Clapham, A. R. 1929. The estimation of yield in cereal crops by sampling methods. *The Journal of Agricultural Science.* 19 (2), pp. 214-235.

THE ESTIMATION OF YIELD IN CEREAL CROPS BY SAMPLING METHODS.

By A. R. CLAPHAM, B.A.
(*Rothamsted Experimental Station, Harpenden.*)

(With Five Text-figures.)

It has long been felt that the use of a reliable sampling method would be highly advantageous for the estimation of yields in experimental work with cereals. Experiments were made during the summer of 1928 with a view to determining the accuracy with which such estimates may be made, and to deciding on a satisfactory sampling technique.

In all cases the number of samples taken from plots whose size varied from about one-seventeenth to one-fifth of an acre, was 30, each sample consisting of the total produce from a metre-length of a single drill. Each sample was tied with string and labelled on the field. Later it was weighed as a whole, after which the heads were cut off and threshed by hand. The grain was then weighed, and the weight corrected for moisture content, so that all calculations were based on dry weight.

Two, and in one case three, sampling methods were tried.

(*a*) The plots, all of which were narrowly rectangular in shape, were divided transversely into three equal parts, and ten samples were taken at random from each part.

(*b*) Six sets, each of which comprised a succession of five contiguous metre-lengths, disposed symmetrically within the plot, were cut as samples, each metre-length being tied separately. A somewhat similar scheme—the "Rod-Row Method,"—has been adopted by American agronomists (1), and it was thought desirable to compare it with a method based on random sampling.

(*c*) Six metre-lengths at equal intervals along the plot were cut from each of five drill-rows chosen at random. This scheme was only tried on one occasion, with wheat. Tables of the primary data will be found in Appendix I.

I. RESULTS WITH BARLEY.

These plots, of areas one-seventeenth to one-eleventh of an acre, and of three different varieties, were each sampled by methods (*a*) and (*b*) above. Total produce and grain were weighed for each metre-length, and results were calculated for straw as well as grain.

The statistical technique known as the "Analysis of Variance" was devised by R. A. Fisher, and first published in its complete form in 1923 (5). The principle of the method is that the total variation between the individual results in a set of data, if measured in terms of the sum of squares of deviations of these results from their general mean, may be analysed into a number of parts by the application of a well-known algebraic identity. This allows of the apportioning of fractions of the total sum of squares to various known causal factors, leaving a residual fraction due to unknown or uncontrolled factors. This latter fraction provides a logical basis for an estimate of the errors of an experiment. Fisher has further shown that the mean value of the fraction ascribable to any factor—the "variance"—is obtained by dividing that fraction by the number of "degrees of freedom" on which it is based, where "degrees of freedom" is used in the sense of "independent comparisons." Thus between n quantities whose mean is fixed there are in general $n - 1$ independent comparisons or degrees of freedom.

In the following pages the experimental results are treated in turn by this method. There are in all cases 29 degrees of freedom, since 30 samples were taken from every plot. In Method (*a*), since each plot was divided into three parts, 10 samples being taken from each part, the total variance may be analysed into a portion representing differences between the mean yield of the parts, and a residue representing differences between metre-lengths within the same part. The former portion may fairly be eliminated as being due to differences in mean fertility between the parts; the latter is used for the estimation of experimental error, representing as it does variance due to smaller differences in fertility within each part, to errors of measurement of the metre-lengths, to loss of grain in threshing, to errors in weighing, etc., etc. In a precisely similar manner the variance of results of Method (*b*) may be divided into two fractions, one due to differences between 5-metre-lengths or "sets," the other due to differences between metre-lengths within the same set. The former would be the basis of the estimate of error were the sets cut and weighed as wholes.

ANALYSES OF VARIANCE OF WEIGHTS PER METRE-LENGTH, IN GRAMS.

1. *Variety* "824."

Method (*a*). *Random sampling.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	2072·50	—	—
Remainder	27	6491·46	240·42	15·51
Total	29	8563·96		

Diminution* of variance = 18·59 per cent.

Standard error of a single metre-length = 15·51 gm.

Hence standard error of mean of 30 = $\frac{15{\cdot}51}{\sqrt{30}}$ = 2·83 gm.

Mean = 47·29 gm.: hence standard error of mean = 5·99 per cent.

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	5068·27	—	—
Remainder	27	12331·93	456·74	21·37
Total	29	17400·20		

Diminution of variance = 23·88 per cent.

Standard error of a single metre-length = 21·37 gm.

Hence standard error of mean of 30 = $\frac{21{\cdot}37}{\sqrt{30}}$ = 3·90 gm.

Mean = 72·97 gm.: hence standard error of mean = 5·35 per cent.

Method (*b*). *Systematic sampling: symmetrical method.*

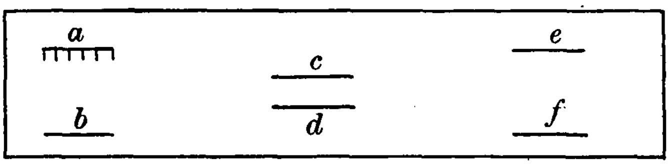

Fig. 1. Plan showing position of "sets" of metre-lengths in Method (*b*).

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	1247·10	249·42	(15·79)	0·476
Intra-set	24	2350·75	97·95	—	
Total	29	3597·85	124·06	(11·14)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 15·79 gm.;

(*b*) as calculated from individual values = 11·14 gm.

* See p. 230.

