Trans. Nat. Acad. Sci. & Tech. (Phils.)
1987.9:77-95

SAMPLE SIZE DETERMINATION IN POST HARVEST EXPERIMENTS

Mariano B. de Ramos* and Aleli B. Olea**
Statistical Laboratory, Institute of Mathematical Sciences and Physics
***Post Harvest Training and Research Center, U.P. at Los Baños, College, Laguna, Philippines*

ABSTRACT

Nine sets of data representing nine characters or variables were analyzed with the objective of determining the optimum sample size for post harvest experiments on mango. The estimated values of the variance components σ_E^2 and σ_S^2 were used to relate the precision of a treatment mean with the number of replications *r* and number of subsamples *s*. For any desired degree of precision one can refer to the graph obtained to determine the sample size n which is the product of *r* and *s*. On the other hand, if the cost of an experiment is known or can be estimated, then the optimum numbers *r*, *s* and n can be determined from the tabular values.

Introduction

One very important consideration that a research worker has to consider in the planning of his experiment is sample size. In the case of comparative experiments, sample size may refer to the number of experimental units used per treatment which is also called replications as in a completely randomized design or it may refer to the combination of the number of replications and number of subsamples per replication as in a completely randomized design with subsampling. Researchers can be guided in this particular problem by examining and using the results of the statistical analysis of past experiments in which the variance components due to identified sources can be estimated. By using the estimated values of those variance components, the researcher can determine the precision of a treatment mean from which the sample size of future experiments can be determined. Thus, in this study, the focus of the analysis were some data from past experiments of the Post Harvest and Training Research Center (PHTRC) at the University of the Philippines at Los Baños with the aim of determining the appropriate sample size for future experiments.

In many of the experiments that have been conducted at PHTRC, the statistical design used is usually the completely randomized design with subsampling. In such design, the experimental error variation of the data comes from two sources, namely, from the differences between the experimental units treated alike and from the differences between the sampling units within experimental units. Thus, if the experimental error variance of a character x is denoted by σ_x^2, then

$$\sigma_x^2 = \sigma_E^2 + \sigma_S^2$$

where σ_E^2 is the variance component due to the experimental units or replications and σ_S^2 is the variance component due to the sampling units. Without any knowledge of the magnitudes of these two variance components, the researcher would not know the appropriate sample size to use in order to obtain reliable and precise experimental results, hence, he may just utilize whatever materials are available. If the sample size used happened to be too small, then the experimental results may not be able to detect real treatment differences, while if the sample size used happened to be too large, the results of the experiment might be more precise than what would be required statistically.

In the light of the problems stated above, this study was conducted with the main objective of determining the appropriate and optimum sample size for mango post harvest experiments. The specific objectives of the study were: (i) to obtain estimates of the variance components due to experimental units and due to sampling units for various mango post harvest characters, (ii) to obtain the appropriate sample size for mango post harvest experiment that will yield results with certain degree of precision, and (iii) to obtain the optimum sample size for mango post harvest experiments that will give optimal results for a given fixed cost per treatment.

Review of Literature

Anderson (1947) used the analysis of variance to test the significance of variance components that affects the prices of hog meat in two markets. Marcuse (1949) obtained an estimate of the reciprocals of n_1, n_1n_2, and $n_1n_2n_3$. Anderson and Bancroft (1952) utilized a general estimation procedure for the variance components, such as the method of maximum likelihood.

Kempthorne (1952) derived the optimum number of secondary sampling units and optimum number of primary sampling unit for sampling in field experiments. Goldsmith and Gaylor (1970) used the three stage nested design for the estimation of variance components, however unbalanced the arrangements may be. Sahai (1976) studied various estimators of the variance components for the balanced three stage nested design.

In sampling for laboratory brix, Solivas (1978) found that in raw and adjusted sugar rendement, the variance components among the rows and the experimental error variance component were significantly greater than zero. In the study of the avocado fruit characters, Ledesma (1983) found various variance components that gave higher contribution to the total variation. In sampling for coconut characters, Alforja (1983) found that the sample size n considered optimum varies depending with the uniformity of the cultivars. He also found that 24 palms is sufficiently enough to obtain reliable information for nuts per tree estimation.

Materials and Methods

The data

The data used in this study were obtained from the past post harvest experiments on mango that were conducted at PHTRC. Table 1 shows the statistical description of the experiments. There were four experiments and the number of treatments range from four to eight, the number of replications from three to four, and the number of subsamples per replication from three to nine. Nine post harvest characters or variables were measured, two from experiment I, two from experiment II, four from experiment III and one from experiment IV. These characters were color index at day 0 and 5, total soluble solids, titratable acidity, percent cumulative weight loss, firmness, disease incidence, PH and visual quality rating.

The statistical design and model

All the four experiments were conducted in a completely randomized design with subsampling. A typical example of such design is where the treatments are heating temperatures, the experimental units or replications are boxes of fruits, and the sampling units are the individual fruits in the boxes. An experiment may then involve, say, t treatments, r boxes of fruits per treatment, and s fruits per box. Thus, if a character x is measured on sampling unit, the statistical model is of the form

$$x_{ijk} = \mu + \tau_i + e_{ij} + d_{ijk} \qquad (2)$$

$$i = 1, 2, \ldots, t$$

$$j = 1, 2, \ldots, r$$

$$k = 1, 2, \ldots, s$$

Table 1. Description of the four post harvest experiments on mango from which the data of the study were obtained

Experiment No.	*No. of Treatments (t)*	*No. of Replications (r)*	*No. of Sub-Samples (s)*	*Characters Measured*
I	4	4	8	Color index at day 0 and 5
II	6	3	9	Total soluble solids and titratable acidity
III	8	4	3	Percent cumulative weight loss, firmness, disease incidence and pH
IV	6	4	4	Visual quality rating

Source: Post Harvest Training and Research Center, U.P. at Los Banos.

