THE MOVING AVERAGE AS A BASIS FOR MEASURING CORRELATED VARIATION IN AGRONOMIC EXPERIMENTS [1]

By Frederick D. Richey

Agronomist in charge of Corn Investigations, Office of Cereal Crops and Diseases Bureau of Plant Industry, United States Department of Agriculture

INTRODUCTION

In comparing varieties of cereals it is necessary to determine (1) which are more productive on a given soil in a given year and (2) which are more productive on this same soil over a period of years. Even this information is of value only for the particular soil type on which the experiments were conducted. To make the information of general application, the comparison must be repeated in different places to determine which varieties are more productive on soils of different types or levels of productivity. Each of these determinations is a problem in random sampling. Consequently, the use of the standard error, or its constant function, the probable error, is essential if the results are to be completely evaluated. Moreover, each element of the comparison is more or less distinct. A failure to obtain an accurate evaluation of a difference for a single soil and season can not be compensated for by repetition on different soils or in different seasons. It therefore becomes necessary to determine the probability that a difference in any individual experiment was or was not due to chance, and it is with this phase of the problem that the present paper primarily is concerned.[2]

In varietal experiments, the variation in yield of a variety on different plats may be divided into that due to variation in the sampling of (1) the variety, (2) the soil, and (3) all other conditions, which may be termed variation due to the experiment. The extent to which each of these is a factor is not known, but it generally is conceded that soil heterogeneity is many times as important as both of the other factors combined. The recognition of this importance led to the use of check or control plats, and more and more replications. The use of check plats constituted an effort to measure the extent of soil variation directly. The larger number of replications constitutes an effort to obtain a reliable mean without correcting for soil heterogeneity.

Check plats require land and labor that otherwise could be devoted to additional replications or varieties. Moreover, adjusting yields to checks involves an arbitrary assumption that may or may not be warranted in any given case. On the other hand, increasing the number of replications not only tends to make the experiment more reliable in itself but also reduces the probable error of the mean,

[1] Received for publication Nov. 11, 1925; issued June., 1926.

[2] Throughout the discussion, it is assumed that systematic errors due to competition, differences in stand, and the like were avoided by proper arrangement of the experiment, leaving only the variation due to random sampling to be measured. The probable error gives no information as to the cause of a difference but only the probability that it is significant, i. e., that it is, or is not, due to chance.

Journal of Agricultural Research, Washington, D. C.

Vol. XXXII, No. 12
June 15, 1926
Key No. G-528

provided always that the probable error of a single result is not increased by extending the experiment onto other soil. Unfortunately, the number of replications usually is limited, and the reduction in the probable error of the mean decreases only as the square root of the number of replications increases. Moreover, in practice, increasing the area frequently increases the absolute variation in the experiment. The limit to the reduction in the probable error by replicating therefore soon is reached. Finally, no matter how much of the probable error of the mean may be eliminated by replicating, a large part of the remaining error will be due to soil heterogeneity. This part in no way lessens the accuracy of the results, yet a satisfactory way to eliminate it would be well worth while. With this in mind, the writer (*13*)[3] suggested regression on a moving average as a means of correcting for soil heterogeneity.

At about the same time "Student" (*3*) presented a formula for obtaining a generalized standard error for a difference between any two of *m* varieties in a varietal experiment. There were certain essentials in which this resembled the moving-average method. The two methods also differed fundamentally. It therefore seemed of interest to ascertain just what the differences were. This was done, and the relations found suggested a modification of the moving-average method. It is the purpose of the present paper to present this modification, together with certain simplifications of the moving-average method, and to consider the similarities and differences between it and the method presented by "Student." These similarities and differences depend largely upon the relations between the standard errors of the means of two variables, the standard error of their difference, and the coefficients of correlation and regression. Therefore, before taking up a detailed consideration of the two methods it is desirable to review certain aspects of the standard error of a difference.

THE STANDARD ERROR OF A DIFFERENCE

Using the convenient term "variance" which Fisher (*4, 5*) has applied to the squared standard deviation (σ^2), the variance of a difference between two variables is given by the equation—

$$\sigma^2_{1-2} = \sigma^2_1 + \sigma^2_2 - 2r_{12}\sigma_1\sigma_2 \qquad (1)$$

If the two variables are uncorrelated, the last term on the right of this equation becomes zero, and the variance of the difference may be had from—

$$\sigma^2_{1-2} = \sigma^2_1 + \sigma^2_2 \qquad (2)$$

The value obtained by using equation 2 in any case is identical with that which would be obtained by taking the differences between the two series in every possible way and computing the variance of all such differences. Although equations 1 and 2 give identical results when $r=0$, they differ fundamentally in the philosophy that underlies them. The use of equation 1 implies that facts are known which suggest the possibility of correlation between the variables. The use of equation 2 may imply that there is no cor-

[3] Reference is made by number (italic) to "Literature cited," p. 1175.

relation between the variables, although it more usually implies that no reason is known for making comparisons between the series in any specific order.

