A COMPARATIVE STUDY OF THE STATISTICAL METHODS MOST COMMONLY USED IN AGRICULTURAL RESEARCH

By BERNARDO G. CAPÓ

INTRODUCTION

In statistical works the term *population* stands for a large group of individuals which possess one or more characteristics in common. These individuals may be persons, animals, numbers, leaves, etc., constituting respectively populations of persons, animals, numbers, leaves, etc.

The individuals themselves are different from one another and therefore the values which measure the intensity of any one of the characteristics common to all the individuals of a given population fluctuate between two limits: an upper and a lower one. These measurements, however, do not distribute themselves uniformly throughout all the range included between the limits, but on the other hand, their tendency is to distribute themselves in such a way that the number of values increase gradually from a very small number of them near the extremes to a very large number near the center of the distribution range.

The history of mathematical statistics has witnessed numerous attempts to explain by means of some mathematical formula or equation the exact way in which these values distribute themselves in naturally occurring populations. Among the curves represented by these equations the *normal curve* or *normal curve of error* has played a major role. Though inexact in itself, it is a close approximation to the exact distributions known as Binomial Law and Poisson Series, which in common with many other distribution curves, approach the normal curve as a limit, under certain conditions.

To graph the curve which represents any population of values normally distributed it is necessary to know the values of two *constants* or *statistics,* known as the *mean* and *standard deviation.*

The *mean* is equal to the sum of all the values divided by their total number, or in another way:

$$M = (X_1 + X_2 + X_3 + \ldots\ldots) / n$$

where M is the mean of the values that constitute the population, X_1, X_2, X_3, etc., are the individual values; and n is the number of values.

For example the sale-prices of some grapefruit lots sold at public auction in New York were as follows: 2.50, 3.37, 3.62, 4.00, 3.87, and 3.75. The mean of these sale-prices is found by adding them together and dividing their sum by 6, which gives as result $21.11/6 = 3.52$. This is of course, the mean of these six sale-prices, which is also an estimate of the true mean of the population of sale-prices represented by them.

The *standard deviation* or *standard error* of the individal values of a population is equal to the square root of the quotient obtained on dividing the sum of the squares of the differences between each value and the mean, by one less than the number of values, or in another way:

$$D = \sqrt{(X_1-M)^2 + (X_2-M)^2 + . \text{-----} / (n-1)}$$

where D is the standard deviation of the values of the population and the other signs mean the same as in the previous formula.

Therefore, the estimate of the standard deviation of one of the sale-prices of the sale-price population represented by those of the preceding example is found as follows:

Values	Deviations from the mean (Differences between the values and the mean)	Squares of the Deviations
2.50	-1.02	1.0404
3.37	-0.15	0.0225
3.62	0.10	0.0100
4.00	0.48	0.2304
3.87	0.35	0.1225
3.75	0.23	0.0529
21.11 = Sum.		Sum of squares = 1.4787

Mean $= 21.11/6 = 3.52$ $\qquad$ $1.4787/5 = 0.2957$

Standard Deviation $= \sqrt{0.2957} = 0.54$.

Another way to find this estimate of the standard deviation which will be used on illustrating some of the statistical methods which will be described further on, is as follows:

	Values	Squares	
	2.50	6.2500	Correction $= (21.11)^2/6 =$
	3.37	11.3569	
	3.62	13.1044	$= 445.6321/6 = 74.2720$
	4.00	16.0000	
	3.87	14.9769	
	3.75	14.0625	Standard deviation $= \sqrt{0.2957} = 0.54$
Sums	21.11	75.7507	
		$-$ 74.2720	
		1.4787 / 5 = 0.2957	

The square of the standard deviation has been called by R. A. Fisher *variance* and since this term will be used frequently in the future, it is convenient to note its relation to the standard deviation. In the preceding example the estimate of the variance of a single value of the population is equal to 0.2957.

In what follows it will be assumed that these estimates of the mean and standard deviation are equal to the real mean and standard deviation which could be calculated only from the values of all the individuals comprising the population; theoretically, an infinite number of values.

If twice the standard deviation of one of the values of the population under study is respectively added to and subtracted from the mean, there will be obtained two numbers which, if used as limits, will include between them about 95.45 per cent of the values which constitute the population. Therefore, the probability that any one of these values will fall by chance within these limits is of 95.45/100, where the total number of possibilities is unity.

Thus, then, in the previous example, if two times 0.54 or 1.08 is respectively added to and subtracted from 3.52, there will be obtained the quantities 4.60 and 2.44. If under similar conditions a large number of these sales were effected, it should be expected that 95.45 per cent of the sale-prices in these cases lie between 2.44 and 4.60, or the probability that any one of these sale-prices fall within these limits would be of 95.45/100.

In other words, 4.55 per cent of the values or sale-prices in sales made under similar conditions would be at prices lower than 2.44 or higher than 4.60, and of these, one-half would be at prices lower than 2.44. Thus, then, with this information one can deduct that of a large number of sales made under similar conditions, 2.28 percent will be made at prices lower than 2.44 and 97.72 per cent at prices higher than 2.44. It can also be deducted that under similar conditions the probability that a sale be made at a price higher than 2.44 against the probability that it be made at a price lower than 2.44 will be of 97.72 to 2.28 or of 43 to 1. Accentuating still more, it should be expected that 43 out of every 44 sales made under conditions similar to these will be made at prices higher than 2.44.

On the other hand, if it is desired to know the two limits (at equal distances from the mean) between which a certain percentage of the population is to be found, for example, 70 per cent, one must proceed as follows: If 70 per cent of the values fall within these limits, then 30 per cent of the values of the population will fall outside of them; or in other words, the deviations from the mean of 30 per cent of the values of the population will be larger than that which is sought, that is, than the deviation of the limits from the mean.

