3rd March, 1956.

Dear Mr. Dunn,

I have just seen your letter of March 1st, and by a curious coincidence I have recently got out explicit formulae for the probability of d̲ exceeding any particular value for small even numbers of sample size, the two numbers of degrees of freedom being both odd.

In your particular case with two samples of four so that the degrees of freedom are three and three, the probability of exceeding any positive value d̲, i.e. the single tail frequency, is

$$P = \frac{1}{2} - \frac{1}{\pi}\left\{\alpha + \sin\alpha\cos\alpha + \frac{3\tan\theta}{(1+\tan\theta)^2}\cdot\frac{2}{3}\sin\alpha\cos^2\alpha\right\}, \quad \text{where } d = \sqrt{3}(\sin\theta + \cos\theta)\tan\alpha, \quad \tan\theta = s_1/s_2$$

This distribution, and the others like it, have not been tabulated numerically, and I have been engaged, in such little time as I have to spare, in working some of them out. The use of d̲, as defined by Sukhatme as the test criterion, makes the values vary rather little for different ratios of $\underline{s}_1$ to $\underline{s}_2$, but because the major variation due to these causes has been largely neutralized, the series, if I have calculated them rightly, sometimes seems a little capricious. As examples, I have for 10% (5% in each tail)

θ	
0°	2.353
15°	2.386
30°	2.446
45°	2.559

these values not being properly checked. It does look as though the 45% value, which must be the centre of a symmetrical curve, i.e. 30° is the same as 60°, has risen rather sharply out of line as compared with the others. For the 5% points (2½% in each tail) I have similarly 3.182, 3.191, 3.110, and 3.244, in which the 30° point looks surprisingly low. However, it is only a matter of getting the arithmetic done to get these and corresponding values checked with professional precision.

I should like to see a set of ten tables from (1,1) to (7,7) at four levels of significance, 10%, 5%, 2%, 1%, for the combined ~~values~~ tails, and either 15% or 5% intervals for the angular variation. The ten tables on a seven by four scale could be accommodated conveniently, with the other tables appropriate to Behrens' problem, in Statistical Tables.

You might like to think whether it would be worth the time and money of your firm to meet the cost of getting this job done expeditiously through one of the computing agencies available. I dare say someone in Frank Yates' Department might programme it for electronic computation, but of course it is not at all beyond manpower methods.

As to your questions, for number one the method of the paper you referred to would not be accurate for such small samples.

Your second point should be answerable by checking the <u>Example</u> on page 161. Is it done wrong?

Sincerely yours,

x [CP 181]