An excellent hypothetical example of this difference may be had from an experiment with fruit trees. Assume that two treatments are compared, each treatment being applied to a group of 10 plats, the plats receiving treatment A not being adjacent to those receiving treatment B. Much of the variation among the plats of each treatment probably would be due to soil heterogeneity, the higher-yielding plats being on the more fertile soils, and *vice versa.* In the absence of more definite information there would be no basis on which to pair the plats in the order of their relative production and equation 2 would be used. If, on the other hand, data obtained on the yields of the same plats under uniform treatment prior to beginning the experiment had shown the inherent productivity of the plats to be in this (or any other) order, it would be very desirable to pair the plats as indicated by the preliminary data. Obtaining the variance of these differences would be equivalent in effect to using equation 1.[4] The same logic applies, of course, if the series of plats receiving the two treatments were known to be more comparable in any specific order because of juxtaposition or for any other reason.

It is rare that there is not some obvious reason for pairing certain plats in field experiments with cereals. Not to eliminate the correlated portion of the variance is to make the probable error larger and the results less conclusive than need be. Indeed, as has been pointed out (*12*), one objection that has been raised to the probable error by agronomists in the United States has been that differences which clearly were significant agronomically, frequently were not significant in terms of their probable errors. Much credit is due to Love for calling attention (*8, 9, 11*) to the fact that in these cases the direct computation of the probable error from successive differences could be used more properly, and so obtaining a wider usage of the pairing method. It is unfortunate, however, that Love has failed to distinguish sometimes between the effects of (1) pairing observations and (2) using Student's tables of probability (*1*) in applying what has been called "Student's method."

Thus, Love and Brunson (*12*) have used both the tables for the normal curve and for Student's curve in computing the probable significance of the differences in a number of experiments. In doing this, however, they have used the standard error computed by equation 2 in connection with the Gaussian curve and the standard error computed from successive differences with Student's curve. As there was more or less correlation between the series this really constituted a comparison between Student's probability tables applied to the correct standard error and the more usual tables applied to an incorrect standard error of the difference. Again, Love (*11, p. 88*) states, "The application of Student's method presupposes that experiments are so conducted that pairing may be accomplished."

The result has been to confuse the facts to an unfortunate degree. For example, Wiggans (*15, p. 6*) states that, "By the determination

[4] It should perhaps be noted that when only two items are being compared, equation 1 rarely will be used in practice, as the direct computation from successive differences is easier. The underlying theory, however, is the same.

of the probable error according to Bessel's or Peter's formula as an aid in interpreting the significance of a given lot of data, the consistency of differences in both amount and direction is often completely overshadowed by large seasonal variations. Seasonal differences, which enter so largely in the final expression of the determination of the probable error, affect very little the final results as determined by Student's method and expressed in odds." Again, it is stated in a recent publication (*7*, *p. 6*) that: "Bessel's method has been generally employed rather than Student's because the small number of plants used has made it possible to maintain similar soil conditions and because the comparisons averaged have been made, usually, in a single season, so that systematic errors were not important."

Inasmuch as Student's earlier contribution (*1*) primarily was the determination of probability values for use with small samples, rather than a method for pairing observations to eliminate correlated variation, the extent of the confusion is manifest. Thus, in the first example (*15*), Student's tables were better suited to the data than the ordinary ones because of the small number of seasons sampled. The standard error of the successive differences would have been the same, regardless of which tables were used, and the probability based on this standard error and the ordinary tables would have been unaffected by seasonal variation. In the second example (*7*), the standard error of the difference obtained with equation 2 is correct, but the use of Student's tables, or the more convenient ones published by Love (*10*) from which the odds may be read directly, would have been preferable because of the small number of plants used. In view of these facts, it is interesting that another of Student's early papers (*2*) considered correlation between adjacent plats, and that the method presented in his more recent paper (*3*) is essentially a short method for eliminating the average correlated among the yields in a varietal experiment.

NOTATION

For convenience, the symbols for the constants needed in the moving-average method and that presented by Student will be defined and illustrated in connection with a hypothetical experiment.

An experiment is assumed in which four varieties are replicated five times. To permit a continuous moving average without eliminating end plats, the plats are arranged in a circle. The assumed data are shown in Table 1.

The plat numbers in the first column of this table of course are arbitrary, as the plats have been assumed to be in a circle, with Nos. 1 and 20 adjacent. The plat yields are unimportant. The deviations from the means, in the fourth column, are the basis for all computations, except in the later illustration of Student's full 1923 formula. The assumed circular arrangement of plats permits using all deviations an equal number of times in computing the moving average, or index of productivity, shown in the fifth column. Thus, the I value of 2.0 units for plat 1 is the mean of the deviations of plats 19, 20, 1, 2, and 3. In the last column, the means of the deviations within each replication are the basis for Student's σ^2_G when the experiment is considered as consisting of five replications or groups.

TABLE I.—*Assumed data showing method of (1) adjusting to the regression on a moving average, and (2) adjusting on the basis of replications*

Plat No. [1]	Variety	Assumed yield	F Deviation from varietal mean	I Moving average of deviations	G Mean of replicate deviations
1	2	3	4	5	6
1	A	24	4	2.0	−0.5
2	B	22	0	.4	
3	C	22	−2	−.8	
4	D	22	−4	−2.4	
5	A	18	−2	−3.2	−3.0
6	B	18	−4	−3.2	
7	C	20	−4	−3.2	
8	D	24	−2	−3.2	
9	A	16	−4	−2.0	−1.0
10	B	20	−2	−1.2	
11	C	26	2	−.8	
12	D	26	0	.4	
13	A	20	0	.8	1.0
14	B	24	2	.8	
15	C	24	0	1.2	
16	D	28	2	2.0	
17	A	22	2	2.4	3.5
18	B	26	4	3.2	
19	C	28	4	3.6	
20	D	30	4	3.2	

[1] Plots are assumed to be in a circle with Nos. 1 and 20 contiguous.