If one looks in the table of probabilities based on the normal curve, at the end of this article, it will be seen that in 30 per cent of the cases the deviations of the values from the mean exceeds 1.04 times the standard deviation or that the limits outside of which 30 per cent of the population lies are at distances from the mean equal to 1.04 times the standard deviation of the population. As in this case the standard deviation is 0.54, the limits will lie at distances from the mean equal to $1.04 \times 0.54 = 0.56$, and therefore they will be $3.52 - 0.56$ and $3.52 + 0.56$, or 2.96 and 4.08, respectively. In other words, 70 per cent of the population will lie between 2.96 and 4.08.

The relation which has given practicability to these formulas is the following fundamental postulate in all kinds of statistical methods: If a number of values are normally distributed around their mean with a standard deviation equal to D, then means of groups of n of these values will be distributed around their mean with a standard deviation equal to $D/\sqrt{n}$.

The standard deviation of the mean in the case of reference will then be $0.54/\sqrt{6} = 0.54/2.45 = 0.22$. This means that on subtracting from and adding respectively to the mean two times 0.22 or 0.44, the quantities 3.08 and 3.96 will be obtained, between which 95.45 per cent of the means of groups of 6 sales made under similar conditions lie. This postulate has given practicability to these formulas, not only for the fact that, as has been seen, the standard deviation of the mean of a number of values is smaller than the standard deviation of any one of them, and thus the value that another mean of the same number of values may have is fixed within narrower limits, but also because the mean represents the whole population.

The table of probabilities based on the normal curve, however, can be used only when the population is distributed normally and when the *exact* values of the mean and its standard deviation are known. However, as "Student" pointed out in 1908, with the necessarily limited number of values with which the research worker must work, be it in the laboratory or in the experimental field, it is impossible to determine exactly both the mean of the population which said values represent, and its standard deviation, being still more difficult to determine if the population is normally distributed or not. Therefore, if it is not possible to know exactly these facts, the use of the table of probabilities based on the normal curve becomes academic and loses all its utility.

For these conditions, however, "Student" developed another distribution curve which he called the distribution of "z", which holds for the estimates of the mean and standard deviation as obtained from only a few values of the population, and, therefore, conclusions having a known degree of accuracy may be derived, even with some few values of observations, as in the case of a large part of the experiments performed by the investigators all over the world.

With these new tables, moreover, this can be done not only in those cases in which it is known that the means of groups of values distribute themselves normally with respect to their mean and standard deviation, but also in cases where the original values are distributed in a non-normal way, provided the departure from normality be not too wide. Therefore, the table of "z" has a very wide application, its use being of incalculable aid for the research workers in all branches of science.

R. A. Fisher has modified this table of "z" and obtained the distribution of $Z/\sqrt{n-1}$ which he has called "t". In Fisher's table of "t" the letter n means not the total number of values, but what he has called the *number of degrees of freedom.* This number of degrees of freedom is the number of values from a series of them which can be altered arbitrarily without altering their mean.

In the series of sale-prices the values have 5 degrees of freedom, since the total number of observations is 6, and if arbitrary values are assigned to 5 of them, the sixth value must have only one magnitude, if the mean is to remain equal.

With these new tables the problem of comparing the mean of the values which constitute a population with a certain fixed value and to determine the probabilities with which the mean will exceed said value is thus simplified. Thus, then, one can determine the probability with which a given fertilizer or cultivation treatment will produce yields that surpass a certain limit or with which the yield of a given variety of plants will exceed a given fixed value.

However, the object of the agricultural experiments in general is not to obtain varieties or treatments which will produce at least a certain fixed yield, but rather to determine which of any given number of these is the best under certain environmental conditions, or in case that there were two or more almost equal and better than the rest, which are these.

At present there are two general ways for making this selection in both of which the population whose mean is considered best, is compared successively with each of the other populations.

In the first of these ways one calculates the probability with which the difference between the means of the two treatments or varieties to be compared is positive. This procedure is employed in what are known as Bessel's Method, Fisher's Method for the Analysis of Variance and Hayes Deviation from the Mean Method.

In the second of these ways one calculates the probability with which the mean of the differences between corresponding values of the two treatments or varieties to be compared is positive. This procedure is employed in what is known as "Student's" Method of Paired Results.

In what precedes the ways of estimating means and their standard deviations have been described. But now it has been seen that the estimates of the standard deviations of differences must be calculated also in the application of all those methods which calculate the probability with which the difference between the means is positive.

In order to do this the following relation is made use of: The standard deviation of the difference between any two means is equal to the square root of the difference between the sum of the variances of both means and twice the product of their standard deviations by the coefficient of correlation between the means they represent.

This is expressed mathematically as follows:

$$D_d = \sqrt{V_1 + V_2 \quad - 2D_1D_2C}$$

where D_d is the standard deviation of the difference between the means, V_1 is the variance of one of the means and D_1 its standard deviation, V_2 is the variance of the other mean and D_2 its standard deviation, and C is the coefficient of correlation between the populations which the said means represent.

The method can be used then only to compare populations whose coefficient of correlation can be estimated with a certain degree of accuracy, for it cannot be determined exactly, due to the same reasons which hold for the values of the means and their standard deviations.

The labor which is needed to calculate this coefficient of correlation, in addition to the fact that in order that it may be calculated the experiment must comply with certain requisites, has made that the majority of the investigators use modifications of this method, of which there are some. In fact, the differences between these methods consist precisely in the different ways of modifying the application of this formula. These modifications will therefore be taken up in connection with the descriptions of the methods.

