Analysis of Combined Data for Identical Replicated Experiments

By E. B. Roessler and L. D. Leach, *University of California, Davis, Calif.*

Frequently in order to test the merits of different varieties or treatments at each of several locations and also to investigate variety or treatment differences for the combined data, identical replicated experiments are conducted at the several locations.

If one is interested only in variety or treatment differences for these particular locations, not considered as a sampling region, then only the error variance for each location and variety or treatment totals are necessary for a combined analysis. In this paper is outlined a short method for making such an analysis.

The analysis of variance is based on the assumption that the experimental errors to which the data are subject are independently and normally distributed with the same variance, the latter restriction being the most important. Sets of measurements may be validly combined to estimate an experimental error for the pooled data if, and only if, all sets may reasonably be expected to have the same variance, that is if their variances form a homogeneous group. If the variances of the original sets are homogeneous, then the sets may be considered as random samples from the same population; but if the variances are not homogeneous, the original sets may not be considered as belonging to the same population; and, therefore, the pooled or average error is not really applicable to any of the original sets. Methods for determining homogeneity of sets of data have been discussed by Neyman and Pearson (1), Bartlett (2), and Roessler (3); and methods of combining sets of data are given by Hayes and Immer (4) and Snedecor (5).

If identical replicated experiments to test the merits of various treatments have been conducted at different locations, loss of information may result unless the data are subjected to one or more combined analyses. If only the effect of treatment is being considered and the variation due to location and that due to interaction are of no interest, the method of analysis consists of combining into a single analysis of variance those experiments whose variances form a homogeneous set, and then selecting from the complete analysis that portion necessary for testing the effect of treatments. The analysis can be divided into the following parts:

I. Determination of the homogeneity of the set of location variances.

II. Determination of significant treatment differences by calculating an appropriate F-value.

III. Determination of the difference between means of treatments required for significance.

The procedure is outlined below.

Procedure

k = number of locations

s = number of seed treatments including control

r = number of replications of each treatment at each location
v_i = error variance at each location
m = number of degrees of freedom for each location error variance[1]
$\bar{v}$ = average error variance for all locations.

I. The average variance v may be used to represent all the location variances from which it is determined if, and only if, the location variances form a homogeneous set. Homogeneity of the set of location variances may be ascertained by calculating the following approximate expression for chi-square, (X^2):

$$X^2 = 2.3026\ (m - \tfrac{1}{3})\ (k \log_{10}\bar{v} - \Sigma \log_{10} v_i)$$

with $k - 1$ degrees of freedom.

If this leads to a significantly high value of X^2, extremely high or low location variances or both must be rejected until a non-significant value of X^2 is reached. Locations whose error variances have been rejected cannot be included in the combined analysis. If many variances must be rejected, it is better to group the data into two or more sets, the variances in each set being homogeneous. In this way all the data are included.

II. Calculate the F-value for each group whose variances form a homogeneous set as follows:

$$F = \frac{v_s}{\bar{v}} \begin{matrix} \text{with } s - 1 \text{ degrees of freedom} \\ \text{with } km \text{ degrees of freedom} \end{matrix}$$

where v_s is the variance between the means of treatments in the combined analysis of variance and is given by the expression

$$v_s = \frac{kr\left[\Sigma M^2 - \dfrac{(\Sigma M)^2}{s}\right]}{s - 1}$$ where M is each treatment mean for the combined data.

III. If the F-value obtained above is not significant, no treatment differences exist. If F is significant, calculate for the combined data the difference between means of treatments required for significance, *i. e.*

$$d = t\sqrt{\frac{2\bar{v}}{kr}}$$ where the value of t depends upon km degrees of freedom.

As an illustration the method will be applied to the 1942 data of the cooperative lima bean-seed-treatment tests compiled by Dr. J. C. Walker and conducted under the auspices of the committee for coordination in Cereal and Vegetable Seed Treatment Research of the

[1]If one or two location variances are based on slightly different numbers of degrees of freedom, no appreciable error will be made in assuming for them the number of degrees of freedom of the rest of the set.

American Phytopathological Society. The seed protectants Thiosan[2] and Spergon were applied to seeds of the Fordhook variety. Tests were conducted at 25 locations in 20 states. Each test consisted of five replications of 100 seeds of each treatment and one untreated check. In Table I are given for each location the average seedling emergence, the error variance, and its common logarithm.