$m=4.0$	$s=5.0$	$n=5.0$
$\sigma^2_T=13.0$	$\sigma^2_I=5.184$	$\sigma^2_G=4.7$
$\sigma^2_R=5.0$	$r^2_{FI}=0.8226$	$r^2_{FG}=0.5875$
$\sigma^2_F=8.0$	$\text{Regr.}_{FI}=1.1575$	$\text{Regr.}_{FG}=1.000$

Using Student's notation, σ^2_T is the total variance of the experiment, i. e., the variance obtained when the deviation of each plat is taken from the grand mean without reference to variety or grouping. The variance of the group means is σ^2_G, and σ^2_R is the variance of the varietal or race means. Student further designates as σ^2_e "the casual error, which is the only part subject to random sampling" (*3, p. 282*).

The use of the deviations from the varietal means is necessary in connection with the moving-average method. These deviations, which may be considered as due to field variation apart from varietal influence, will be designated *F*. Their variance, σ^2_F may be computed either directly, or from the equation—

$$\sigma^2_F=\sigma^2_T-\sigma^2_R \qquad (3)$$

The correlation of *F* with *G* necessitates using one *G* value with each of the component plats to obtain the product moment, p_{FG}. Finally, the symbol p_{FF} indicates the product moment, when each *F* value is multiplied by each of the others in the same replication, and the sum of these is multiplied by $\frac{2}{mn(m-1)}$ to obtain the mean. In this terminology, $\frac{\sigma^2_{F-F}}{n-1}$ is the variance for the difference between the means of any two varieties, $(n-1)$ being used instead of n to correct for the small number, in keeping with Student's usage.

ELIMINATING GROUP VARIATION

The method presented by Student in 1923 (*3*) is essentially one for eliminating the average correlated variation among the varieties in the different replications from the average variance of all varieties in the experiment. The equation for accomplishing this, for the proof of which Student gives credit to R. A. Fisher, is—

$$\frac{\sigma^2_e}{n}=\frac{m(\sigma^2_T-\sigma^2_R-\sigma^2_G)}{(m-1)\ (n-1)}.$$

Twice this,

$$\frac{2\ \sigma^2_e}{n}=\frac{2m\ (\sigma^2_T-\sigma^2_R-\sigma^2_G)}{(m-1)\quad (n-1)}, \tag{4}$$

gives the average variance for the differences among the varieties in the experiment.

Student shows that the value for the variance of a comparison from this very simple formula is the same as that obtained by taking the successive differences between A and B, C, D, etc.; B and C, D, etc.; etc.; and averaging these $\frac{m\ (m-1)}{2}$ variances, a method that would be very laborious.

Using this equation with the values obtained from Table I,

$$\frac{2\sigma^2_e}{n}=\frac{2\times4\ (13.0-5.0-4.7)}{3\times4}=2.2$$

This is identical with the value obtained by computing the six variances for the successive differences between A, B, C, and D, averaging them, and dividing by $(n-1)$. Moreover, it is almost 50 per cent less than the variance for a difference obtained by equation 2, thus,

$$\frac{\sigma_{F-F}^2}{n-1}=\frac{8+8}{4}=4.0$$

It may be noted in passing that equation 4 may be used also to eliminate the correlated variation due to seasonal fluctuation from the average standard error for the differences among several varieties or treatments compared in a number of seasons. In such a comparison σ^2_G becomes the variance of the seasonal means. As the final value obtained is the mean of the variances for all possible comparisons, care should be used to include only data that are reasonably comparable. Thus, if one of the treatments might be expected to vary much more widely than the others with differences in seasonal conditions, the use of this equation scarcely would be warranted. It should be borne in mind, moreover, that the variance based upon the deviations of the annual mean values measures only the variation due to fluctuation in soil and climate from year to year. The true variance should measure also the variation within the individual experiments.

REGRESSION ON A MOVING AVERAGE

The method previously suggested by the writer (*13*) for adjusting for soil heterogeneity consisted in (1) expressing each plat yield as a percentage of the mean yield of the variety grown in that plat, (2) computing a moving average based on a group of several contiguous plats, (3) determining the regression of the individual plats on the moving average, and (4) adjusting the individual plat yields on the basis of this regression and the deviations of the moving average.

In using this method it became evident that it was unnecessarily complicated. The use of deviations of the individual plat yields from the mean yields of the varieties grown in them obviated the computation of percentage yields. It also was unnecessary to adjust each individual yield. The corrected variance may be had by multiplying the gross variance by $(1-r^2_{FI})$, where r_{FI} represents the correlation between the individual and the moving-average deviations,or, symbolizing the variance of F for constant I by ${}_I\sigma^2_F$,

$$ {}_I\sigma^2_F = \sigma^2_F\,(1-r^2_{FI}) \tag{5}$$

Finally, the mean of each variety may be adjusted to the mean of all moving-average values for that variety. This eliminates further computation, and usually gives essentially the same final result. Even if varieties differing extremely in yield were included in the experiment, it is questionable whether a corrected error based on percentage deviations would be more reliable than one based on actual deviations.