Before proceeding with the detailed descriptions of the methods there is still one point to be taken up, which is as follows. In order that uniform conclusions may be derived from the interpretation of the same data by different investigators it has been accepted to consider as a statistically significant difference, that difference which has a probability in favor of its being positive of not less than 21/22. Since this is the probability that should be used on making the interpretation of agricultural experiments the tables of "z" and "t" for this probability have been calculated and reproduced at the end of this article.

These preliminary considerations, which seemed to be necessary for the practical use of the different methods in use at present being finished, it is now time to describe the different methods in order.

Bessel's Method

This method, as has been already said, belongs to those methods in which one calculates the probability with which the difference between the means of two populations is positive, that is, with which one of the populations will be better than the other, as witnessed by the means of the characteristic under study.

In this method the formula for the standard deviation of a difference between two means is modified by discarding altogether the term which contains the coefficient of correlation converting it into:

$$D_d = \sqrt{V_1 + V_2}$$

In order that this modification be correct, it is necessary that the conditions under which the experiment which is going to be interpreted in this way be such that the value of C, or coefficient of correlation between the populations to be compared, be so small that the term $2D_1D_2C$ be negligible. Under these conditions one has to calculate the means of the two populations to be compared and their variances and apply this formula.

To approximate these conditions as much as possible, it is the custom to distribute by chance throughout all the experimental field the different replications or repetitions of the treatments to be compared.

The diagram which follows represents the plan of an experiment performed under the direction of Mr. Fernando Chardón at Aibonito, P. R. The object of the experiment was to determine which of the seven fertilizer treatments tested would help to produce more tobacco leaf in a certain variety of tobacco at that region and under the then existing climatic conditions. The size of the plots was 1/50 of a *cuerda* (1 cuerda is 0.97 of an acre) and the yields are in hundredweights per *cuerda*.

DIAGRAM No. 1

B 8.38	D 9.00	A 7.45	G 8.03	F 9.69	E 9.50	C 7.92
F 9.64	G 8.86	D 9.84	E 9.55	A 8.91	C 8.44	B 8.29
D 9.39	A 9.73	E 10.00	C 10.34	G 8.75	B 8.06	F 7.61
G 10.03	E 10.30	B 8.36	F 8.41	C 9.59	A 7.58	D 8.44
C 9.22	F 7.95	G 7.06	A 8.36	B 9.70	D 9.45	E 8.98
E 7.98	C 8.70	F 8.69	B 8.97	D 9.09	G 7.69	A 7.63
A 7.30	B 8.23	C 6.98	D 8.22	E 9.13	F 7.86	G 8.03

Even though the distribution was such that there should be a repetition of each treatment in each file and column of the experimental field, that is, in the form of a Latin Square, as the repetitions for each treatment were distributed throughout all the experimental field, and within most columns and files the treatments were distributed by chance, it is to be expected that the correlation should not be too high and that the method of the modified formula may have here application. As these same results are to be used to illustrate the application of all the methods, one can assume for the present that the distribution was done by chance and that the results obtained were as appear in the diagram.

The means, their variances and standard deviations are calculated as explained in the preceding pages:

Treatment A—

Values	*Squares*	
7.45	55.5025	
8.91	79.3881	Mean $= 56.96/7 = 8.14$
9.73	94.6729	Correction of the sum of squares $= (56.96)^2$
7.58	57.4564	$= 3244.4416/7 = 463.4917$
8.36	69.8896	Variance of one value $= 4.9247/6 = 0.8208$
7.63	58.2169	Variance of the mean $= 0.8208/7 = 0.1173$
7.30	53.2900	Standard deviation of the mean $= \sqrt{0.1173} = 0.34$
56.96	468.4164	Yield of A $= 8.14 \pm 0.34$
	−463.4917	
	4.9247	

In the same way these values are found for the other treatments.

The yields of the different treatments in order as regards amount, are:

Treatment—E:	9.35 ± 0.29	Variance of the mean $= 0.0822$
Treatment—D:	9.06 ± 0.22	Variance of the mean $= 0.0468$
Treatment—C:	8.74 ± 0.45	Variance of the mean $= 0.1994$
Treatment—B:	8.56 ± 0.22	Variance of the mean $= 0.0483$
Treatment—F:	8.55 ± 0.32	Variance of the mean $= 0.1010$
Treatment—G:	8.35 ± 0.51	Variance of the mean $= 0.2628$
Treatment—A:	8.14 ± 0.34	Variance of the mean $= 0.1173$

As each one of the means of these treatments is based on 6 degrees of freedom there being 7 observations of each population, it is evident that the difference between any two of these means will be based on 12 degrees of freedom. Looking up in the table of "t" the ratio which corresponds to 12 degrees of freedom, for a probability of 21/22, one finds 1.83.

Now then, as the square root of the sum of the squares of two numbers must be greater than any one of the two numbers, and furthermore, as the difference between the means to be compared is to be compared with a number that is 1.83 times its standard deviation, it is evident that all the means of yields smaller than the largest but lying below it by not more than 1.83 times its standard deviation, will not give yields smaller than its yields in 21 cases out of 22.

Thus then, as the standard deviation of the mean 9.35 is 0.29, all the treatments whose means be less than 9.35 but more than $9.35 - 1.83 \times 0.29$ or $9.35 - 0.53 = 8.82$, will not be different in yield from treatment E, in at least 21 cases out of 22. That means that the method indicates preliminarily that treatment E gives not higher yields than treatment D in at least 95.45% of the cases, for the mean of the yield of D exceeds the inferior limit 8.82 already fixed.

The differences between the yields of treatment E and the other ones, are proved as follows as regards their statistical significance:

Between E and C:

Mean of E $= 9.35$	Variance of the mean of E	$= 0.0822$
Mean of C $= 8.74$	Variance of the mean of C	$= 0.1944$
Diff. 0.61	Variance of the diff.	$= 0.2766$

Standard deviation of the difference $= \sqrt{0.2766} = 0.52$

Amount which the difference must exceed to be significant

$$= 1.83(0.52) = 0.95$$

Since the difference does not exceed this limit, it is not significant.

