Telephone
Telegrams } Harpenden 1621 (3 lines)

Railway Station: Harpenden (L.M.R.)

ROTHAMSTED EXPERIMENTAL STATION
(LAWES AGRICULTURAL TRUST)

Director: Dr. W. G. OGG, M.A.

DEPARTMENT OF STATISTICS
(Research Statistical Service)

Head of Department:
F. YATES, Sc.D., F.R.S.

HARPENDEN

HERTS.

8th May, 1950.

My dear Ron,

I think an addition on the lines you indicate is well worth making. I like the distinction between definite and indefinite categories, but I am not sure whether the last paragraph of p.2 covers all contingencies, and also whether the phrase "appropriate portions" should not be amplified.

There are three separate problems which I have encountered in analysis of the results of groups of experiments.

(a) The errors in different experiments are often very different

(b) The interactions of different experimental effects with other categories vary considerably, even when the effects are of the type that are normally pooled in the analysis of variance e.g. varietal differences within a set of varieties.

(c) The different types of comparison which are of interest require different combinations of the various interactions of definite and indefinite comparisons.

(a) is a somewhat separate problem, which can well be omitted, thought it does arise in practice. The lines on which I tackle it are given in the enclosed paper.

(b) some amplification of the insertion on p.2 might help the noviciate. In practice, the complication is extremely tiresome, because if the analysis is broken down so as to isolate each effect of interest, and its interactions, the number of degrees of freedom for the interactions is usually very small. Thus in the Woburn ley experiment there are four main treatments (1) 3 years ley, (2) 3 years lucerne, (3) 2 years arable and 1 year ley, (4) 3 years arable, all followed by two "indicator" crops, potatoes and barley. There are 5 phases each with two replicates (later differentiated) of each treatment in a single block. The analyses of say the potatoes for the first cycle is thus

Years	4
Treatments	3
Treatments x Years	12
Error	20
Total	39

Treatments x Years was in fact clearly significant, and considerably greater than the plot error. This is obviously the more relevant "error" to use when drawing conclusions from the experiment, but I suspect when further data are available that it will prove non-homogeneous, there being greater variation from season to season for the contrast 1 + 2 - 3 - 4 than from 1 - 2, and 3 - 4. For the moment, however, we can only use the pooled estimate.

(c) Suppose an "effect" is measured in p places in each of q seasons, there being r replicates in each place-season test. The analyses of variance is

	D.F.	Mean value of M.S.	
Effect	1		
Places	p-1	P	
Seasons	q-1	Q	T
P x S	(p-1)(q-1)	I	
Exp[l] error	(r-1) pq	E	
Total	pqr		

The estimate of error variance of the mean effect due to experimental error is E/pqr. If places are regarded as indifinite the variance is P/pqr, which gives a test for the effect over a generality of places in the seasons of the experiment. Similarly if seasons are regarded as indefinite the variance is Q/pqr. But if both are regarded as indefinite I make the variance (P+Q-I)/pqr. If in this case we test the effect against the combined mean-square T for places, seasons and their interaction we shall in general over estimate the reliability of the results.

Thus if seasons have no effect, so that Q and I both equal E, my expression gives a variance of P/pqr, as in the first case considered, but use of T will give a variance of $\{(p-1)P+p(q-1)E\}/pqr(pq-1)$. Thus the pooling of two indefinite categories does not appear to be justified for this purpose.

On the other hand I agree that to a first approximation T will give an unbiased estimate of the variation to be expected in another season x and another place w. The mean square of the difference between the mean effect and the value we may expect to obtain from an experiment of the same accuracy as the previous ones is $\{(pq-1)T+2(P+Q-I)\}/pqr$ which is approximately $T\{1+1/pq\}/r$. This is probably what you had in mind in drafting the last paragraph.

Yours sincerely,

Frank.

<u>Professor R.A.Fisher.</u>

I am returning your draft herewith.