CHAPTER 25

SAMPLING OF EXPERIMENTAL PLOTS

25.1 The usefulness of sampling

25.1.1 Sampling of experimental plots for estimation of yield is not nearly as common as it once was. At one time the problem of threshing the reaped produce from co-operative small-grain cereal experiments led to the taking of small samples which could be removed to a research station for threshing. This, however, tended to bring in fresh difficulties (cf. § 23.2.5). The advent of the combine harvester brought about a return to the complete harvesting of plots even though the use of such a large piece of equipment entailed revision of plot size and shape (cf. § 12.11.7) and was not without technical difficulties.

25.1.2 The complete harvesting of each net plot must be regarded as normal for most crops and is, of course, the ideal. However, with experiments in which large grass plots are cut for hay, for example, the bulk of the harvested material may be too great for convenient handling and weighing. Sampling may also be resorted to when different treatments ripen at different times so that it is inconvenient to use machinery, or in co-operative trials when the experimental plots are ready for harvesting before the farmer's surrounding crop. In grazing experiments wire-netting "cages" placed over small areas of each plot for a short period as a temporary protection from grazing permit estimates to be made of the growth of grass during the period and of the consumption by the animals.† This is, of course, a sampling procedure, and sampling is inevitable in this case.

25.1.3 There is virtually no alternative to sampling if measurements have to be made on a field experiment prior to harvesting. Examples are tiller counts on wheat, measurement of height of plants, etc. The labour of complete counting and measurement would be prohibitive and unwarranted. Also, it is necessary to minimize damage caused to the plants and the effect of treading. Fear of the latter may suggest sampling only the easily accessible portions of the plots (e.g. the ends), but such samples would not be representative. Where repeated measurements of the same type are to be taken on the plots during the growing period, it is advisable to use the same S.U.'s on each occasion, these being marked for re-identification. This obviously cannot be done

† See, for example, Lynch, **P. B.**, *N.Z. Journal of Science and Technology* 28, 385–405, 1947.

where the samples have to be removed; for example, in a greenfeed experiment it may be desired to measure protein content at various stages of growth.

25.1.4 Sampling the produce of the plot after harvesting is required for such purposes as chemical determinations, e.g. dry matter content, sugar content of beets, etc., since clearly it would be an impossibility to make these determinations on the entire produce of an experimental plot of normal size.

25.2 Technique of sampling

25.2.1 For the purposes of sampling, an experimental plot is regarded as a finite population of non-overlapping sampling units (S.U.'s), each of which can be identified. The principles of sampling a finite population discussed in Chapter 3 therefore apply. For avoidance of bias random selection of S.U.'s (e.g. by random numbers) should be followed, and for reduction of sampling error, stratification should not be overlooked. For reasonable precision of yield determinations the sampling fraction (S.F.) should not be less than 5%, but on the other hand a fraction in excess of about $\frac{1}{8}$ may mean that complete harvesting (if feasible) may be less laborious (besides being free from sampling error), since sampling usually entails more troublesome details than complete hervesting. The sampling and the population of S.U.'s are, of course, confined to the net plot, and here sampling may have an advantage, since separation of the border areas from the net plot is not always easy.

25.2.2 Many types of S.U. are encountered, and these may be natural or artificial (cf. § 3.7.4). Among the latter, small rectangular areas are common, these being known as quadrats. In this case it is unnecessary to delineate all the S.U.'s on a plot, only those actually randomly selected. In addition, extreme accuracy in the location of the selected S.U.'s is hardly necessary, and rough measurement by pacing may be considered reasonable. However, if pacing is used, there is a possibility of bias if the observer were, for example, to lengthen his pace in order to avoid an area of patchy growth in favour of a higher-yielding area. Bias is also possible in the delineation of the S.U.'s, since decisions have to be constantly made whether a particular plant is to be included in the sample or not, when, for example, its base is inside the area but the ear lies outside. To some extent this may be avoided, when the size of S.U. has become standardized with a particular crop, by use of a rectangle made of metal pipe. This ensures accurate delineation of the area but does not overcome all possibilities for bias. Unfortunately it is found that observers tend to be soft-hearted and include plants in cases of doubt which should really be excluded. This results in an upward bias, which is proportionately less with large S.U.'s.

25.2.3 In the botanical analysis of pasture the task of counting species in even a small quadrat is formidable, and so with the Levy point-analyser (cf. § 4.1.6) the quadrat is reduced to a point. It should be noted, however, that the points are not the S.U.'s since they are not independently located; if the positions of the rack are randomly determined, then the S.U.'s will be the clusters of 10 points.

25.2.4 When crops like small-grain cereals are planted in rows, small lengths of rows may be taken as S.U.'s. Equal lengths of adjacent rows are often taken for convenience, this being equivalent to a quadrat. This may be further extended to patterns such as that in the diagram,

the idea being to help ensure representativeness. The S.U. is the pattern of 4 sections of row, not the individual sections. However, a S.U. of this sort does not strictly comply with the requirement of non-overlapping S.U.'s covering the entire net plot, and the same applies to a S.U. such as a circle in the case of broadcast crops, though this may not be of great importance in practice.

25.2.5 With suitable crops such as maize or cabbages the individual plant is a natural S.U. A common procedure here which saves the trouble of preparing large numbers of random digits and which will usually ensure greater representativeness is to take (say) every 10th plant in a plot. This is actually a systematic sample, but, provided the initial plant is randomly selected (an essential precaution), it can be considered as the selection of one large patterned S.U. out of a population of 10 (corresponding to the 10 possible initial plants). While this method is convenient, it has an element of danger. Obviously, the number of plants per row should not be a multiple of k if every k^{th} plant is selected, since this would mean that the selected plants would be in the same position in every row and decrease the representativeness. Moreover, there should be no periodicity of the material which coincides with the interval of sampling. This is unlikely in experimental plots and so the sample should be effectively random (Yates called it "quasi-random") with a strong stratification effect.

25.2.6 Another fairly common type of S.U. is the cylindrical core used for soil sampling. This would be used in experiments in which it is part of the objective to study treatment effects on soil characteristics such as structural or nutrient properties or its degree of insect infestation. The diameter of the core varies, but in insect investigations should be at least 3 inches since smaller diameters might cause too many fractional bodies!

