24 July 1934.

Professor E.B. Wilson,
Dept. of Vital Statistics,
School of Public Health,
Van Dyke Street,
Boston, 17.
Massachusetts, U.S.A.

Dear Professor Wilson,

Many thanks for supporting Neyman's form of application.

I was interested to see Sterne's papers, as I had previously seen a letter of his to "Nature". As you expected, there is very little new in their actual content, but they are written with admirable clarity and should be valuable when spreading the light among astronomers. I have made in the margins a few notes, which, perhaps, I ought to include here, in case you care to hand on this letter to Sterne.

1) On notation. The author takes the sum of the squares of residuals divided by \( \bar{n} \) for his value \( s^2 \). As \( s \) is regarded as an estimate of \( \sigma \) it would be better to use \( (n-1) \) or \( (n-m) \) at this stage. The estimates are then comparable whenever the number of degrees of freedom \( (n-1) \) or \( (n-m) \) are the same, irrespective of the value of \( \bar{n} \).
The sampling variance of the mean is then estimated by dividing $s^2$ by $n$ and not, as Sterne does, by $(n - 1)$.

The author ought to substitute $t$ for $z$ from page 4 of his first note to conform to the notation used by "Student" and myself. $z$ has unfortunately been used in a different sense, this same problem by "Student", and the new sense in which it is used by Sterne is, therefore, liable to be confusing.

The "probable error" of page 6 and Table I of the first note is the quartile distance of $t$ distribution and is given to three decimal places in Table IV of "Statistical Methods" in the column headed $T = .50$. The practical use of the probable error is, however, to be multiplied by some factor, such as 3, in order to find a value exceeded with a known small probability. As the form of the distribution changes with $n$, this factor changes, and it is simplest and best to use the actual values of $t$ for 5% or 1% or whatever the probability chosen may be, i.e. to use one of the later columns in Table IV.

The second note is interesting in making a new proposal which amounts to using the fiducial mean value of $\epsilon$ as an estimate. The author should, I think, consult and compare the method developed in a paper of mine in 1938 (Proc. Roy. Soc. A 139, 143-173). The fiducial mean, like the fiducial mode, is
is open to the objection that neither of these are invariant for functional transformations, whereas, for a large class of such transformations, the median and other placentiles are invariant. These give what Neyman and his Polish students call the "confidence intervals". In any case, the author ought to make clear that his \( \tilde{\mu} \) is the fiducial mean. The median and placentiles are, of course, available for the cases \( n = 2 \) and \( n = 3 \).

The third note is a very clear exposition of what I call the \( \tau \) test for regression coefficients. I think it would be clearer if \( n \) and \( m - n \) were interchanged, as explained above. While the fourth note provides material for an explanation of what I have called the degrees of freedom of an observed value \( \chi^2 \), a point on which the explanation might, perhaps, be usefully expanded. I think the author is much to be congratulated on going ahead with the solution of these really fundamental problems, and could really do valuable work in making them known to astronomers, a somewhat conservative body, the advances which have taken place in other fields. In particular, a great many statistical discussions in astronomy could be made very much clearer and easier to read, and, I suspect, also to write, if the arithmetical arrangement, which I call the analysis of variance, based on just these ideas,
were familiar to the authors.

Historically, it is interesting that the distribution of \( S^1 \) was discovered by an astronomer, Helmert, as early as 1875, but its importance was obviously overlooked, owing to the widespread use of the mean error (Peters' formula) in astronomical work. Perhaps the more so, as Peters was the editor of the journal in which Helmert published his results, which appeared as insignificant excrescences of papers relating principally to the sampling errors of Peters' formula.

I am returning Sterne's papers herewith.

Yours sincerely,