Hence standard error of mean:

$$(a) = \frac{15{\cdot}79}{\sqrt{30}} = 2{\cdot}88 \text{ gm.} = 7{\cdot}20 \text{ per cent. of mean;}$$

$$(b) = \frac{11{\cdot}14}{\sqrt{30}} = 2{\cdot}04 \text{ gm.} = 5{\cdot}10 \text{ per cent. of mean.}$$

It is interesting to note that, had the sets been cut as a whole, the standard errors would have been considerably overestimated, owing to the greater variability between than within sets. (Since the arrangement of the metre-lengths was systematic and not random, the standard errors obtained do not provide, in either case, valid estimates of the error of the mean.)

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	6666·08	1333·22	(36·51)	0·9730
Intra-set	24	4570·83	190·45	—	
Total	29	11236·91	387·48	(19·68)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 36·51 gm.;

(*b*) as calculated from individual values = 19·68 gm.

Hence standard error of mean:

$$(a) = \frac{36{\cdot}51}{\sqrt{30}} = 6{\cdot}67 \text{ gm.} = 10{\cdot}27 \text{ per cent. of mean;}$$

$$(b) = \frac{19{\cdot}68}{\sqrt{30}} = 3{\cdot}59 \text{ gm.} = 5{\cdot}54 \text{ per cent. of mean.}$$

As before there is a much higher variation between than within sets. The significance of this difference is easily found by R. A. Fisher's "*z*" test. "*z*" is half the difference between the natural logarithms of the two variances, and its standard error depends only on the number of degrees of freedom on which the variances are based. Tables have been provided (4) showing the value of "*z*" which must be attained for two different levels of significance, the 5 per cent. and the 1 per cent. points. If the 5 per cent. point of "*z*" is reached, it is to be understood that as great a difference between the two variances as was actually observed, would only occur by chance, from homogeneous material, once in 20 samples. Taking the 5 per cent. point, then, as a convenient minimum level for significance, the difference here found is hardly significant in the case of the grain, but highly significant with straw. The 5 per cent. point of "*z*" is 0·4817, and the 1 per cent. point, 0·6799.

2. *Variety "Spratt Archer."*

Method (*a*). *Random sampling.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	1195·27	—	—
Remainder	27	6778·17	251·04	15·84
Total	29	7973·44		

Diminution of variance = 9·70 per cent.

Standard error of a single metre-length = 15·84 gm.

Hence standard error of mean $= \frac{15 \cdot 84}{\sqrt{30}} = 2 \cdot 89$ gm. = 5·72 per cent. of mean.

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	2871·63	—	—
Remainder	27	19112·13	707·86	26·61
Total	29	21983·76		

Diminution of variance = 6·62 per cent.

Standard error of a single metre-length = 26·61 gm.

Hence standard error of mean $= \frac{26 \cdot 61}{\sqrt{30}} = 4 \cdot 86$ gm. = 6·09 per cent. of mean.

Method (*b*). *Systematic sampling: symmetrical method.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	1568·31	313·66	(17·71)	0·2914
Intra-set	24	4203·37	175·14	—	
Total	29	5771·68	199·02	(14·11)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 17·71 gm.;

(*b*) as calculated from individual values = 14·11 gm.

Hence standard error of mean:

$$(a) \quad = \frac{17 \cdot 71}{\sqrt{30}} = 3 \cdot 23 \text{ gm.} = 6 \cdot 57 \text{ per cent. of mean;}$$

$$(b) \quad = \frac{14 \cdot 11}{\sqrt{30}} = 2 \cdot 58 \text{ gm.} = 5 \cdot 24 \text{ per cent. of mean.}$$

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	2852·58	570·51	(23·89)	0·2773
Intra-set	24	7864·37	327·68	—	
Total	29	10716·95	369·55	(19·22)	

Standard error of a single metre-length:

(a) as calculated from whole sets = 23·89 gm.;

(b) as calculated from individual values = 19·22 gm.

Hence standard error of mean:

$$(a) = \frac{23\cdot89}{\sqrt{30}} = 4\cdot36 \text{ gm.} = 6\cdot95 \text{ per cent. of mean;}$$

$$(b) = \frac{19\cdot22}{\sqrt{30}} = 3\cdot51 \text{ gm.} = 5\cdot59 \text{ per cent. of mean.}$$

Here, although the inter-set is greater than the intra-set variance, the difference is not great, and falls short of significance when tested by the "z" method.

3. *Variety "Plumage Archer."*

Method (*a*). *Random sampling.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	2870·52	—	—
Remainder	27	9319·32	345·16	18·58
Total	29	12189·84		

Diminution of variance = 17·89 per cent.

Standard error of a single metre-length = 18·58 gm.

Hence standard error of mean $= \frac{18\cdot58}{\sqrt{30}} = 3\cdot39$ gm. $= 7\cdot47$ per cent. of mean.

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	13502·00	—	—
Remainder	27	24293·65	899·76	30·00
Total	29	37795·65		

Diminution of variance = 30·96 per cent.

Standard error of a single metre-length = 30·00 gm.

Hence standard error of mean $= \frac{30\cdot00}{\sqrt{30}} = 5\cdot477$ gm. $= 8\cdot31$ per cent. of mean.

Method (*b*). *Systematic sampling: symmetrical method.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	5733·95	1146·79	(33·86)	0·9842
Intra-set	24	3845·30	160·22	—	
Total	29	9579·25	330·32	(18·17)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 33·86 gm.;

(*b*) as calculated from individual values = 18·17 gm.