Table 2. Format of the analysis of variance for the nine post harvest characters of mango

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*Expected Mean Square*
Treatment	$t-1$	$rs \sum_i (\bar{x}i\,..-x\,..)^2$	MS(Tr)	$\sigma_S^2 + s\,\sigma_E^2 + \sum_i \tau_i^2 / (t-1)$
Experimental error	$t(r-1)$	$s \sum_i \sum_j (\bar{x}_{ij\,..} - \bar{x}_i\,..)^2$	MSE	$\sigma_S^2 + s\,\sigma_E^2$
Sampling error	$tr(s-1)$	$\sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij}\,.)^2$	MS(SE)	σ_S^2
Total	$trs-1$	$\sum_i \sum_j \sum_k (x_{ijk} - \bar{x}\,...)^2$		

where x_{ijk} is the observed value in the kth sampling unit of the jth replication and ith treatment, μ is the general mean effect common to all observations, τ_i is the effect of the ith treatment, e_{ij} is the random error effects due to the ith experimental unit of the ith treatment, and d_{ijk} is the sampling error effect due to the kth sampling unit of the jth replicate. For each of the nine characters used in the model, it was assumed that the treatment effects τ_i are fixed and $\Sigma_i \tau_i = 0$. Also, the experimental error e_{ij} were assumed to be normally and independently distributed with mean 0 and variance σ^2_E or $e_{ij} \sim$ NID $(0, \sigma^2_E)$, and the sampling error d_{ijk} were assumed to be normally and independently distributed with mean 0 and variance σ^2_S or $d_{ijk} \sim$ NID $(0, \sigma^2_S)$.

Estimation of variance components

The variance components estimated for each character were those due to the differences between experimental units or replications which was denoted by σ^2_E and due to the differences between the sampling units within the experimental units denoted by σ^2_S.

The method used in estimated these variance components was by analysis of variance. Essentially, the steps involved in the estimation of σ^2_S and σ^2_E were:

(1) construction of the analysis of variance table (Table 2), and

(2) equating the actual mean square and the expected mean square for the sampling error and experimental error. Hence, if

(i) $\text{MS(SE)} = \sigma^2_S$

(ii) $\text{MSE} = \sigma^2_S + s\,\sigma^2_E$

then

$$\hat{\sigma}^2_S = \text{MS (SE)} \tag{3}$$

$$\hat{\sigma}^2_E = (\text{MSE} - \text{MS (SE)}\,)\,/\,s \tag{4}$$

where MSE is the mean square due to experimental error defined as

$$\text{MSE} = s \sum_{i}^{t} \sum_{j}^{r} (\bar{x}_{ij.} - \bar{x}_{i..})^2 \;/\; t(r-1), \tag{5}$$

and MS (SE) is the mean square due to the sampling units defined as

$$\text{MS (SE)} = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij.})^2 \;/\; tr(s-1) \tag{6}$$

In these formulas, the quantities

$\bar{x}_{...} = \sum_i^t \sum_j^r \sum_k^s x_{ijk}/trs$, the grand mean

$\bar{x}_{i..} = \sum_j \sum_k x_{ijk}/rs$, the treatment mean, and

$\bar{x}_{ij.} = \sum_k x_{ijk}/s$, the replication mean.

Determining the sample size for a given degree of precision

The basis of determining the sample size which in this study is the combination of the number of replications r and the number of subsample s or $n = rs$ was by the use of the precision of a treatment mean. In a completely randomized design with subsampling, the variance of treatment mean $\bar{x}_{i..}$ is given by the formula

$$\text{var}(\bar{x}_{i..}) = \text{MSE}/rs \tag{7}$$

In terms of the estimated variance components $\hat{\sigma}^2_S$ and $\hat{\sigma}^2_E$, the variance of a treatment mean is

$$\text{var}(\bar{x}_{i..}) = \frac{\hat{\sigma}^2_E}{r} + \frac{\hat{\sigma}^2_S}{rs} \tag{8}$$

Therefore, the standard error of a treatment mean is

$$\text{s.e.}(\bar{x}_{i..}) = \sqrt{\frac{\hat{\sigma}^2_E}{r} + \frac{\hat{\sigma}^2_S}{rs}} \tag{9}$$

In terms of the coefficient of variation, the precision is

$$\text{CV}(\bar{x}_{i..}) = \text{s.e.}(\bar{x}_{i..})/\bar{x}_{...}$$

Thus, the final form of the precision formula used in this study is

$$\mathrm{CV}\,(\bar{x}_{i\,.\,.}) = \sqrt{\frac{\hat{\sigma}_E^2}{r} + \frac{\hat{\sigma}_S^2}{rs}} \Big/ \; \bar{x}_{.\,.\,.} \tag{10}$$

The values of the CV $(\bar{x}_{i\,.\,.})$ were then computed and their graphs were drawn against r and s for values of $r = 2, 3, 4, 5$ and $s = 1, 2, ..., 10$.

Determining the optimum sample size at a given experiment cost per treatment

Kempthorne (1952) defined the information on each treatment as

$$I = \frac{rs}{\hat{C}_S + s\,\hat{\sigma}_E^2} \tag{11}$$

By assuming a cost function of the form

$$C_0 = r\,(C_E + s\,C_S) \tag{12}$$

where C_0 is the cost of the experiment per treatment, C_E is the cost per experimental unit, and C_s is the cost per sampling unit. Solving for r in (12) and substituting the result in

(11) gave the formula for information as

$$I = \frac{(s)\,(C_0)}{(C_E + sC_S)\,(\hat{\sigma}_S^2 + s\,\hat{\sigma}_E^2)} \tag{13}$$

Minimizing (12) with respect to s gave

$$S\sqrt{\left(\frac{C_E}{C_S}\right)\left(\frac{\hat{\sigma}_S^2}{\hat{\sigma}_E^2}\right)} \tag{14}$$

The optimum value of r was then found to be

$$r = \frac{C_0}{\sqrt{C_E + C_E\,C_S\;\hat{\sigma}_S^2\,/\,\hat{\sigma}_E^2}} \tag{15}$$

Since estimates of C_E and C_S were not available, various ratios of C_E to C_S were assumed and then the values of s and r were computed by formulas (14) and (15) using the known values of $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$ for each character.

Results and Discussion

The analysis of variance

The analysis of variance for the nine post harvest characters of mango showing only the degrees of freedom (DF) and mean square (MS) for treatment, experimental error and sampling error are given in Table 3. In these results the estimates for the mean square error, MSE and the sampling error mean square, MS (SE) may be considered as stable since these were based on sufficient degrees of freedom.

The mean squares for the treatment, MSTr are marked to indicate that they are either, significant *** or not significant (NS) as compared with the mean square error. Out of the nine characters, seven are identified for which the treatment effects were significant. Only two characters, color index at day 0 and 5 did not show the significant effects of treatments.

With respect to the magnitude of the mean error and mean square sampling error, it was noted that in all but one character the values of the former are larger than the latter. This would indicate that the error variance component in such characters are all positive. The only character which showed a negative estimate for the error variance component was pH. However, the test of significance for the experimental error variance component σ_E^2 resulted into only two significant mean square error and those were for color index at day 5 and total soluble solids. In case of the non-significant error mean squares, pooling of the mean square error and mean square sampling error may be in order.

