A REVIEW OF THE APPLICATION OF STATISTICAL THEORY TO AGRICULTURAL FIELD EXPERIMENTS IN INDIA*

BY

P. C. MAHALANOBIS

Statistical Laboratory, Presidency College, Calcutta

I. General review of the present position

A REVIEW of the application of statistical theory to agricultural field trials in India in recent years is largely a story of the triumph of methods devised by R. A. Fisher at the Rothamsted Experimental Station. The new developments also bear remarkable testimony to the scientific vision of Sir John Russell, Director of the Rothamsted Experimental Station, who had recognized as early as 1919 the need of the application of statistical theory to agricultural research, and had not only persuaded Fisher to take up this subject but had given him full scope and freedom for working out appropriate statistical methods in his own way.

The basic principles of the new method are now well known and need not be discussed in detail. In order to appreciate the revolutionary advance brought about by the introduction of the new technique, let us however consider for a moment the contrast between experiments of the old and new type.

The old type of field experiment

Suppose we wish to compare the yield of say six varieties or the effect on yield of six kinds of manures. In the old type of experiment the field would be divided into six plots, and a single plot would be allotted to each treatment. As Fisher [1935] explains 'the treatment giving the highest yield would of course appear to be best, but no one could say whether the plot would not in fact have yielded as well under some or all of the other treatments'. It is known that within the same field wide differences exist in the fertility of the soil. Even when the soil fertility is uniform, there are innumerable other causes which affect the yield. How can we be sure that the observed differences in yield are due to the difference in the treatments, and not to soil heterogeneity ? How can we be sure that they are not due to chance fluctuations ? This is the basic problem. In order to solve it we must eliminate the effect of soil heterogeneity, and make an unbiased estimate of the magnitude of errors due to chance so that we may be sure that the observed effect is significant in comparison with the size of such chance errors.

The Fisherian technique

Let us now see how Fisher solved the problem. Consider the same experimental field which had been originally divided into six portions.

* Presented before the meeting of the Board of Agriculture and Animal Husbandry in India held in Lahore in December 1937

Fisher simply further sub-divided each portion* into a number of plots of smaller size; and within each portion (or block as he called it) he assigned one plot to each treatment but strictly in a random manner. We have now the randomized block in its modern form. Using the principle of block division in two directions symmetrically we get the well-known Latin square.

Results governed by laws of chance

The important point to be noticed is that the results will be now governed entirely by the laws of chance. There are innumerable causes which produce differences between the plots, and we know from the conditions of the experiment that it is impossible in practice to secure that the plots will be all alike. But the validity of the estimate of error is now guaranteed by the process of randomization, namely 'the provision that any two plots, not in the same block, shall have the same probability of being treated alike, and the same probability of being treated differently in each of the ways in which this is possible' [Fisher, 1935]. The calculus of probability and the apparatus of the statistical theory of sampling distribution can now be used with complete confidence. The logical foundations of scientific inference were thus made secure, and agricultural experiments were placed for the first time on the same footing as experiments in other sciences. In actual fact the statistical theory of exact distribution in terms only of actual observations, popularly known as distributions in 'Studentized form', achieved a good deal more. It made possible general conclusions being drawn with logical rigour from particular observations. But this is a topic of statistical rather than agricultural interest and must be passed over here.

Elimination of soil differences

The second point to be observed is that by the technique of block division the problem of soil heterogeneity was solved at the same time. As each block contains all the treatments once and once only, differences between the total yields of the different blocks could safely be ascribed, apart from errors of sampling, to soil differences; and could be eliminated by suitable statistical methods. This of course led to a great improvement in the precision of the comparisons. When we remember that in particular experiments in India as much as 90 per cent of the total variation is sometimes caused by soil differences, the importance of eliminating its effect will be easily appreciated.

The analysis of variance

The third point to be emphasized is the close connexion between the field procedure and the procedure of statistical analysis in the Fisherian technique. In fact they are merely two aspects of the same problem; and to quote Fisher [1933] 'once the practical field procedure was fixed, only a single method of statistical analysis could be valid. . . The specification of the particular process of randomization carried out, determined in advance the correct statistical analysis of the results'.

*I need scarcely add that the experimental field may be divided into any number of convenient portions each of which is further sub-divided into a number of plots.

To sum up then, replication, randomization, and block division (or local control) were the principles of design introduced by Fisher [1923] at Rothamsted. Replication is essential because it is the sole source of the estimate of error, while randomization is necessary to guarantee the validity of the estimate, i.e. to ensure that the estimate will be unbiased. The purpose of block division is to increase the precision of the comparisons by elimination of soil differences, while replication is also useful in securing the same object by diminishing the experimental error. Finally the analysis of variance* gives a convenient and valid method of extracting the information contained in the observations. As Wishart has pointed out, the Fisherian technique 'was something in the nature of a revolution,' and altered the subsequent course of agricultural experiments throughout the world.

Previous conditions in India

It took some time before the new technique was introduced into this country. Seven or eight years ago in India the control used to be almost always repeated, but the treatments were usually laid down without replication. Even when replication was used, it was of the systematic type and inadequate in number. In interpreting the results, the usual practice was to compare the means of the various treatments. In a few cases probable errors of means were calculated. The ordinary formula in the classical theory of errors was used for this purpose. This was inexact for the twofold reason that the observed variance was substituted for the corresponding population value, and the effect of using small samples was ignored.† Besides in the absence of randomization, such estimates were not unbiased, and could not be validly used for purposes of comparisons. Finally there was no attempt to eliminate the effect of soil differences. It is no wonder therefore that many of the inferences drawn from the old experiments were unreliable. Even when the results were true, this could not be asserted with scientific precision. A fair idea of these old-type experiments, which used to be conducted in India a few years back, can be obtained from 'Analysis of Manurial Experiments in India' by Vaidyanathan [1934].

Introduction of the new technique in India

Like most other recent movements in agriculture in India, we owe the introduction of statistical methods to the Royal Commission on Agriculture (report, pages 617-8), which had made definite recommendations on this point in 1928. In actual practice the modern period of field experiments began in India, I believe, with the foundation of the Imperial Council of Agricultural Research in 1929 on the recommendation of the Royal Commission.