Hence standard error of mean:

$$(a) \;= \frac{33{\cdot}86}{\sqrt{30}} = 6{\cdot}18 \text{ gm.} = 15{\cdot}77 \text{ per cent. of mean;}$$

$$(b) \;= \frac{18{\cdot}17}{\sqrt{30}} = 3{\cdot}32 \text{ gm.} = 8{\cdot}46 \text{ per cent. of mean.}$$

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	7924·23	1584·85	(39·81)	0·8012
Intra-set	24	7661·56	319·23	—	
Total	29	15585·79	537·44	(23·18)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 39·81 gm.;

(*b*) as calculated from individual values = 23·18 gm.

Hence standard error of mean:

$$(a) \;= \frac{39{\cdot}81}{\sqrt{30}} = 7{\cdot}27 \text{ gm.} = 13{\cdot}78 \text{ per cent. of mean;}$$

$$(b) \;= \frac{23{\cdot}18}{\sqrt{30}} = 4{\cdot}23 \text{ gm.} = 8{\cdot}03 \text{ per cent. of mean.}$$

The 5 per cent. of "*z*" is 0·4817, and the 1 per cent. point, 0·6799. The significance of the difference between the intra- and inter-set variance therefore exceeds 1 in 100 both for grain and for straw. The effect of this is seen in the very much higher estimate of standard error obtained from whole sets as compared with individual metre-lengths.

II. RESULTS WITH WHEAT.

Variety "Red Standard."

Method (*a*). *Random sampling.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	1562·55	—	—
Remainder	27	8476·48	313·94	17·72
Total	29	10039·03		

Diminution of variance = 9·31 per cent.

Standard error of a single metre-length = 17·72 gm.

Hence standard error of mean $= \frac{17 \cdot 72}{\sqrt{30}} = 3 \cdot 24$ gm. = 8·08 per cent. of mean.

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error
Blocks	2	1122·02	—	—
Remainder	27	74295·45	2751·68	52·46
Total	29	85524·47		

Diminution of variance = 6·69 per cent.

Standard error of a single metre-length = 52·46 gm.

Hence standard error of mean $= \frac{52 \cdot 46}{\sqrt{30}} = 9 \cdot 58$ gm. = 8·72 per cent. of mean.

Method (*b*). *Systematic sampling: symmetrical method.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	1469·41	293·88	(17·14)	0·4044
Intra-set	24	3141·50	130·90	—	
Total	29	4610·91	159·00	(12·61)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 17·14 gm.;

(*b*) as calculated from individual values as previous = 12·61 gm.

Hence standard error of mean:

$$(a) = \frac{17 \cdot 14}{\sqrt{30}} = 3 \cdot 13 \text{ gm.} = 8 \cdot 09 \text{ per cent. of mean;}$$

$$(b) = \frac{12 \cdot 61}{\sqrt{30}} = 2 \cdot 31 \text{ gm.} = 5 \cdot 96 \text{ per cent. of mean.}$$

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-set	5	10036·45	2007·29	(44·80)	0·4732
Intra-set	24	18699·00	779·13	—	
Total	29	28735·45	990·88	(31·48)	

Standard error of a single metre-length:

(*a*) as calculated from whole sets = 44·80 gm.;

(*b*) as calculated from individual metres = 31·48 gm.

Hence standard error of mean:

$$(a) = \frac{44\cdot80}{\sqrt{30}} = 8\cdot18 \text{ gm.} = 9\cdot04 \text{ per cent. of mean;}$$

$$(b) = \frac{31\cdot48}{\sqrt{30}} = 5\cdot75 \text{ gm.} = 6\cdot35 \text{ per cent. of mean.}$$

In this case the differences between inter- and intra-set variances just fail to reach the 1 in 20 level of significance ("z" = 0·4817) for straw, and is smaller for grain. There is therefore little difference between standard errors based on the two variances.

Method (*c*). *Systematic sampling: Random Row method.*

1. Grain.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-row	4	21922·83	5480·71	(74·03)	1·4109
Intra-row	25	7154·37	286·17	—	
Total	29	29077·20	1002·66	(31·66)	

Standard error of a single metre-length:

(*a*) as calculated from whole rows = 74·03 gm.;

(*b*) as calculated from individual metres = 31·66 gm.

Hence standard error of mean, as calculated from individual metre-lengths $= \frac{31\cdot66}{\sqrt{30}} = 5\cdot78$ gm. = 11·33 per cent. of mean.

2. Straw.

Fraction	Degrees of freedom	Sum of squares	Mean square	Standard error	"z"
Inter-row	4	92529·83	23132·46	—	1·4250
Intra-row	25	33452·49	1338·10	—	
Total	29	125982·32	4344·22	(65·91)	

Standard error of a single metre-length = 65·91 gm.

Hence standard error of mean $= \frac{65\cdot91}{\sqrt{30}} = 12\cdot03$ gm. = 9·67 per cent. of mean.

As the very high values of "z" indicate, the variation between rows has been very much greater than within rows. This was largely due to the fact that an edge-row was sampled, and exaggerates somewhat the danger of systematic sampling of this type. Partial choking of drill-coulters, nearness of rows to field-drains, and many other factors, do, however, tend to make rows as a whole differ widely from their neighbours, and add weight to the case for random sampling.

RÉSUMÉ OF RESULTS.