Estimates of variance components

The two variance components, σ_S^2 and σ_E^2 were estimated by formulas (3) and (4) using the values of MSE and MS (SE) given in Table 3. These estimates of variance components, $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$ are given in Table 4, and they are expressed in absolute form or as percentage of their total. For instance, the values of $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$ for color index at 0 are .20 and .0055, respectively, or they are 97% and 3%, respectively, of the total variance component 0.2055. The results of these estimated variance components for the nine characters indicate that the variance component due to sampling units, $\hat{\sigma}_S^2$ were very much larger than the variance component due to experimental unit, $\hat{\sigma}_E^2$ by as much as 5 to 36 times. This would only show that most of the variability in post harvest characters of mango comes from the differences between the sampling units and very little comes from the differences

Table 3. Results of the analysis of variance for the nine postharvest characters of mango

Source Of	CI(0)		CI (5)		TSS		TA		WL		F		DI		PH		VQR	
Variation	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS
Treatment	3	.19^{ns}	3	2.59^{ns}	5	4.5**	5	10.03**	7	.7**	7	.08*	7	2.39**	7	.43**	5	6.23
Experimental error	12	.25^{ns}	12	1.15*	12	.56**	12	.18^{ns}	24	.24^{ns}	24	.03^{ns}	24	.52^{ns}	24	.04^{ns}	18	1.49^{ns}
Sampling error	128	.20	128	.57	162	.18	162	.11	96	.07	96	.02	96	.44	96	.06	96	1.29

CI (0) – Color index at day 0
CI (5) – Color index at day 5
TSS – Total soluble solids
TA – titratable acidity
WL – Weight loss
F – firmness
DI – Disease incidence
VQR – Visual quality rating
DF – degrees of freedom
MS – Mean square error

Table 4. Estimate of experiment mean ($\bar{x}$), coefficient of variation (CV), and variance components for the nine postharvest characters of mango

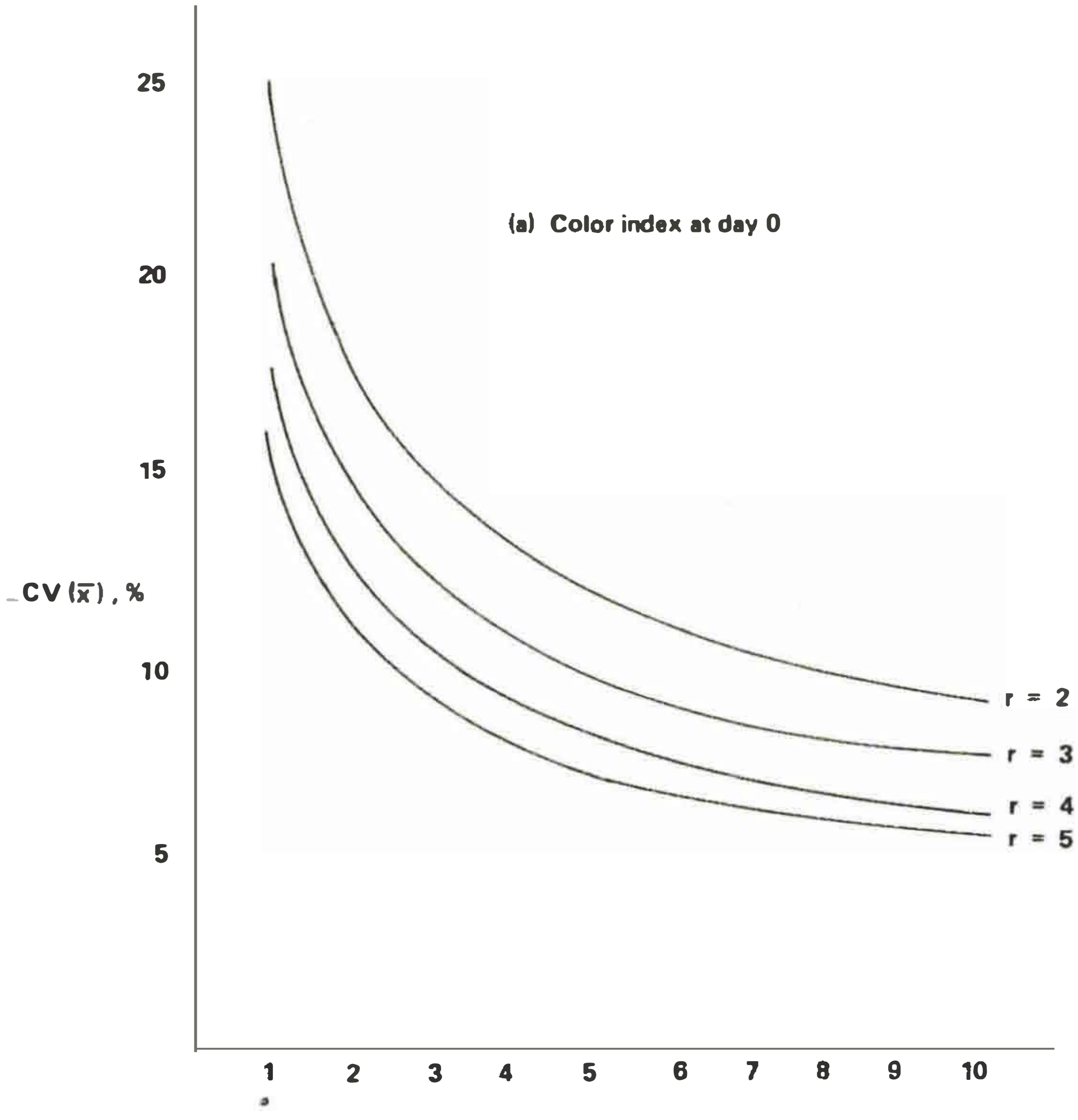
Character	*MEAN*	*CV (x)*	*Variance components*					
	($\bar{x}$)	*(%)*	$\hat{\sigma}_S^2$	%	$\hat{\sigma}_E^2$	%	$\hat{\sigma}_x^2 = \hat{\sigma}_S^2 + \hat{\sigma}_E^2$	%
Color index at day 0	1.29	38.7	0.20	97	0.0055	3	0.2055	100
Color index at day 5	4.44	24.2	0.57	90	0.064	10	0.634	100
Total soluble solids	6.74	11.1	0.18	83	0.038	17	0.218	100
Titratable acidity	3.33	12.7	0.110	94	0.007	6	0.117	100
% Weight cumulative loss	.99	35.0	0.07	85	0.0125	15	0.825	100
Firmness	.99	17.5	0.02	89	0.0025	11	0.0225	100
Disease incidence	.42	171.1	0.44	96	0.02	4	0.46	100
PH	4.54	4.44	0.66	100	0	0	0.06	100
Visual quality rating	6.25	19.50	1.29	97	0.04	3	1.33	100

between the experimental units. In such cases, the implication is that in those experiments that were conducted, very little control have been given to keep the sampling units more uniform, such as using more uniform fruits with respect to weights or size, etc. The use of other experimental design, such as randomized complete block design with sub-sampling may even bring about a more efficient results.