So much for these changes which are but simplifications. The real modification, suggested by Student's (*3*) presentation, is the inclusion of a further term, $\frac{s}{s-1}$, in equation 5. The final equation for obtaining the variance of a mean difference, after adjusting to the regression on a moving average, then is—

$$\frac{2\,s\,\sigma^2_F(1-r^2_{FI})}{(s-1)\,(n-1)}, \tag{6}$$

where s is the number of plats used in computing the moving average. The reason for including the term $\frac{s}{s-1}$ will be discussed later. It is desirable first to apply equation 6 to the data in Table I,

$$\frac{2\times 5\times 8\ (1-0.8226)}{4\times 4}=0.8870$$

The variance for a difference between the means of any two of the varieties in the hypothetical experiment shown in Table I, therefore, would be 0.8870 after the yields were adjusted to a moving average of 5 adjacent plats. This is materially smaller than the variance of 2.2, obtained by equation 4; and that of 4.0 obtained by equation 2. Accordingly, it is preferable if it can be justified.

COMPARISON OF METHODS

As previously noted, there are certain fundamental similarities and differences between the methods. One important difference was eliminated by introducing the term $\frac{s}{s-1}$ into the equation for correcting the variance by the moving-average method. The effect of this term is the same as that of $\frac{m}{m-1}$ in equation 4, namely, to compensate for the fact that the individual plat deviations are components of the averages with which their correlation was determined and therefore that the coefficient obtained is partly the result of the correlation of the individual deviations with themselves. This may be shown most easily through equation 1.

Modifying equation 1 so that it applies to the variance of a difference between any two of the m variables,

$$\sigma^2_{F-F} = \sigma^2_F + \sigma^2_F - 2r_{FF}\sigma_F\sigma_F = 2(\sigma^2_F - p_{FF}). \tag{7}$$

But, as the means of the deviations of the plats in the replications constitute the deviations of the replications, $\sigma^2_G = p_{FG}$, and

$$p_{FF} = \frac{m\,\sigma^2_G - \sigma^2_F}{m-1}.$$

Substituting the latter equality for p_{FF} in equation 7,

$$\sigma^2_{F-F} = 2\left(\sigma^2_F - \frac{m\,\sigma^2_G - \sigma^2_F}{m-1}\right).$$

or, simplified and with $\frac{1}{n-1}$ introduced to obtain the variance for a mean difference,

$$\frac{\sigma^2_{F-F}}{n-1} = \frac{2m(\sigma^2_F - \sigma^2_G)}{(m-1)(n-1)}. \tag{8}$$

In view of the equality in equation 3, this is equivalent to equation 4.

With the terms $\frac{m}{m-1}$ and $\frac{s}{s-1}$ included, the only essential difference between the methods is that one is based on the regression on a moving average, whereas the other is based on the regression on the means of the replications. Thus, the variance of F for constant G is given by the equation—

$$_G\sigma^2_F = \sigma^2_F(1 - r^2_{FG}), \tag{9}$$

which is analagous to equation 5.

But,

$$r^2_{FG} = \frac{p^2_{FG}}{\sigma^2_F\,\sigma^2_G},$$

and

$$p_{FG} = \sigma^2_G.$$

Therefore,

$$r^2_{FG} = \frac{\sigma^2_G}{\sigma^2_F}.$$

Substituting this in equation 9, and simplifying,

$$_G\sigma^2_F = \sigma^2_F - \sigma^2_G, \qquad (10)$$

which, with the term $\frac{m}{m-1}$ included, and multiplied by $\frac{2}{n-1}$ to give the variance of a mean difference, gives a result identical with that obtained by equations 4 or 8.

Both methods, therefore, consist essentially in obtaining the average variance of the individual plats for a constant group value, and using this in place of the total variance in determining the standard error of a difference.

So much for the resemblance between the methods. The question then is whether the replication average or the moving average is a better basis for measuring the variation due to soil. It seems reasonably clear that the moving average is better in any case where the number of plats used in computing the moving average approximates the total number of strains being compared. In the hypothetical case under consideration, for example, strain A always occurs adjacent to strain D in the previous replication, whereas it is two plats removed from strain D in the same replication. With a 5-plat moving average, the corrected variance for A is based on its correlation with B in the same replication, with C in the same and in the previous replications, and with D in the previous replication. The corrected variances for the other sorts are on a corresponding basis. That is, the correlated variation for which correction is made with the moving-average method is that between plats of one strain with the nearest plats of the other strains, rather than with the plats of the other strains occurring in the same more or less arbitrary replication groups.

In order to compare the methods in an actual experiment, the moving-average method has been applied to the data presented by Student (*3*). These are the yields of individual plats of barley in Beaven's No. 1 yield experiment of 1913. The data comprise a comparison of eight strains of barley compared in square plats arranged in chessboard fashion in a strip of 5 plats wide by 32 plats long. Each plat was formed by sowing eight rows 4 feet long and 6 inches apart, the seed being spaced 2 inches apart in the row. One row on each side and 6 inches at each end of the other rows of each plat were rejected at harvest to avoid the effects of competition. This left the yields of plats 1 yard square as the basis of comparison. The yields and planting order are shown in Table II, being the same as shown by Student in Diagram 1 (*3, p. 277*). The mean yields of the strains were computed first (Table IV, column 2), and the deviation of each plat from the mean of the strain grown in it was obtained (Table II, columns 3). A moving average of the deviations of 4 or 5 adjacent plats then was computed. For plats 2 to 31 inclusive in Series II, III, and IV, the 5 deviations averaged were those of a central plat with the plats to the north, south, east, and west of it. The same principle was followed around the perimeter of the field, but only four plats were available for averaging. The average value for a corner plat was the mean deviation of the four plats in that corner. The moving-average values, or indexes of productiveness, for all plats are shown in Table II, columns 4.