The most important result that emerges is that with plots having an area of about one-sixteenth of an acre, a "random sampling" method will provide an estimate of yield with a standard error of less than 6 per cent. when 30 samples of metre-length of drill are taken. This would indicate that with plots one-fortieth of an acre in area, the average standard error should be not more than 5 per cent. It is customary at Rothamsted to have experimental plots of about this area, and since the standard error of such plots arising from causes other than sampling errors has been shown to be about 8–10 per cent., the additional inaccuracy introduced by the use of the sampling method described, will be quite small. Thus a standard error of 8 per cent. is increased to 9·4 per cent., and one of 10 per cent. only to 11·2 per cent. by the superposition of a further error of 5 per cent.

As the figures for "Percentage of variance eliminated" show very clearly, it is of great advantage to divide the area to be sampled into a small number of parts within each of which an equal number of samples is taken. By this means, and by the use of R. A. Fisher's statistical technique, the Analysis of Variance, a substantial reduction in the standard error may be effected.

Certain disadvantages of the systematic methods tried also stand out clearly. In the first place, the "rod-row" method used extensively in America is shown to suffer from the grave defect that the unit is often too coarse. This will be referred to later. Secondly, any attempt to reduce the labour of sampling by taking samples only from a small number of rows, whether these be chosen systematically or at random, is liable to lead to an increased estimate of error owing to the difference between rows as wholes—*i.e.* to the greater variation between than within rows. On the one occasion on which the method was adopted, the intra-row correlation, easily calculated from the ratio of intra-row and total variance (see R. A. Fisher (4) p. 191), is $+0·6635$, and highly significant. This indicates that the location of samples was very far from random, and the validity of the estimate of error correspondingly prejudiced.

Table of results.

	Crop ... Variety ... Area ...	Barley "824" 0·06 acre		Barley "Spratt Archer" 0·06 acre		Barley "Plumage Archer" 0·09 acre		Wheat "Red Standard" 0·2 acre	
		Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
Method (*a*) Random sampling	Mean weight per metre in grams	47·29	72·97	50·60	79·75	45·41	65·94	40·04	109·92
	Standard error of mean (%)	5·99	5·35	5·72	6·09	7·47	8·31	8·08	8·72
Method (*b*) Systematic sampling: symmetrical method	Mean weight per metre in grams	40·01	64·91	49·25	62·78	39·20	52·74	38·70	90·44
	Standard error of mean calculated from whole sets (%)	7·20	10·27	6·57	6·95	15·77	13·78	8·09	9·04
	Standard error of mean calculated from individual metres (%)	5·10	5·54	5·24	5·59	8·46	8·03	5·96	6·36
Method (*c*) Systematic sampling: "Random row" method	Mean weight per metre in grams	—	—	—	—	—	—	51·00	124·40
	Standard error of mean (%)	—	—	—	—	—	—	11·33	9·67

Discussion.

Sampling methods have hitherto been employed on numerous occasions for estimating the yields of cereal plots, but in extremely few cases is it possible to gain any idea of the accuracy of the methods used. Perhaps the fullest available sets of data are those of Arny and Garber (1), and of Arny and Steinmetz (2) and it will be instructive to examine these in some detail.

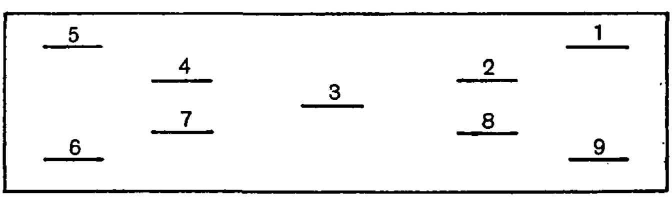

Fig. 2.

Arny and Garber employed the "rod-row" method, cutting nine symmetrically disposed rod-lengths of drill from each plot, and comparing the resulting estimates of yield with those obtained by harvesting the whole plots. The area of the plots was in all cases a tenth of an acre.

The position of the samples was as shown in Fig. 1.

The yield was estimated not only from the mean of all the nine samples but also from four only (samples 2, 4, 7 and 8), and from five only (samples 1, 3, 5, 6 and 9).

In order to obtain some idea of the additional errors introduced by sampling, the figures for "Marquis" wheat grown at the Morris Substation, and for "Haynes Bluestem" wheat grown at University Farm, St Paul, Minn. (Tables III and V in the original paper), have been grouped together, and an analysis made of the total variance. Of the sampling figures only those for nine and for four samples have been considered.

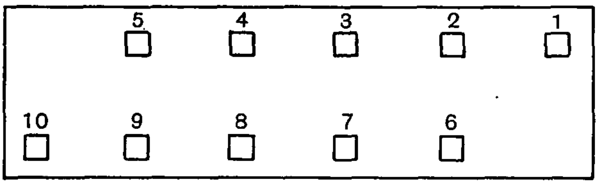

Fig. 3.

Each of the two experiments involved 18 plots, there being six different treatments in triplicate. It is therefore possible to eliminate that portion of the variance due to treatments, thus obtaining a residual variance representing differences between plots similarly treated.

1. Analysis of figures for total yields.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	11	160·23	14·57	—	—	—
Remainder	24	116·73	4·86	0·7907	2·205	6·68
Total	35	276·96				

2. Analysis of figures for nine rod-row samples.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	11	399·17	36·88	—	—	—
Remainder	24	163·11	6·796	0·9583	2·607	7·63
Total	35	562·28				

3. Analysis of figures for four rod-row samples.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	11	299·42	27·220	1·5513	2·861	8·54
Remainder	24	196·50	8·187	—	—	—
Total	35	495·92				

The data of Arny and Steinmetz have been similarly treated. Here the sampling units were square yards instead of rod-rows. Their arrangement is shown in Fig. 3.