The precision of a treatment mean as a function of sample size

The precision of a treatment mean may be expressed as a function of the number of replications r and number of sampling units s after having obtained the estimates of $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$. By using equation (11) and the values of $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$ given in Table 4, the values of coefficient of variation of a treatment mean, CV $(xi\ldots)$ were computed for values of r ranging from 2 to 5 and s ranging from 1 to 10. The graphs of the CV $(\bar{x}i\ldots)$ values versus r and s were drawn and they are shown in Figs. 1 (a) to 1 (h). The graphs for each character is the function

$$\mathrm{CV}\ (\bar{x}\,i\ldots) = \sqrt{\frac{\hat{\sigma}_E^2}{r} + \frac{\hat{\sigma}_S^2}{rs}}\ /\bar{x}\ldots)$$



(a) Color index at day 0

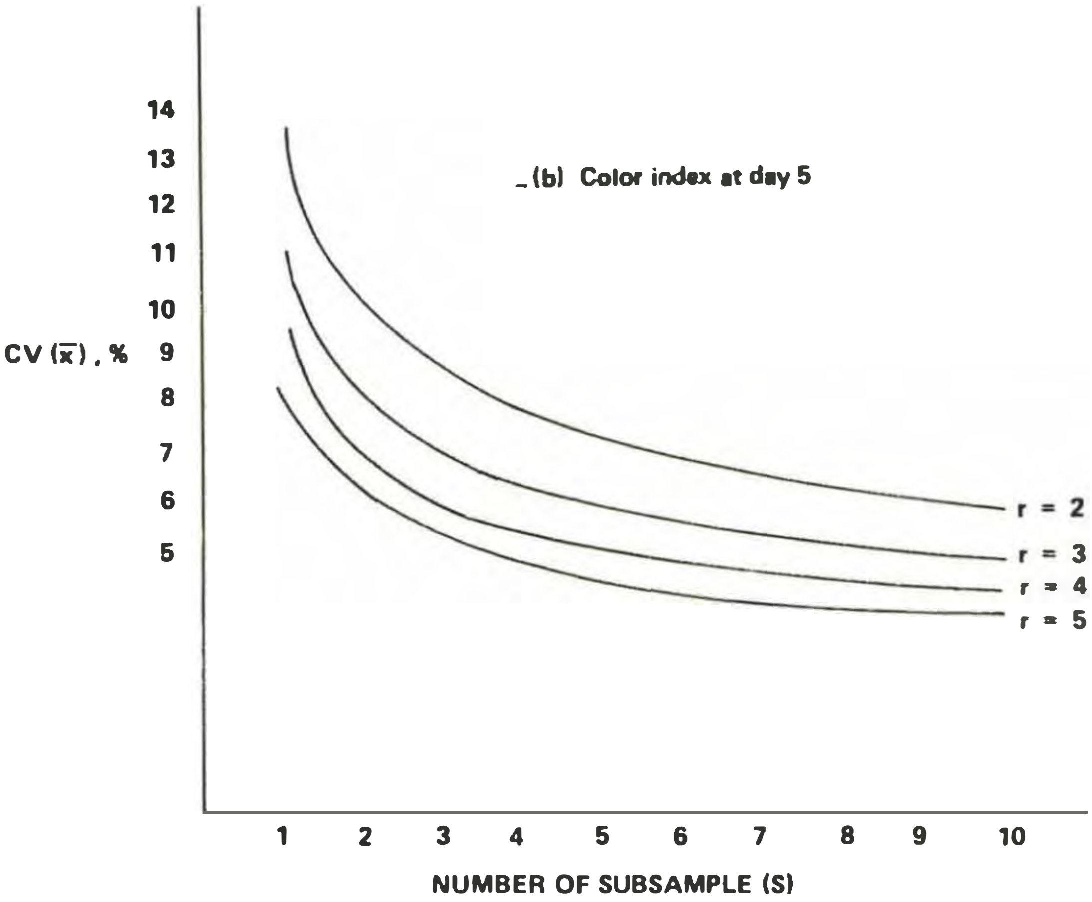
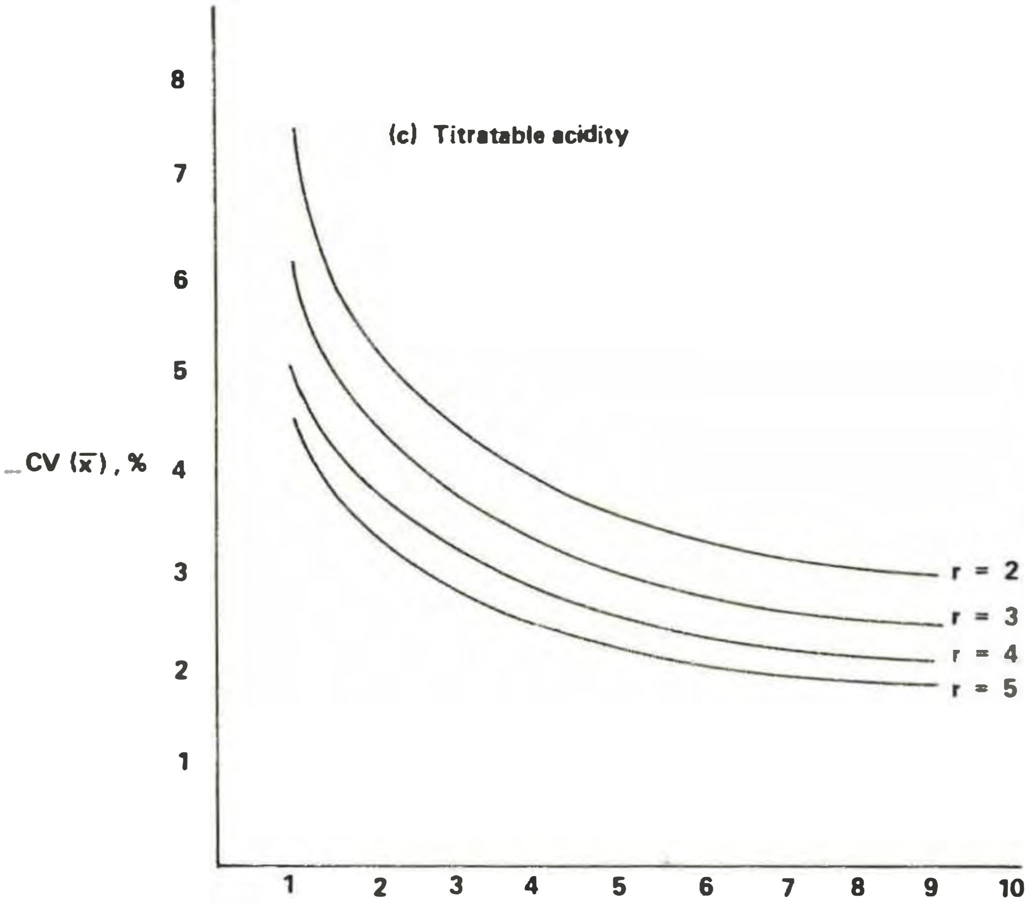