TABLE I—SUMMARY OF LIMA BEAN TESTS CONDUCTED ON FORDHOOK VARIETY — 1942

Location	Test No.	Average Number of Seedlings			Error Variance	Log v_i
		Thiosan	Spergon	Check		
N. Y. (Staten Island)	16	62.2	—	66.6	211.4	2.3251
Texas	22	43.0	45.8	29.0	149.7	2.1752
Virginia	24	80.4	84.6	70.0	78.4	1.8943
Iowa	6*	76.6	72.2	48.4	73.7	1.8675
California	1	79.6	81.2	74.0	70.5	1.8482
Connecticut	2*	62.4	66.0	37.8	64.2	1.8075
Miss. (Holly Springs)	11*	33.4	25.0	17.2	55.4	1.7435
Illinois	5*	38.8	49.4	27.8	47.5	1.6767
S. C. (Edisto Station)	19*	44.6	54.2	25.0	42.5	1.6284
Virginia	23*	79.6	72.4	57.0	38.0	1.5798
S. C. (Charleston)	20*	80.4	87.8	43.0	30.8	1.4886
Miss. (State College)	10*	62.4	64.6	36.2	30.0	1.4771
Wisconsin (Madison)	28*	81.6	80.0	61.6	29.4	1.4684
Minnesota	8*	14.0	23.8	8.2	22.5	1.3522
Oklahoma	18*	73.8	73.6	63.6	20.2	1.3054
New Jersey	12	67.2	71.8	74.6	18.6	1.2695
Ohio	17	72.0	76.2	75.0	16.3	1.2122
Miss. (Chrystal Spring)	9	87.2	88.4	85.8	15.6	1.1931
N. Y. (Geneva)	13	73.0	67.4	67.4	14.5	1.1614
Florida	3*	79.2	87.0	78.8	13.7	1.1367
N. Y. (Long Island)	14	89.0	84.6	85.4	10.6	1.0253
West Virginia	26*	85.6	90.8	84.0	9.9	0.9956
Georgia	4	81.6	76.8	73.0	7.1	0.8513
Maryland	7	88.2	89.0	86.8	3.9	0.5911
South Dakota	21	−8.0	7.0	11.2	1.8	0.2553
Totals	—	—	—	—	1076.2	35.3294

*Significant differences between treatments were indicated by the analyses at these locations.

PROCEDURE

I. *Test for homogeneity of variance.*

In the case of all the data, $m = 8$, $k = 25$, $s = 3$, $r = 5$.

$$v = \frac{1076.2}{25} = 43.0, \log v = 1.6340, \Sigma \log v_i = 35.3294$$

$$X^2 = 2.3026\ (m - \tfrac{1}{3})\ (k \log v - \Sigma \log v_i)$$
$$= 2.3026\ (8 - \tfrac{1}{3})\ [25\ (1.6340) - 35.3294]$$
$$= (17.65)^*\ (5.521)$$
$$= 97.4 \text{ with 24 degrees of freedom}$$

Since this value exceeds the tabular value of 36.4 at the 5 per cent level of significance, the variances are not homogeneous and the data cannot be analyzed as a unit.

[2]The material designated as Thiosan in these tests was later introduced commercially under the name of Arasan.

*It should be noted that in the approximation for X^2 the term $2.3026\ (m-{}^1/_3) = 17.65$ is a constant for all tests with $m = 8$.

An examination of the variances in Table I shows that those of test 16 and 22 are considerably larger than those of the other tests. If these two be rejected, the remaining tests may be analyzed for homogeneity. In this case m = 8, k = 23, s = 3, r = 5.

$$v = \frac{715.1}{23} = 31.1, \log v = 1.4926, \Sigma \log v_i = 30.8291$$

$X^2 = (17.65)\ (34.3298 - 30.8291) = 61.8$ with 22 degrees of freedom.

For 22 degrees of freedom at the 5 per cent level the value of X^2 not to be exceeded is 33.9; therefore the above set of variances is still not homogeneous. The 11 top variances must be rejected for the remaining set to be homogeneous whereas the omission of tests 16 and 22 at the top and 26, 4, 7, and 21 at the bottom leave for the remaining tests a value of $X^2 = 27.1$ with 18 degrees of freedom. The value of 5 per cent level which must not be exceeded is 28.9.

Therefore, in order to have one homogeneous set, the fewest variances which can be rejected are the two at the top and four at the bottom. The averages for the remaining 19 tests are shown in Table II.