In the same way are calculated the amounts to be exceeded by the differences between the mean of E and the means of B, F, G and A, concluding that treatment E is superior to treatments B, F and A.

Adding up, if one calculates the experiment under study by this method, that is, without making any correction for the error introduced on discarding the term which includes the coefficient of correlation in each comparison, one arrives at the conclusion that the treatments inferior to treatment E are treatments A, B and F and that there is no evidence as to statistically significant differences between treatment E and treatments D, C and G.

Note: There is a method which tries to eliminate the effects of soil heterogeneity by introducing check plots every three or four treatment plots. The fertility of the field is assumed to vary uniformly from one check plot to the other and the yields of the treated plots are then corrected by subtracting from each yield the respective yield which it would have produced if it were a check plot. The residues are then treated as above.

Once illustrated the preceding method, we will proceed to illustrate the method that uses the formula for the standard deviation of a difference without modifications of any sort. In this method, as has been already pointed out, one must calculate the coefficients of correlation between the different pairs of populations to be compared.

As the correlation of the yields is the result of the heterogeneity of the soil to the extent that the nearer two plots are from one another the larger will be the similarity between their autochthonous fertilities, and therefore, the larger the correlation between the yields of any pair of treatments applied in them, it is the custom when the said coefficient is to be used, to distribute the replications systematically throughout all the experimental ground in order that the replications of each treatment be within given distances from the respective replications of any other treatment at least in a major part of the replications.

To calculate the coefficient of correlation between the two populations or series, there must be known first of all with which value of one of the series each value of the other one is respectively related. In other words, the individual values of both populations must be related in some specific manner and this relation must be known in order to calculate said coefficient.

As our interest rests in presenting the most common methods of interpreting the results of the agricultural field experiments, and not the different ways of carrying into effect the application of these methods, an exposition of but one way of calculating the coefficient of correlation between two populations is made.

In the example which occupies our attention the related values of the populations to be compared are those which lie in the same file and in the same column, for if diagram No. 1 is observed, it will be seen that each value of one population is related in this form with two values of every other population.

The coefficient of correlation between two populations is found by making use of the following formula where C is the coefficient of correlation between the two populations "1" and "2".

$$C = \frac{S(xy)}{(n-1)\ d_1 d_2}$$

where the numerator is the sum of the products obtained on multiplying the respective deviations of the associated values of the two series from their respective means; n is the number of pairs of correlated values, and d_1 and d_2 are respectively the standard deviations of the individual values of both series.

If one allows that N be the number of observations in which each mean is based and D_1 and D_2 the standard deviations of said means, the preceding formula may be modified in this way:

$$C = \frac{S\ (xy)}{(n-1)\ \dfrac{d_1}{\sqrt{N}}\ \dfrac{d_2}{\sqrt{N}}\ N} = \frac{S\ (xy)}{(n-1)\ D_1 D_2\ N}$$

The advantage of the use of this last formula is obtained on using it to modify the formula for the standard deviation of a difference in this way:

$$D_d = \sqrt{D_1^2 + D_2^2 - 2\ CD_1\ D_2}$$

$$= \sqrt{D_1^2 + D_2^2 - \frac{2\ D_1\ D_2}{(n-1)}\ \frac{S\ (xy)}{D_1\ D_2\ N}}$$

$$= \sqrt{D_1^2 + D_2^2 - \frac{2\ S\ (xy)}{(n-1)\ N}}$$

To apply then this formula to the comparison between the means of the two populations desired, one would need only to know the value of S(xy) between the said populations. That value between the populations of yields of the treatments E and D is found as follows:

E Values	E Deviation from the mean x	D Values	D Deviation from the mean y	xy
9.50	0.15	9.00	-0.06	-0.0090
9.55	0.20	9.84	0.78	0.1560
10.00	0.65	9.39	0.33	0.2145
10.30	0.95	8.44	-0.62	-0.5890
8.98	-0.37	9.45	0.39	-0.1443
7.98	-1.37	9.09	0.03	-0.0411
9.13	-0.22	8.22	-0.84	0.1848
7.98	-1.37	9.39	0.33	-0.4521
10.30	0.95	9.00	-0.06	-0.0570
10.00	0.65	9.84	0.78	0.5070
9.55	0.20	8.22	-0.84	-0.1680
9.13	-0.22	9.09	0.03	-0.0066
9.50	0.15	9.45	0.39	0.0585
8.98	-0.37	8.44	-0.62	0.2294
130.88 / 14 = 9.35		126.86 / 14 = 9.06		S(xy) = -0.1169

As in this case the correlation was negative, without existing any apparent cause for such an event, one should disregard the term containing the coefficient of correlation, interpreting the results as previously done.

By this same process one finds that the value S(xy) between the populations corresponding to treatments E and C is equal to 0.9107. The standard deviations of the difference between the means 9.35 and 8.74, subject to their respective variances 0.0822 and 0.1994 is then

$$D_d \begin{array}{l} = \sqrt{0.0822 + 0.1994 - 2(0.9107)/13(7)} \\ = \sqrt{0.2816 - 0.0200} = \sqrt{0.2616} = 0.51. \end{array}$$

The amount which the difference between the means must exceed is now $1.83 \times 0.51 = 0.93$. As the difference is only 0.61, it is evident that it is not significant.

In the same way are calculated the standard deviations of the differences between the means of the yields of E on the one hand and those of B, F, G and A on the other, concluding that treatment E is superior to treatments B, F and A.

General Considerations

The two methods which precede, however, have been criticized as inefficient on pointing out the fact that in a large number of cases the standard deviations of the differences between the means are based on too few observations, and therefore, the amounts which said differences must exceed to be significant are too large.