(b) Color index at day 5
CV (x̄), %
14
13
12
11
10
9
8
7
6
5
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10
NUMBER OF SUBSAMPLE (S)

Figure 1


(c) Titratable acidity
CV (x̄), %
8
7
6
5
4
3
2
1
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10

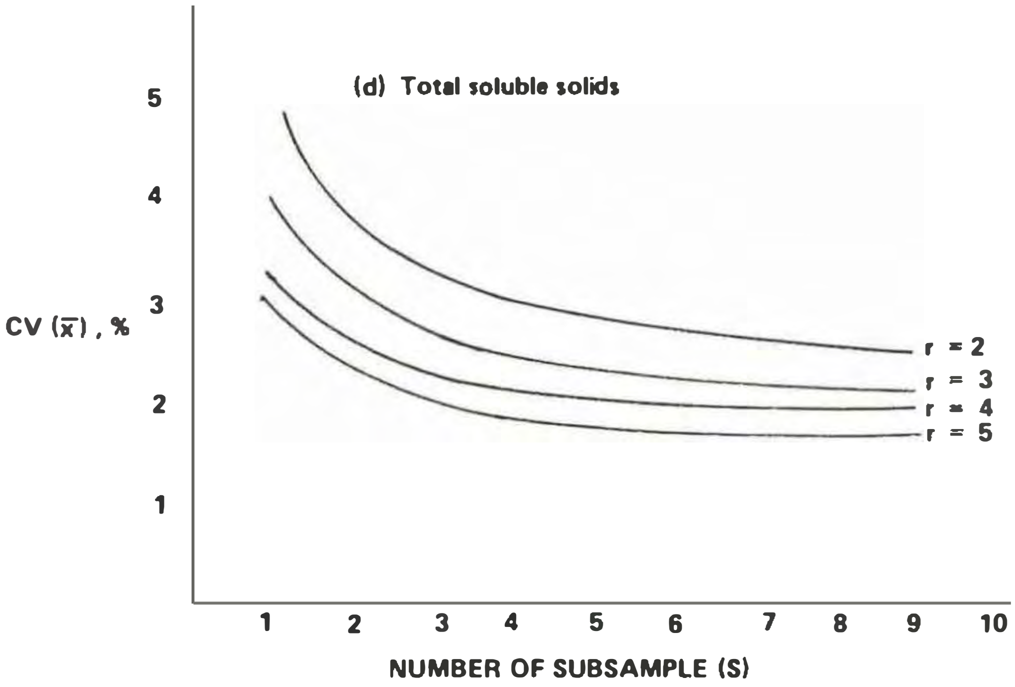
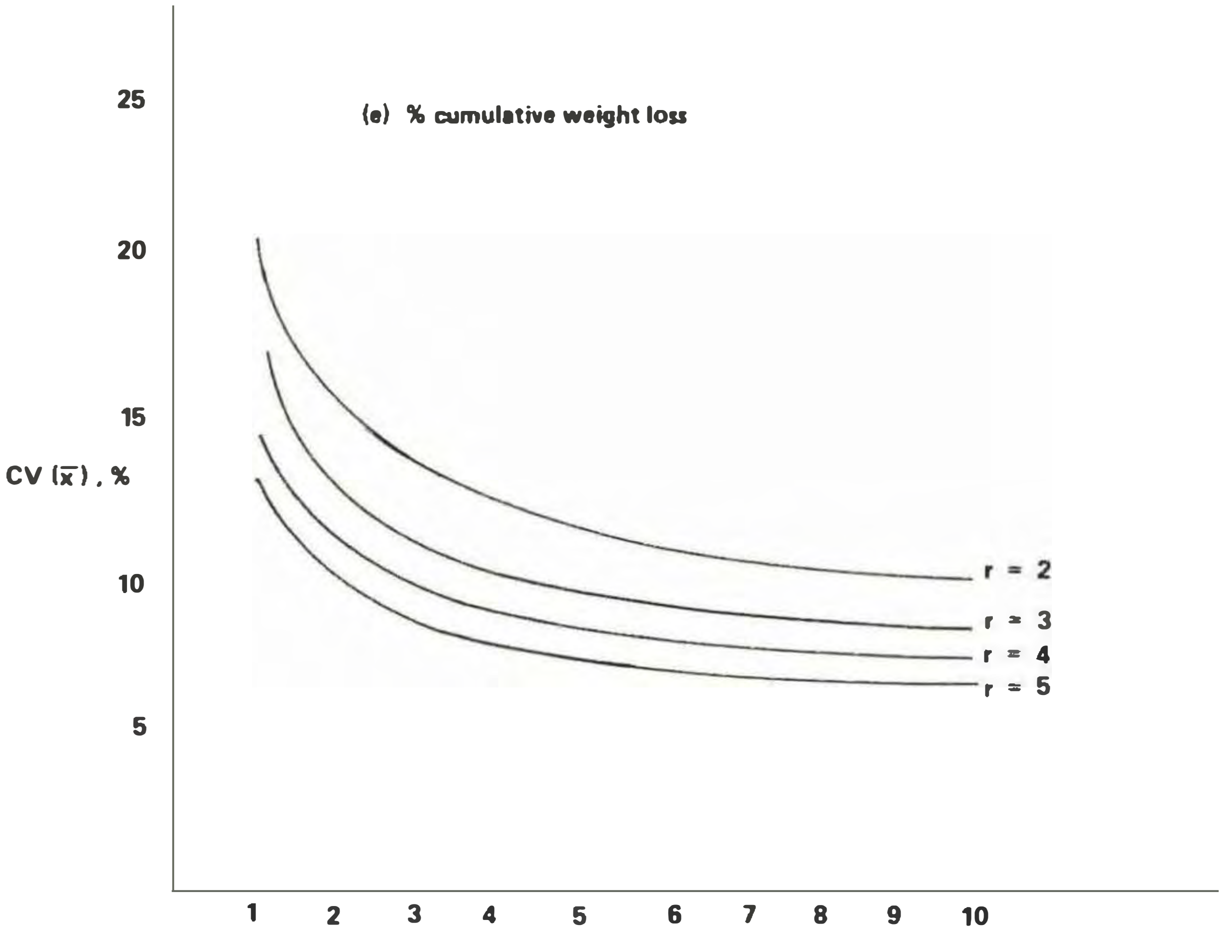

