

ROTHAMSTED EXPERIMENTAL STATION
(LAWES AGRICULTURAL TRUST)

Director: Sir WILLIAM G. OGG

DEPARTMENT OF STATISTICS
(Research Statistical Service)

Head of Department:
F. YATES, Sc.D., F.R.S.

HARPENDEN
HERTS.

10th February, 1956

My dear Ron,

Thanks for your letter concerning "Statistical Tables". I have no strong feelings about Stevens's suggestion for the exponential function. Inverse linear interpolation in our tables seems to be about as good as Stevens's interpolation by multiplication, and there is probably not much in it as far as trouble is concerned, since we give the proportionate differences. I don't think there is any need to add anything in the introduction.

I am, however, still rather keen to include tables for the logit and log log transformations. I don't dispute that the algebraic expressions for the logit (r, Z) transformation are very simple but the relevant values can only be obtained from Table VII with a good deal of fiddle. In practice anyone wanting to make a maximum likelihood adjustment with this transformation would go elsewhere for his tables. Even as a straight table of the transformation Table VII is rather unsatisfactory, requiring the steps

- (a) $r = 2p - 1, p > \frac{1}{2}$, or $r = 1 - 2p, p < \frac{1}{2}$.
- (b) Inverse use of Table VII with change of sign if $p < \frac{1}{2}$.

For final adjustment for a given provisional value Z the following steps are necessary:-

- (a) Look up the R corresponding to Z from Table VII.
- (b) Calculate the weighting value $= Z + \frac{1}{1+R} - p \frac{2}{1-R^2}$
- (c) Calculate the weighting coefficient $1 - R^2$.

This sort of thing just gives the computer who is not particularly familiar with the whole process too much to think about. Unless

he does the calculation on a slide rule he will also be put to a good deal of extra work.

The functions for the log log transformation are much more tiresome and certainly require tabulation if the transformation is to be used. As I said in my last letter, I have recently found both these transformations useful. Finney also considers them sufficiently valuable to tabulate in his "Statistical method for biological assay". I am particularly keen to include them here as there are many problems outside biological assay for which they are useful. Inclusion in our tables will make them available to a class of reader who will not have Finney's book.

I have now looked at the question of arrangement. The logit transformation, which we may define as

$$Y = \frac{1}{2} \log_e \frac{p}{q}$$

(omitting the 5) will go easily on a single page if we tabulate from 50(x 0.1%) 99.9% to three decimal places. On balance, I feel the 5 is more of a nuisance than it is worth. It would certainly require a table covering the whole range 0 - 100% to make it worth while.

The table for special adjustments will also easily go on a page. As we shall have the room I am disposed to use an interval of 0.05 in Y covering the range 0.0(x 0.05) 3.5, say. This is on the assumption that we leave Table VII undisturbed. The alternative would be to scrap Table VII and Table VIII which I don't believe anyone uses, although I am fond of it. The table for final adjustments could then go on the same page as the correlation coefficient, Table VI, with a bit of a squeeze. However, they would make rather strange bedfellows.

The log log transformation is more tiresome since it is not symmetrical about 5%. To get the transformation on one page I would like to be a little radical and tabulate at intervals of 0.2%. This, though rarely done, is practically convenient since it is very easy to take the mean of two numbers mentally. The current conventions require changes in the tabular interval ~~(to be)~~ by a factor of 10, which is much too big a ratio and often results in excessively wide or close intervals. The table for final adjustment in the log log transformation will conveniently go on one page with an interval of 0.1 in Y.

I propose we tabulate what I have defined as the "Complementary log log transformation", i.e.

$$\log_e \left\{ - \log_e (1 - p) \right\}$$

instead of $\log_e (-\log_e p)$. The complementary transformation is just as likely to be required as the other, and has the advantage that it increases with p so that the general formulae already given by you are applicable.

With regard to the extension to the Behrens-Fisher table, I agree that the table is not likely to be very frequently used and on this ground might possibly be better published in a scientific journal. However, inclusion in the Tables would give further emphasis to the difference of opinion and lead people to think about the general issues. If you would like it included the problem of page arrangement arises. It occurs to me that the table will just go in the space occupied by Table VII which could well be dispensed with if the inverse table is to be included. This would fit in very happily with the other Behrens-Fisher tables. Table VI could be moved to the bottom of the page. I suggest we might economise half the entries by giving the columns for 0° , 15° , 30° and 45° only. If θ is $> 45^\circ$ and n_1 and n_2 can be interchanged.

I am off to India at the end of next week so I shall be very glad if we can get matters settled before I go so that I can arrange for the tables to be prepared and sent to Oliver and Boyd. We can then deal with any modifications to the introduction at leisure as setting up the introduction takes far less time.

I have passed on Healy's proposal form to Finney and will deal with the form for Box as you suggest. I take it you are also putting up Owen?

Yours sincerely,



Professor Sir Ronald Fisher, F.R.S.