This is due to the fact that the standard deviations of these differences have to be multiplied by relatively high values of "t" for these "t" values correspond to too small numbers of degrees of freedom.

The differences in yield of the different plots are due, in an experiment like the one which occupies our attention, to three main causes: first, to the difference in autochthonous fertility between the plots; second, to the difference of the treatments used; and third, to the fact that all the plants used in the experiment will be different and each will behave, of course, differently to its environment.

On finding the standard deviation of the difference between the means by using the formula used above, part of the sum of the variances of the means is caused by the heterogeneity of the soil and the other by the different way in which the plants behave to their environment. This sum of the variances is diminished by the term $2CD_1D_2$ that represents the effect of the heterogeneity of the soil, leaving a quantity which is due solely to the different behaviour of the plants with respect to their environment, errors in weighing the yields, etc., which are considered experimental or indeterminable errors. It is with respect to this variance that the assumption is made that said difference distributes itself in the required way to make use of the values appearing in the table of "t".

Fisher and Hayes have assumed that the variance due to the differences in yields of the plots treated in a different way, once corrected to eliminate all effects that will not be constitutional differences between the plants and other indeterminable errors, are equal or almost so, and that all these variances may be added together obtaining from their total sum a generalized variance for the means of the yields of the different treatments.

This is more exactly true in those cases in which the treatments or varieties to be compared in the experiment be not fundamentally different among them, that is, when the difference among the treatments be one of degree only and not fundamental. Even when the difference between the treatments were fundamental, there are cases in which the variances though not qual, would not at least differ by much.

As the generalized variance is obtained from the total number of observations, it is evident that it will be subject to a much larger number of degrees of freedom than the variances of the means obtained individually in each case. The value of "t" by which the standard deviation corresponding to any difference between the means must be multiplied will be smaller, therefore, and any difference would have to exceed a smaller quantity in order to be significant. In this way these investigators have increased the number of observations available for the estimation of the standard deviation of any one of the populations to be compared.

Fisher's Analysis of the Variance Method, Hayes' Deviation from the Mean Method, and Bessel's Method (where the term $2CD_1D_2$ is dropped), are applied when the plots where the different treatments are to be located are selected by chance, in order that any repetition may fall by chance, anywhere within any section of the experimental field or of the whole field.

Fisher's Analysis of the Variance Method has three phases. The first one is known as the "Unrestricted Method", where the plots are distributed by chance throughout all the experimental field. Hayes' Deviation from the Mean Method is almost identical with this phase of the Analysis of Variance Method, differing from it only in that the standard deviation is expressed in percentage of yield, there being therefore, different standard deviations of the means of the treatments when they are calculated in terms of weight, while in Fisher's Unrestricted Method, the standard deviation is fixed without taking account of the differences in amounts of the means of the yields of the different treatments.

The second phase of the Analysis of the Variances Method is known as the "Method of Randomized Blocks". To apply this method the plots are distributed by chance within some blocks or sections in which the experimental field is divided such that within each block there may be one and only one application of each treatment.

The third phase is known as the "Method of Least Squares". This one differs from the preceding one in that instead of dividing the experimental field into blocks, it is divided into files and columns and the treatments are located in such a way that in each file and in each column there may be but one repetition of each treatment.

In a general way, Fisher's Analysis of the Variance Method postulates that the total variance of the observations of any experiment is due to various causes, and that it can be divided in accordance with these causes, assigning to each cause that intensity of the variance which corresponds to it.

In Fisher's Unrestricted Method, and Hayes' Deviation from the Mean Method it is assumed, and this has been mathematically proved by "Student", that when it is assumed that the variabilities within the different treatments are equal, then the total sum of squares consists of the sum of squares between the treatments plus the sum of squares within the treatments.

The variation between the treatments is due precisely to the differences among them, for, being different, they will produce in general different yields.

The variation within the treatments is due to the differences in constitution of the plants that are used in the experiment and to the difference in autochthonous fertility of the plots where the repetitions of the same treatment are planted, to errors on weighing the crop, etc.

The manner of calculating the variances due to these causes is demonstrated in the explanation which follows:

Fisher's Unrestricted Analysis of the Variance Method

This method, as has been pointed out, attempts to reduce the quantity to be exceeded by the difference between two means of treatments compared in an experiment in order that it may be considered significant by means of a theoretical increase in the number of observations that may be used in its determination. Instead of calculating the standard deviation of the mean of the yields of the different replications of any treatment by using the observations in the plots of this treatment only, in this method the variances of the means of all the treatments in the experiment are averaged together and a generalized standard deviation which is used for all at a time is calculated.

Following may be found the calculations necessary for the interpretation of the results of the experiment which is occupying our attention, assuming, as in the illustration of the application of the modified Bessel's Method, that the different repetitions of the different treatments were located by chance throughout all the experimental field.

Taking diagram No. 1 as example, there are two ways of calculating the standard error or standard deviation due to experimental error. The variance from which the standard deviation is to be calculated is obtained by dividing by the proper number of degrees of freedom, the sums of the squares obtained either by adding the sums of squares due to error within the different treatments or by subtracting the sum of the squares between the treatments from the total sum of squares. Both methods are illustrated in what follows:

To find the sum of squares due to error by adding the sums of squares within the different treatments.

These sums of squares within the different treatments are found just as was done in the method which uses the modified formula where the following sums of squares were found:

Sum of squares within treatment A	4.9247
Sum of squares within treatment B	2.0285
Sum of squares within treatment C	8.3745
Sum of squares within treatment D	1.9639
Sum of squares within treatment E	3.4505
Sum of squares within treatment F	4.2466
Sum of squares within treatment G	11.0361
Total sum of squares within the treatments	**36.0248**

The variance within each treatment is subject to 6 degrees of freedom for there are 7 observations and only the mean of each treatment is fixed. Thus, as there are 7 treatments, the total sum of squares within the treatments is subject to $7 \times 6 = 42$ degrees of freedom. The variance of any single value due to error will then be $36.0248/42 = 0.8577$.