As before yields were calculated not only from the 10 samples from each plot, but also from 9, 8, 5, 4 from the centre, and 4 from the ends. Analyses were made of figures for total yields, for estimates from 10 samples, and from 5 samples. Series II, III and IV, were grouped for this purpose.

4. Analysis of figures for toal yields.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	17	2293·73	134·925	—	—	—
Remainder	36	119·38	3·316	0·5994	1·821	6·62
Total	53	2413·11				

5. Analysis of figures for ten square-yard samples.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	17	1116·93	65·702	—	—	—
Remainder	36	186·75	5·188	0·8232	2·278	9·80
Total	53	1303·68				

6. Analysis of figures for five square-yard samples.

Fraction	Degrees of freedom	Sum of squares	Mean square	½ log mean square	Standard error	Standard error %
Sets	17	1046·06	—	—	—	—
Remainder	36	275·30	7·647	1·0172	2·765	12·01
Total	53	1321·36				

The additional standard error per plot due to sampling may now be estimated for the values of the mean squares given above:

9 rod-rows	3·67 per cent.
4 ,, ...	5·32 ,,
10 square yards	7·23 ,,
5 ,, ...	10·03 ,,

The first three of these standard errors are fairly small, and, since a rod is roughly equivalent to 5 metres, of about the same order as those obtained in the random sampling method. It must be noted, however, that these estimates are subject to very large sampling errors, owing to the fact that they are calculated from the differences between variances.

The standard error of the difference between two variances based on N degrees of freedom is given by the formula:

$$s_d = \sqrt{\frac{2}{N}(\sigma_1^4 + \sigma_2^4 - 2r_{\sigma_1^2\,\sigma_2^2}\,\sigma_1^2\,\sigma_2^2)},$$

where σ_1^2, σ_2^2 are the two variances, and $r_{\sigma_1^2\,\sigma_2^2}$ is the correlation between them in samples.

Now it has recently been shown by Wishart ((6) p. 43), that if $\rho_{1.2}$ is the correlation between the two variates in the original population $r_{\sigma_1^2\,\sigma_2^2} = \rho^2_{1.2}$, exactly.

Then, substituting for the population parameters σ_1^2, σ_2^2, $\rho_{1.2}$, the observed values s_1^2, s_2^2, $r_{1.2}$, which are the best available estimates, we have:

$$s_d = \sqrt{\frac{2}{N}\left(s_1^4 + s_2^4 - 2r^2_{1.2}s_1^2 s_2^2\right)}$$

is in each case the residual correlation between the estimates of yield obtained by the two methods—*i.e.* the correlation calculated from the "Remainder" variances and covariances[1]. The covariances must therefore be analysed in precisely the same manner as the variances, before the correlation coefficients can be obtained.

(*a*) *Total yields and* 9 *rod-rows.*

Fraction	Covariance	Correlation coefficient
Sets	150·733	+0·5960
Remainder	89·056	+0·6464
Total	239·789	+0·6076

(*b*) *Total yields and* 4 *rod-rows.*

Fraction	Covariance	Correlation coefficient
Sets	154·669	+0·7061
Remainder	105·855	+0·6989
Total	260·524	+0·7030

(*c*) *Total yields and* 10 *square yards.*

Fraction	Covariance	Correlation coefficient
Sets	1528·066	+0·9547
Remainder	88·286	+0·5913
Total	1616·352	+0·9113

(*d*) *Total yields and* 5 *square yards.*

Fraction	Covariance	Correlation coefficient
Sets	1451·438	+0·9370
Remainder	106·399	+0·5869
Total	1557·837	+0·8724

Using the underlined correlation coefficients, the following results are obtained for the accuracy of the difference between the residual variances:

(*a*) $1{\cdot}932 \pm 1{\cdot}878$,

(*b*) $3{\cdot}324 \pm 2{\cdot}077$,

(*c*) $1{\cdot}872 \pm 1{\cdot}199$,

(*d*) $4{\cdot}331 \pm 1{\cdot}699$.

Hence only the last can be said to differ significantly from zero, and in no case is the value of the sampling error established with any approach to certainty.

The same result may be arrived at more simply by the use of R. A. Fisher's "*z*" transformation, by means of which the significance of a difference between variances may be tested directly.

[1] The "covariance" is the average value of the sum of products of deviations from the mean, and is obtained by dividing that quantity by the appropriate number of degrees of freedom.

1. Arny and Garber's data.

	Total yield and 9 rod-rows	Total yield and 4 rod-rows	5 % point of "z"	1 % point of "z"
"z"	0·1676	0·2606	0·3425	0·4890

2. Arny and Steinmetz' data.

	Total yield and 10 square yards	Total yield and 5 square yards	5 % point of "z"	1 % point of "z"
"z"	0·2238	0·4178	0·2596	0·3702

Here again only the last result attains even the 5 per cent. point of "*z*"—*i.e.* only in the last case would such a difference between variances occur as infrequently as once in 20 samples from a homogeneous population. This confirms the conclusion that little reliance can be placed on the calculated values of the errors due to sampling, and illustrates in a striking manner one of the great disadvantages of a systematic as compared with a random sample. A random sample gives a direct estimate of the errors due to sampling, an estimate, therefore, of far greater accuracy than the indirect estimate obtained as above. It is of great importance that such an estimate should be arrived at, since the improvement of experimental technique depends on a knowledge of the causes of inaccuracy. Data such as those of Arny and Garber do not distinguish adequately between errors due to insufficient replication, and errors of sampling within the plots. If the main sources of error were the former, the taking of a greater number of samples from each plot would do little towards increasing the accuracy of the experiment, and vice versa.

