

COMMENT ON THE NOTES BY NEYMAN, BARTLETT, AND WELCH IN THIS JOURNAL
(VOL. 18, No. 2, 1956)

By R. A. FISHER

It is gratifying that none of the three commentators on my note on Table 11 of *Biometrika Tables* now claims that Behrens' solution of the problem is incorrect, or is capable of improvement by the alternatives that have been put forward. This is an important point gained, both for those who wish to make such tests, and for the clarification of the principles on which such composite hypotheses should be tested.

I was, of course, aware that both Bartlett and myself had given the essential mathematics of my note twenty years ago. It was in the logical interpretation of the mathematics that we differed, for whereas I regarded them as confirming Behrens' result decisively, Bartlett has twice put forward alternatives, to each of which, however, serious objections have had to be made.

Dr. Welch explains, carefully and at some length, that he did just what he intended to do. The doubt, of course, is only whether his intention was well advised. His work was introduced to the world by Bartlett as supplying a better test of significance than Behrens'; and this could scarcely now be claimed, for Behrens' test uses the true probabilities observable within the subset of cases in which s_1/s_2 is the same as that actually observed. Cases not of this subset are irrelevant. Whatever Bartlett may mean by his phrase "There is no logical reason to regard s_1/s_2 as fixed", it is at least certain that this ratio of observables is *known* to the experimenter as a fact, and that all other ratios are known not to be facts. I doubt if logic could provide any such reason as that for which Bartlett seems to ask. The non-sufficiency of the ratio s_1/s_2 as an estimate of σ_1/σ_2 also, which causes him some anxiety, is only a consequence of assuming as a *datum* that the two populations have indeed the same mean, and this is exactly what is under enquiry.

If Professor Neyman were in the habit of learning from others he might profit from the quotation he gives from Yates (although he has left out more than a page of his explanation), for Yates is there warning his readers against a pitfall into which it would seem Neyman has himself fallen. I refer to the fallacy that a statement of fiducial probability about the population from which an observed sample has been drawn, refers not simply to that population, but to an imaginary, or hypothetical, aggregate of populations from which, at some anterior stage, it was picked at random. Even if such an aggregate existed, which is scarcely axiomatic, it is obvious that nothing could be known of it, beyond what we can learn about that one of its members which has been sampled.

References

- BARTLETT, M. S. (1956), "Comment on Sir Ronald Fisher's paper: 'On a test of significance in Pearson's *Biometrika Tables* (No. 11)'"', *J. R. Stat. Soc.*, B, 18, 295-296.
 NEYMAN, J. (1956), "Note on an article by Sir Ronald Fisher", *J. R. Statist. Soc.*, B, 18, 288-294.
 PEARSON, E. S. & HARTLEY, H. O. (eds.) (1954), *Biometrika Tables for Statisticians*. I. Cambridge University Press.
 WELCH, B. L. (1956), "Note on some criticisms made by Sir Ronald Fisher", *J. R. Statist. Soc.*, B, 18, 297-302.
 YATES, F. (1939), "An apparent inconsistency arising from tests of significance based on fiducial distributions of unknown parameters", *Proc. Camb. Phil. Soc.*, 35, 579-591.