To find the sum of squares due to error by difference between the total sum of squares and the sum of squares between treatments.

The total sum of squares is found by taking all the observations and finding the sum of squares by treating them as if they belonged to a single treatment, and one were to find the sum of squares within that treatment, thus:

Copying from page 11:

Treatment	Values	Sum of squares of the values
A	7.45, 8.91, 9.73, 7.58, 8.36, 7.63, 7.30	468.4164
B	8.38, 8.20, 8.06, 8.36, 9.70, 8.97, 8.23	514.6014
C	7.92, 8.44, 10.34, 9.59, 9.22, 8.70, 6.98	543.2625
D	9.00, 9.84, 9.39, 8.44, 9.45, 9.09, 8.22	576.7303
E	9.50, 9.55, 10.00, 10.30, 8.98, 7.98, 9.13	615.2210
F	9.69, 9.64, 7.61, 8.41, 7.95, 8.69, 7.86	515.9641
G	8.03, 8.86, 8.75, 10.03, 7.06, 7.69, 8.03	499.0936
	Sum of all values: = 425.22	3733.2893
		− 3690.0099
		43.2794 = Total sum of squares

$$\text{Correction:} = \frac{(425.22)^2}{49} = 3690.0099$$

The sum of squares between the treatments is found thus:

Treatment	Sum of values	Squares of the sums of the values	
A	56.96	3244.4416	
B	59.90	3588.0100	
C	61.19	3744.2161	
D	63.43	4023.3649	
E	65.44	4282.3936	
F	59.85	3582.0225	
G	58.45	3416.4025	
Sum	425.22	25,880.8512 / 7	= 3697.2645 −3690.0099
			7.2546

$$\text{Correction: } \frac{(425.22)^2}{49} = 3690.0099$$

Sum of squares between the treatments = 7.2546

Analysis of Variance		Degrees of Freedom
Total sum of squares	43.2794	48
Sum of squares between the treatments	7.2546	6
Sum of squares due to error	36.0248	42

The total sum of squares is subject to 48 degrees of freedom, there being 49 observations and the only fixed value is the general mean, 8.68; and the sum of squares between the treatments is subject to 6 degrees of freedom, for there are 7 treatments. Thus there remain 42 degrees of freedom for the sum of squares due to error and therefore, for the variance due to error, that is, within the treatments.

The variance of a single value due to error would be equal to $36.0248/42 = 0.8577$, that is, the same as previously found; and the standard deviation of a single value is equal to $\sqrt{0.8577} = 0.93$. The standard deviation of a mean of 7 values is equal to $\sqrt{0.8577/7} = \sqrt{0.1225} = 0.35$, and the standard deviation of the difference between any two means of 7 values each would be $\sqrt{2(0.1225} = \sqrt{0.2450} = 0.50$. In order that the difference in yields between any two means were significant they would have to exceed $1.74 \times 0.50 = 0.87$, for 1.74 is the value of "t" that corresponds to once in 22 cases that this value be exceeded by chance with 30 degrees of freedom, since 30 is the value most near to 42 in the list. If the means on page 9 are looked over, it will be seen that by using this same reasoning treatment E is superior to treatments G and A for $9.35 - 8.35 = 1.00$, which exceeds the value 0.87, and $9.35 - 8.14 = 1.21$, which also exceeds the value 0.87.

Hayes' Deviation From the Mean Method

In the application of this method one follows the same procedure that was followed under Fisher's Unrestricted Analysis of the Variance Method until the standard deviation of the mean of any treatment is found, which was there found to be equal to 0.35. Finding then what per cent this standard deviation is of the general mean of all the treatments, 8.68, one finds:

$$0.35 \times 100\% / 8.68 = 4.03\%$$

Multiplying now this generalized standard deviation in percent by each mean the standard deviations corresponding to each mean are obtained; these appearing at their sides respectively in the list which follows:

Treatment	Mean	Standard deviation of the Mean	Variance of the Mean
E	9.35	0.38	0.1444
D	9.06	0.37	0.1369
C	8.74	0.35	0.1225
B	8.56	0.34	0.1156
F	8.55	0.34	0.1156
G	8.35	0.34	0.1156
A	8.14	0.33	0.1089

Applying the same reasoning that was used on page 207, one deducts that all the treatments whose means be higher than $9.35 - 1.74 \times 0.38 = 9.35 - 0.66 = 8.69$, will not be statistically different from treatment E. One can say then, that treatment E does not differ statistically from treatments D and C.

The standard deviation of the differences between the mean E and the means of the other treatments are found as follows:

Between E and B:

$$D_d = \sqrt{0.1444 + 0.1156} = \sqrt{0.2600} = 0.51$$

The amount to be exceeded by the difference is $1.74 \times 0.51 = 0.89$. As $9.35 - 8.56 = 0.79$ it is evident that treatment E does not differ significantly from treatment B.

In the same way the differences between the mean of E and those of F, G and A are tested concluding that treatment E is superior only to treatments G and A.

METHOD OF RANDOMIZED BLOCKS

To apply this method, as has been said before, the field must be divided into sections of various plots each, and the repetition of each treatment within any one of these sections or blocks must be placed by chance.

The procedure to be followed in order to calculate the standard deviation due to error can then be modified by reducing the sum of squares due to error by that quantity which corresponds to the variation between blocks, and which therefore corresponds to differences in fertility between the different sections or blocks into which the experimental field has been divided. The variation due to error remains in this way reduced if the variation between blocks is larger than the variation due to experimental error.