* See note on 'variance', 'standard error', 'covariance', etc. (Appendix II)

† It is of some personal interest to me to recall here that to this particular problem I owe my contact with agricultural work. In 1924 my attention was drawn by Dr. W. Burns (Agricultural Commissioner with the Government of India), then working in Bombay, to an experiment in which six varieties of rice were laid out in ten replicates systematically arranged side by side in long stripes. On the assumption of a systematic variation in soil fertility, it was possible to eliminate the soil differences by graduation, and it was found that the precision of the comparisons could be considerably increased. At that time I was quite unfamiliar with the Rothamsted work, but Dr. Burns' problem soon made me get acquainted with the Fisherian method, and made me realize its great value.

The earliest experiment of the new type, a varietal trial on rice with a 12 × 12 Latin square was reported in the *Indian Journal of Agricultural Science* in 1931. The Imperial Council of Agricultural Research from its inception laid emphasis on statistical methods, created a statistical section at headquarters with a whole-time statistician at its head, and gave a grant to the Statistical Laboratory, Calcutta, for advanced studies and researches in statistics. In fact I believe it was soon made a condition of all Imperial Council of Agricultural Research schemes that the experimental designs should be of the approved type. The Statistician to the Imperial Council of Agricultural Research gives his advice on all standard schemes in the province, and personally visits a large number of farms every year. Help is also available, especially on the research side, from the Calcutta Statistical Laboratory. In the course of this work a series of 'Statistical Notes for Agricultural Workers' was started of which 24 numbers have been published so far. In 1932 arrangements were made in Calcutta for giving special courses of instruction in statistical methods to officers who were sent there on deputation for this purpose. During the last five years such training has been given to over 75 agricultural officers from all over India, which, I believe, has helped materially in raising the general standard of work. The lead given by the Imperial Council of Agricultural Research in all these ways has resulted in the Latin square and randomized block designs being used with great success all over India. It is probably no exaggeration to say that no important experiment in India is now laid out on an old type design. This must be considered to be a solid achievement.

Factorial (complex) experiments

We may now consider some further developments of the new techniques. As early as 1926 Fisher had advocated the use of factorial designs in which two or more types of treatments were laid out on the same field.

Suppose we wish to compare three varieties, and the effect of three manurial treatments on each of these varieties. If we conduct the experiments separately, and use six replications, we shall require for the varietal trial 3 × 6 = 18 plots. For the manurial portion we shall require three experiments, dealing respectively with the three varieties. With six replications we shall therefore require 54 plots for the manurial investigations and 18 plots for the varietal comparison or 72 plots altogether.

Instead of simple experiments, suppose we combine them in one factorial (or complex) design. First of all, for nine combinations (3 varieties × 3 manures) we can then afford to give eight (instead of six) replications each in the same field of 72 plots Secondly, we shall have no less than 24 replications available for the varietal or manurial comparisons; so that, if the standard error per plot remains the same, the accuracy of the main comparisons will be increased four times. Finally, the three manurial treatments cannot be directly compared in the separate experiments; but in the factorial design the comparisons would be completely valid. In other words, the differential manurial requirement of particular varieties, i.e. the interaction between varieties and manures, if any, can be investigated only if the experiment is designed in the factorial form. With three or four factors the amount of information obtained is proportionately even greater. Besides the main effects we can not only study the differential effect (or interaction) of the factors two by two, but also the response of one factor in the presence or absence of two or more of the other factors.

A factorial experiment is thus not only more efficient in the sense that with the same number of plots all the factors can be studied with greater precision, but is also more comprehensive and will give information about differential response which could not possibly have been obtained by any number of experiments of the simple type. This is why Fisher [1926] definitely rejected the orthodox principle of varying the factors only one at a time and said : 'No aphorism is more frequently repeated in connexion with field trials than that we must ask Nature few questions, or ideally, one question at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire ; indeed, if we ask her a single question, she will often refuse to answer until some other topic is answered.'

Before leaving this topic it is perhaps worth while pointing out a third advantage of the factorial design. In the orthodox method all the factors except one are deliberately kept approximately constant. In the result, information is obtained only for a narrow range of controlled conditions. In the factorial design on the other hand a number of factors are allowed to vary at the same time, so that conclusions drawn from such an experiment have a much wider basis for induction.

In India the first factorial experiment with three varieties of potato under three manurial treatments was laid down at the instance of the Statistical Laboratory at the Visvabharati Institute of Rural Reconstruction at Sriniketan in 1931. During the last four or five years similar two-factor experiments have become quite common all over India. Designs with three or four factors are also being used with success. As an example I may mention the four-factor cultivation experiment with rice (three varieties, five dates of planting, three spacings, and three numbers of seedlings per hole) designed at the Statistical Laboratory and conducted under the Imperial Council of Agricultural Research rice research scheme at Chinsurah for the four seasons 1933-37. The summary of results shown in Appendix I will give some idea of the wealth of information which can be obtained only from designs of this type.

In spite of their efficiency and comprehensiveness certain objections have been raised against the use of factorial designs which may be briefly considered here. It has been pointed out that the main effects are obtained with greater precision than the interactions ; also that the experiment includes many combinations which are never likely to be used in practice. This is quite true but inevitable. When we have no knowledge as to what particular combinations are likely to be useful, it is desirable that we should seek to survey the whole range of all the factors. But an extensive field of survey inevitably implies a lower level of accuracy. However as experience is gathered, the field of enquiry can be narrowed by reducing the number of combinations, with an automatic increase in the precision.

A second objection is more serious. With an increase in the number of combinations, the size of the block becomes too large for adequate elimination of soil-heterogeneity with consequent increase in the residual error. The difficulty has been admirably got over recently by the 'splitting of plots' and the 'confounding of interactions'

Split-plot and 'confounded' designs

In the factorial design complete information about all the combinations can be obtained at the cost of accuracy. We can, however, increase the precision by sacrificing a portion of the information. This is just what is achieved in the 'confounded' design. The whole array of treatment combinations is therefore not included in the same block, but deliberately distributed over two or more balanced sub-blocks. Experience has shown that high-order interactions are often insignificant, or even when statistically significant are not of much practical importance. In the confounded design information about such high-order interactions is usually sacrificed to increase the precision of other comparisons. If we like we can, however, arrange to obtain some information about all the interactions, but inevitably at a lower level of precision, by 'partial confounding'.