25.2.7 When samples are taken for chemical determinations from the produce either of a completely harvested plot or of S.U.'s selected for yield estimation, a common procedure is to mix up the harvested material as well as possible and then to follow the procedure known as "quartering". One form of this would be to divide the pile into equal quarters and to choose one randomly, the procedure being repeated on the selected quarter until a sample of the required size is obtained. A simpler but chancier method is the "grab sample"; the grab or grabs should be made with averted eyes, however!

25.2.8 The question arises whether, when a plot has been harvested by a sampling method, samples for chemical determinations should be taken from each S.U., or whether the S.U.'s should be bulked or "composited" and the samples taken from the bulked S.U.'s. The answer is that, if one is carrying out an investigation of a particular method of sampling plots to ascertain what contribution to the over-all sampling error is made by the various stages of the sampling procedure, then it may well be necessary to sample each S.U. separately in duplicate. On the other hand, for routine experimental work bulking is perfectly satisfactory and duplicate chemical determinations may have no virtue other than as a check against gross errors. In fact, in routine experimentation there is no need to preserve the identity of the S.U.'s even for yield determinations. A single weighing of the bulked produce is all that is necessary.

25.2.9 A related point is whether it is necessary to determine (say) dry matter percentage for each and every plot in an experiment for the purpose of calculating dry weights from fresh weights. Strictly speaking, this should be done, but in practice the variation between plots of the same treatment of such a laboratory determination may be relatively so small that Finney suggests as an economy that replicates be pooled in sets of 2 or 4 so that only $\frac{1}{2}$ or $\frac{1}{4}$ of the determinations have to be done. Then the mean percentage for each treatment should be applied to the fresh plot weights of the appropriate replications of that treatment to give the required dry weights. To carry this to the extreme, however, and make only one determination for each treatment from the bulked produce of all replications would lead to seriously incorrect results if the single determination were appreciably inaccurate for any treatment.

25.3 The effect of sampling on the analysis of the data

25.3.1 The fact that experimental data are obtained by means of a sampling procedure entails no change in the method of analysis. Since the S.F. is kept the same for all plots, there is no need even to apply the raising factor to get estimates of plot yields as in §§ 3.12.3 and 3.13; this is taken care of in the conversion factor to standard units, where the net plot area is equated to the area actually covered by the S.U.'s taken in a plot. The sample data for each plot are therefore totalled or meaned, as may be appropriate, and the ordinary analysis performed except for the alteration to the conversion factor, if appropriate.

25.3.2 Nevertheless, because we are no longer dealing with the data relating to the complete plots but with sample estimates, an additional component of error is introduced. This is included in the experimental error, increasing it, and causing a loss of precision. However, this does not affect the validity of the *F*-test, for example, since the Treatments and Error M.S.'s are equally affected; other tests of significance are similarly unaltered.

25.3.3 There is always the question, however, of the extent to which sampling error has affected precision, and whether the sampling method

could be improved in order to increase precision or whether the sampling may not have been unnecessarily intensive. Answers to such questions may be provided by an auxiliary "Within plots" analysis. For this analysis to be valid an essential requirement is that separate observations be available for a minimum of two S.U.'s selected at random within each plot, or, where there is stratification, within each stratum of each plot. Thus, if we wished to take 8 S.U.'s per plot in an experiment, we might divide the plot into quarters and choose two S.U.'s at random from each quarter. However, it is not necessary to estimate the sampling error on every occasion once a satisfactory method has been established. It would probably be more precise to divide the plot into 8 strata and take one S.U. at random from each, but then there can be no unbiased estimate of sampling error (cf. §§ 3.13 and 3.14).

25.3.4 We consider first the case of simple random sampling of each plot. Table 25.1 shows the yields of each S.U. in a randomized blocks experiment with 4 replications of 3 varieties of maize. The net plot was 10 yds. × 125 yds. and 3 S.U.'s each 5 yds. square were randomly selected on each plot. The S.F. was therefore $\frac{3}{50}$ or 6%.

Table 25.1: Sampled yields in a maize experiment. (Data from Saunders and Rayner, *Statistical methods with special reference to field experiments*)

S.U. No.	Block 1			Block 2			Block 3			Block 4		
	A	*B*	*C*	*A*	*B*	*C*	*A*	*B*	*C*	*A*	*B*	*C*
1	34	26	24	31	26	20	34	30	28	31	24	21
2	33	29	23	32	25	22	35	30	24	33	20	20
3	31	27	23	34	22	23	33	27	26	34	26	23
Totals	98	82	70	97	73	65	102	87	78	98	70	64

It is apparent that the form of the data is similar to that of a split-plot experiment, the S.U.'s corresponding to sub-plots and the plots to whole-plots. The difference is that there are no sub-plot treatments, the variation between S.U.'s within a plot being attributable to random causes only.

25.3.5 We therefore adopt the model

$$y_{ijk} = \mu + \beta_j + \tau_i + \epsilon_{ij} + \xi_{ijk}, \qquad [25.1]$$

where y_{ijk} represents the yield of the k^{th} S.U. on the plot receiving the i^{th} treatment in the j^{th} block. The first part of the model is the same as [13.6]; ϵ_{ij} is a component common to all S.U.'s of the $(i, j)^{\text{th}}$ plot with zero mean and variance V_p, and ξ_{ijk} is a component peculiar to the k^{th} S.U. of the $(i, j)^{\text{th}}$ plot and has zero mean and variance V_s. The latter represents the sampling variance, i.e. the variance of a single S.U. due to sampling error, and V_p is the component of variance due to experimental error. Implicit in the above is that the ϵ_{ij} and the ξ_{ijk} are variate-values from infinite populations; we further assume that all ϵ_{ij} and ξ_{ijk} are mutually uncorrelated. What this amounts to is that we regard each observed y_{ijk} as one variate-value from an infinite number of possible values which might have been

observed for that S.U. under the same conditions. The mean of these is $\mu + \beta_j + \tau_i$ and the observed deviation is the combined result of two uncorrelated random sources. It is very probable that adjacent S.U.'s within a plot will tend to be positively correlated (cf. § 12.1.4), though competition could induce negative correlation. The validity of the assumption of no correlation among the ξ_{ijk} rests on the randomness of the selection of the S.U.'s and the fact that we do not identify the observed variate-values within a plot with any particular S.U.'s.