To illustrate the method use is made of diagram No. 1, assuming that the blocks are constituted by all the plots in horizontal rows, that is, block No. 1 consists of the plots where the treatments are found in the order that follows: B, D, A, G, F, E, and C.

The total sum of squares is the same as in the Unrestricted Method: 43.2794. To this quantity one has to subtract now the sums of squares due to the variation between the treatments and to variation between the blocks in order to obtain the sum of squares due to error.

The sum of squares corresponding to the variation between the treatments is equal to that in the unrestricted method, 7.2546, and therefore the sum of squares due to the variation between blocks and that due to error is now 36.0248.

To find the sum of squares due to the variation between blocks, one proceeds as was done to find the sum of squares due to the variation between the treatments, thus:

Block	Sum of values	Squares of the sums of the values	
1	59.97	3596.4009	
2	63.44	4024.6336	
3	63.88	4080.6544	
4	62.71	3932.5441	
5	60.72	3686.9184	
6	58.75	3451.5625	
7	55.75	3108.0625	
Sum	425.22	25880.7764 / 7	= 3697.2538
			-3690.0099
			7.2439

$$\text{Correction} = \frac{(425.22)^2}{49} = 3690.0099$$

Sum of squares due to blocks $= 7.2439$.

The sum of squares due to error is reduced now to $36.0248 - 7.2439 = 28.7809$. This sum of squares, however, is subject to 36 degreees of freedom, for the variation between the blocks is subject to 6 degrees of freedom, there being 7 blocks.

The variance of a single value due to error is now $28.7809/36 = 0.7995$ and the difference between any two means must exceed now the quantity $1.74 \times \sqrt{2(0.7995)/7} = 1.74 \times \sqrt{1.5990/7} = 1.74 \times \sqrt{0.2284} = 1.74 \times 0.48 = 0.84$.

By this method, then, treatment E results superior to treatments A and G.

Latin Square Method

If the true distribution is now used, and the Analysis of the Variance Method is applied to it, one will obtain a more clear idea of the method in general. The latin square is the distribution for which it might be said that the method was developed, although there are experiments where various blocks are used each of which is by itself a latin square, and if the experiment were repeated for a number of years, that part of the sum of squares due to differences in climatic effects would have to be subtracted also.

To complete the analysis of variance according to the original distribution, one has only to subtract to the sum of squares assumed to correspond to error in the preceding method, the sum of squares corresponding to the variation between the columns, if the sum of squares corresponding to the blocks is assumed to correspond now to the files.

As previously, the sum of squares due to the columns is found as follows:

Column	Sum of values	Squares of the sums of the values	
1	61.94	3836.5636	
2	62.77	3940.0729	
3	58.38	3408.2244	
4	61.88	3829.1344	
5	64.88	4209.4144	
6	58.58	3431.6164	
7	56.81	3227.3761	
Sum	425.22	25882.4022 / 7	= 3697.4860
			-3690.0099
			7.4761

$$\text{Correction: } \frac{(425.22)^2}{49} = 3690.0099$$

The sum of squares due to error is reduced to $28.7809 - 7.4761 = 21.3048$, with 30 degrees of freedom. The variance due to error is then $21.3048/30 = 0.7102$. The standard deviation of the difference between any 2 means is now $\sqrt{2(0.7102)/7} = \sqrt{1.4204/7} = \sqrt{0.2029} = 0.45$. The difference between the means must exceed now the quantity $1.74 \times 0.45 = 0.78$.

Treatment E is thus proved superior to treatments B, F, G and A.

The following table resumes the Analysis of the Variance in this Experiment:

Variance due to:	Degrees of Freedom	Sum of squares	Variance
Columns	6	7.4761	1.2460
Files	6	7.2439	1.2073
Treatments	6	7.2546	1.2091
Error	30	21.3048	0.7102
Total	48	43.2794	

With this ends the illustration of the methods which determine if the difference between the means of the yields of 2 treatments is statistically significant or not, and we will now proceed to illustrate the method which determines if the mean of the differences between the correlated results of two populations is statistically significant or not.

"STUDENT'S" METHOD OF PAIRED RESULTS

Although this method is generally known as the "Student's" Method, Goulden says with a good deal of justice in "Statistical Methods in Agronomic Research" that to call this method in such a way tends to obscure "Student's" real contribution to statistical knowledge.

The best explanation that can be given here of this method is to illustrate it by applying it to the experiment under study. The paired results have to be the corresponding observations of the two series to be compared. For example, as each value of any treatment is found on the same file and in the same column with other two values of any other treatment, it is evident that 14 pairings may be done with the values of any two treatments, just as it was done on finding the coefficients of correlation in a previous method.

Following may be found the illustration of the pertinent comparations to do the interpretation of the results of the experiment by this method:

Between E and A:

Values of E	Values of A	Differences	Squares of the differences
9.50	7.45	2.05	4.2025
9.55	8.91	0.64	0.4096
10.00	9.73	0.27	0.0729
10.30	7.58	2.72	7.3984
8.98	8.36	0.62	0.3844
7.98	7.63	0.35	0.1225
9.13	7.30	1.83	3.3489
7.98	7.30	0.68	0.4624
10.30	9.73	0.57	0.3249
10.00	7.45	2.55	6.5025
9.55	8.36	1.19	1.4161
9.13	8.91	0.22	0.0484
9.50	7.58	1.92	3.6864
8.98	7.63	1.35	1.8225
		Sum = 16.96	30.2024
			-20.5458
			9.6566

Correction = $(16.96)^2/14 = 20.5458$
Mean = 16.96 / 14 = 1.21
Sum of squares/n = 9.6566/14 = 0.6898.
$Z = 1.21/\sqrt{0.6898} = 1.21/0.83 = 1.45.$

For n = 14, this value of "*z*" falls above 0.50 which is that corresponding to a probability in favor of their being different equal to 21/22 and therefore, the two treatments are statistically different. (In the *z* table, *n* is equal to the number of pairs of values compared.)