The split-plot lay-out is a simple example of confounding in which the main effects of one of the factors are confounded. This design is particularly useful when some of the treatments are such that they cannot be conveniently applied to small plots. These main treatments are therefore laid out in a randomized block or Latin square design, but each whole-plot is divided into a number of sub-plots which are allotted at random to the different sub-treatments. The residual variance between sub-plots gives the appropriate error for the comparison of sub-treatments, while the residual variance between plots gives the error for the whole-plot treatments.

The split-plot design is being extensively used in India, but the confounded design has so far not attracted much notice.* As far as I know, one elegant design prepared by Yates has been laid down at the Tocklai Tea Experimental Station, and one design has been supplied by the Statistical Laboratory to the Dacca University for the Imperial Council of Agricultural Research scheme.

The designing of confounded lay outs is an interesting exercise, and in skilled hands it has attained a high degree of efficiency. I would draw the attention of all agricultural workers interested in this subject to the discussion by Fisher [1935]. Yates [1937] and that in the *Rothamsted Reports* for the last few years.

Complete factorial designs, as we have seen, are both efficient and comprehensive. But they need great care at every stage of the work, and with a large number of factors require blocks which are inconveniently large in practice. There is, therefore, a limit to the usefulness of this type of design depending on the heterogeneity of the land, the number of factors and nature of the problem, the skill and experience of the investigator, etc. The split-plot design is very convenient in problems in which knowledge about the main treatments is already available. But I am of opinion that it is the confounded design which has the greatest possibilities in India, both on account of its flexibility as well as its economy of cost. Caution is needed, however, both

* This review was originally written in December 1937. Since then the principle of confounding is being increasingly used in India.

in designing the experiment and in carrying out the statistical analysis. In the beginning, therefore, it will be desirable to use standard patterns under the guidance of statistical workers.

Interpretation of results

Before leaving this subject I would like to add a few words in regard to the interpretation of the results. I have found that many agricultural workers are able to reduce the data correctly and complete the arithmetical part of the analysis of variance without, however, being able to draw the necessary inferences. 'Significance' and 'non-significance' are purely technical terms with the exact implication of which every experimenter should be familiar.

Suppose we are working on the five per cent level of significance. Then the rule is that any effect which is likely to occur by pure chance less than once in twenty times on an average will be called 'significant'. On the other hand, effects which are likely to occur more frequently than once in twenty trials will be called 'non-significant'. Let us see the application of this rule in a concrete case. Suppose we have an experiment in which the treatments do not in fact produce any effect. Even then, with the present rule, the effect will appear to be significant about once in twenty trials, and in the remaining 95 per cent of cases we shall quite correctly decide the effect to be nil. The risk of considering an effect to be real, when in fact it does not exist is thus limited to just five per cent. Similarly working at one per cent level of significance, we limit the risk of our accepting a spurious effect as real to one per cent. In other words we work with odds of 99 to 1 in our favour.

I may point out at this stage a peculiar property of statistical inference. Suppose we are working on five per cent level. We have seen that even when the effect is nil, we shall judge it to be real once in 20 trials. In other words, if statistical theory is right, we must be wrong in our judgment in five per cent of the cases. The possibility, or rather, the certainty of error is thus inherent in the structure of statistical inference. This knowledge is a salutary check against an exaggerated sense of our own infallibility.

The experimenter must, therefore, be careful in attaching undue importance to an isolated result which may appear to be statistically significant and yet does not fit in with general agricultural experience. Such results should not be ignored, but should neither be accepted until corroborated by further experiments. On the other hand, results statistically insignificant should not be always neglected. If they appear to be plausible from other considerations, further investigations should be made with increased precision of comparison.

In short, the experimenter must use his critical judgment and discretion in the final interpretation of the results. Statistics is both indispensable and invaluable, but it cannot replace the human mind.

Precision of Indian experiments

Having reviewed the broader features of the new technique, it will be of some interest to examine the precision attained in Indian experiments. I am sorry, in the limited time at my disposal, I was unable to collect relevant information from the different provinces of India. I shall, therefore, discuss this point with the help of materials from Bengal and Assam which were readily available in our Laboratory.

The average standard errors per plot (expressed as percentages of mean yields) for four or five careful series of varietal trials with *aus* and *aman* rice at Chinsurah Farm during the five seasons 1932-33 to 1936-37 are shown below. (The figures within brackets give the number of experiments on which the average is based.)

Bengal: Chinsurah Farm varietal tests

(Standard error per plot as percentage of mean)

Year.	Rice crop	
	Aus	*Aman*
1932-33	11·61 (3)	10·10 (5)
1933-34	9·68 (2)	10·06 (5)
1934-35	8·38 (3)	10·21 (6)
1935-36	..	19·10 (5)
1936-37	9·19 (2)	9·74 (5)

Similar figures for recent rice and sugarcane experiments in Assam for the period 1932-33 to 1935-36 are given below.

Assam

(Standard error per plot as percentage of mean)

Centre	Crop	Variety	Manure	Complex
Karimganj	Rice	6·78 (24)	8·00 (4)	8·25 (6)
Titabar	Rice	10·50 (20)	7·98 (4)	..
Jorhat	Sugarcane	8·21 (9)	7·71 (7)	15·40 (4)

It will be noticed that in Bengal and Assam in the case of rice and sugarcane, a standard error of 8 or 10 per cent per plot is quite usual.

Comparative figures for English experiments are quoted below from the Report of the Rothamsted Experimental Station for 1935.

English stations

Crop	Latin square	Randomized block	All arrangements
Potato	6·8	9·2	..
Sugarbeet	6·1	7·9	..
Swedes	..	..	6·9
Mangolds	..	..	8·2
Kale	..	..	7·7

Wishart and Sanders [1936], are of opinion that a standard error of 5 per cent for root crops and of 10 per cent for cereals may be considered satisfactory. Judged by English standards, the work in Assam and Bengal is therefore not unsatisfactory. I have no reason to think that careful work in other parts of India is in any way less accurate.

Latin square v. *randomized block*

Owing to the possibility of eliminating soil differences in two directions, one would expect the Latin square to be more accurate than the randomized block, and English experience has generally borne this out. I am not in possession of enough data to judge the position in India. My general impression is that the Latin square has been given preference here in small-scale varietal work. For large-scale work, I think on the whole the randomized block has been used more extensively in India, no doubt on account of its greater flexibility. One advantage of the randomized block is that an estimate of error can be calculated separately for each comparison. Yates [1935] has pointed out that 'this is of great value when handling new and unknown material, or treatments which may produce large differences and even partial or complete failures. In such cases the assumption of constancy of error variance is entirely unjustified, but in a randomized block experiment any treatment or treatments may be excluded and the analysis carried out on the remainder. This is not true of either the Latin square or of confounded arrangements'.