TABLE II.—*The yields, deviations from the means, and indexes, in grams, for 160 plats of barley grown in Beaven's No. 1 yield experiment of 1913* [1]

Plat No.	Series I				Series II				Series III				Series IV				Series V			
	Strain	Yield	Deviation	Index	Strain	Yield	Deviation	Index	Strain	Yield	Deviation	Index	Strain	Yield	Deviation	Index	Strain	Yield	Deviation	Index
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	A	236.5	−61.6	−57.6	B	265.9	−34.8	−54.4	C	265.6	−53.1	−48.5	D	255.9	−39.4	−61.1	E	230.1	−76.3	−60.7
2	F	210.4	−66.1	−55.8	G	236.7	−67.9	−47.0	H	205.0	−66.8	−64.6	A	222.6	−75.5	−58.2	B	249.3	−51.4	−48.9
3	C	291.1	−27.6	−35.3	D	295.8	+0.5	−42.9	E	246.7	−59.7	−41.4	F	218.7	−57.8	−49.2	G	312.2	+7.6	−33.4
4	H	223.9	−47.9	−36.5	A	238.3	−59.8	−28.4	B	277.4	−23.3	−55.1	C	258.2	−60.5	−34.4	D	263.4	−31.9	−29.9
5	E	295.6	−10.8	−20.1	F	265.0	−11.5	−30.4	G	232.3	−72.3	−23.0	H	273.3	+1.5	−40.8	A	263.4	−34.7	−17.1
6	B	290.5	−10.2	−9.3	C	320.9	+2.2	−5.5	D	286.0	−9.3	−25.9	E	268.2	−38.2	−9.7	F	273.4	−3.1	−20.0
7	G	286.1	−18.5	−14.2	H	272.9	+1.1	−7.0	A	286.1	−12.0	−13.4	B	301.1	+0.4	−14.5	C	314.9	−3.8	−20.1
8	D	266.2	−29.1	−15.5	E	298.7	−7.7	−15.1	F	229.3	−47.2	−36.7	G	285.7	−18.9	−44.6	H	198.0	−73.8	−43.8
9	A	291.6	−6.5	−14.4	B	307.9	+7.2	−21.6	C	221.2	−97.5	−64.4	D	212.0	−83.3	−55.9	E	227.9	−78.5	−77.7
10	F	247.3	−29.2	−9.1	G	301.3	−3.3	−35.5	H	170.7	−101.1	−55.6	A	296.7	−1.4	−59.9	B	225.7	−75.0	−48.5
11	C	321.4	+2.7	−24.9	D	244.2	−51.1	−32.3	E	231.8	−74.6	−66.2	F	237.7	−38.8	−32.8	G	265.4	−39.2	−61.3
12	H	249.9	−21.9	−18.5	A	263.1	−35.0	−38.4	B	235.5	−65.2	−41.0	C	308.8	−9.9	−50.6	D	203.0	−92.3	−50.7
13	E	286.5	−19.9	−28.7	F	257.5	−19.0	−18.9	G	284.4	−20.2	−41.4	H	225.0	−46.8	−30.8	A	236.9	−61.2	−62.2
14	B	246.6	−54.1	−19.5	C	318.2	−0.5	−13.9	D	239.7	−55.6	−17.6	E	290.4	−16.0	−34.0	F	228.0	−48.5	−37.7
15	G	301.0	−3.6	+10.0	H	331.7	+59.9	+15.5	A	302.2	+4.1	+1.5	B	297.5	−3.2	−2.0	C	293.8	−24.9	−27.7
16	D	333.0	+37.7	+18.9	E	324.0	+17.6	+23.7	F	278.8	+2.3	+7.3	G	334.4	+29.8	−3.6	H	237.6	−4.2	+0.2
17	A	322.0	+23.9	+17.1	B	301.7	+1.0	+9.4	C	301.2	−17.5	−7.7	D	282.6	−12.7	+21.2	E	336.6	+30.2	+3.2
18	F	282.4	+5.9	+33.2	G	326.8	+22.2	+9.6	H	260.3	−11.5	+18.2	A	374.3	+76.2	+21.1	B	330.2	+29.5	+38.7
19	C	399.3	+80.6	+43.9	D	325.5	+30.2	+27.3	E	328.0	+21.6	+17.5	F	300.4	+23.9	+40.0	G	323.5	+18.9	+29.5
20	H	330.8	+59.0	+29.8	A	280.2	−17.9	+26.6	B	324.1	+23.4	+17.5	C	377.9	+59.2	+36.7	D	341.0	+45.7	+38.2
21	E	303.9	−2.5	+34.4	F	315.0	+38.5	+1.3	G	305.9	+1.3	+21.6	H	303.0	+31.2	+33.1	A	327.0	+28.9	+34.4
22	B	343.3	+42.6	+2.0	C	305.9	−12.8	+15.8	D	308.7	+13.4	−0.6	E	351.5	+45.1	+33.1	F	308.1	+31.6	+28.0
23	G	285.2	−19.4	−0.3	H	269.2	−2.6	−3.0	A	248.0	−50.1	+0.2	B	344.7	+44.0	+11.7	C	324.9	+6.2	+27.2
24	D	273.5	−21.8	+16.0	E	376.4	+70.0	+11.1	F	272.7	−3.8	+6.9	G	317.7	+13.1	+29.5	H	298.7	+26.9	+18.7
25	A	333.4	+35.3	+11.6	B	314.2	+13.5	+39.1	C	324.2	+5.5	+21.2	D	362.4	+67.1	+29.6	E	334.8	+28.4	+34.6
26	F	295.9	+19.4	+42.5	G	376.0	+71.4	+52.2	H	295.6	+23.8	+16.3	A	331.8	+33.7	+32.5	B	316.6	+15.9	+7.9
27	C	362.6	+43.9	+68.6	D	428.4	+133.1	+60.6	E	253.5	−52.9	+22.4	F	298.7	+22.2	+0.9	G	258.2	−46.4	−1.0
28	H	349.8	+78.0	+79.6	A	405.7	+107.6	+78.2	B	286.7	−14.0	+31.3	C	366.4	+47.7	+22.8	D	299.8	+4.5	+11.7
29	E	395.3	+88.9	+88.3	F	362.9	+86.4	+82.4	G	372.6	+68.0	+44.2	H	325.6	+53.8	+56.2	A	338.9	+40.8	+35.9
30	B	400.6	+99.9	+77.6	C	379.9	+61.2	+70.3	D	322.3	+27.0	+58.3	E	377.0	+70.6	+50.0	F	321.0	+44.5	+38.6
31	G	364.9	+60.3	+75.9	H	348.6	+76.8	+65.6	A	362.9	+64.8	+54.5	B	354.7	+54.0	+41.0	C	317.2	−1.5	+23.0
32	D	361.8	+66.5	+67.1	E	371.3	+64.9	+64.5	F	326.3	+49.8	+49.2	G	321.8	+17.2	+29.0	H	266.8	−5.0	+16.2