25.3.6 If the sampling error were of no interest, the plot totals would be analysed as a randomized blocks design as explained in § 25.3.1. To estimate the sampling variance, the analysis is carried a step further in the manner of a split-plot analysis, but without sub-plot treatments. The difference between the Total S.S. (for all S.U.'s) and the S.S. for plot totals (Whole-plots S.S.), gives a S.S. for S.U.'s within whole-plots. It is usual, as in a split-plot analysis, to carry out the calculations on the basis of a single S.U. as unit, i.e. all the S.S.'s in the ordinary analysis described in the first sentence of this section would be divided here by an extra 3. Sometimes, however, the whole-plot analysis is done on a whole-plot basis and the sub-plot analysis brought up to the basis of the whole-plot analysis by multiplying the S.S.'s in the sub-plot analysis, here by 3. The analysis of variance with the former method is given in Table 25.2.

Table 25.2: Analysis of variance for the data of Table 25.1

Source of Variation	D.F.	S.S.	M.S.
Blocks	3	86·0	
Varieties	2	612·2	306·10
Error	6	21·8	$3{\cdot}633 = s_1^2$
Total, between plots ..	11	720·0	
Sampling units, within plots..	24	72·0	$3{\cdot}000 = s_2^2$
Total	35	792·0	

An alternative version of the skeleton analysis of variance might be:

Source	D.F.	M.S.
Blocks	3	
Varieties	2	
Error (experimental)..	6	s_1^2
Error (sampling) ..	24	s_2^2
Total	35	

The F-test for the Varieties M.S. (which requires the assumption that the ϵ_{ij} and ξ_{ijk} are N.D.) is made against the M.S. for experimental error, which in this example does not comply with the acceptable minimum number of D.F. This M.S. estimates the residual variance between plots and is therefore appropriate to comparisons between varieties.

25.3.7 The term "sampling error", like "experimental error" (cf. § 10.3.18),

is used variously. It can stand for the M.S. for S.U.'s within plots as an estimate ($\hat{V}_s$) of the sampling variance, or for its square root, the S.E. of a single S.U., or for the latter expressed as a percentage of the mean per S.U. (cf. also § 25.5.4). We shall try to distinguish between these, while retaining "sampling error" as a general term for the variability introduced by the sampling procedure over and above that which would have been present with complete harvesting. The phrase "loss of information due to sampling error" (cf. § 25.4.7) is a typical example of this usage. It goes without saying that sampling error may not be contracted to "S.E.".

25.3.8 Data of the above type, in which each of a set of experimental units is divided into sub-units in such a way that there is no relationship between the sub-units of different main units is referred to as a **nested classification** (as opposed to the two-way classification of Blocks and Treatments in a randomized blocks design or of two factors in a factorial experiment). If we had an experiment with a simple random sample taken from each plot and each S.U. sub-sampled for duplicate chemical determinations (say), this would constitute a further stage of nesting or a two-fold nested classification within plots. The skeleton analysis of variance of § 25.3.6 would then have been as in Table 25.3:

Table 25.3: Skeleton analysis of variance for a randomized blocks design with 3 S.U.'s per plot and 2 sub-samples per S.U.

Source	D.F.	M.S.
Blocks	3	
Varieties..	2	
Error	6	s_1^2
S.U.'s within plots	24	s_2^2
Sub-samples within S.U.'s ..	36	s_3^2
Total	71	

The 36 D.F. for the last category arises from the fact that there are 2 sub-samples (1 D.F.) within each of 36 S.U.'s.

The nested classification is also called the hierarchal or hierarchical classification. Examples were encountered earlier in Exercise 10.3 (iii) and in § 11.7.1.

25.3.9 *Stratification.* Suppose that in the example of Table 25.1 each plot had been divided into 2 halves (strata) and 3 S.U.'s had been randomly selected within each stratum. The effect would be to add a further stage of nesting as in the skeleton analysis of variance of Table 25.4.

25.4 Use of sampling data to investigate the efficiency of a sampling procedure

25.4.1 In Table 25.2 the Varieties M.S. is tested by $F = 306 \cdot 10/3 \cdot 633$, and is highly significant. However, it might not have been, and the question might have arisen whether non-significance was due to the sampling, and whether in a repetition of the experiment it might be possible to lower the

Table 25.4: Skeleton analysis of variance for a randomized blocks design with 3 S.U.'s selected from each of 2 strata per plot

Source	D.F.	M.S.
Blocks	3	
Treatments	2	
Error	6	s_1^2
Strata within plots	12	
S.U.'s within strata	48	s_2^2
Total	71	

experimental error appreciably by increasing the S.F. With F actually significant the question might be whether the S.F. could be lowered without increasing experimental error too much.

25.4.2 To answer these questions we must first estimate V_s and V_p (cf. § 25.3.5), using the results of § 10.7.4. In Table 25.2, since there are 3 S.U.'s per plot, s_1^2 is an estimate of $V_s + 3V_p$, and also s_2^2 is an estimate of V_s. Hence

$$\hat{V}_s = 3{\cdot}000$$
$$\hat{V}_s + 3\hat{V}_p = 3{\cdot}633,$$

i.e. $\hat{V}_p = 0{\cdot}211$. With n_s S.U.'s per plot, s_1^2 would have been an estimate of $V_s + n_s V_p$, i.e.

$$s_1^2 = \hat{V}_s + n_s\hat{V}_p. \qquad [25.2]$$

From [25.2] we may now estimate what the Error M.S. would have been with different numbers of S.U.'s per plot:

Table 25.5: Changes in precision with variation of S.F.

No. of S.U.'s per plot	Error M.S. (single S.U. basis)	Estimated variance of treatment mean per single S.U.
1	$3{\cdot}000 + 0{\cdot}211 = 3{\cdot}211$	$\frac{1}{4}(3{\cdot}211) = 0{\cdot}803$
2	$3{\cdot}000 + (2 \times 0{\cdot}211) = 3{\cdot}422$	$\frac{1}{8}(3{\cdot}422) = 0{\cdot}428$
3	$3{\cdot}000 + (3 \times 0{\cdot}211) = 3{\cdot}633$	$\frac{1}{12}(3{\cdot}633) = 0{\cdot}303$
4	$3{\cdot}000 + (4 \times 0{\cdot}211) = 3{\cdot}844$	$\frac{1}{16}(3{\cdot}844) = 0{\cdot}240$
	etc.	etc.