Complex or factorial designs in India apparently have a slightly higher standard error per plot (of the order of 10 or 15 per cent of mean yield) than the simple Latin square or randomized block. This is probably due to the experimental difficulties in managing more than one set of factors on the same plot and to large block sizes.

Uniformity trials

In a randomized block design the greater the homogeneity of plots within blocks the greater the accuracy of the experiment. In practice this can be secured experimentally only to a limited extent. But sometimes it is possible to increase the precision of comparison very considerably by suitable statistical adjustments. Suppose, for example, that the initial fertility of plots is known from a previous uniformity trial in which the same variety is planted on all the plots and given the same manuring, and the relative fertility of individual plots remains fairly stable; then the yields in a succeeding year will be appreciably correlated with the yields in the uniformity trial. In this situation it is possible, with the help of the analysis of covariance, to make allowances for the initial differences in fertility among plots within each block. The use of such adjusted yields can then be used for the final comparison. This method has been sometimes known to have increased the precision even ten or twelve times.

It should not be imagined, however, that this is always or even generally possible. In fact with annual crops the fluctuations in fertility of the same plot from year to year are usually so great that the increase in precision obtained by this method is not in general commensurate with the expense or the delay of one season involved in a uniformity trial.

It is therefore usually unprofitable to conduct a uniformity trial as a preliminary to the main experiment with a view to increasing its precision. Repeating the actual experiment twice would most often give more information. This is why we have for a long time discouraged the adoption of a uniformity trial as a routine practice. It may be noted, however, that there are special circumstances in which such trials may be very useful indeed, for example in the case of horticultural experiments.

Size and shape of plots

We may now consider the question of size and shape of plots. As early as 1910, Hall harvested a wheat and a mangold field in small units, and found that the variation between plots was appreciably reduced until the size reached was about 1/40 acre. It was therefore concluded, and the conclusion was corroborated by other uniformity trials, that the optimum size in England was somewhere about plots of 1/40 acre.

We have had occasion to examine the results of a number of uniformity trials in India, and we found that for varietal trials in many cases the plot could be reduced, so far as precision is concerned, to a very small size of the order of 1/140 acre. Plots of size 1/80 acre or 1/40 acre also give quite good results and can be safely recommended for convenience of agricultural operations. Given the area, the question of shape or orientation comes in. Christidis [1931] showed from theoretical considerations as well as experimental data that long plots placed parallel to the fertility gradient gave the best results. Our experience in India is also more or less similar. Sugarcane experiments at Pusa and other places in North Bihar show that strips, the length of which is 10 or 15 times greater than the width, give more accurate results. Rice experiments show the same tendency but to a smaller extent.

Size of blocks

The final precision of an experiment does not depend only on the best selection of plot size. What is needed is a choice of the optimum combination of the size of both blocks and plots. The best results will be obtained when the blocks are fairly homogeneous, (i.e. all the plots within the same block have nearly the same fertility), but differ appreciably as a whole between themselves. It is obviously not possible to give any limits for the block size. If the soil is fairly uniform, it is possible to work with blocks of a large size; on the other hand, if the fertility gradients are steep, the size of the blocks must be kept small. I had the opportunity of studying in detail the variation in soil fertility of a field of about one acre under rice at the Chinsurah Farm, which was harvested at my request in 7040 units of 9 inches by 90 inches (1/7744 acre). We tried many combinations of block and plot sizes, and found that a low standard error of about 3·5 per cent of mean yield per plot was obtained with a block size of 80 ft. × 44 ft. (about 1/12 acre) with 8 plots each of size 20 ft. × 22 ft. (1/100 acre). But considerably larger size of blocks 160 ft. × 44 ft. (or about 1/6 acre), or 160 ft. × 88 ft. (1/3 acre) with 8 or 16 plots each could be used with only a moderate increase in the error to about 5 per cent of mean yield per plot.

Success of the new technique

From the brief review given above, I think it can be stated without hesitation that in India wherever the Fisherian technique has been used on proper lines in field trials, it has been found entirely satisfactory in every way and has given excellent results. The working procedure is very flexible so that it can be adapted to suit the most diverse problems and conditions of work.

A good deal of valuable information regarding soil differences and the relative accuracy of different types of experimental designs is also fast accumulating in India. It is desirable in designing a new experiment that each experimenter should utilize all available information relating to his own work. In this way he would be often able to get a good idea of the type of design likely to give the best results, and also to safeguard himself against too large a margin of error by using an adequate number of replications or other methods of controlling soil differences.

Concomitant variations and correlational analysis

I have already considered the use of uniformity trials, and I may now briefly refer to certain other methods of increasing the precision of field trials by using concomitant measurements and the analysis of covariance. The underlying principle is simple. In a field trial there are many other factors besides yield which can be studied, and it often happens that some of these factors are correlated with the yield in the sense that variations in such factors cause (or are associated with) variations in the yield. It then becomes possible to separate and eliminate that portion of the variation in the yield which may be ascribed to these factors. In this way the precision of the experiment can be often increased very considerably. For example, it may happen in a field trial that the yields of different plots are disturbed by variations in the number of plants which have established themselves. When such disturbances are due to causes which have no connexion with the treatments under trial, it is clear that there can be no objection to making allowances for such variations. In the present example, by counting the number of plants in the different plots we can easily eliminate the variations in yield due to variations in plant number, and hence increase the precision of the experiment.

Similar methods may be used for eliminating the influence of varying intensity of attacks of pests and insects in different plots. It would be most undesirable to reject some of the yields simply because they appear to be too low. As Wishart and Sanders [1936] have remarked, 'once a start is made in rejecting actual figures, there is no knowing where to stop . . . . a little skill in the game will lead to very significant, but quite untrustworthy results. There is no wish to impugn the reader's honesty, but no man is so virtuous that he can afford to treat temptation with disdain'. The position would be quite different if some observations were recorded on the intensity of the insect attack in the different plots before the crop is harvested. Such records can then be used for making adjustments without bias.

The method of correlation or analysis of covariance can also be used with great advantage in other ways. If records of growth of the plant (height, girth, tillering, etc.), are kept at different stages, such records can be correlated with the final yield, and may be utilized to furnish valuable information on many points. These methods deserve greater attention than they have received so far in India.

