September 7, 1938

Dear Cochran,

Thanks for your letter. I had made a few trials to satisfy myself that the t test would work well, and I am glad to see that you have done the same. The exact test of significance is, of course, that based on permutations, and from this point of view the scores are only conventional.

I think I arrived at them first by considering the problem: given the ordinal series for two variates drawn from uncorrelated normal distributions, what weighting would give an efficient estimate of a suspected correlation. The correlation of [illegible] is just the sum of products of tabulated values divided by the sum of squares given in the second table. This also, except for coarse grouping in small samples, naturally agrees well with the correlation from normal deviates.

The computation of the table is considerably simplified by expanding the series of values for each value of n in odd orthogonal polynomials. The series terminates, e.g., for n = 2 or 3 the linear term is exact; for n = 4 or 5 the cubic term is exact, and so on; but the diminution of the numerical values is also very helpful, e.g., I suppose that the 13th degree is very nearly right, apart from the last value, for numbers up to 20 or 25. The coefficients of these polynomials are a good deal simpler than the original integrals, being free from the variable r.

Years ago - I write from memory - I found that the loss of information due to replacing variates by ordinal values is not very large, if it is remedied by replacing ordinal values by mean variates. My impression is that the sum of squares of the true mean deviates, which differs a little from what we have tabulated, is ~~in~~ $(n - 1)$ times the efficiency. If this is so the percentage loss of information ~~x~~ falls off proportionately to $\frac{\log n}{n}$ and is moderately small when n exceeds 10.

If you are interested you might look in some time and I will try to hunt up what I have in the way of algebra, as I have a strong impression that a good many pretty, though intricate, results could be obtained by proper ~~numerical~~ algebraical treatment.

Yours sincerely,