Probable error in field Experiments.

Every member of this Department is working, more or less directly, to bring about improvements in Indian Agriculture. All improvements, whatever their nature, must be worked out by actual experiments. In the great majority of cases these experiments finally resolve themselves into field trials in which different *manures*, *methods of culture*, *varities of plant* etc., are compared to decide which gives the greatest crop or monetary return. Thus practically the whole work of the Department is bound up, in the end, in field experiments.

It is very commonly thought, by those who have not tried, that to carry out a field experiment is a very simple matter. Many of you have had experience of such work and have no doubt begun to realize, how difficult it is to obtain definite reliable results, in all cases, except those in which the differences to be measured are distinctly large. This is purely a mutter of the relative magnitude of the experimental error on the one hand and the difference to be measured on the other.

It would appear natural that, before setting out to make comparisons by means of field trials, some attempt should be made to estimate the error that may be expected in the type of experiment contemplated. This error is a constant factor in all field experiments and yet I have seen no reference to any attempt at its estimation in any of the many reports on Agricultural experiments in India.*

My main object in this paper is to call attention to the importance of this subject and to show the value of work definitely designed to throw light on it, for I feel that no experimenter is justified in drawing conclusions from his results, and using these as a basis for practical advice, unless he has a definite idea of the degree of accuracy of his methods.

I will now try to give, very briefly, some idea of what is meant by *probable error*. You are all no doubt familiar with the use of what is called a *frequency curve* for showing the variations of a number of different determinations of any given value. Figure I shows such a

*Dobbs (Agricultural Journal of India Vol. VI Part I page 59) gives a method of reducing the error due to variation in fertility of the land.

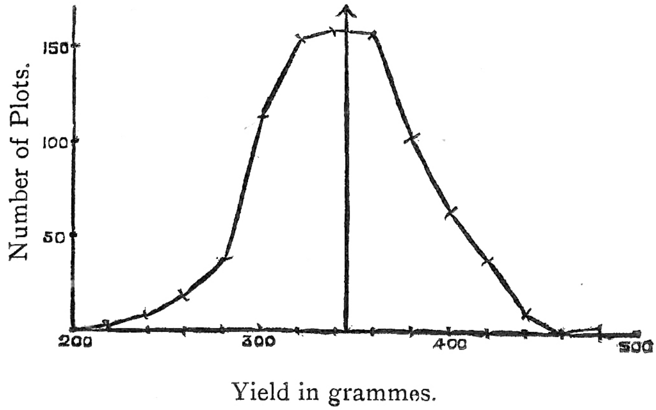

Fig. 1.—Variation in yield of 900 small plots of paddy.

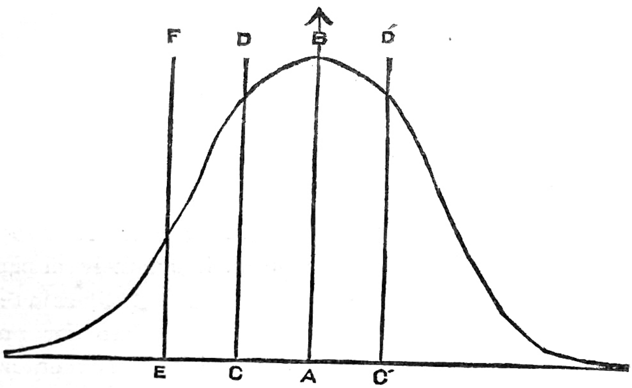

Fig 2.—Typical frequency curve.

curve for the variations in yield of grain from about 900 very sm plots of paddy of equal area. The horizontal line shows yield grammes and the vertical height of any point on the curve represen the number of plots giving that particular yield. You will notice th there are comparatively few plots of very low yield, that as the yiel increases the number of plots increases to a maximum, at about th average, from which point it decreases as the yield increases beyon the average. In other words, the nearer the yield approaches th average the greater will be the number of plots and *vice versa.*