From Table 25.5 we can calculate that the significant differences (at the 5% level) detectable with 1, 2, 3, 4, . . . S.U.'s per plot would have been 3·10, 2·26, 1·90, 1·70, . . . respectively, where, for example $3{\cdot}10 = \sqrt{2 \times 0{\cdot}803} \times t_{.05}$ (6 D.F.). Hence in the present experiment where the treatment means per S.U. are

A	B	C	General mean
32·9	26·0	23·1	27·3 ,

two S.U.'s per plot would have been sufficient. With 1 S.U. per plot we could have detected as significant a difference of 11·4% of m, the mean per S.U.;

with 2 S.U.'s per plot a difference of 8·3% would have been detectable, and so on.

25.4.3 It might be argued that Table 25.5 seems to falsify the position in that the variance of a treatment mean *per S.U.* is considered, whereas a treatment mean would still be based on only the 4 replications of each treatment. This argument is, of course, incorrect, but in any case the effect of varying the S.F. may be alternatively demonstrated by calculating the C.V. on a plot basis (cf. § 20.8.3). If we were to analyse *totals* of n_s S.U.'s per plot on a plot basis, the mean yield per plot would be estimated as $n_s m$; also in the analysis of variance there would be no division of S.S.'s by n_s (cf. § 25.3.6), so that the Error M.S. would be an estimate of $n_s(V_s + n_s V_p) = n_s V_s + n_s^2 V_p$. Thus, the C.V. with n_s S.U.'s per plot would be estimated as

$$\frac{100\sqrt{n_s \hat{V}_s + n_s^2 \hat{V}_p}}{n_s m} = \frac{100}{m}\sqrt{\frac{1}{n_s}(\hat{V}_s + n_s \hat{V}_p)}. \qquad [25.3]$$

Using [25.3], we can draw up Table 25.6:

Table 25.6: Effect of varying S.F. on C.V.

No. of S.U.'s per plot	Mean	Error M.S.	Esimated C.V.
1	27·3	3·211	6·56%
2	54·6	6·844	4·79%
3	81·9	10·899	4·03%
4	109·2	15·376	3·59%

It will be observed that the figures in the last column of this table are in the same proportion as the square roots of the figures in the last column of Table 25.5. The gains in precision through taking more than 2 S.U.'s per plot do not seem to be very great when the extra work is taken into consideration.

25.4.4 With an additional nested classification as in Table 25.3 a similar procedure is followed. Let V_d stand for the variance of duplicate determinations within a S.U., V_s the component of variance due to S.U.'s within plots, and V_p the component of variance due to experimental error. An extra random component due to differences of sub-samples within a S.U. would be added to Model [25.1]. Then, in general, if there are n_d sub-samples per S.U., and n_s S.U.'s per plot in the observed data,

$$s_3^2 = \hat{V}_d,$$
$$s_2^2 = \hat{V}_d + n_d \hat{V}_s,$$
and
$$s_1^2 = \hat{V}_d + n_d \hat{V}_s + n_d n_s \hat{V}_p.$$

These equations give $\hat{V}_d$, $\hat{V}_s$, and $\hat{V}_p$, which can then be used to investigate the precision of various alternative sampling procedures. Thus with n'_d sub-samples per S.U. and n'_s S.U.'s per plot the Error M.S. would have been $\hat{V}_d + n'_d \hat{V}_s + n'_d n'_s \hat{V}_p$ and the variance of a plot mean per sub-sample would have been

$$\frac{1}{n'_d n'_s}(\hat{V}_d + n'_d \hat{V}_s + n'_d n'_s \hat{V}_p) = \frac{\hat{V}_d}{n'_d n'_s} + \frac{\hat{V}_s}{n'_s} + \hat{V}_p. \qquad [25.4]$$

The effect of varying n'_d and n'_s may therefore be studied. For example, with 8 S.U.'s per plot and 1 sub-sample per S.U., [25.4] would become

$$\tfrac{1}{8}(\hat{V}_d + \hat{V}_s) + \hat{V}_p.$$

The more realistic question of what the variance would have been if the S.U.'s had been bulked to give a single composite sub-sample per plot is answered by

$$\hat{V}_d + \tfrac{1}{8}\hat{V}_s + \hat{V}_p,$$

the coefficient of $\hat{V}_d$ being unity because there is only one ultimate determination. From these examples the general procedure should be clear. The usual situation is that $\hat{V}_p$, $\hat{V}_s$, and $\hat{V}_d$ are of diminishing importance, so that, for example, little is achieved by increasing n'_d. Similarly, the additional precision arising from increasing n'_s tends to get swallowed up in the plot variability.

25.4.5 This brings up the question of whether, if increased precision is sought, it might not be better obtained by increasing the number of replications; this would, for the same S.F., automatically increase the number of S.U.'s in the same ratio, but there is nothing to stop consideration of an increase in the number of replications accompanied by a decrease in the S.F. Where [25.4] is relevant, the corresponding variance of a treatment mean with r' replications is estimated by

$$\frac{\hat{V}_d}{n'_d n'_s r'} + \frac{\hat{V}_s}{n'_s r'} + \frac{\hat{V}_p}{r'}, \qquad [25.5]$$

which permits the desired comparisons. In the simpler case with which Table 25.2 deals, the variance of a treatment mean is estimated from [25.5] as

$$\frac{1}{n_s r}(\hat{V}_s + n_s \hat{V}_p) = \frac{\hat{V}_s}{n_s r} + \frac{\hat{V}_p}{r}. \qquad [25.6]$$

(We do not need the notation n'_s, r' here because we had specific numerical values of n_s and r to start with; thus in Table 25.5 r is always equal to 4, the number of actual replications.)