Missing yields of plots

Before leaving the subject of field trials I may refer briefly to another question which occasionally arises in practice. Owing to accidents or negligence on the part of the subordinate staff, it sometimes happens that the yields of one or more plots are missing or get mixed up. In such cases it is often possible to reconstruct approximately the missing yields by purely statistical methods, and thus recover much of the information which would have been otherwise thrown away. Formulæ for certain simple cases were given for this purpose originally by Allan and Wishart [1930] and a general solution was subsequently given by Yates [1933]. Additions to the theory have been made in the Statistical Laboratory and have been used with success for certain types of mistakes which had actually occurred in India. It cannot, however, be emphasized too much that such procedures are at best make-shift arrangements, and the damage done by careless work cannot be repaired by such methods. In any case these methods must be used with very great caution.

Use of random samples

The use of concomitant measurements usually involves a great deal of labour, which can be often reduced very considerably by adopting the method of random sampling. Consider an ordinary field trial. Suppose for any reason, such as shortage of labour or inclement weather or some other difficulty, it is found impracticable to measure the complete yield of each plot. In this situation we may take one or more random samples from each plot and measure the yield of these samples. Or consider the measurement of the height of plants at different stages of growth in the case of a field trial. For ordinary crops the number of individual plants in each plot is very large, and it is practically impossible to measure separately every plant in each plot. We may here take a random sample of the same number of plants in each plot, and measure only those plants which are included in these random samples. Sometimes complete enumeration is not only impracticable but even theoretically impossible. For example, if the dry weight of plants is sought to be studied at different stages of growth under different treatments, it is obvious that necessary measurements cannot possibly be carried on the same plants, but only on portions of the material under observation. In such situations there is no alternative but to have recourse to random sampling.

Fortunately, when used judiciously, this method is quite efficient, and the additional error introduced by this method is usually small. Thus, for instance, when the yields in a field trial are obtained by random sampling, the effective error variance will be simply the sum of the variance between plots and the variance, due to sampling. The latter being considerably smaller than the former, the increase in the variance will consequently be small. The method of random sampling* has great possiblities which should be more fully explored in India.

Other applications of statistical theory

I have considered field trials at some length as this is the main topic for discussion. But statistical methods have also been used with success in other types of work in this country some of which may be briefly mentioned at this stage.

The principle of randomized replication has been used in pot culture, animal nutrition studies, experiments on incidence of pests, horticultural experiments, etc. Very recently it has also been used in sylvicultural experiments at the Forest Research Institute at Dehra Dun.

Correlational analysis has been used in a number of investigations on the influence of rainfall and other weather conditions on crop out-turn, and although valuable results have been obtained, the scope of such studies has been unfortunately much restricted in India by the paucity of reliable crop data extending over a large number of years. A good deal of valuable work is being done, chiefly under the auspices of the Meteorological Department, in agricultural meteorology in which statistical methods are being used extensively. Modern statistical methods are being increasingly applied in linkage and genetic studies at the Indore Institute of Plant Industry and elsewhere.

More advanced statistical technique has been occasionally used for the investigation of special problems, such as detailed studies of frequency distributions of cotton fibre in the Cotton Technological Laboratory at Matunga; the use of composite tests of significance in plant physiological work; the use of quantitative measures of group divergence for the scientific classification of varieties, etc. In the limited times at our disposal it will not be possible to discuss with profit such recent developments in detail.

On the whole, it may be said that agricultural workers in India have shown great readiness in using statistical methods, and have fully responded to the lead given by the Imperial Council of Agricultural Research in this matter. Given necessary guidance and facilities, there is every reason to hope that the use of such methods will steadily extend in India.

II. The future programme

Problems of special importance to India

We may now consider the future programme. It was only natural that in the pioneer stage, much of the work in India followed closely the agricultural practice at Rothamsted and other English stations. But with the valuable background of experience gained during the last six or seven years, and with

* To be quite pedantic it should be called 'random sampling from random samples' For, all statistical work is necessarily based on random samples. The plot yields in a field trial, for example, are considered to be random samples from the hypothetical population of similar yields from the same plots under similar treatments in similar circumstances. Complete enumeration here merely means measurement of complete yields of all the plots which together constitute one single random sample. In the method of 'random samples', smaller random samples are taken from the plots.

the better organization of statistics in India, the time has come for using statistical theory and developing suitable methods for the study of problems of special interest to our country. A few suggestions in this connexion may be useful as a basis for discussion.

Rainfall and irrigation in relation to agriculture

We are all familiar with the essential facts. Agriculture is the basic occupation, and the prosperity of trade, commerce, and industry are more dependent on it in India than in most countries of the world. The seasonal rainfall is concentrated within a comparatively short period, but fluctuates widely both in total amount as well as in distribution from year to year. A good monsoon with well-distributed rain usually means good crops and general prosperity, while a bad monsoon still causes, over a vast area, failure of crops and widespread distress.

Water conservation, irrigation, and drainage naturally constitute a subject of overwhelming importance, and I hope to be excused if I dwell at some length on this question. I have had the opportunity of studying in some detail the problem of rainfall and floods in Bengal and in Orissa. This has made me realize how great is its direct and indirect bearing on agriculture. In most parts of India, we have enough rainfall to produce sufficient foodstuff for our present population. Our real problem is to conserve the water, prevent waste, distribute the available supply in the most efficient way over different areas and at different times according to agricultural requirements for the production of the optimum crop, and finally to drain away the excess without causing any mischief. Viewed in this way irrigation, drainage, flood control, and agriculture are merely different aspects of the same fundamental problem.

We possess fairly satisfactory data about rainfall, owing to the activities of an efficient Meteorological Department. We also have some, though neither enough nor quite reliable, data relating to rivers. But unfortunately the chief gap is in the agricultural knowledge.