The only manner in which direct estimates of the sampling error of a systematic method can be obtained, is by making a series of observations in which at least duplicate sets of samples are taken from each plot. These sets must further be such that they form a random sample from the whole population of possible sets. Only under these circumstances can a valid estimate of error be made. Hence there is the initial condition that such a population exists, for if it does not exist, no random sample can be made from it, nor can a standard error be calculated. In the present instance it is very difficult to see how such populations can be constructed. It is, however, possible, to devise systematic methods which do admit of the calculation of a valid standard error. The method used by Engledow (3), is a case in point. Here, in one variant of the method, 1-ft. samples are cut as in Fig. 4.

AB represents the width of the area to be sampled and is measured parallel with the drill-rows. 1, 2, 3 ... represent the 10th, 20th, 30th ...

drill rows. Samples are taken successively from these rows, there being a constant lateral shift from sample to sample. When a complete traverse of the area has been effected, a fresh start is made from the far side, as on the 70th drill-row in the figure.

If n samples are taken in passing from side to side of the area, there may be considered to be n possible starting-points along the base line AB. It would then be possible to get a valid estimate of error by taking

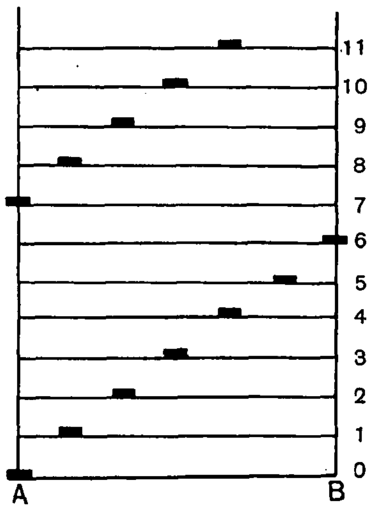

Fig. 4.

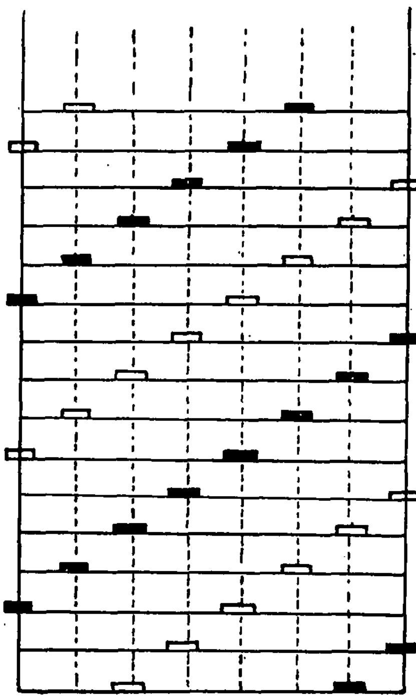

Fig. 5.

duplicate sets from each area to be sampled, the starting-points of the sets being chosen at random from among the n possible points. Thus in one of the areas, the samples might be taken as in Fig. 5.

A systematic arrangement, such as that of Engledow, samples the area very effectively, but it can scarcely be maintained that this advantage outweighs the disadvantage that the samples do not in themselves yield an estimate of their standard error, as would be the case with a random sample. Thus it should be noted that n separate *sets* of systematic samples would be required to yield the same information as to sampling errors as a single set of n samples distributed at random.

This is true whatever the number of samples in the systematically arranged set. It can be ensured moreover that the random method does sample the whole area, by dividing the area into a number of subdivisions. Then if the same number of samples is taken from each part of the area, there is secured both an effective distribution of samples, and an arrangement which permits of the elimination of that portion of the total variance which is due to differences in the mean fertility of the subdivisions of the area. This is, of course, equally possible with a systematic method, but although the necessary statistical technique—the Analysis of Variance—has been available for some years, no such use has previously been made of it.

The advantage to be gained by the subdivision of the area to be sampled has already been referred to. It may be useful, however, to collect the figures which bear on the point. Below are given the percentages of the total variance which are eliminated as being due to difference in soil fertility over subdivisions of the areas sampled, and which, consequently, represent percentage reductions of the variance on which the standard errors are based.

		% reduction in variance	
Crop	Variety	Grain	Straw
Barley	"824"	18·59	23·88
,,	"Spratt Archer"	9·70	6·62
,,	"Plumage Archer"	17·89	30·96
Wheat	"Red Standard"	9·31	6·69

It will be seen that in no case is less than 6·5 per cent. of the variance removed by this procedure, the mean reduction being 15·46 per cent.

With regard to the size and nature of the sampling unit, our results show conclusively that the rod-row is too coarse a unit. The significant intra-class correlations obtained when separate weighings are made of the five metre-lengths in each sampling unit of method (*b*), are as shown below:

		Intra-class correlations (where significant)	
Crop	Variety	Grain	Straw
Barley	"824"	—	+0·7966
,,	"Spratt Archer"	—	—
,,	"Plumage Archer"	+0·6886	+0·3856
Wheat	"Red Standard"	—	+0·1866*

* $p =$ about 0·053.

Except where otherwise stated the level of significance has been taken as $p = 0{\cdot}050$, where p is the probability that so high a value could be obtained by chance.