As an example of the use of [25.6], if the experimenter had been dissatisfied with the precision actually obtained in the variety trial of Table 25.1 (which is unlikely with a C.V. of only 4·03%), he could consider a comparison of, say, (i) doubling the number of S.U.'s taken per plot for the same number of replications, and (ii) doubling the number of replications while keeping the S.F. the same. In case (i) the estimated variance of a treatment mean would have been $\frac{1}{24}\{3{\cdot}000 + (6 \times 0{\cdot}211)\} = 0{\cdot}178$; in case (ii) it would have been $\frac{1}{24}(3{\cdot}633) = 0{\cdot}151$. The relative information provided by the two methods is therefore (cf. § 22.12.2) 0·178/0·151 or 118%. There is thus a relative gain of information of 18% in doubling the replications rather than the S.F., but this ignores the increase in Error D.F., which could be regarded as essential here. The above comparison involves a constant number of S.U.'s (72), but we might also consider taking $r = 8$, $n_s = 1$, which gives an estimated variance of $\frac{1}{8}(3{\cdot}211) = 0{\cdot}401$. This combination is clearly a loser for the present data, which are, however, somewhat atypical in that $\hat{V}_s$ is so large relative to $\hat{V}_p$.

In practice the relative labour and costs involved in increasing the number of replications as compared with increasing the S.F. would have to be considered, and there is a method of finding n_s and r so as to minimize the overall cost for a required degree of precision.†

25.4.6 Should s_1^2 be less than s_2^2 in Tables 25.2 or 25.3, or s_2^2 be less than s_3^2 in Table 25.3, then under Model [25.1] and the corresponding model relevant to Table 25.3 (cf. § 25.4.4), we would assume that this was due to chance, since on average $s_1^2 \geqslant s_2^2 \geqslant s_3^2$ according to this model. We would therefore take the component of variance concerned equal to zero in calculations such as those above (cf. §§ 10.7.7 and 20.6.11). Yates and Zacopanay‡ in their fundamental paper presented an amended model which allows for competition between S.U.'s within plots which could cause s_1^2 to be less than s_2^2 (cf. § 20.4.3). If s_1^2 were markedly less than s_2^2 this would, under the amended model, be evidence of competition, but, since competition effects cannot be isolated from the other components of variance, no further progress can be made. If $s_1^2 \geqslant s_2^2$ calculations made on the same lines as above are still valid under the amended model.

25.4.7 *Loss of information due to sampling.* Under Model [25.1], it is possible to estimate what the Error M.S. would have been with complete harvesting. For the data of Table 25.1 with all 50 S.U.'s on each plot harvested, this is (by [25.2]) $3{\cdot}000 + (50 \times 0{\cdot}211) = 13{\cdot}550$, which, it must be noted, still contains a component due to the sampling variance, even though the sampling error must now be zero in view of the complete harvesting. The corresponding efficiency criteria of Tables 25.5 and 25.6 would be:

$$\text{Estimated variance of treatment mean per single S.U.} = \tfrac{1}{200}(13{\cdot}550) = 0{\cdot}068$$

$$\text{Estimated C.V.} = \frac{100}{27{\cdot}3}\sqrt{\frac{1}{50}(13{\cdot}550)} = 1{\cdot}91\%$$

The information, relative to complete harvesting, supplied by the sampling procedure actually followed in the experiment (3 S.U.'s per plot) is therefore $0{\cdot}068/0{\cdot}303$ or $22{\cdot}4\%$, i.e. the loss of information due to sampling is $77{\cdot}6\%$, which is relatively high since $\hat{V}_s$ is large relative to $\hat{V}_p$.

In general, if the total number of S.U.'s per plot is N_s and if n_s S.U.'s were actually selected, the variances of a single plot mean per S.U. would be estimated from [25.2] as

$$(\hat{V}_s + N_s\hat{V}_p)/N_s = \hat{V}_p + \hat{V}_s/N_s \text{ (complete harvesting)} \qquad [25.7]$$

and

$$(\hat{V}_s + n_s\hat{V}_p)/n_s = \hat{V}_p + \hat{V}_s/n_s \text{ (sampling).} \qquad [25.8]$$

The relative information is therefore

$$\frac{\hat{V}_p + \hat{V}_s/N_s}{\hat{V}_p + \hat{V}_s/n_s} \qquad [25.9]$$

† Cf., for example, Kempthorne, O., *Design and analysis of experiments*, p. 214.
‡ *Journal of Agricultural Science* 25, 545–577, 1935.

and the loss of information is 1 minus this, or

$$1 - \frac{\hat{V}_p + \hat{V}_s/N_s}{\hat{V}_p + \hat{V}_s/n_s} = \frac{\hat{V}_p + \hat{V}_s/n_s - \hat{V}_p - \hat{V}_s/N_s}{\hat{V}_p + \hat{V}_s/n_s}$$

$$= \frac{(1 - n_s/N_s)\hat{V}_s/n_s}{\hat{V}_p + \hat{V}_s/n_s}$$

$$= \frac{(1 - n_s/N_s)\hat{V}_s}{\hat{V}_s + n_s\hat{V}_p}, \qquad [25.10]$$

multiplied by 100, if desired, to bring to a percentage. In [25.10], n_s/N_s is the S.F., commonly denoted by f. The use of f to stand for the S.F. is a debatable convention in view of our use of f for D.F., but actually there is little likelihood of confusion since their spheres of application are very distinct. Formula [25.10] may also be expressed directly in terms of s_1^2 and s_2^2 by means of [25.2], i.e. as

$$(1 - f)\frac{s_2^2}{s_1^2}. \qquad [25.11]$$

In the above it is always assumed that harvesting of all the S.U.'s of a plot separately would be the same as complete harvesting of the plot as a single operation.

25.4.8 *Stratification.* If stratified samples are drawn from each plot as described in § 25.3.9, it would be usual to consider as alternative sampling schemes to be investigated only those based on the same strata. In this case effects representing strata introduced into [25.1] would be regarded as fixed effects, e.g.

$$y_{ijkl} = \mu + \beta_j + \tau_i + \epsilon_{ij} + \gamma_{ijk} + \xi_{ijkl},$$

where γ_{ijk} represents the effect of the k^{th} stratum in the $(i, j)^{\text{th}}$ plot, and ξ_{ijkl} represents the effect of the l^{th} S.U. in the $(i, j, k)^{\text{th}}$ stratum. We would take $\sum_k \gamma_{ijk} = 0$ for each plot. With this model no stratum effects enter into the expected values of any M.S.'s of Table 25.4 other than the M.S. for Strata within plots. In fact, s_1^2 and s_2^2 in Table 25.4 are estimates of the same quantities as in Table 25.3, and so stratification has no effect on the calculations illustrated above with the numerical results of Table 25.2, provided the same strata are retained.

