFFICIENTS OF VARIATION FOR FIVE VARIABLES

CV for X stratum		2	3	4	5
X_1	.33348	.38894	.36191	.45658	.41414
X_2	.21850	.19586	.20640	.19720	.19708
X_3	.20958	.21380	.20022	.19784	.21044
X_4	.12024	.06486	.04039	.09303	.08407
X_5	.05429	.09509	.07205	.06518	.05804

Variables:

X_1 — Leaf angle mean
X_2 — Panicle number
X_3 — Tiller number
X_4 — Plant height
X_5 — Grain yield

Strata:

1 — P/2 Semi-dwarf Line
2 — P/3 Semi-dwarf Line
3 — P/4 Semi-dwarf Line
4 — P/3 Tall Line
5 — P/4 Tall Line

TABLE 2. COMPUTED VALUES OF n and n_h (FOR X_1) FOR THE THREE METHODS

	n_1	n_2	n_3	n_4	n_5	n
Method I	13.04	15.21	14.15	17.85	16.19	76.44
Method II	7.98	9.30	8.66	10.92	9.91	46.77
Method III	11.17	13.02	12,12	15.29	13.87	65.47

TABLE 3. CONFIDENCE INTERVALS FOR n_h FOR THE THREE METHODS (FOR X_1)

	CI n_1		CI n_2		CI n_3		CI n_4		CI n_5	
Method I	10.4	15.70	12.1	18.3	11.3	17.0	14.2	21.5	12.89	19.5
Method II	3.9	12.0	4.7	14.0	2.3	15.0	5.6	12.3	3.4	16.7
Method III	5.6	16.7	6.7	19.3	2.2	22.0	8.3	22.3	4.2	23.5

The upper and lower bounds restrictions become:

$$1/n_{iu} \leqslant U_i \leqslant 1/n_{iu}; \qquad i = 1, \ldots L$$

Let the desired precision be:

$$RV_{01} = .1125$$
$$RV_{02} = NV_{03} = .05$$
$$RV_{04} = RV_{05} = .0045$$

Confidence intervals for n_i, i = 1, ... L, are computed for all 5 variables and are shown below in Table 4. (Only confidence limits obtained by using Method I will be employed in this illustration.)

The upper and lower bounds restrictions are then:

$$.075793 \leqq U_1 \leqq .326156$$
$$.115569 \leqq U_2 \leqq .354195$$
$$.132074 \leqq U_3 \leqq .357730$$
$$.097954 \leqq U_4 \leqq .351791$$
$$.108685 \leqq U_5 \leqq .351989$$

Making a simple transformation of variables, $Y_i = U_i - a_i$, where a_i, i = 1, ... L are the lower limits of U_i, and adding the necessary slack variables (Y_6, ... Y_{10}) to the structural constraints and (X_1 ... X_5) to the upper bounds restrictions, the resulting equations are:

TABLE 4. CONFIDENCE INTERVALS FOR n_h FOR ALL CHARACTERISTICS

X_j	CI n_1		CI n_2		CI n_3		CI n_4		CI n_5	
X_1	4.61	6.98	5.38	8.14	5.01	7.57	6.32	9.55	5.73	8.66
X_2	3.15	5.72	2.82	5.13	2.98	5.40	2.84	5.16	2.84	5.16
X_3	3.07	5.58	3.13	5.70	2.93	5.34	2.89	5.27	3.08	5.61
X_4	8.23	13.19	4.49	7.12	2.80	4.43	6.44	10.21	5.84	9.20
X_5	3.38	4.94	5.91	8.65	4.48	6.56	4.05	5.93	3.61	5.28

