31st May, 1935

F.Garwood Esq.,
Building Research Station,
Garston,
Nr.Watford,
Herts

Dear Mr. Garwood.

One Dr. Behrens, who seems to have disappeared (for I can get no answers to letters), gave a mathematical solution of the problem, but never went so far as to provide the tables needed for its use. These would be troublesome, in any case, being of quadruple entry, since they need two numbers to specify the sizes of the samples, and one more for the ratio of the estimated variances.

The principle of solution is simple, though fundamentally important, and in giving it you I shall make no attempt to reconstruct Behrens' line of argument.

If $\bar{x}_1$ and $\bar{x}_2$ are the means obtained from samples of n_1 and n_2 observations, respectively, while s_1 and s_2 are the standard deviations as estimated for the means of these samples, then, if the true common mean of populations is μ, we may define:

$$t_1 = \frac{x_1 - \mu}{s_1}$$

and

$$t_2 = \frac{x_2 - \mu}{s_2}$$

where t_1 and t_2 are independently distributed in Student's distribution.

It follows, then, that the difference observed between the means is

$$s_1 t_1 - s_2 t_2$$

and the line

$$s_1 t_1 - s_2 t_2 = d$$

will divide the bi-variate distribution into two unequal portions. Given s_1 and s_2, the slope of this line is determined by their ratio, and its distance from the origin by the value of d. So that the test of significance is completed by calculating for what value of d, given n_1, n_2, s_1 and s_2, the portion of the distribution which lies outside the permissible locus. is an assigned fraction of the whole

Tabulation would not seem at all impossible if properly planned; but it would be laborious, and although I had some correspondence on it at the time, I believe that both Behrens and others came to the conclusion that the problem was not of sufficient importance to justify putting any considerable amount of labour into making the test available.

Yours sincerely,