(d) Total soluble solids
CV ($\bar{x}$), %
5
4
3
2
1
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10
NUMBER OF SUBSAMPLE (S)

Figure 1


(e) % cumulative weight loss
CV ($\bar{x}$), %
25
20
15
10
5
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10

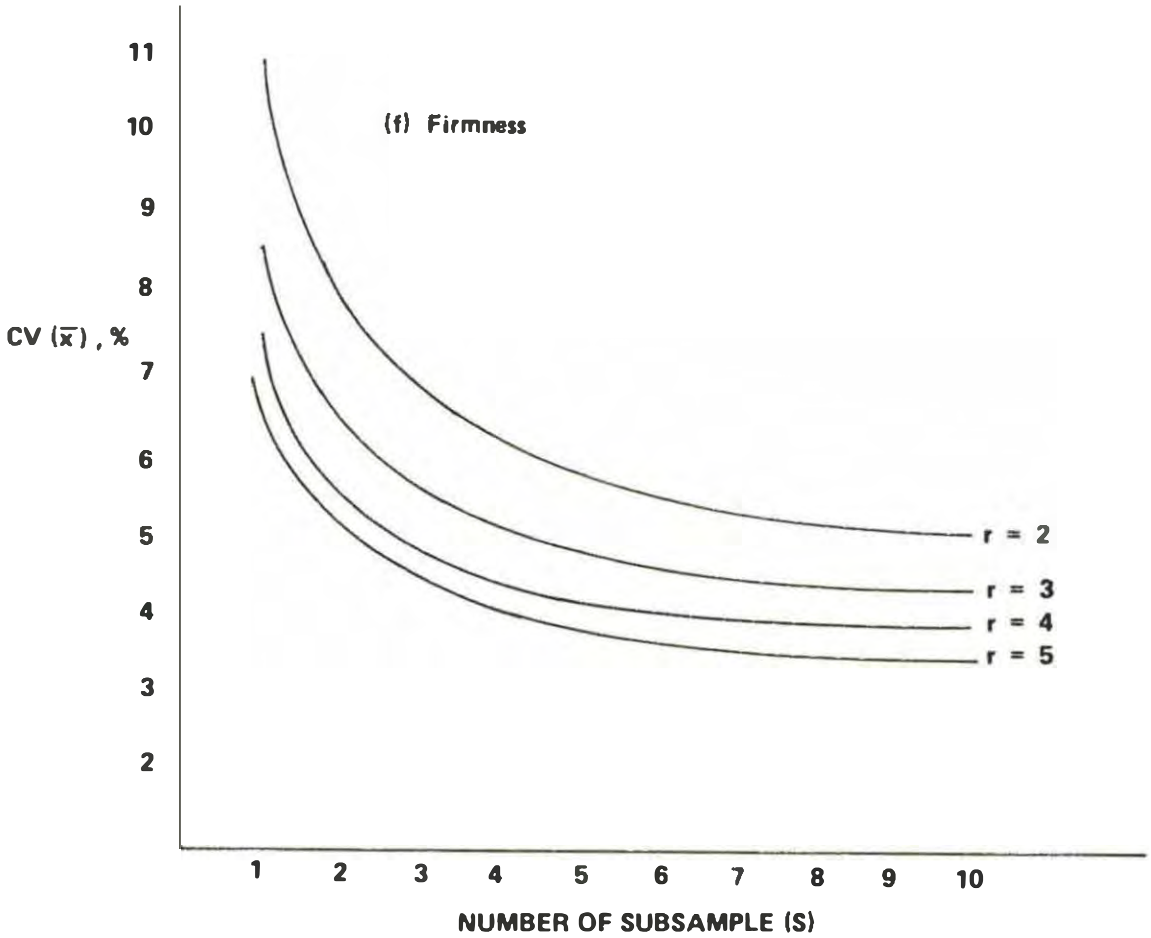
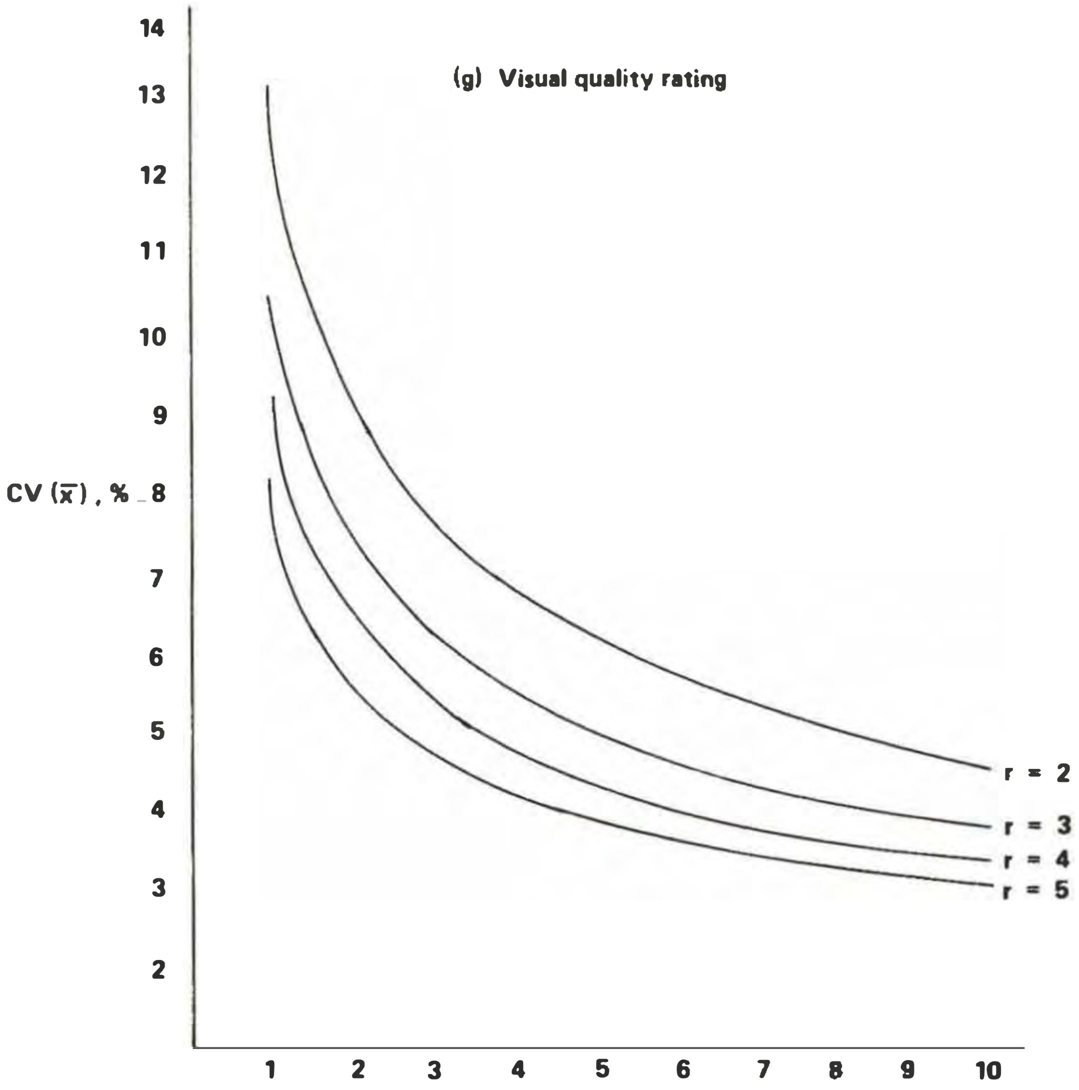