It is always possible to consider what the experimental error would have been without stratification, but this is a rather academic matter since stratification is such a simple and untroublesome precaution that it would seldom be wise to discard it if it was considered worth introducing in the first place. If the number of strata per plot is small (as it usually is), then it is hardly likely that sampling of strata (as initial large S.U.'s) would be considered. However, if it were, stratum effects would have to be considered random instead of fixed and Table 25.4 would be the same in principle as Table 25.3.

25.5 S.E. of the mean of a sample from a finite population

25.5.1 The variance of the mean of a sample of size n from an infinite

population with variance σ^2 is σ^2/n. But this is clearly inappropriate when the population is a finite one consisting of N S.U.'s. Intuitively we would expect the variation of the means of repeated simple random samples of n S.U.'s out of a population of N to be less about the mean of the finite population the closer n is to N. Finally, with $n = N$ the sample and population means coincide, and there can be no variation at all.

25.5.2 The discrepancy between the finite and infinite cases is reconciled in the following theorem:

Theorem 25.1 The S.E. of the mean ($\bar{x}$) of a sample of n variate-values (x_j) from a population of size N is given by

$$\sqrt{(1-\frac{n}{N})\frac{s^2}{n}}, \qquad [25.12]$$

where $s^2 = \Sigma(x - \bar{x})^2/(n-1)$.

It is seen that s^2 is the same as the unbiased estimate of variance which would be calculated if the sample were actually from an infinite population. This suggests the following proof, which is not favoured in finite population theory since it involves an assumption which is quite unnecessary, but which nevertheless fits in with the assumptions of Model [25.1], for example.

25.5.3 *Proof of Theorem 25.1* We assume that the finite population is itself a random sample of size N from an infinite population with variance σ^2. Let the N variate-values of the finite population be $x_1, x_2, \ldots x_N$, so that $\bar{x}_N$, the finite population mean $= (x_1 + x_2 + \cdots + x_N)/N$, or in vector notation

$$\bar{x}_N = \left[\frac{1}{N} \quad \frac{1}{N} \ldots\ldots \frac{1}{N}\right]\boldsymbol{x},$$

where $\boldsymbol{x}$ represents the vector of x's. Now suppose that the sample consists of the first n variate-values. (The sampling is simple random, but we can re-arrange the population values so that the n selected values come first.) Then we may write the sample mean as

$$\bar{x}_n = \left[\underbrace{\frac{1}{n} \quad \frac{1}{n} \ldots\ldots \frac{1}{n}}_{(n \text{ terms})} \vdots \underbrace{0 \ldots\ldots 0}_{(N-n \text{ terms})}\right]\boldsymbol{x}.$$

Hence

$$\bar{x}_n - \bar{x}_N = \left[\underbrace{\frac{1}{n}-\frac{1}{N} \quad \frac{1}{n}-\frac{1}{N} \ldots \frac{1}{n}-\frac{1}{N}}_{(n \text{ terms})} \vdots \underbrace{-\frac{1}{N} -\frac{1}{N} \ldots -\frac{1}{N}}_{(N-n \text{ terms})}\right]\boldsymbol{x},$$

and the variance of the sample mean about the finite population mean is therefore

$$\begin{aligned}
\text{Var.}(\bar{x}_n - \bar{x}_N) &= \left\{n\left(\frac{1}{n}-\frac{1}{N}\right)^2 + (N-n)\left(\frac{1}{N}\right)^2\right\}\sigma^2 \\
&= \frac{N-n}{N}\left\{\frac{N-n}{nN} + \frac{1}{N}\right\}\sigma^2 \qquad \text{(by Theorem 6.5)} \\
&= \left(1-\frac{n}{N}\right)\frac{\sigma^2}{n}.
\end{aligned}$$

(Note that σ^2 is the variance of the hypothetical infinite population, not of the finite population, which has not been defined.)

But now, since the sample of size n is being regarded as a sample of the random sample of size N from the infinite population, the sample of size n is effectively a random sample from the infinite population also (cf. § 3.6.5). Hence σ^2 is estimated from the sample as $\sum_{j=1}^{n}(x_j - \bar{x}_n)^2/(n-1)$, and the result follows.

25.5.4 Formula [25.12] is usually written as

$$\sqrt{(1-f)\frac{s^2}{n}}, \qquad [25.13]$$

where f is the S.F. and $1-f$ is known as the **finite population correction** (f.p.c.). Formula [25.13] or [25.12] may be termed the "sampling standard error" of the mean, to distinguish it from the S.E. of the mean of an infinite population $\sqrt{s^2/n}$. "Sampling standard error" is often contracted to "sampling error", which is difficult in view of the remarks of § 25.3.7, but, provided "sampling error of the mean" is used, the term is clear enough. Clearly [25.13] meets the requirements mentioned in § 25.5.1, having a limiting value of zero when $f = 1$. The f.p.c. makes very little difference if the S.F. is less than 5%, and is often discarded for S.F.'s of up to 10%.

25.5.5 The f.p.c. should not be used in tests of significance which require an assumption that the sample is a random sample from an infinite population.

25.6 Application of Theorem 25.1 to the sampling of experimental plots

25.6.1 *Theorem 25.2* The estimated variance of the mean of a single plot per S.U. is equal to the estimated variance with complete harvesting plus the estimated variance due to sampling.

Proof: In deriving the numerator of [25.10] it was seen that

$$(\hat{V}_p + \hat{V}_s/n_s) - (\hat{V}_p + \hat{V}_s/N_s) = (1 - n_s/N_s)(\hat{V}_s/n_s)$$

i.e.

$$\hat{V}_p + \hat{V}_s/n_s = (\hat{V}_p + \hat{V}_s/N_s) + (1 - n_s/N_s)(\hat{V}_s/n_s), \qquad [25.14]$$

which in view of [25.8], [25.7] and [25.12] proves the theorem.