These figures indicate that there has been a considerable loss of information over that provided by the same number of metre-lengths arranged at random over the area. In certain other investigations by the author it has even appeared that the metre may be too long, significant correlations having been obtained between successive half-metre-lengths of drill. In view of this experiments were tried with a dissected 4-ft.-length, each foot being separated from its neighbour by 2 ft. of unsampled corn. No significant intra-class correlations were obtained with this method, even when "neighbouring" foot-lengths were compared. Engledow uses the foot-length as his unit, and points out that smaller lengths would be impracticable owing to the increased importance of end-errors. There being no intra-class correlation between the parts of a dissected 4-ft.-length, it is better to use 30 of such units rather than 120 separately located 1-ft.-lengths, since the location of each of the former units fixes 4 ft. at once, thus reducing the labour involved in sampling.

What has been said of using the rod-row as a unit will be equally true of the square yard. In fact it seems highly probable that the loss of information would be even greater in the case of five metres lying side by side than if they were end to end.

In conclusion, I wish to thank Dr R. A. Fisher of this Station for valuable criticism and advice; and Messrs H. J. Johnson and T. W. Simpson of Armstrong College, for carrying out almost the whole of the experimental work.

SUMMARY.

1. Cereal plots were sampled by three different methods; two systematic, and one involving a random location of sampling units.

2. The disadvantages of the systematic methods as compared with random sampling, emerged clearly.

3. These disadvantages were further emphasised in an analysis of earlier data on sampling methods. For this purpose the methods and results of certain recent contributions to statistical theory were used.

4. By the use of a random sampling method, the variance due to sampling errors may be made a satisfactorily small fraction of the total variance of cereal plots one-fortieth of an acre in area.

REFERENCES.

(1) ARNY, A. C. and GARBER, R. J. "Field Technic in Determining Yields of Plots of Grain by the Rod-Row Method." *J. Amer. Soc. Agron.* (1919) **11**.

(2) ARNY, A. C. and STEINMETZ, F. H. "Field Technic in Determining Yields of Experimental Plots by the Square Yard Method." *J. Amer. Soc. Agron.* (1919) **11**.

(3) Engledow, F. L. "A Census of an Acre of Corn." *J. Agri. Sci.* (1926) **16**.
(4) Fisher, R. A. *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh (1925). 2nd edit. 1928.
(5) Fisher, R. A. and Mackenzie, W. A. "Studies in Crop Variation. II. The Manurial Response of Different Potato Varieties." *J. Agri. Sci.* (1923) **13**.
(6) Wishart, J. "The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population" *Biometrika* (1928) **20**.

APPENDIX I.

Tables of primary data.

(The figures are in grams per metre: the grain figures in grams of dry weight per metre.)

I. Barley.

1. *Variety* "824."

Method (*a*). *Random sampling.*

Block A		Block B		Block C	
Grain	Straw	Grain	Straw	Grain	Straw
28·47	61·3	56·18	90·5	22·56	34·3
78·83	101·7	64·04	95·2	41·99	62·3
93·87	137·9	45·96	62·6	14·45	37·9
35·40	58·1	55·17	88·7	28·90	50·8
53·31	51·9	69·62	127·6	33·04	50·9
55·34	73·5	36·08	79·3	27·88	63·0
65·99	35·9	54·07	84·0	41·99	53·3
44·61	72·2	35·74	78·7	49·09	81·9
61·42	86·3	47·82	92·4	45·37	63·3
29·40	50·2	50·53	90·2	51·45	73·1
546·64	729·0	515·21	889·2	356·72	570·8

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1418·57	2189·0	47·286	72·967

Method (*b*). *Systematic sampling: symmetrical method.*

Set *a*		Set *b*		Set *c*		Set *d*		Set *e*		Set *f*	
Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
13·10	23·5	22·73	37·1	35·82	60·6	41·15	63·3	43·85	45·1	37·68	62·4
34·81	45·8	34·98	51·6	47·91	86·3	53·40	85·8	52·89	82·4	28·98	50·7
36·67	59·6	28·22	39·6	27·97	67·9	38·61	68·3	58·38	94·9	40·89	69·6
44·36	34·1	46·89	61·5	44·19	77·7	57·96	92·4	47·23	80·1	44·35	68·5
31·78	58·5	24·76	37·7	58·21	88·1	49·51	93·4	29·57	85·0	43·34	75·7
160·72	221·5	157·58	227·5	214·10	380·6	240·63	403·2	231·92	387·5	195·24	326·9

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1200·19	1947·2	40·006	64·907

2. *Variety "Spratt Archer."*

Method (a). Random sampling.

Block A		Block B		Block C	
Grain	Straw	Grain	Straw	Grain	Straw
63·84	122·4	80·94	117·8	62·01	86·7
62·40	108·2	47·12	70·5	43·41	65·2
34·06	69·0	78·57	109·8	42·05	61·9
41·73	80·3	37·30	39·9	51·40	73·1
35·24	64·5	61·38	82·5	35·16	55·6
48·95	79·2	68·67	109·3	51·40	68·1
30·49	61·5	79·75	119·3	34·14	54·9
85·30	162·3	57·50	64·4	57·66	81·2
48·55	69·7	34·45	95·5	36·51	50·9
23·68	65·1	48·39	40·9	35·80	62·8
474·24	882·2	594·07	849·9	449·54	660·4

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1517·85	2392·5	50·595	79·75

Method (b). Systematic sampling: symmetrical arrangement.

