St John's College
Cambridge.

1938 Sept. 22.

Dear Fisher,

Thanks for letter and your remarks on 'Student'. I shouldn't be much surprised if K.P. was actually the first to get the exact form algebraically and associate it with the distribution of chance ; these things have such a way of being got piecemeal, so that formulae in their present-day forms are known by the names of people that never saw them. It doesn't seem to be in Poisson or De Moivre, but I usually quail at the thought of trying to find anything in Laplace. The algebra is easy enough and must, I should think, have been used as a school example ; my Hall and Knight and C.Smith haven't it, but they have a lot of harder ones that should come out by similar methods. Anyhow I don't think the mathematical difficulty of the problem is such that it need affect anybody's reputation much.

I mention the point about a varying standard error giving something like a Type VII distribution in my Phil. Trans. paper ; I think Eddington mentioned it somewhere but can't remember where. But it is really an academic point, because we don't know which of the observations are those with the larger and which with the smaller standard errors, and there is nothing for it but to do our best with the distribution as we find it.

I shall try that letter on Biometrika, possibly through Yule. My main feeling about grouping corrections, however, is that they don't begin to matter until the whole of the data

are out into about 3 or 4 groups.

I like your note on Student. Curiously, when I rediscovered his rule in 'Sci. Inf.' he was only a name to me ; and when you went for me in the Proc.Roy.Soc. I thought that you were attacking this rule ! But there are a few points that I don't quite agree with. On p.2, near the foot, I should not call the rule a test of significance, but one of pure estimation. What it gives, in either way of looking at it, is the probability of an error of the mean being more than so much, in terms of the observations, if there is nothing otherwise to call attention to one possible true value rather than another. As I use the term 'significance' I am definitely considering cases where the last condition is not satisfied. It might be well to emphasize the estimation side of the matter a bit more.

p.3. I don't think it matters much whether you divide by n or n-1. I got into the habit early on of using 'standard error' for what you would call the 'population standard deviation' and 'standard deviation' for $\sqrt{S(x-\bar{x})^2/n}$. The use of different words prevents confusion of the population and sample variances, and once it is recognized that they are not the same, but that each has a definite probability distribution in terms of the other, the question of n or n-1 is only one of convenience. I haven't succeeded yet myself in deciding which is the more convenient.

p.8, first half. This is just what I expected ; and it would go over exactly into my result in the case where the median and the scale of the law are initially unknown.

p.14. Top. I think this may be, and indeed has been, misinterpreted. If randomization was necessary I think you would consider

assigning all 25 plots at random in a 5 x 5 square as the ideal procedure, getting two of a treatment in one row and none in another. 'Student' misses the point of the Latin square, I think, when he says that it is both balanced and random. I should say that it is balanced for ground effects known to be often large (possibly indeed more balanced than is really necessary) and random for others that might mount up if the design was repeated in different squares ; the balance and the randomness refer to different features of the ground variation. The real question, it seems to me, is, what ground effects matter in practice ? If there is no correlation at all between plots it doesn't matter what you do. If there are correlations and systematic variations it is legitimate to arrange the work so as to estimate them and allow for them, also allowing of course for the degrees of freedom lost on the way. If they are doubtful, deliberate randomization will provide an alternative treatment and easier arithmetic. I am not sure, but I think I have seen discussions on the point that don't notice that if there is anything to be dealt with by a systematic design, its proper analysis will involve the separation of additional degrees of freedom, the variance associated with which is neither treatment nor error variance. The row and column variance in a Latin square is a particular case ; so would be the fertility gradient along the row in a randomized block experiment, which Student suggests would be worth taking into account. That is, I think that in the Latin square design you admit his main contention ; but I have not gone into the matter enough to have any opinion about whether he has got the best way of analysing the results of other systematic

designs. It is a matter of stating explicitly what terms in the fertility are to be considered, and working out the maximum likelihood way of eliminating them. Neither Student's paper nor that of E.S. Pearson is sufficiently explicit on this point to convince me that they have got to the root of the matter. I should say that any design will give a valid estimate of error if the results are analysed correctly.

When I sent 'Student' a copy of my paper 'The relation between direct and inverse methods..' I got a short note in reply, calling attention to the fact that in the title of his table the words 'a unique sample' occurred. I should like to acknowledge this, because it shows that the condition that I said he had assumed without mentioning it, that the sample constituted the only relevant information, was in fact in his mind. If there are several samples from the same population, everything in the direct argument is still holds, but nobody would use such a set separately, so that the extra condition comes in in the transition from the argument to the use that is made of it.

Yours

Harold Jeffreys.

I think you had better see the enclosed paper, which is the one I tried on Bimetallism. The part that is relevant is from p. 9 to the end. It states the problem a bit more clearly than has been done, I think, but I haven't the data to give a clear answer to it. But it may suggest something to you.