25.6.2 In the example of Table 25.1, the estimated variance of a plot mean with complete harvesting is $\frac{1}{50}(13 \cdot 550) = 0 \cdot 27$, which is 4 times the corresponding figure for a treatment mean, viz. 0·068 (cf. § 25.4.7). The relationship of Theorem 25.2 may be illustrated as follows:

Table 25.7: Numerical illustration of Theorem 25.2 with varying S.F.

No. of S.U.'s per plot (n_s)	Estimated variance of plot mean due to sampling $\left(\frac{50 - n_s}{50} \cdot \frac{3.000}{n_s}\right)$	Estimated total variance of plot mean $\left(\frac{\hat{V}_s + n_s\hat{V}_p}{n_s}\right)$
1	2·94	3·21
2	1·44	3·422/2 = 1·71
3	0·94	3·633/3 = 1·31
4	0·69	3·844/4 = 0·96

In each case the difference of the last two columns is 0·27, this figure representing the estimated minimum variance obtainable when the sampling error is zero.

25.6.3 Equation [25.14] is still valid when N_s is infinite or when n_s/N_s is small enough to neglect, for then the two terms on the R.H.S. become $\hat{V}_p$ and $\hat{V}_s/n_s$ respectively.

25.6.4 Another useful relation is given by the following:

Theorem 25.3 If the experimental error (C.V.) with complete harvesting is $a\%$ and if the sampling standard error of a single plot mean is $b\%$, then the C.V. with sampling would be estimated as $\sqrt{a^2 + b^2}\%$.

Proof: Multiply all terms of [25.14] by $(100/m)^2$, where m is the mean per S.U. The last term is $(100/m)^2(1 - n_s/N_s)(V_s/n_s)$, which is b^2. From [25.3] the C.V. with full harvesting would be $(100/m)\sqrt{\hat{V}_p + \hat{V}_s/N_s}$, which is a. Hence the first term on the R.H.S. is a^2. Since the L.H.S. is also, by [25.3], the square of the C.V. with n_s S.U.'s per plot, the result follows.

For the data of Table 25.1 the C.V. is 4·03% (Table 25.6), while the percentage sampling standard error is $\sqrt{(0{\cdot}94/27{\cdot}3)100} = 3{\cdot}55\%$ (Table 25.7). Hence the C.V. with full harvesting may be estimated by Theorem 25.3 from

$$(4{\cdot}03)^2 = a^2 + (3{\cdot}55)^2,$$

whence $a = 1{\cdot}91\%$, as previously obtained in § 25.4.7.

25.6.5 The usefulness of Theorem 25.3 arises when we have carried out an investigation into a particular sampling method and have established its sampling standard error as $b\%$ of the mean per S.U. Then to estimate the experimental error as a percentage of the mean (C.V.) when we have this sampling error, we use any previous knowledge of the experimental error (with complete harvesting of plots) in the type of experiment proposed. Failing any such knowledge we may estimate it as 10% of the mean, usually regarded as a reasonable upper limit in field experiments. Then the C.V. with sampling is $\sqrt{100 + b^2}\%$. If the sampling error is also 10% the experimental error would be 14%. Hence a sampling error of 10% causes only an increase of 4% in the C.V.

25.6.6 If it should be thought that finite and infinite theories jostle as uneasy bed-fellows in §§ 25.5 and 25.6, it is reassuring that all the results of this chapter can be derived with no assumptions about the S.U.'s within a plot other than that they are equal in size and that n_s are randomly selected out of a total of N_s. For a valid F-test we would need to assume that means per S.U. of each plot were N.D., which should be approximately correct, a result used in the theory of sampling finite populations. This may also be relevant in connection with possible transformation of the variate (cf. § 24.4.5).

25.7 Size and shape of S.U.

25.7.1 In the above discussions we have assumed that the S.U. remained

unchanged in a given problem, but the problem may be to find the optimal S.U., which is similar to finding the optimum size and shape of plot in field experiments.

25.7.2 *For a given number of S.U.'s to be sampled per plot*, it is generally found that the larger the unit, the more precise it will be, because the S.F. is greater, and because larger S.U.'s will tend to vary less, proportionally, than small. However, one does not want the S.F. to be too large (cf. § 25.2.1).

25.7.3 On the other hand, *for a given S.F.*, the smaller S.U. will in general be more precise since there will be more small units than large. Thus, with soil sampling, for a given volume of soil 2-inch cores will give more precise estimates than 3-inch cores, and similarly for quadrat S.U.'s in field experiments. Nevertheless there are limitations to the minimum size of S.U. (cf. § 25.2.2).

25.7.4 Precision is not the only determinant. For example, suppose 6 large S.U.'s are found equivalent in precision to 10 S.U.'s half the size of the large units. In such a case it may be much more convenient from the point of view of the labour involved to handle the larger units even though the bulk involved is larger.

25.7.5 Some idea of the optimal S.U. may be gained by carrying out a preliminary sampling investigation. Thus a wheat plot may be harvested in a large number of very small areas (as in a uniformity trial), so that the relative precision of various sizes and shapes of S.U. based on area may be compared. If it were desired to base the unit on length of row, various small lengths of row would be harvested separately, and so on.

EXERCISES

25.1 The following are the plan and yields of a 2^4 experiment with factors l, p, k, n on wheat. The yields consist of 2 sampling units per plot and are given in kilograms. The sampling unit was a pattern of 4 equal rectangles each 2 ft. × 35 in., representing $\frac{1}{16}$ of the total effective plot area. The sampling fraction was therefore $\frac{1}{8}$.

Analyse the data and present the results in bags per morgen given that 1 bag = 200 lb., 1 lb. = 0·4536 kg., 1 morgen = 10,244 sq. yds.

Estimate what the experimental error would have been if only one sampling unit had been taken per plot, and also with complete harvesting of each plot.

Calculate the sampling error of a plot mean with 2 sampling units per plot and hence estimate the coefficient of variation if the whole plot had been harvested and thus confirm your previous estimate of the experimental error with the whole plot harvested.

(N.B.: The design is the same as that of Exercise 22.3.)