11
10
9
8
CV ($\bar{x}$), %
7
6
5
4
3
2
(f) Firmness
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10
NUMBER OF SUBSAMPLE (S)

Figure 1


14
13
12
11
10
9
CV ($\bar{x}$), % 8
7
6
5
4
3
2
(g) Visual quality rating
r = 2
r = 3
r = 4
r = 5
1
2
3
4
5
6
7
8
9
10

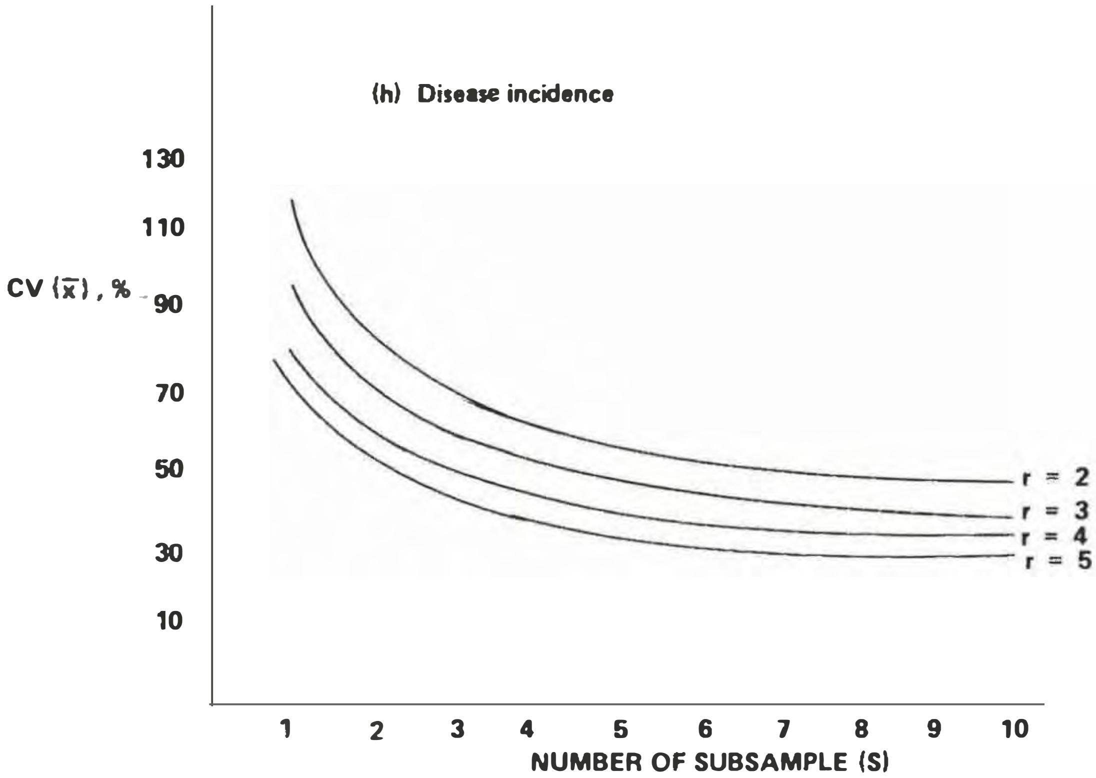


Figure 1

For total soluble solids, for instance, the graph shown in Fig.

1(d) is the function

$$CV\ (\bar{x}i\ldots.) = \sqrt{\frac{.038}{r} + \frac{0.18}{rs}}\quad / (6.74)$$

where $r = 2, 3, 4, 5$ and $s = 1, 2, \ldots\ldots, 10$.

The graph of CV ($\bar{x}i$ . . . . .) against the sample size s for a given number of replications r is like a negative exponential with maximum value at ∞ when $s = 0$, and a minimum value at $\sqrt{\frac{\hat{\delta}^2 E}{r}}$ at $s = \infty$. By proper choice of r and s, one can bring down the value of CV ($\bar{x}i$ . . .) to any prescribed percentage, such as 10% or 5%. By using the graph for any character, the desired precision level can be set and then simply locate proper combination of r and s. For example, using Fig. 1 (a), one can make the precision of the treatment mean equal to 10% for choices of (r) (s) as (2) (7), 3(5), (4) (4) and (5) (3). This choices on the average led to a sample size $n = 15$.

