November
Twenty-eighth
1927.

Dear Sir,

I am sorry I have not had time to give to your problems the attention they deserve. Perhaps a few notes will be better than nothing.

It seems to me that the proportion alive is a suitable variable to use in your data where most of the counts used are fairly extensive, and the percentage variation is not great. (I have not, however, the data for the control). You wish to express this percentage by some such formula as

$$p = a_1 + c\,d$$

where d is the density, a_1, a_2, a_3, a_4 are constants for the 4 fluids and c is a general constant. Perhaps you could choose a better form than this; in any case the observation that in a certain case 165 died and 365 lived gives a term

$$365 \log p + 165 \log (1-p)$$

in the general quantity, made up of such terms, to be maximised. Differentating by a_1, we have for all ~~summed~~ [illegible] observations

$$S\left(\frac{365}{p} - \frac{165}{1-p}\right) = 0 \quad \text{or} \quad S_1\left(\frac{L}{p} - \frac{D}{1-p}\right) = 0$$

and 3 similar equations for the other fluids. Differentiating by c, we have

$$S\left\{ d \left(\frac{L}{p} - \frac{D}{1-p}\right)\right\} = 0$$

for all twigs.

Now you have all the work ready done for obtaining good approximate values of a_1, a_2, a_3, a_4 and c. Using any such approximation find

$$\underset{1,2,3,4}{S}\left(\frac{L}{p}-\frac{D}{1-p}\right)=A_{1,2,3,4} \quad \text{and} \quad S\left\{d\left(\frac{L}{p}-\frac{D}{1-p}\right)\right\}=C$$

which will not of course be exactly zero for the approximate solutions, the modifications in the approximate solutions needed will then be given by the 4 equations of the form

$$b_1\,da_1+b_2\,da_2+b_3\,da_3+b_4\,da_4+b_5\,dc=A_1$$

where

$$b_1=-\frac{\partial}{\partial a_1}S_1\left(\frac{L}{p}-\frac{D}{1-p}\right)=S_1\left(\frac{L}{p^2}+\frac{D}{(1-p)^2}\right) \text{ for all Sunqoo twigs}$$

$$b_2=-\frac{\partial}{\partial a_2}S_1\left(\frac{L}{p}-\frac{D}{1-p}\right)=0=b_3=b_4$$

$$b_5=-\frac{\partial}{\partial c}S_1\left(\frac{L}{p}-\frac{D}{1-p}\right)=S_1\left\{d\left(\frac{L}{p^2}+\frac{D}{(1-p)^2}\right)\right\} \quad ''$$

The matrix of quantities b, gives at once the standard error of each estimate, and so supplies a direct test of significance for the difference a_1-a_2 etc., for if

$$\Delta=\begin{vmatrix} S_1\left(\frac{L}{p^2}+\frac{D}{(1-p)^2}\right) & 0 & 0 & 0 & S_1\{d(\;)\} \\ 0 & S_2\left(\frac{L}{p^2}+\frac{D}{(1-p)^2}\right) & 0 & 0 & S_2\{d(\;)\} \\ 0 & 0 & S_3(\;) & 0 & S_3\{d(\;)\} \\ 0 & 0 & 0 & S_4(\;) & S_4\{d(\;)\} \\ S_1\{d(\;)\} & S_2\{d(\;)\} & S_3\{d(\;)\} & S_4\{d(\;)\} & S\left\{d^2\left(\frac{L}{p^2}+\frac{D}{(1-p)^2}\right)\right\} \end{vmatrix}$$

then $\sigma^2_{A_1}=\frac{\Delta_{11}}{\Delta}$ where Δ_{11} is the minor found by deleting the first row and columns. The theory is given in Mathematical Foundations of Theoretical Statistics, Phil. Trans. A, 222, p. 309.

The method is quite general and would apply equally to Mr Ackermann's data.

Yours sincerely

PS. Thanks for the chart of $\check{\chi}$ it shows it very well. Certainly you may use my table in this way. Yes, I had spotted the error in Table VI, but I think no one else has. I shall put further tables of in the new edition coming out early next year.