[1] Yield data and arrangement taken from Student, on testing varieties of cereals (*3, p. 277*).

The coefficient of correlation between the deviations of the individual plats and the indexes was determined from the correlation table (Table III). The average variance for a single plat, and for the mean of 20 plats, was computed from the same table.

TABLE III.—*Correlation table for the deviations in yield (in grams) of individual plots of barley, and the indexes for such plots computed as a moving average for values in columns 3 and columns 4 of Table II*

INDEXES

Class centers	−75	−65	−55	−45	−35	−25	−15	−05	+05	+15	+25	+35	+45	+55	+65	+75	+85	Total
INDIVIDUAL PLATS																		
−105			1															1
−95		1	1															2
−85			1															1
−75	1	2	1	2		1												7
−65		2	2	2	1													7
−55				4	1	1	2		1		1							10
−45					4			1										5
−35		2	1		2	1	1	1	1									9
−25			1	1	1	1	2	1		1								8
−15				1	2	2	3	3		2	2	1						16
−05			2		1	2	4	2	2	1	1	1						16
+05				2	1	2	1	2	3	1	3	1						16
+15								1	1		4	1	1					8
+25								1	2	5		4		1				13
+35									2	2	2	2						8
+45									1	1	1	4	1		1			9
+55										1	1	1	1	1				5
+65										1	1		1	1	2	2		8
+75											1		1	1	1	1		5
+85													1				2	3
+95																1		1
+105																1		1
+115																		0
+125																		0
+135															1			1
Total	1	7	10	12	13	10	13	12	13	15	17	15	6	4	5	5	2	160

Individual plats: $\sigma_F=46.165$ gm, $\sigma^2_F=2{,}131.25$ gm. Indexes: $\sigma^2_I=38.389$ gm, $\sigma_I=1{,}473.73$ gm. $r=0.8386$, $r^2=0.7032$, $(1-r^2)=0.2968$. Regression of individual plats on indexes=1.009.

Letting F represent the deviations of the individual plats, and I represent the moving-average values as before, the multiplication of the gross mean variance for the 20 plats of a strain by $\frac{s}{s-1}(1-r^2_{FI})$, or $\frac{4.5}{3.5}(0.2968)$,[5] gives the corrected variance. The corrected variances obtained in this way are shown in column 4 of Table IV. The sum of the variances for any two strains gives the variance for a difference between them. If only the average corrected variance is desired, it may be obtained directly from the average gross variance. Thus,

$$112.2\times\frac{4.5}{3.5}(0.2968)=42.8,$$

the average variance for any strain mean. Twice this, 85.6, is the variance for a difference between any two of the strain means. The square root of this (9.25 gms.) times 0.6745 gives 6.2 gms. as the average probable error for a difference, or about 2.1 per cent of the

[5] s is taken as 4.5, the unweighted average of the 4 and 5 plats used in computing the moving average.

mean yield. This is materially lower than the error for a difference of 3.4 per cent based upon the gross variance of 112.2 gms. It also is 12.5 per cent less than the error of 2.4 per cent reported by Student (*3*), obtained by the equivalent of the pairing method.