This is the ordinary type of simple variation curve that is obtained where the same set of conditions affects all the determinations. If a very large number of determinations are made, and in the absence of any disturbing factor, the curve will be symmetrical about a vertical line through the average. Such a symmetrical curve is shown diagrammatically in Figure 2. The vertical height of any point on the curve represents the number of individuals for that particular value; therefore the area between the curve and the base line, which is proportional to the sum of all the vertical heights, may be taken to represent the total number of individuals. Similarly the areas falling on either side of any vertical line will represent the numbers of individuals falling on either side of whatever value that line represents. Thus the median line AB, through the average, separates those which are less than the average from those greater. These two groups are equal in number and the respective areas are equal. Now suppose a line CD is drawn dividing the left half of the area into equal parts ; as the curve is symmetrical a corresponding line C'D', at the same distance from the median, will divide the right half into equal parts in the same manner. The area between CD and C'D' is then half the total area, and is equal to the sum of the two triangular areas outside these two lines. In other words half the total number of individuals will give values lying between two points C and C', at equal distances below and above the average respectively, and the remaining half will be further removed from the average. It is thus an even chance that any individual will fall inside or outside the area between CD and

C'D', or in other words that it will differ from the average by less or more than the value represented by CA. This value CA is called the *probable error* of the experiment; it has nothing to do with the average error but simply means that, on the whole, half the individuals will differ from the average by more and half by less than that amount.

It must be noted that the difference from the average may be on either side. If we consider the individuals which are *greater* than the average by more than the probable error we see that they occupy only the triangular area on the right, cut off by C'D', that is one-fourth of the total numbering as compared with three-fourths for the remainder. Thus the odds are 3:1 that any result will not *exceed* the average by more than the probable error. Similarly the odds are 3:1 that any result will not be less than the average by more than the probable error.

Since this is a well-known type of curve, represented by a definite mathematical formula, it is possible to calculate the relative areas of the two parts cut off by a vertical line in any given position. This means that it is possible to determine the chances of obtaining any degree of variation from the average. If the variation from the average is put into terms of the probable error the odds will be the same for all curves of this type irrespective of the actual value of the probable error in each case. In Figure 2 the distance EA is twice CA, *i. e.*, twice the probable error. It is found mathematically that the small triangle to the left of EF is one-tenth of the remainder, thus the odds are 10:1 that any result will not be less than the average by more than twice the probable error. Table I shows the odds against any result being less or greater than the average by various multiples of the probable error. It will be noticed that the odds increase very rapidly as the factor increases. A glance at the curve will show this since the triangular area obviously decreases rapidly as the vertical line rece des from the average.

Table I.

Difference from average in terms of PE.	Odds against such difference occuring.
1·00	3 to 1
2·00	10 ,, 1
2·48	20 ,, 1
2·70	30 ,, 1
3·00	44 ,, 1
3·44	100 ,, 1
4·00	290 ,, 1
5·00	2,700 ,, 1

Thus if the probable error of a given type of experiment is known the results can be worked out in definite probabilities. For determining the probable error definite experiments are necessary, except in cases where the type of experiment has been carried out many times before and the results obtained from duplicates are available. The method is fully described in a paper by Wood and Stratton [1] but for the convenience of those who have not read this paper I will give, very briefly, an account of the method with an example.

Suppose a number of varieties, or strains, are to be compared on 5 cent plots and it is required to find the probable error of the experiment. A number of plots are put down with any one variety, under the conditions that will be maintained in the experiment. The yields, though they are all from the same variety, show variation due to the experimental error involved. The yields are averaged and the difference of each plot from the average is noted. The differences are squared and from the sum of the squares, $\sum d^2$, the probable error is calculated by the following formula :—

$$PE = 0{\cdot}67\sqrt{\frac{\sum d^2}{n-1}}$$

where 'n' is the number of plots taken and PE is the probable error of one plot compared with average. This is a well-known formula, embodying what is known as the method of *least squares*, and it is unnecessary here to go into the mathematical principles involved.

1 Wood, T. B., and Stratton, F. J. M., Jl. Agr. Science, Vol. III [1910], P. 415.

As an example of the method I will give last year's yields of paddy from the standardization plots on block N of the Central Farm. These plots were laid down for another purpose but they are also useful for this estimation of PE. There were 20 plots, varying in area from 3·4—7·9 cents, planted and treated uniformly. The yield per acre of the average is put at 100 and the individual plots calculated accordingly. The figures, together with the calculation of the PE, are given below :—