Block 1*A*

kn	*lpn*	*l*	(1)	*pn*	*pk*	*lpk*	*lkn*
1·20	1·25	1·08	1·16	1·26	1·13	1·15	1·29
1·16	1·07	1·12	1·20	1·28	1·31	1·11	1·22

Block 1*B*

n	*lpkn*	*p*	*lk*	*ln*	*k*	*lp*	*pkn*
1·19	1·20	1·15	1·40	1·20	1·12	1·35	1·24
1·21	1·18	1·20	1·30	1·28	1·18	1·27	1·25

Block 2*A*

p	(1)	*ln*	*lk*	*lpn*	*kn*	*pkn*	*lpk*
1·10	1·07	1·32	1·26	1·12	1·18	1·32	1·01
1·22	1·03	1·29	1·24	1·28	1·14	1·16	1·17

Block 2*B*

pk	*n*	*lpkn*	*k*	*lkn*	*pn*	*lp*	*l*
1·19	1·26	1·20	1·16	1·18	1·15	1·10	1·19
1·28	1·22	1·21	1·11	1·30	1·06	1·27	1·24

(Adapted from an experiment conducted by the Department of Agriculture, New Zealand)

25.2 (From F. R. Immer, *Applied Statistics*) The following data are yields in lb. from 6 S.U.'s per plot in a randomized blocks experiment with 4 varieties of crested wheat grass:

Block	Variety *A*	*B*	*C*	*D*
1	0·52	0·68	0·48	0·58
	0·49	0·62	0·55	0·58
	0·59	0·70	0·46	0·61
	0·36	0·70	0·58	0·63
	0·28	0·62	0·51	0·65
	0·49	0·66	0·38	0·71
2	0·52	0·85	0·27	0·42
	0·42	0·77	0·61	0·51
	0·66	0·46	0·44	0·58
	0·57	0·81	0·51	0·54
	0·59	0·58	0·61	0·66
	0·56	0·33	0·41	0·58
3	0·61	0·77	0·44	0·68
	0·49	0·91	0·48	0·48
	0·52	0·89	0·49	0·75
	0·56	0·95	0·61	0·71
	0·57	0·77	0·58	0·65
	0·49	0·77	0·41	0·68
4	0·42	0·70	0·55	0·58
	0·31	0·37	0·72	0·39
	0·47	0·33	0·68	0·44
	0·50	0·66	0·65	0·66
	0·35	0·64	0·48	0·65
	0·26	0·33	0·56	0·41

The S.U.'s were quadrats of 1 square yard in area and they were chosen randomly from the plots. The fraction of the net plot area sampled was $\frac{1}{10}$.

Analyse the varietal trial and present the results in bags per acre given that 1 bag = 200 lb. and 1 acre = 4,840 sq. yd.

Estimate what the experimental error would have been if only 2 quadrats had been taken per plot, and also with complete harvesting of each plot.

Calculate the sampling error of a plot mean with 6 S.U.'s per plot and, using Theorem 25.3, estimate the C.V. with complete harvesting, confirming your previous estimate of the experimental error with complete harvesting.

CONCLUDING REMARKS

It would be difficult to improve, either in wit or in wisdom, on D. J. Finney's "Valediction" in his *Introduction to statistical science in agriculture*, which is recommended to the attention of the reader.

The student who has completed this course (read this book and conscientiously worked through a set of distinctive exercises) can scarcely have avoided deriving the following minimal benefits: (1) an improvement in arithmetic, (2) a changed mode of thinking. It is to be hoped that both will be lasting.

Where a piece of research involves the collection of numerical data, especially by comparative experimentation, a much more confident approach should now be possible. The key here is, of course, the experimental design, which stipulates both how the data should be collected and how it should be analysed statistically. By following such a relatively foolproof method of procedure, the harsher criticisms to which a research worker can be exposed will be avoided.

In ordinary life a more sceptical attitude may have been engendered towards the "scientific tests" and "convincing graphs" which feature in much contemporary advertising.

On the other hand, the student should not imagine that he is now a qualified biometrician! Apart from the absence of the mathematical statistical theory, there has been much that could not be covered in a first course. There is, for example, the whole field of sample surveys and of biological assay. Even in the area of design and analysis of experiments only a basic coverage has been attempted, and the student may make the dismaying discovery that in practice more complicated designs are common. Mention has been made in the text of some of these developments. For the field experimentalist the most important are probably: (i) the more complex confounded designs for factorials, (ii) the lattice designs (also a type of incomplete block design) which are used with large numbers of non-factorial treatments, especially in plant breeding, and (iii) rotation experiments for evaluating sequences of crops and/or treatments. Mention of the last also brings to notice that the text has been concerned almost exclusively with annual experiments, and that little has been said about the repetition of an experiment in different seasons or on different sites, or about the continuation of an experiment on the same plots for several years. Nevertheless, with the present course as basis, the study of these more advanced types of design and analysis should be facilitated.

Rather should there now be a keener appreciation of the need to consult a biometrician, who, *if approached at a suitably early stage* may be able to point out unnoticed difficulties in a proposed experimental design and lay-

out and who, *if advised of all the facts*, may be able to suggest considerable improvements. In such consultations the reader will be pleasantly surprised at his ability to hold his own by virtue of his mastery of what in the early chapters must have seemed rather like a "new language", viz. the biometrical phraseology. The same applies to the reading of scientific papers. This advantage, apparently modest and perhaps not even fully realized, should not be under-estimated.

A First Course in Biometry for Agriculture Students

ERRATA

Title page and following page: The date of publication should be 1969, not 1967.

Page 6 (4 lines from foot of page): For "preceeding" read "preceding".

Page 11: The two diagrams in § 1.9.6. should be labelled (1) and (2).

Page 62: Formula 5.3 should read:

$$\bar{x} = \sum_j x_j \left(\frac{n_j}{n}\right)$$

Page 344, second line of Note K: For $\Sigma\xi_1 n$ * read $\Sigma\xi_1 n_1$*.

Page 376, Table 17.2: The last M.S. should be $s^2_{y,x}$.

Page 408: Value of x for Treatment D, Block 4, should be 62.2, not 62.3.

Page 486, line 13: For "D.F. of s^2" read "D.F. of $s_a{}^2$".

Page 507, end of first line of Note B: For "works" read "work-"; end of line 4 of same paragraph: "analysi-" to read "analysis".

Page 537: Data acknowledgement at end of Exercise 22.2 should read "Department of Agricultural Technical Services".

Page 570, line 11 of § 25.2.1: For "hervesting" read "harvesting".

Page 577, Table 25.6: Heading of last column should read "Estimated C.V.".