TABLE IV.—*The mean yield in grams, the variance of the mean before and after adjusting to its regression on a moving average, and the correction for regression and for position, for eight strains of barley*

Variety	Mean yield [1]	Gross variance [1]	Variance after adjusting [2]	Correction for position [3]	Correction for regression [4]	Corrected yields on regression basis
1	2	3	4	5	6	7
A. English Archer	298.1	130.8	49.9	+2.3	+0.625	298.7
B. 145	300.7	95.0	36.3	+1.7	+2.125	302.8
C. 145/146	318.7	96.6	36.9	+1.0	+.125	318.8
D. Plumage	295.3	153.4	58.5	+.3	+.125	295.4
E. Early Archer	306.4	133.9	51.1	−.3	−1.375	305.0
F. Irish Archer	276.5	76.0	29.0	−1.0	−2.375	274.1
G. 7 A	304.6	80.2	30.6	−1.7	+2.125	306.7
H. Biffin	271.8	131.3	50.1	−2.3	−1.375	270.4
Average	296.5	112.2	42.8	0	0	296.5

[1] The means and variances in columns 2 and 3 differ slightly from those shown by Student (*3*), due to the fact that the means were computed directly without grouping. As all other computations made in this paper were based upon deviations from these means, similar discrepancies occur elsewhere.

[2] Raw variance multiplied by $\frac{4.5}{3.5}$ (0.2968), or 0.3816.

[3] Student's values, although their use was not recommended by him.

[4] Computed from class values shown in Table III, and corrected for mean of all index values, which was 0.125.

This difference is due largely to the fact that the correlated variation removed by grouping according to replication is not as great as when the moving average forms the basis. Thus, the plats of A always occur at some distance from the plats of D within replications. There are however 19 possible comparisons between A and D in adjacent plats, i. e., A in one replication with D in the preceding replication. When the comparisons are made in these two ways, the results are:

Comparison within replications A–D $= 2.8 \pm 7.82$.

Comparison of adjacent plats A–D $= 9.6 \pm 7.15$.

The variation in the mean difference is due to the exclusion of A in Series I, plat 1 from the comparison of adjacent plots. The reduction in the probable error is 8.6 per cent for this comparison, as compared with 12.5 per cent for the experiment as a whole. Strains A and D were among the most variable in the experiment. Adjusting the means and variances of A and D on the moving-average basis gives a difference of 3.3 ± 7.02. This difference is more in accord with that obtained when all plats are averaged, while the variance approximates that obtained by comparing only adjacent plats.

If the yield of each plat is adjusted to its index, the mean yield as well as the variance may be modified. As already noted, however, it is unnecessary in practice to adjust the yields of the individual plats, as the mean yield of each strain may be adjusted to the mean of the indexes for that strain. Thus, the means of the indexes for each strain in Beaven's experiment, with the signs changed, are shown in column 6 of Table IV. The regression of F on I is given by $r_{FI}\frac{\sigma_F}{\sigma_I} = 1.009$, so that the means may be corrected directly as indicated by the heading of column 6, "Correction for regression." The mean of the deviations of any strain of course is 0.[6] The mean of the indexes for a strain, therefore, is unaffected by the fact that the deviations of that strain were components of the indexes. Consequently no term comparable to $\frac{s}{s-1}$ has been used in the regression. Student (*3*) presented values for correction for position based on the facts that the plats became more productive toward one side, and that the varieties occurred in the same relation throughout. Student's values are shown in column 5 of Table IV. It is only fair to note, however, that Student did not recommend such correction for position. The correction values obtained in the two ways are in agreement as to sign except for strain G. The corrections for regression fail strikingly to present the orderly change from the maximum minus to the maximum plus of the other method. None of the corrections are important. Nevertheless, the mean index values furnish additional evidence on the accuracy of the experiment that is of some value. They would have been more reliable, however, if the order of the varieties had been changed in the different replications.

CONCLUSION

Yields obtained in comparisons of varieties or strains of cereals are relative, not absolute. They fluctuate from plat to plat, and from season to season, as a result of variation in soil, climate, plants, and other conditions of the experiment. Granted that gross or systematic errors have been avoided, the reliance that can be placed in the data depends upon the adequacy of random sampling. The best measure of this is the standard error, or its constant function, the probable error. From this may be derived a quantitative expression of the probability that a given result is, or is not, due to chance. The probable error, therefore, becomes a valuable tool in agronomic experiments. Unfortunately, however, it has certain limitations, and it has been misapplied and misinterpreted so frequently in the past as to make some investigators question the desirability of using it at all.

A frequent fault in using the probable error has been to draw conclusions that the odds were greatly in favor of a certain treatment or variety being superior, when as a matter of fact the probable error gave no information as to this. Thus, for example, the individual plants in two rows have been noted for some character, the probable error for the mean of each row and for the difference between the two means computed, and the conclusion drawn, in terms of this probable error, that there was a significant difference between the varieties or strains grown in the two rows. Evidently no conclusion is warranted as to the cause of the difference in such a case. It may have been due to soil, or seed, or both. The blame for a wrong conclusion, however, rests on the investigator making it rather than

[6] This is not strictly true in the present case as the influence on the index of any given plat in the perimeter is one-fourth its value, whereas in the body of the experiment it is only one-fifth.

on the probable error. Fortunately, such misuse of the probable error is becoming less with a better understanding of its real functioning.