The optimum sample size

The formulas for determining the optimum number of subsamples s and optimum number of replication r are given as

$$s = \sqrt{\left(\frac{C_E}{C_S}\right)\left(\frac{{}^2S}{{}^2E}\right)}$$

$$r = \frac{C_o}{C_E + \sqrt{C_E C_S \hat{\sigma}_S^2 / \hat{\sigma}_E^2}}$$

If one knows the cost of the experiment per treatment (C_o), the cost per experimental unit (C_E), and the cost per sampling unit (C_S), then the optimum value for s and r can be computed since $\hat{\sigma}_S^2$ and $\hat{\sigma}_E^2$ are already known. To see the behavior of the estimates of s and r, certain ratio of the cost estimate as C_E C_S were given and then the values of s and r computed for each character. For example, if the ratio is 4:1, say, then

$$C_o = r' \ (4 + s' \ (1)$$

where r ' and s ' were the numbers of replications and samples in the actual experiment. Therefore, the value of

$$s = \sqrt{\left(\frac{4}{1}\right)\left(\frac{\hat{\sigma}_S^2}{\hat{\sigma}_E^2}\right)}$$

$$\text{and r} = \frac{\text{r'}\,(4 + \text{s'})}{4 + \sqrt{(4)\,(\hat{\sigma}_S^2 / \hat{\sigma}_S^2 / \hat{\sigma}_E^2}}$$

The computed values of s and r for the cost ratios 1:4, 1:2 1:1, 2:1 and 4:1 are given in Table 5. Thus for a cost ratio of 4:1, say, the optimum numbers r, s and n were computed as 3, 11, 33 for color index at day 0; 5, 6, 30 for color index at day 5; 5, 4, 20 for total soluble solids; 4, 7, 28 for titratable acidity; 3, 5, 15 for percent cumulative weight loss, 3, 8, 24 for firmness, 2, 10, 20 for disease incidence; and 2, 11, 22 for visual quality rating. As Kampthorne had pointed out, these numbers maximizes the information on each treatment mean for a given cost per treatment C_o.

Table 5. Optimum numbers of replications (r), subsamples (s) and sample size ($n = rs$) at a given cost ratio ($C_E{:}C_s$) for the eight post harvest characters of mango

Character	*SIZE*	*COST RATIO*, $C_E{:}C_s$				
		1:4	*1:2*	*1:1*	*2:1*	*4:1*
Color index of day 0	r	12	8	6	4	3
	s	3	4	6	8	11
	n	36	32	36	32	38
Color index at day 5	r	21	15	10	7	5
	s	2	2	3	4	6
	n	42	30	30	28	30
Total soluble solids	r	23	16	10	7	5
	s	1	2	2	3	4
	n	23	32	20	21	20
Titratable acidity	r	14	10	7	5	4
	s	2	3	4	7	7
	n	28	30	28	35	28
% Cumulative weight loss	r	9	6	5	4	3
	s	1	2	2	3	5
	n	9	12	14	12	15
Firmness	r	8	6	4	3	3
	s	1	2	3	4	8
	n	8	12	12	12	24
Disease incidence	r	5	3	3	2	2
	s	2	3	5	7	10
	n	10	9	15	14	20
Visual quality rating	r	6	4	3	2	2
	s	3	5	6	8	11
	n	18	20	18	16	22

Summary and Conclusions

Nine sets of data obtained from the experiments conducted at the Post Harvest Training and Research Center of U.P. Los Baños were analyzed using a completely randomized design with subsampling model for main purpose of obtaining estimates for two variance components that will be used for determining optimum sample size for post harvest experiments on mango. One set of data represent one post harvest character or variable and those characters were color index at day 0, color index at day 5, total soluble solids, titratable acidity, percent cumulative weight loss, firmness, disease incidence, pH and visual quality rating.

The analysis of variance of the characters showed that the mean square error (MSE) were larger than the mean square sampling error (MS (SE)) except for the characters pH. Those results meant that the estimates for the variance component due to the experimental unit or replication, $\hat{\sigma}_E^2$ were all positive. Further tests of significance, however, revealed that only two characters indicated significant estimates of $\hat{\sigma}_E^2$.

From the results of the analysis of variance, the value of MSE and MS (SE) were used to estimate the two variance components, $\hat{\sigma}_E^2$ and $\hat{\sigma}_S^2$ the variance component due to sampling units. The comparison of the two estimated variance components indicated that $\hat{\sigma}_S^2$ represents from about 85 to 97 percent of the experimental error variance among the nine post harvest characters. In terms of ratio, values of $\hat{\sigma}_S^2$ were larger than the values of $\hat{\sigma}_E^2$ by as much as 5 to 36 times.

By expressing the precision of a treatment mean, s.e. $(\bar{x}_i)$ in terms of the coefficient of variation of a treatment mean, *cv* $(\bar{x}_i)$, a function relating the *cv* $(\bar{x}_i)$ with the number of replications *r* and number of subsamples s was obtained for each character using the estimated values of $\hat{\sigma}_E^2$ and $\hat{\sigma}_S^2$. The graph of each function was then drawn for each character by varying the values of *r* from 2 to 4 and s from 1 to 10. The graphs showed that for a particular value of *r*, the values of the $cv(\bar{x}_i)$ decreases exponentially with increasing s and the points of inflection where somewhere between 4 and 6. Thus, if one wishes to obtain the right combination of *r* and *s* that will give the desired precision, he would simply refer to the graph of a particular character.

With respect to the determination of optimum sample size, the formulas derived by Kempthorne were used. Various ratios of the cost per experimental unit, C_E to the cost per sampling unit, C_S were used in the formula to get the optimum number of subsamples *s* and optimum number of replications *r* for each character. Thus, for any given cost ratio that is within the cost ratios used in this study, one simply refer to the tabular values of *r* and *s* to get the optimum sample size $n = rs$.

Literature Cited

Alforja, L. M. 1983. Sample size determination for eight coconut cultivars, MS Thesis, U.P. Los Baños, College Laguna.

Anderson, R. L. 1974. The use of variance components in the analysis of hog prices in two markets, *Jour. of Am. Stat. Assoc.*, 42: 612-634.

Anderson, R. L. and T. A. Bancroft. *Statistical Theory in Research.* Mc Graw-Hill Book Co., Inc.

Goldsmith, C. H. and D. Gaylor 1970. Three Stage Nested Design for Estimating Variance Components. *Technometrics* 12: 487-498.

Kempthorne, O. 1952. *The Design and Analysis of Experiments* John Wiley and Sons, Inc.

Ledesma, D. R. 1983. Optimum sample size for different fruit characters in avocado, M.S. Thesis, U.P. Los Baños, College, Laguna.

Marcuse, S. 1949. Optimum allocation and variance components in nested sampling units with an application to chemical analyses, *Biometrics* 5: 189-206.

Sahai, H. 1976. A Comparison of estimators of variance components using mean squared error criterion *Jour. of Amer. Stat. Assoc.*, 71: 435-444.

Solivas, E. S. 1978. Field plot and sampling techniques for sugar cane experiments, M.S. Thesis, U.P. Los Baños, College, Laguna.