Let me give a concrete example. I had occasion recently to examine a large irrigation scheme in Bengal which had the dual purpose of supplying water for crops at times of deficient rainfall, and of flood flushing the area as an anti-malarial measure. The future health, prosperity, and happiness of one million people depended on the success or failure of the scheme. We could estimate from past records with reasonable accuracy what deficiencies in rainfall were to be expected in future. We could also calculate how much water could be supplied from the Damodar river at different parts of the season. But unfortunately the agriculturists were quite unable to supply reliable information regarding the optimum water requirement of paddy. It was not possible therefore to make any estimate with confidence of the increased yield which might be reasonably expected with irrigation from the available supply. And yet this was the critical factor on which everything hinged. If the increase in production was sufficiently large the scheme would succeed; otherwise it would fail. The effect of a wrong decision either way would be disastrous. If the scheme were abandoned when in fact it might have succeeded, a great opportunity would be lost. If it were proceeded with but failed, such failure would jeopardize for at least one generation the initiation of other schemes even when the prospects of success were great. I had to make the best of a bad job, and tried to get round the difficulty by using statistical methods in a rather speculative fashion.

But that is a different story. To come back to our topic, we need then careful studies of the influence of rainfall and other climatic factors on crops. Such studies would be useful in two ways. First, for supplying badly needed information about the water requirement of crops. Secondly, for purposes of forecasting; even when a failure of crops cannot be prevented, early information may often enable ameliorative action being taken in time.

Permanent climatic series

It will be desirable therefore to start well-designed experiments in different parts of the country for studying the influence of rainfall and weather conditions on the yield of standard varieties of crops. The experiments will be definitely of the long-range type, and will be continued for many years. Arrangements will also be made for recording a number of carefully selected meteorological elements. In planning this series, the needs of the country as a whole will be naturally kept in mind, and the work will be standardized sufficiently to enable valid comparisons being made between results obtained at different stations.

Phenological observations

The question of starting systematic phenological observations (such as earliest and latest dates of flowering of well-known plants, passage of migratory birds, advent of seasonal pests and insects, etc.), may also be given careful consideration in the same connexion. Such observations are likely to prove useful in many ways, not only in the study of seasonal variations of the weather, but in the control of pests and blights, and in throwing light on the behaviour of plants to environmental conditions.

Irrigation experiments

Well-designed experiments will also have to be laid down for the direct study of water requirement and the growth of crops under irrigation. In the first stage it will probably be desirable to conduct such experiments under conditions in which both the supply and the drainage of water can be controlled at desired levels. As experience is gathered, it will no doubt be possible to approximate more closely to field conditions

In certain parts of India waterlogging and floods are often of almost as great importance as the lack of water. Carefully designed experiments are therefore needed for studying questions of seepage, waterlogging, etc., under actual agricultural conditions.

Soil erosion is a problem of importance in many regions. This question is closely connected with run-off and drainage, and requires to be studied in relation to irrigation. The possibilities of using agricultural methods, such as planting of suitable crops of trees for controlling soil erosion, deserve investigation in the same connexion.

All these irrigation experiments, to give the best results, require the active cooperation of the engineer and the agriculturist; while the scope of using statistical methods is practically unlimited

Soil studies

Another problem of great importance is the study of the soil, and of the changes in its condition, in different parts of the country. As regards progressive deterioration, the Royal Commission on Agriculture [1928] was of opinion : ' While paucity of records of crop out-turn throughout India over any long period of time makes the matter impossible of exact proof, we are of opinion that the strong presumption is that an overwhelming proportion of original lands of India long ago reached the condition to which experimental data point'

Permanent manurials

Careful experiments are needed to study whether soil deterioration is still progressing, and if so at what rate, and also to study the influence of different types of manures to prevent such deterioration and maintain the soil in a healthy condition. The time has therefore come to lay down a series of permanent manurials on modern lines at a number of selected stations. Where practicable, the manurial series may be suitably combined with advantage with the climatic series.

Multiple experiments

Multiple experiments offer great advantages for the study of climatic, varietal, manurial, and other questions. In this plan a number of experiments of the same type would be laid down with the same or similar groups of varieties or treatments in different parts of the country. Owing to the large differences in soil and climatic factors, not only between different provinces but even in different districts of the same province, these experiments would be conducted under widely varying conditions.

The work will have to be planned as a whole. When the same set of varieties or manures or other treatments cannot be used in all the experiments of a given series, it should be still possible to link up the work by providing overlapping treatments through which comparisons can be made with confidence. Standardization will obviously be necessary, but sufficient flexibility must be retained to adapt the work to suit local needs.

If the multiple series is designed as a whole, it will be often possible to conduct a joint analysis of the results, and to study the influence of the variations in the different factors. In this way valuable information might be obtained in a few years which would otherwise take a very long time to collect. In 1931 I had pointed out the need and scope of such multiple experiments under Indian conditions and had pleaded for their adoption at the joint session of the Agriculture, Physics and Mathematics sections of the Indian Science Congress at Nagpur. Six or seven years ago the time was probably not ripe for undertaking such experiments on a large scale. But the Fisherian technique has now become so familiar that it should not be difficult to start them in the immediate future.

Cultivation and rotation experiments

Other problems of special importance in India are connected with methods of cultivation and rotation of crops. Given soil and water, the basic problem is to secure the greatest return to the cultivator. A wide outlook is desirable in designing such experiments. When we compare different methods of cultivation, for example, it is obviously not sufficient merely to concentrate attention on which method gives the largest yield of crop. It is also necessary to take into consideration the question of relative costs, the real aim being to find out which method will secure the largest net return to the cultivator. Similarly in rotation experiments it is not enough to concentrate attention on merely the influence of a particular crop in one year on the yield of another crop in a succeeding year. The object should be to find out that particular sequence of crops which, after making allowances for differences in the cost of cultivation, would on an average secure the highest profit over a number of years.

Crop-cutting experiments

Although crop-cutting experiments do not fall under field trials, I would like to point out how such experiments may be made to supplement the information obtained from field experiments. Consider any given region. In an adequately designed crop-cutting experiment this region will be divided into a number of homogeneous zones with more or less the same type of soil, climate, irrigation facilities, type of crop, method of cultivation, etc. Suppose we now arrange to conduct the crop-cutting work at a number of spots selected strictly at random (but so arranged as to include all the varieties or conditions we desire to study) within each zone. The experiment as a whole will then resemble, on a very large scale, a field trial with a design of the randomized block type. I am not suggesting that in practice it will be possible to preserve the analogy in detail. But I think it should not be difficult to plan a crop-cutting experiment as a whole in such a way as to supply useful information regarding the performance of different varieties or treatments under actual cultivating conditions on a large scale.