Since "t" $= z\sqrt{n-1}$ where "n" is the total number of paired observations and therefore "n − 1" is the number of degrees of freedom, it is obvious that the list of "t" may be used to determine the probabilities in connection with this method also.

Using the table or list of "t" to do the interpretation of the results by the use of this method we have:

Variance of a single difference: 9.6566/13 = 0.7428.
Variance of the mean of the differences = 0.7428/14 = 0.0531.
Standard deviation of the mean of the differences = $\sqrt{0.0531}$ = = 0.23.

As there are 14 differences, there are 13 degrees of freedom. The quantity to be exceeded by the mean of the differences is equal to $1.82 \times 0.23 = 0.42$. As the mean of the differences is 1.21, it is therefore significant and the two treatments differ statistically.

In this same way we find the values of "z" for the differences between the treatments E on the one side and treatments D, B, C, F, and G on the other, obtaining the following values for "z":

Between E and D : 0.33
Between E and B : 0.77
Between E and C : 0.55
Between E and F : 0.67
Between E and G : 0.91

As the value of "z" to be exceeded is equal to 0.50 it is evident that treatment E is superior to all the other treatments except treatment D.

TABLE OF "z" AND "t"

"n"	"z"	"t"
1	....	6.95
2	6.95	3.09
3	2.18	2.46
4	1.42	2.22
5	1.11	2.08
6	0.93	2.01
7	0.83	1.96
8	0.74	1.92
9	0.68	1.89
10	0.63	1.87
11	0.59	1.85
12	0.56	1.83
13	0.53	1.82
14	0.50	1.81
15	0.48	1.80
16	0.46	1.79
17	0.45	1.79
18	0.43	1.78
19	0.42	1.78
20	0.41	1.77
21	0.40	1.77
22	0.39	1.76
23	0.38	1.76
24	0.37	1.76
25	0.36	1.75
26	0.35	1.75
27	0.34	1.75
28	0.34	1.75
29	0.33	1.74
30	0.32	1.74
Infinity	0.30	1.69

The values of "z" and "t" in the above table are those which correspond to a probability of 21/22 in favor of said difference being statistically significant. The "n" in the table of "z" means "the number of pairs of values compared" and the "n" in the table of "t" means "degrees of freedom" to which the difference is subject.

NORMAL CURVE PROBABILITIES

Per cent	X	Per cent	X	Per cent	X	Per cent	X
1%	2.58	15%	1.44	50%	0.6745	80%	0.25
2%	2.33	20%	1.28	55%	0.60	85%	0.19
3%	2.17	25%	1.15	60%	0.52	90%	0.13
4%	2.05	30%	1.04	65%	0.45	95%	0.06
5%	1.96	35%	0.93	70%	0.39	96%	0.05
10%	1.64	40%	0.84	75%	0.32	97%	0.04
						98%	0.03
						99%	0.01

The percentage represents the probability which the deviation from the mean of a value taken at random from the whole population has of exceeding the standard deviation of the population by "X" times.

Description of Contents

This paper comprises a brief and simple explanation of the fundamental principles underlying the use of the statistical methods most commonly used in agricultural research. Among the methods compared are Bessel's method, the deviation from the mean method, Fisher's method for the Analysis of Variance (including the following phases: Unrestricted method, method of Randomized Blocks, method of Latin Square), and "Student's" method for Paired Observations. An example for the application of each method and new tables devised to test exclusively whether a difference is statistically significant or not are also presented.

References:

1 **Fisher, R. A.**, Statistical Methods for Research Workers, 4th Ed. Edinburg, 1932.

2 **Forsyth, C. H.**, An Introduction to the Mathematical Analysis of Statistics, New York, 1924.

3 **Fry, Thornton C.**, Probabilities and its Engineering Uses, New York, 1928.

4 **Goulden, C. H.**, Statistical Methods in Agronomic Research, Ottawa, 1929.

5 **Hayes, H. K.**, Control of soil Heterogeneity and the Probable Error Concept in Plant Breeding, Minnesota Agric. Expt. Sta. Tech. Bull. **30**, 1925.

6 **Maskell, E. J.**, Experimental Error. Trop. Agri. (Trinidad) **6**(1) : 5–11; (2) : 45–48; (4) : 97–99, 1929.

7 **Merriman, Mansfield**, A Textbook on the Method of Least Squares, 8th Ed. New York, 1911.

8 **Rietz, H. L.**, Statistical Methods. In Davenport, Principle of Breeding, New York, 1907.

9 **"Student"**, The Probable Error of a Mean, Biometrika, 6(1) : 1–25, 1908.

10 **Turner, P. E.**, 1.—Manurial Experiments with sugar cane I Burying versus non-burying of sulphate of ammonia in relation to its effect on the yield of first ratoon canes. Fourth Cong. Int. Soc. Sugar Cane Tech., 1932. Bull. **14**, 5 p., 1932.

11 ———— Cultivation experiments with sugar cane. I. The effect of type of preparatory tillage on yield of plant canes. Fourth Con. Int. Soc. Sugar Cane Tech. 1932. Bull. **15** 7 p., 1932.

12 ———— Manurial experiments with sugar cane. II. A comparison of the effect of pen manure and single and double applications of sulphate of ammonia, nitrate of soda and nitrolim on the yield of first ratoon canes. Fourth Cong. Int. Soc. Sugar Cane Tech. 1932, Bull. **16**, 7 p., 1932.