$$1112.1077 \ Y_1 + 1512.7466 \ Y_2 + 1309.8195 \ Y_3$$
$$+ 2084.6130 \ Y_4 + 1715.1431 \ Y_5 + Y_6 = 302.2838$$

$$477.4024 \ Y_1 + 383.6209 \ Y_2 + 426.0298 \ Y_3$$
$$+ 388.8670 \ Y_4 + 388.4213 \ Y_5 + Y_7 = 282.9094$$

$$439.2446 \ Y_1 + 457.1132 \ Y_2 + 400.8799 \ Y_3$$
$$+ 391.4253 \ Y_4 + 442.8478 \ Y_5 + Y_8 = 274.4618$$

$$144.5808 \ Y_1 + 42.0737 \ Y_2 + 16.3162 \ Y_3$$
$$+ 86.5443 \ Y_4 + 70.6741 \ Y_5 + Y_9 = 10.8659$$

$$29.4740 \ Y_1 + 90.4254 \ Y_2 + 51.9083 \ Y_3$$
$$+ 42.4828 \ Y_4 + 33.6892 \ Y_9 + Y_{10} = 17.6371$$

$$Y_1 \qquad + \ X_1 \qquad\qquad = \ .2503632302$$
$$Y_2 \qquad\quad + \ X_2 \qquad\qquad = \ .2386267554$$
$$Y_3 \qquad\qquad + \ X_3 \qquad\quad = \ .2256560783$$
$$Y_4 \qquad\qquad\quad + \ X_4 \qquad = \ .2538367670$$
$$Y_5 \qquad\qquad\qquad + \ X_5 \quad = \ .2433034905$$

$$Y_j, \ X_i \ \geqq \ O \ ; \qquad\qquad i = 1, \ \dots \ 5$$
$$j = 1, \ \dots \ 10$$

Applying Charnes and Lemke's method to the above problem.

The solution arrived at gives the following sample size allocations:

$$n_1 \ = \quad 7.7$$
$$n_2 \ = \quad 8.6$$
$$n_3 \ = \quad 3.2$$
$$n_4 \ = \quad 10.2$$
$$n_5 \ = \quad 9.2$$

6. Conclusions.   The results obtained indicate that the values of the sample size n and stratum size n_h (in stratified sampling), and its confidence limits are in proportion to the values of the corresponding coefficient of variation.

If a comparison is made among the three methods proposed for obtaining confidence intervals for n_h, Method III is observed to give the widest interval in all the strata and for all the characteristics, while Method I, the least. It has been likewise noted that in all cases, Method II has the lowest confidence limits, whether in the lower or upper bounds.

Charnes and Lemke's Method was used to get a compromise allocation in a multipurpose survey where several different allocations are possible. Herein, because of the upper and lower bounds constraints, not only is the over-allprecision satisfied, but each stratum sample size n_h also meets the precision requirements.

BIBLIOGRAPHY

Chung, An-min. 1963. "Linear Programming." Charles E. Merrill Books, Inc., Columbus, Ohio, U.S.A.

Dalenius, T. 1957. "Sampling in Sweden: Contributions to the Methods and Theories or Sample Survey Practice." Almqvist and Wiksell, Stockholm, Sweden.

Gutierrez, Jose S. 1965. "Some Theoretical Considerations on the Uses of the Coefficients of Variation in Sample Surveys." The Philippine Statistician, XIV no. 3.

————————. 1958. "On the Comparison of the Ratio of Squared Coefficients of Variation and Rattio of Variances as Indicators of Relative Efficiency." Master's Thesis, Department of Experimenttal Statistics, North Carolina State College, Raleigh.

Hansen, Hurwitz, and Madow. 1966. "Samples Survey Methods and Theory." 2 Vols. Johan Wiley and Sons, Inc., New York.

Neyman-Splawa, J. 1925. "Contributions to the Theory of Small Samples drawn from a. Finite Population." Biometrika Vol. XVII.

Norris, N. 1936. "Some Efficient Measures of Relative Dispersion." Annals of Mathematical Statistics.: 9:214.

Sinsioco, C. S. 1969. "The Use of the Coefficient of Variation in Sample Size Determination and Allocation." Master's Thesis, Statistical Center, University of the Philippines.

Sukatmen, P. V. 1958. "Sampling Theory of Surveys with Applica-