Apart from such considerations, a crop-cutting experiment of course has its very important primary function of supplying information about the total out-turn of crop. As the only method available here is that of random samples, this question offers great scope for the application of statistical theory. Valuable pioneer work has been done in the United Provinces in this connexion, but the subject is of sufficient importance to deserve systematic and sustained study in other provinces.

Place of statistics in agriculture

Before concluding I would like to make a few general remarks about the place of statistics in agriculture. It is I hope sufficiently clear from the previous discussion that the first function of statistical theory is to supply an adequate technique for collecting the primary data in such a way that valid inferences may be drawn from them. The use of the principle of randomized replication in some form or other is indispensable for this purpose. The second function is to extract the whole of the information contained in the data in the most efficient way. It has been already pointed out that the appropriate method for this purpose will depend entirely on the particular way in which the process of randomization is introduced.

We have seen how successfully these principles have been used in the case of field trials. It is essential that the same principles should also be applied in the case of experiments of all other kinds. There is great scope for work in this direction in India. For, I am afraid, the need of statistical methods in experiments other than field trials has not yet been sufficiently recognized in this country. Much effort and time have been wasted in consequence.

Need of definite statistical objects

In fact it would be a salutary practice in most experimental studies to refrain from taking any measurements or recording any observations whatsoever until one was satisfied that these could be utilized in a valid manner for some useful purpose. In any case, it would be a safe rule to carry out a trial analysis with available material at the earliest opportunity. If this was done, it would often reveal gaps in the data or defects in the method of collecting them which could be often put right in time. If one waited until the end it would usually be too late.

Indeed in India it is tragic to see the enormous amount of statistical material which is collected at considerable expense but which is never used, or which can never be used in any way except as food for white ants. It would save a great deal of labour and money if no measurements or observations were recorded without a definite statistical purpose in view.

As already pointed out, in order to secure this end, the process of randomization and the projected method of analysis must be such that it would be possible to make precise statements as to the significance or non-significane of the results. When, as is usually the case, some previous information is already available, it is further necessary that the experiment should be designed in such a way that the expected precision is adequate for the purpose in view.

Statistics as a tool and not the end

Even this is not sufficient, something more is necessary. Before starting an experiment each worker should satisfy himself that, if the experiment is successful, something will be gained which is worth the time, labour, and money spent on it. I frankly confess that I have sometimes wondered whether this condition had been really fulfilled in the case of some of the agricultural experiments which I have had occasion to examine. I know that this is treading on dangerous grounds, but I do not think it can be emphasized too much that statistics is merely a means and not an end in itself. Wishart and Sanders [1936] have wisely remarked : 'In these days it is difficult, but very important to keep a sense of proportion over the question of experimentation. The statistical side has been given so much prominence in recent years that there is a real danger of statistics being regarded as the main interest in experimentation'.

Safeguards against statistical excesses

Agriculturists must not therefore allow statistics to degenerate into a kind of mysterious cult. The fundamental principles are easy to understand, and there is no reason why the experimenter should not take an intelligent interest in the designing of experiments. The statistician, owing chiefly to constant practice, is more skilled in handling certain technical tools which can be safely left to him. But it is the experimenter who is in a better position to judge the value of the experiment as a whole in its wider aspects.

Fortunately statistics itself may be used as a check against its own excess. It is possible, and possible only by statistical methods, to determine with scientific precision the marginal (or additional) cost in money and human labour for obtaining any given amount of additional information or increased accuracy. In this way a kind of scientific cost accounting of experiments can be made possible, so that the experimenter may be guided in his decision by rational considerations.

Cooperation between agriculturists and statisticians

In India there is always a danger of our not being able to see the wood because of the trees. The only corrective is to keep the basic problem prominently in view. The experimenter should constantly remind the statistician (and also himself perhaps occasionally) that improving the standard of living of the 350 millions of our countrymen by increasing the produce of the earth is the ultimate aim of all agricultural experiments ; and that how far progress is made in this direction is the final test by which all work will have to be judged. This is the only way in which the agriculturist will be able to breathe life into the dry bones of statistics. I therefore plead for a close, friendly, active, and fruitful cooperation between the agriculturist and the statistician in this task.

APPENDIX I

Mean squares in analysis of variance : Chinsurah complex experiment on rice, 1933-36

—	Degrees of freedom	1933	1934	1936
Block	2	144098	359084	75533
Date of planting	4	2389142**	499687**	5417487**
Error	8	30461	40555	23907
Variety	2	708517**	564263**	385807**
Error	4	12455	19494	16339
Planting × Varieties	8	29608**	37276**	26112**
Error	16	4259	8181	3018
Spacing	2	64848**	4061	10809**
Seedling	2	32691**	5733	2028
Spacing × Seedling	4	1396	3221	905
Planting × Seedling	8	2484	5570	1236
Planting × Spacing	8	5320*	1814	970
Variety × Seedling	4	1112	10710*	1065
Variety × Spacing	4	1306	4898	1331
Planting × Variety × Seedling	16	3991*	2889	1305
Planting × Variety × Spacing	16	2952*	5347	1751
Planting × Seedling × Spacing	16	1442	1585	848
Variety × Seedling × Spacing	8	1198	4634	1050
Planting × Variety × Seedling × Spacing	32	947	2289	1779
Error	240	1364	2847	1882

Note.—The results are given for the three seasons 1933, 1934, and 1936 as the experiment failed in 1935 owing to drought. The test of significance indicates that the effects of the three varieties of rice, of the five dates of planting, and of their mutual interaction of the first order were appreciable in all the three seasons under observation. Variation of spacing also showed significant differences in 1933 and 1936 but not in 1934. Other effects were either insignificant or significant only for one season. A detailed examination showed that the variety called *bhasamanik* gave the highest yields at Chinsurah during all the three seasons. The first week of August was found to be the best date of planting; in fact the yield showed a distinct tendency to be lower if the transplanting was finished earlier or delayed by a fortnight. A close spacing and an increased number of seedlings per hole were necessary to insure against late transplantings, particularly in years of adverse rainfall distribution. But under a favourable monsoon, small variations in spacing or seedling numbers produced practically no differences in yield. Finally the superiority of one variety over another was not identical for all the dates of planting but was found to be significantly associated with the time of transplanting.

Appendix II

Note on variance, standard error, covariance, etc.