It is with two other difficulties, however, that the present paper is concerned. One of these is in the fact that the limitations of land and labor frequently prevent enough replications of any one comparison to permit a strict statistical interpretation of the standard error in terms of probability. The method of computing a generalized error for the experiment, based on the average coefficient of variation, as suggested by Hayes (*6*), or on the average variance, as shown by Student (*3*) and used by the writer (*14*) and others, helps to overcome this difficulty. The other difficulty is that the standard error is a measure of the gross or total variation. In most field experiments a certain amount of the variation is concomitant. The productivity of the soil does not change abruptly but more or less gradually, so that adjacent plats are more nearly alike than plats that are farther apart. In general, moreover, varieties tend to respond similarly to these differences in soil productivity, as well as to climatic changes. Not to eliminate correlated variation from the standard error before using it to determine probability is to make the results seem less conclusive than they really are.

Where only two items are involved, pairing observations is a simple, direct way of eliminating correlated variation. When several items are being compared, however, this becomes unwieldy in the extreme. The method presented by Student (*3*) provides a quick, satisfactory means for eliminating the average correlated variation and providing a generalized error for the experiment. When many strains are being compared, as frequently is the case, the replication groups are large, the intergroup correlations are likely to be small, and correlated variation within these groups, of which no account is taken in this method, becomes more important. This seems to the writer to be a serious drawback. The moving average is a more flexible measure, the number of plats used in the average being changeable to meet the conditions of the experiment. It provides, furthermore, an index of the average productivity of the soil on which each strain was grown, whether this be used for adjusting yields or not. Even with few strains in the experiment, it would seem that the moving average provides a better basis for measuring correlated variation than the replication average. The difference would be slight, however, and probably would be more than offset by the greater convenience of the other method. On the other hand, Student's method (*3*) is a straightforward statistical process involving no assumptions whatever. This is true of the moving-average method only when the number of plats used in computing the moving average equals or exceeds the number of strains in the experiment. In other cases, the use of the moving average assumes that such an average gives a fair index of the productivity from plat to plat. This assumption seems entirely justified. The use of the moving average as a basis for eliminating correlated variation consequently seems safe and logical.

The principle involved in the method presented by Student (*3*) would seem to have a very important application in interpreting experiments repeated over a number of seasons. Here the grouping is determined arbitrarily, and the correlations are likely to be large. The use of equation 4 or equation 8, with *G* representing the seasonal means, should do much to eliminate some of the present cumbersome comparisons by pairs.

LITERATURE CITED

(1) ANONYMOUS.
1908. THE PROBABLE ERROR OF A MEAN. By Student. Biometrika 6: 1–25, illus.

(2) ———
1911. NOTE ON A METHOD OF ARRANGING PLOTS SO AS TO UTILIZE A GIVEN AREA OF LAND TO THE BEST ADVANTAGE IN TESTING TWO VARIETIES. By Student. Jour. Agr. Sci. 4: 128–132, illus.

(3) ———
1923. ON TESTING VARIETIES OF CEREALS. By Student. Biometrika 15: 271–293, illus.

(4) ———
1924. NOTE BY STUDENT WITH REGARD TO HIS PAPER "ON TESTING VARIETIES OF CEREALS." Biometrika 16: 411.

(5) FISHER, R. A.
1918. THE CORRELATION BETWEEN RELATIVES ON THE SUPPOSITION OF MENDELIAN INHERITANCE. Roy. Soc. Edinb. Trans. 52 (2): 399–433, illus.

(6) HAYES, H. K.
1923. CONTROLLING EXPERIMENTAL ERROR IN NURSERY TRIALS. Jour. Amer. Soc. Agron. 15: 177–192.

(7) LOOMIS, W. E.
1925. STUDIES IN THE TRANSPLANTING OF VEGETABLE PLANTS. N. Y. Cornell Agr. Expt. Sta. Mem. 87, 63 p., illus.

(8) LOVE, H. H.
1919. THE EXPERIMENTAL ERROR IN FIELD TRIALS. Jour. Amer. Soc. Agron. 11: 212–216.

(9) ———
1923. THE IMPORTANCE OF THE PROBABLE ERROR CONCEPT IN THE INTERPRETATION OF EXPERIMENTAL RESULTS. Jour. Amer. Soc. Agron. 15: 217–224.

(10) ———
1924. A MODIFICATION OF STUDENT'S TABLE FOR USE IN INTERPRETING EXPERIMENTAL RESULTS. Jour. Amer. Soc. Agron. 16: 68–73.

(11) ———
1924. THE RÔLE OF STATISTICS IN AGRONOMIC EXPERIMENTATION. Sci. Agr. 5: 84–92.

(12) ——— and BRUNSON, A. M.
1924. STUDENT'S METHOD FOR INTERPRETING PAIRED EXPERIMENTS. Jour. Amer. Soc. Agron. 16: 60–68.

(13) RICHEY, F. D.
1924. ADJUSTING YIELDS TO THEIR REGRESSION ON A MOVING AVERAGE, AS A MEANS OF CORRECTING FOR SOIL HETEROGENEITY. Jour. Agr. Research 27: 79–90, illus.

(14) ——— and MAYER, L. S.
1925. THE PRODUCTIVENESS OF SUCCESSIVE GENERATIONS OF SELF-FERTILIZED LINES OF CORN AND OF CROSSES BETWEEN THEM. U. S. Dept. Agr. Bul. 1354. 19 p., illus.

(15) WIGGANS, R. G.
1924. EXPERIMENTS IN CROP ROTATION AND FERTILIZATION. N. Y. Cornell Agr. Expt. Sta. Bul. 434, 56 p., illus.