25 April 1932.

Professor R. Summerby,
Macdonald College,
Que., Canada.

Dear Professor Summerby:

It is a great pleasure to hear from you, and to discuss again some of the problems of experimental design and interpretation.

On your main question, the answer is that the pairing method and the variance method as shown on your last sheet are in perfect agreement. You can verify this by noticing that the "correction" 552.33 applied in the pairing method to the sum of squares is just double the entry opposite "treatments" 276.165 in the analysis of variance, and the "correct S.S." of the pairing method 892.58 is just double the entry opposite "error" in the analysis of variance. The reason for the factor 2 is that the analysis of variance works on a single plot basis, while the pairing method is based on the difference between two plots.

As you suspect the discrepancy in the value of P is due to "Student's" table giving the probability in one ~~trial~~ tails only, and my tables giving the probability in both trials tails

together. From "Student's" value .99684, we have .00316 and doubling this .00632, quite appropriate for a value not far beyond the 1 per cent. point.

This also perhaps answers your second point, for when there are several varieties the t-test is equivalent to the pairing method, and therefore measures just as fine differences. It is ~~here~~ but, of course that the physical conduct of the experiment will affect the precision, and that a comparison between two varieties only in 10 pairs of randomised adjacent strips will generally be really finer than the comparison between only two varieties out of 5 in 10 randomised blocks of 5 strips each. But to compare all 5 varieties by the first method all ten ways of comparing pairs must be used, and this requires 200 plots, as many as would be needed for 40-fold replication of the blocks, and then the block method would be very much the more precise.

As to when you should use the P tabled in my book and when the half value, is a point which I might well have discussed more fully. Let ~~us~~ me take an example. A man comes and declares that if we would only use a rotary cultivator we should get much better germination with our mangolds. We might try an unreplicated experiment, with single areas under the two treatments, then if the germination really does better we shall think there is something in the assertion, but if not we will think no more about it.

In circumstances like these when we come to replication and estimates of error, it is clear that the new method scores a significant success only if it differs from the old in the positive direction, consequently if we ~~need~~ decide to use the 5 per cent. point we ought in fairness to give him 5 per cent. in the right hand side of the curve, i.e. everything beyond $t = +1.645$, if n is large, because he is not going to claim any success if the difference is negative, however large it is.

On the other hand with a pair of varieties the fact that when one beats the other, in a single trial would not contribute anything to our decision, unless we already had prior evidence that one was better than the other, and are just testing it further. One or other is bound to win; and so in the test of significance we shall count both ~~trials~~ tails as significant, and for the 5 per cent. point allot $2\frac{1}{2}$ per cent. to each, so counting anything outside the range $t = \pm 1.960$.

I though Hoblyn's pamphlet a very useful one, but he certainly made a slip about the degrees of freedom, which I am very glad you spotted. One of the advantages of calculating a general estimate of error, instead of a special one for each pair of variates lies just in the greater number of degrees of freedom which can be utilised.

I hope to be across in late August for the Genetical Congress at Ithaca; is it possible that you will be there, or that I could get to see you?

Yours sincerely,