Set *a*		Set *b*		Set *c*		Set *d*		Set *e*		Set *f*	
Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
63·87	74·4	27·80	27·1	51·71	68·8	52·97	85·3	59·14	70·0	48·24	59·9
19·86	25·5	47·90	51·3	53·82	63·3	40·89	56·6	46·55	55·9	36·16	64·2
34·81	45·8	29·82	39·7	60·66	66·2	48·58	57·5	57·96	85·4	83·39	99·3
44·10	102·8	47·48	51·8	45·79	58·8	62·61	49·9	63·11	77·3	49·93	61·9
32·95	49·0	53·23	62·0	29·99	34·5	53·74	71·4	52·97	71·3	77·31	96·5
195·59	297·5	206·23	231·9	241·97	291·6	258·79	320·7	279·73	359·9	295·03	381·8

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1477·34	1883·4	49·245	62·780

3. *Variety "Plumage Archer."*

Method (a). Random sampling.

Block A		Block B		Block C	
Grain	Straw	Grain	Straw	Grain	Straw
21·15	34·3	75·70	109·4	45·62	56·0
49·34	77·7	37·94	48·1	32·28	31·8
74·53	118·9	49·93	61·9	8·79	11·6
76·03	124·0	56·10	67·6	22·56	29·3
93·77	167·6	62·35	87·2	23·83	33·8
54·33	90·4	67·25	91·4	16·14	22·9
21·62	28·7	62·10	86·5	36·67	40·6
65·97	119·7	31·35	48·9	48·07	60·1
43·87	85·6	41·57	56·8	42·08	48·2
23·21	53·7	38·36	35·6	39·71	50·0
523·82	900·6	522·65	693·4	315·75	384·3

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1362·22	1978·3	45·407	65·943

Method (*b*). *Systematic sampling: symmetrical method.*

Set *a*		Set *b*		Set *c*		Set *d*		Set *e*		Set *f*	
Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
45·79	53·8	37·26	37·9	41·15	50·3	33·46	56·4	43·77	58·2	28·30	55·5
22·98	37·8	22·64	38·2	71·99	90·8	34·98	32·6	56·61	62·0	24·25	50·3
31·77	35·4	29·32	42·3	75·87	102·2	51·79	71·7	36·67	55·6	14·28	20·1
27·54	38·4	30·75	44·6	83·90	112·7	67·59	91·0	32·53	44·5	6·17	9·7
22·31	22·6	34·39	44·3	42·58	53·6	37·60	53·5	58·04	71·3	29·66	44·9
150·39	188·0	154·36	207·3	315·49	409·6	225·42	305·2	227·62	291·6	102·66	180·5

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1175·94	1582·2	39·198	52·740

II. WHEAT.

Variety "Red Standard."

Method (*a*). *Random sampling.*

Block A		Block B		Block C	
Grain	Straw	Grain	Straw	Grain	Straw
26·50	66·1	36·98	107·7	47·85	154·8
37·14	91·5	37·29	111·3	120·10	344·4
30·49	74·0	42·14	139·1	51·06	147·7
37·14	105·5	34·56	87·8	45·74	125·5
52·39	129·0	35·11	88·1	47·23	110·6
36·04	93·9	41·44	110·0	17·98	62·0
48·01	181·6	37·37	73·2	24·39	64·8
20·33	61·0	46·52	109·5	58·95	143·6
12·12	54·5	25·33	65·6	36·83	103·9
28·07	102·1	37·22	73·4	48·87	115·5
328·23	959·2	373·96	965·7	499·00	1372·8

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1201·19	3297·7	40·040	109·923

Method (*b*). *Systematic sampling: symmetrical method.*

Set *a*		Set *b*		Set *c*		Set *d*		Set *e*		Set *f*	
Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
41·99	88·3	27·52	60·8	38·70	76·5	20·09	47·3	47·38	112·4	29·40	79·4
23·30	51·2	40·66	85·0	46·29	96·8	34·25	79·2	41·28	114·2	45·43	76·9
42·22	92·0	31·67	67·5	30·49	74·0	41·12	100·4	32·37	88·6	34·18	87·3
26·89	65·6	35·18	79·0	60·52	137·6	42·30	94·9	46·60	111·4	36·20	93·7
40·50	83·2	30·81	68·6	88·75	217·5	29·40	66·4	41·83	123·5	33·70	93·9
174·90	380·3	165·84	360·9	264·75	602·4	167·16	388·2	209·46	550·1	178·91	431·2

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1161·02	2713·1	38·701	90·473

Method (*c*). *Random row method.*

Row 1		Row 3		Row 19		Row 20		Row 24	
Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
94·58	184·9	22·08	60·3	30·49	69·5	32·13	163·6	9·53	25·9
90·02	221·2	33·93	80·5	36·59	104·4	29·03	68·2	36·34	96·7
86·41	195·4	33·93	67·5	30·24	61·8	48·88	94·1	42·18	94·9
114·93	267·2	30·24	73·8	50·42	93·3	24·91	110·0	51·02	98·6
148·95	311·6	46·99	121·3	46·39	165·0	51·20	127·4	60·99	136·0
93·55	221·1	30·07	76·0	50·59	133·1	30·07	96·0	43·29	112·6
628·44	1401·4	197·24	479·4	244·72	627·1	216·22	659·3	243·35	564·7

Grand total		Grand mean	
Grain	Straw	Grain	Straw
1529·77	3731·9	50·992	124·397

(*Received December* 6*th*, 1928.)