Plot No.	Yield.	Difference from average.	d^2
1 a	116·4	16·4	269
,, b	97·8	2·2	5
,, c	84·2	15·8	250
,, d	89·0	11·0	121
,, e	83·7	16·3	266
2 a	111·5	11·5	132
,, b	97·1	2·9	8
,, c	113·7	13·7	188
,, d	87·0	13·0	169
,, e	96·0	4·0	16
3 a	115·1	15·1	228
,, b	84·5	15·5	240
4 a	138·5	38·5	1482
,, b	104·1	4·1	17
5 a	101·1	1·1	1
,, b	97·7	2·3	5
6 a	104·0	4·0	16
,, b	96·6	3·4	12
7 a	93·1	6·9	48
,, b	88·5	11·5	132
average.	100	$d^2=$	3605

$$PE=0·67$$

$$=0·67\sqrt{\frac{3605}{19}}$$

$$=\pm 9·2$$

The PE works out at $\pm$9.2 on an average of 100, i. e., $\pm$9·2%. As already explained this means that it is an even chance that any plot, taken at random, will differ from the average by *more* or *less*, respectively, than 9·2%. A glance at the figures shows that out of 20 plots 11 differ by more than 9·2 per cent and 9 by less. This is as nearly equal as could be expected from so small a number as 20 plots. The larger the number of plots the greater will be the accuracy of the result: about 50 plots will give a result sufficiently accurate for ordinary purposes.

In an ordinary field experiment the comparison made is between two plots and for this the PE is $\sqrt{2}$ times that for comparing any one plot with the average. Thus the PE of comparing any two of the above plots is $\sqrt{2} \times 9{\cdot}2$ or 13 per cent. That is to say, that, if plots similar to these are used to compare different treatments or varieties, a difference of 13 per cent. between any two plots is just as likely as not to be due to experimental error and is absolutely no indication of difference due to treatment or variety. If a difference of twice the PE is obtaiaed, i. e., 26 per cent., the chances are 10:1 that this means a real difference. It might be thought at first sight that a 10:1 chance is good enough to rely on but it is undoubtly too low if final conclusions are to be drawn. I would recommend to anyone interested in this subject the paper, already referred to, by Professor Wood of the Cambridge School of Agriculture. He puts a probability of 30:1 as about the least that can reasonably be relied on. In this series of plots it would be necessary to obtain a difference of 35 per cent. for a probability of 30:1. Now 35 per cent. is a very big difference, that could only be expected in exceptional cases, and it would be the height of folly to lay down experiments that wonld not show a difference of less than this. It is true that such a large difference does sometimes occur but, unfortunately, not very often, and the result of laying down experiments with a high PE is, that no definite conclusions can be drawn, as may be seen from many of the Farm reports.

In order to make use of the above plots it would be necessary to repeat each variety over several plots, not in one block but alternating, so that the different varieties might be spread evenly over the whole area. Such repetition reduces the PE; the PE of the average of *n* plots equals the PE of one plot divided by $\sqrt{n}$, thus four repetitions would halve the PE. It would need about nine repetitions, giving one-third of the PE, i. e., about 12%, to make the above plots of any practical value for experimental work.

Repetition of the same experiment for several years reduces the PE in the same way as repetion over several plots. It has the advantage of avoiding results that might be due to special weather conditions in any one year. The time taken to arrive at definite conclusions is correspondingly longer as a general rule, it is desirable to realy as little as possible on such repetition.

I will now give a few examples of the PE shown by actual experiments on paddy of which the results are given in various Farm reports. The figures which are given in Table II are based on the differences between duplicate plots. In addition to these are included the results from the Central Farm standardization plots and the special PE plots that were put down at Manganallur last year in connexion with my selection work.

Table II shows the area of the plots, the percentage PE of comparing any two plots, and the percentage difference that is necessary to ensure a 30:1 chance that the difference is not due to experimental error.

Table II.

Farm.	Experiment.	Area of plot in cents.	%P. E. difference between two.	% difference for 30:1 odds.
Central Farm	Varieties	10–15	10·0	27·0
,,	Patti mannu	6–11	13·4	36·2
,,	Age seedlings	7·3–11	13·6	36·7
Palur.	Manurials	12·5	11·3	30·5
Manganallur	Phosphates	5–7·5	12·6	34·5
,,	Ottadan in Udu Manurials	4	13·1	35·3

Samalkota	Cyanamide	5–8	13·9	37·4
,,	Spacing	6–12	7·7	20·8
Central Farm	N. block standardization.	3·4–7·9	13·0	35·0
,,	B. block ,,	10–15	6·0	16·2
Manganallur	P. E. strips	0·6	5·8	15·7
,,	,,	1·2	5·3	14·3
,,	,,	2·4	5·7	15·4
,,	,,	4×0·6	2·3	6·2

It will be seen from the figures that there is a fairly close agreement between the values of the PE for ordinary farm plots. It varies from about 11-14 percent for comparing two plots, corresponding to a difference of 30-40 per cent for 30:1 odds. The Central Farm varieties experiment, which gives a PE of 10 per cent., no doubt, owes its increased accuracy to the fact that the plots were in the form of long strips across several fields. The Samalkota spacing experiment giving 7·7 per cent. PE, is the most accurate Farm experiment that has yet been worked out. There is no clear indication that the size of the plots, in so far as they vary in these experiments, has any appreciable influence on the PE.