TABLE II—AVERAGES FOR 19 TESTS FORMING A HOMOGENEOUS SET

Number of Locations	Average Number of Seedlings			Average Error Variance
	Thiosan	Spergon	Check	
19	67.1	69.0	54.6	36.4

II. *Calculation of F-value.*

$$F = \frac{v_s}{v} \text{ where } v_s = \frac{kr\left[\Sigma M^2 - \frac{(\Sigma M)^2}{s}\right]}{s-1}$$

$$\Sigma M^2 = (67.1)^2 + (69.0)^2 + (54.6)^2 = 12{,}244.57$$

$$\frac{(\Sigma M)^2}{s} = \frac{(67.1 + 69.0 + 54.6)^2}{3} = 12{,}122.16$$

$$v_s = \frac{19\ (5)\ (12{,}244.57 - 12{,}122.16)}{2} = 5814$$

$$F = \frac{5814}{36.4} = 160 \text{ with 2 degrees of freedom for } v_s \text{ and 152}$$

degrees of freedom for v. From a table of F-values at the 5 per cent level of significance F = 3.06; and, therefore, significant differences between the means of treatments are indicated.

III. *Determination of the difference between means required for significance*

$$d = t\sqrt{\frac{2v}{kr}}$$

For km = 19 (8) = 152 degrees of freedom at the 5 per cent level of significance t = 1.98. Therefore

$$d = 1.98\sqrt{\frac{2\ (36.44)}{19\ (5)}} = 1.7 \text{ seedlings.}$$

At the 5 per cent level both seed protectants are significantly better than the control and Spergon is better than Thiosan.

The above analysis required the rejection of six of the 25 tests. Obviously the fewer tests that need be omitted the more representative are the results. Therefore, in order to avoid rejecting tests we will indicate a second method of procedure in which the tests will be grouped in homogeneous sets. Since the error variance of test 16 is so extremely high, and since some of the data of this test are missing, it will be omitted from the analysis.

The remaining tests will be more or less arbitrarily separated into two groups, one made up of 12 tests with greatest variance (22–28 inclusive in Table I), and the other, the remaining 12 tests (8–21 inclusive). Homogeneity-of-variance tests for the two groups lead to the values $X^2 = 10.8$ and 17.2 respectively with 11 degrees of freedom. The value at the 5 per cent level not to be exceeded for homogeneity is $X^2 = 19.7$; and, therefore, the variances for each group form a homogeneous set. In Table III are indicated the data for the two groups.

TABLE III—AVERAGES FOR DATA SEPARATED INTO TWO HOMOGENEOUS GROUPS

Group	Number of Locations	Average Number of Seedlings			Average Error Variance
		Thiosan	Spergon	Check	
1.....	12	63.57	65.27	43.92	59.2
2.....	12	68.23	69.70	66.15	12.9

For group 1, F = 8448/59.2 = 143, and for group 2, F = 14.8 with 2 degrees of freedom for v_s and 96 degrees of freedom for v. At the 5 per cent level of significance F = 3.09; and, therefore, significant differences between the means of treatments are indicated for both groups. At the 5 per cent level the difference between means required for significance is 2.8 seedlings for group 1 and 1.3 seedlings for group 2. In both cases the increase in the number of seedlings due to the use of either of the seed protectants is significant, but only in the second group is Spergon significantly better than Thiosan.

Conclusion

Either combined analysis indicates that for these particular tests seed treated with either Thiosan or Spergon produced significantly larger stands of seedlings than did untreated seed. In the first analysis and in the second group of the second analysis Spergon appeared to be a better seed protectant than Thiosan at the dosages tested. The second analysis (grouped data) is preferable since it includes all the data (except test 16).

In 12 of 25 locations individual analyses indicated no significant differences between treatments and control, but either combined analysis indicates differences not apparent in the individual tests.

In general, if the entire set of seed-treatment data is not homogeneous, it can be separated into one or more groups having relatively high variances indicating that the infestation has been extremely variable and into one group of relatively low variance indicating fairly uniform but usually a low infestation. A combined analysis for this latter group will often indicate significant differences not apparent in the individual analyses. In any set of identical replicated experiments conducted at different locations such loss of information may result unless the data are subjected to some sort of combined analysis.

Literature Cited

1. Neyman, J., and Pearson, E. S. On the problem of *k* samples. *Bulletin de l'Academie Polonaise des Sciences et des Lettres,* Serie A, 460–481. 1931.
2. Bartlett, M. S. Properties of sufficiency and statistical tests. *Proc. Royal Soc. of London, Ser. A,* 160: 273–275. 1937.
3. Roessler, E. B. Valid estimates of variance in the analysis of pooled data. *Proc. Amer. Soc. Hort. Sci.* 42: 481–483. 1943.
4. Hayes, H. K., and Immer, F. R. Methods of Plant Breeding. pp. 339–350. McGraw Hill Book Co., New York. 1942.
5. Snedecor, George W. Statistical Methods. Third Edition. pp. 240–243. Collegiate Press, Ames, Iowa. 1940.