Let $x_1, x_2 \ldots\ldots\ldots\ldots\ldots\ldots x_n$ be the yields of n plots. Then the average yield is defined as $x = x_1 + x_2 \ldots\ldots\ldots\ldots\ldots x_n$ .... (1)

The 'deviation' (or 'error') of the yield is simply the difference between any individual yield and the average, i.e. $(x_1—x)$, $(x_2—x)$, etc. The 'sum of squares of deviations' is given by $(x_1—x)^2 + (x_2—x)^2 + \ldots\ldots\ldots\ldots\ldots (x_n—x)^2 \ldots\ldots\ldots$ (2).

The 'variance' of the yield is defined as the sum of squares of deviation divided by $(n—1)$, or

$$v = \frac{(x_1—x)^2 + (x_2—x)^2 + \ldots\ldots\ldots\ldots\ldots\ldots (x_n—x)^2}{(n—1)} \ldots\ldots\ldots\ldots (3).$$

Here $(n—1)$ represents the 'degrees of freedom' which can be usually identified with the number of independent comparisons possible in any given case. In the present example, we can clearly have $(n—1)$ independent comparisons between the yields of n different plots.

It will be noticed that 'variance' represents a kind of average of the squares of deviations. The 'standard deviation' or 'standard error' is defined as the square root of the variance (which is sometimes also called the root-mean-square-deviation' or 'root-mean-square-error').

The variance defined in equation (3) is the variance of individual plots, and the corresponding standard deviation or standard error obtained by extracting the square root is the 'standard error per plot'.

The 'variance of the mean' of the yields, i.e. of x is obtained by dividing the variance of individual plots by n, the total number of p ots concerned, i.e. is given by

$$\frac{v}{n} = \frac{(x_1—x)^2 + (x_2—x)^2 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots (x_n—x)^2}{n\,(n—1)} \ldots\ldots\ldots\ldots (4).$$

The 'standard error of the mean' is the square root of the above expression.

When more than one character (or variate) is present, the covariance of any two characters (or any two variates) is similarly defined as the sum of the products of the corresponding deviations divided by the degrees of freedom.

Thus if (x_1, y_1) (x_2, y_2)........ (x_n, y_n) are the n pairs of values of the two characters, and x and y their respective averages, then the covariance is given by

$$\frac{(x_1-x)(y_1-y) + (x_2-x)(y_2-y) + \ldots\ldots\ldots(x_n-x)(y_n-y)}{(n-1)} \quad \ldots\ldots (5).$$

Main effect and interaction.—If we have p treatments (or factors) then the main effect of each treatment (or factor) is the mean value of the relevant treatment (or factor) for all combinations of the other factors. The main effect is thus obtained by taking the average over all plots in which this particular factor occurs.

When two or more factors or treatments are used at the same time as in complex (factorial) experiments, the total effect due to the joint influence of two (or more) factors may or may not be equal to the sum of the effects due to each of the factors taken separately. Interaction between the factors is defined as the magnitude of the departure (if any) from the total effect calculated on the additive hypothesis. When the different factors act independently the joint effect will be necessarily additive, and the interaction will be consequently nil.

In an experiment involving p factors, we can consider the factors in pairs, and we shall have one first-order interaction corresponding to each pair or $\frac{p(p-1)}{2}$ first-order interactions altogether. We can also consider the factors by threes, and in this case, if the first-order interactions are affected by the presence of the third factor, then we shall get second-order interactions. In the same way we can also consider higher order interactions, but usually high order interactions are of little practical importance.

REFERENCES

Allen, R. E. and Wishart, J. (1930). *J. Agric. Sci.* **20** (3), 130, 399

Christidis, B. G. (1931). *J. Agric. Sci.* **21** (1), 14

Fisher, R. A. (1923). *J. Agric. Sci.* **13**, 311

———— (1926). *J. Min. Agric.* **33**, 503

———— (1933). *Rothamsted Experimental Station Report*, p. 5

———— (1935). *The Design of Experiments.* Edinburgh : Oliver and Boyd.

Report of the Royal Commission on Agriculture, 1928

Vaidyanathan, M. (1934). *Report Imp. Coun. Agric. Res.*

Wishart, J. and Sanders, H. G. (1936). *Principles and Practice of Field Experimentation.*

Yates, F. (1933). *Emp. Expt. Agric.* **1**, (2) 129

———— (1935). *J. Roy. Stat. Soc., Supp. II* (2), 181

———— (1937). *Imp. Bur. Soil Sci. Tech. Comm., No.* **35**

A NOTE ON THE ANALYSIS OF 3^3 AND 3^4 DESIGNS (WITH THREE-FACTOR INTERACTIONS CONFOUNDED) IN FIELD EXPERIMENTS IN AGRICULTURE

BY

M. VAIDYANATHAN, M.A., L.T., F.S.S.

AND

S. SUBRAMONIA IYER, M.A., F.S.S.

Imperial Council of Agricultural Research

(Received for publication on 19 December 1939)

INTRODUCTION

On the recommendation of the Joint Committee of Field Experiments in Agriculture of the Imperial Council of Agricultural Research, certain suggestions for replanning of manurial experiments for rice and sugarcane were made to the provinces and states, and most of these have been accepted for adoption from 1939-1940 season. Of the designs suggested by the Committee, what are now known as ' factorial designs ' are important. These are on the lines of the designs adopted at the Rothamsted Experimental Station in England, which the senior author of this paper had the opportunity of examining in detail at close quarters during 1937. As the idea of ' confounding ' the effects with block differences is comparatively new to the agricultural experimenter in India, it may be stated even in the beginning that the object of ' confounding ' with the blocks is to reduce the block-size which will otherwise become huge in a factorial design involving a number of levels of different factors. This is done without disturbing the interpretation of main effects and interactions on which we are interested. Thus with a 3^4 design, (i.e. with 3 levels of each of 4 factors), in the usual randomised block method there would be 81 plots in a block, which is obviously too huge for a correct interpretation ; but if it is possible to confound some of the higher order interactions (say three or four factor interactions) with the blocks, without disturbing the main effects and lower order interactions on which we are interested, the block size can be reduced with only those treatments included in a block giving the desired comparisons. Similarly in a 3^3 design, even 27 plots are too huge for a block and we may confound with the blocks some of the three-factor interactions.

Again, where we are dealing with a 3^4 design, if it is not possible to replicate and find sufficient space for the experiment, Fisher has shown that even with a single set of 81 plots (without replications) it will be a valid interpretation of the data if higher-order interactions are treated as error [See ' *Design of Experiments* ' p. 115 (second edition).]