Of the standardization plots on the Central Farm those on N. Block, which are ordinary farm plots, give about the same PE as such plots in actual experiments. The plots in B block give the very low figure of 6 per cent. and this is undoubtedly due to the fact that they are very long narrow plots on specially even land that has been uniformly cropped for many years.

The PE plots at Manganallur were specially laid down with a view to determining the best type of plot for comparing large numbers of strains arising from my selection work there.

The first three series were in the form of narrow strips 120 links long and 5, 10 and 20 links wide respectively, their areas being 0·6, 1·2 and 2·4 cents. It will be seen that the PE is less than half that of the majority of the ordinary Farm experiments and is about the same as that of the B block standardization plots which were in the form of long narrow plots though of much larger area. The three widths have all given about the same PE showing that it is unnecessary to increase the area by widening the strips. In the last series, however, where the area is increased by the *repetition* of 5 link strips a great reduction of the PE is seen. The area 4×0·6 indicates that four strips of 5 links were taken together by repeating a set of strips four times.

These plots show very clearly the increase in accuracy obtained by using very narrow strips as compared with ordinary plots that are more nearly square in shape. They also illustrate the value of repeating such small strips in an alternating manner instead of inereasing the area by making the strips wider. Using only 2·4 cents of land for each strain or variety it is possible to reduce the PE to about one-fifth of that which is usual for ordinary plots of 5-10 cents. This system has now been adopted on the Paddy Breeding Station where it is specially suitable for comparing large numbers of strains and varieties. It cannot, of course, be adopted without modification for such work as manurial experiments where extremely narrow strips are impracticable.

There is a special point to be noticed in connexion with trials of various manures. Not only is it necessary to decide which gives the best crop; there is also the question of the cost of manuring. Thus it is necessary to decide whether any given manure gives a sufficiently large increase to more than pay for the cost of the manure. This requiries a very much more accurate experiment than in a case where it is only necessary to decide which is the best of several strains or varities.

Now it is possible that some of you may say that all this merely shows that a certain amount of common sense and judgment are required in interpreting results. They are required to the degree of being absolutely necessary, but, in addition, a knowledge of the probable error of the experiment is also an absolute necessity. Suppose in any experiment the duplicates accidentally disagree markedly:

common sense says the results are useless since the treatment cannot be responsible. Suppose again that the duplicates agree moderately well but the treatments show a fair amount of difference: common sense will probably say that this difference is obviously due to the treatment, though an actual knowledge of the PE might show that the difference is just as likely to be accidental as the difference between duplicates is certain to be. In other words common sense alone may show that a result is uncertain, but,, without accurate knowledge, it cannot show that a result is certain.

F. R. Parnell.

Discussion :—

This paper was freely illustrated by diagrams and charts which greatly helped the audience in following the paper. Mr. G. A. D. Sturat said that the paper threw a new light on the method of conducting Field experiments. In his capacity as the Director of Agiculture he often had to prepare short popular summaries of the Scientific reports from the various Agricultnral Stations and he used to find difficulties in presenting results which were not sufficiently conclusive. He asked the audience to remember that even if a small but real increase could be made in the out turn of a crop like paddy the benefits accuring therefrom will be great. He recently made a calculation and found that in paddy an increase in yield of even 1% will put into the pockets of the ryots an additional sum of 50 lakhs of Rupees for this Presidency alone. Mr. R. Thomas remarked that it will take a very long time indeed if they were to wait for the experimental error to be worked out for each individual crop. All that was needed was to select strains which showed fairly large differences and the value of the method is amply vindicated by the success which has attended Mr. Howard's work in Wheat and Mr. Clouston's work in cottons. In the District Farms some work is done on the standardization of plots but the results are not published. The cotton sfrains at Koilpatti were selected for differences of over 30 per cent.