Department of Genetics

c/o

[?] November 1943

Dear Kendall,

Thanks for your letter. I shall certainly be glad to do all I can to help, but you have, even in your letter to me, already adopted part of the language and terminology which Pearson rather maliciously as I thought introduced (after ignoring my work for fifteen years) with a view to misleading the student as to its scope and nature. The so-called "point" or "single-value" methods of estimation constitute the whole problem or subject of estimation in the sense in which this phrase was used for at least fifteen years. When later I introduced the fiducial argument and showed how fiducial limits could be calculated, this did not in my opinion constitute a new method of estimation but rather a new method of constructing probability statements valid in respect of the unknown parameter, obviously quite a closely related topic, but one applicable to a smaller range of cases.

The criteria of estimation originally developed were based in respect of efficiency on the variance of the sampling estimate, a fact which alone shows how misleading is the phrase, "single-value" or "point" estimation, applied to such methods. In respect of sufficiency the exact form of the whole sampling distribution is taken into account or, looked at otherwise, the entire course of the likelihood function.

For several years it is certain that Neyman thought, in speaking of "confidence-intervals", that he was only systematising and developing a new exposition appropriate to the fiducial argument I had put forward and pressed upon his notice. It was perhaps only when he was hoping to get appointed at University College that it occurred to him that Pearson would be willing to believe that the idea had all been Neyman's, Pearson being perhaps himself unaware of what I had already published. It must have been about the same time that Neyman seems to have treated his Polish friend Kolodziejczyk in very much the same way as he treated Pearson (see Annals of Eugenics Vol. xi, p. 143).

The diversity of opinion as to Behrens' problem arises, I think, by an historical ~~argument~~ accident. In our statements of probability, e.g. tests of significance, neither 'Student' nor I took the problem to relate to a population of samples all of the same size and drawn from an identical population. In 'Student's' problem we were both thinking of a population of samples of the same size drawn from populations having their variances fiducially distributed, i.e. that unique distribution of populations for which alone the observed estimate s^2 would be a typical estimate. It so happens, however, that when there is only one unknown parameter, and in a few other cases of special simplicity, the probability which defines the level of significance can be reinterpreted as a probability related to a population of samples drawn from the same population. So to interpret them is to my mind always to wander from the point, but Neyman and Pearson have made this convention the headcornerstone of their exposition and, as the Behrens case shows, are much aggrieved when what is frankly the only possible solution of a practical problem does not conform with their convention. This however will generally be the case whenever tests have to be made in situations involving multiple estimation. It is indeed surprising that Neyman and his followers, who so often discuss a comparison of two samples, should be unable to provide any solution to this rather fundamental and primary problem. I presume, of course, that the method, following Bartlett, of choosing one out of n! equally eligible tests of significance and using its result without reference to that of any of the others, is not to be regarded as a solution of the problem. Moreover, the hope, which I suppose Bartlett originally entertained, that there was an appropriate test which would reject a fixed proportion of samples for all values of the variable parameters when the samples are taken to be drawn from a fixed population, seems now to have been finally abandoned.

One fundamental objection to basing the theory on samples of a fixed size from a fixed population is the obstacle it introduces to the rational use of ancillary information.

Properly speaking, the sample-number, or set of such numbers used
are ancillary statistics, and in particular cases are
to specify the sample, is found to be supplanted by more appropriate ancillary statistics, though of course in other cases such statistics are additional to the sample-number. When, in such cases, estimation is exhaustive, the likelihood function is completely specified by the set of ancillaries, and it can be shown that all other statistical calculations add precisely no information

to that already in hand; it seems absurd to pretend that there is a fundamental distinction between such cases and those which I originally called 'sufficient', which differ from them only in that in these the sample-number alone supplies all the ancillary information required. Yet in the method of exposition adopted by Neyman and Pearson these cases are separated from the first, and indeed the use of ancillary information seems to be almost ignored.

Yours sincerely,