Dear Fisher,

I wonder if you would be kind enough to consider one of two little problems I am up against and to evaluate my suggestions for dealing with them. I should be very grateful if you would do so.

The first is the old question of the continuity of Linsley's data—salinity in a $12 \times 25$ table by months & years. The fact of the months running continuously through all years is forced to give a very high interaction between months & years. As the data are peacefully unexpressly represented in the cells of the $n \times m$ fold table, there is no possibility, with this high interaction of discontinuities yearly effect from monthly effect. My suggestion is to for each consider the last observation as the first of the next, to join them by a straight line, draw a straight line parallel to this through the annual
mean, and calculate the variance from this line instead of from the horizontal line through the mean. The various slopes would absorb degrees of freedom one more in number than the number of years. The difference between the total variance & the variance due to continuity called "variance due to continuity" and, on subtraction of this from the total variance one might expect the annual effect & monthly effect to be additive. If instead of this a regression is fitted, it would absorb some of all of the annual effect wouldn't it? It seems also some of the monthly? It seems to me that one must consider monthly effect as harmonic over each year & if the end points of the monthly harmonic are not at an equal level, this must be ascribed to annual effect. Anything else illogical. Annual effect however, being continuous, some allowance must be made as I suggest. If my suggestion is sound, would
it do equally well if the straight lines were drawn between annual means instead of between end points. Each period would then have two slopes joining at the centre. It would make computation easier if this would do equally well. Simply has so many areas that fitting regressions is really out of the question, I'm afraid. Please consider these suggestions. Simply has considered his means in cells as single observations. He reiterated, getting corrected values for the cells, filling in empty cells, getting corrected values for marginal and total means on the assumption that all cells are equally represented. That each corrected value has equal weight. Nothing further can be done with these, I suppose?

No sort of standard error can be allotted to any of these values. If the simply obtains (for any future area) in a subsequent year, any annual mean can a test be applied, whether this is significantly different from...
the corrected mean of the previous years, allowing that in the new annual mean the months are quite unequally represented. Would it not be preferable to consider the mean in every cell as having equal weight, and as unless to it the mean variance within a cell in every case, the degrees of freedom being the harmonic mean of the degrees of freedom in each cell? — in fact not to correct by reduction at all? How then however would you fill in empty cells? There is no hurry about all this as坎利 is not writing it up yet.

Another thing that worries me a bit is testing significance in a number of parallel cases taken together. It seems to me that if the chance of significance is \( \frac{19}{20} \) you have in parallel cases, the chance of none, one, two &c. being significant is given by the terms of the expansion of \( \frac{(19q)^n}{20^n} (1+q)^2 \).
If the chances are unequal, say the chances of non-significance are \( \frac{1}{n} \), \( \frac{1}{m} \), \( \frac{1}{p} \), \( \frac{1}{q} \) in four samples, the chances of all being non-significant are obviously \( \frac{1}{n} \cdot \frac{1}{m} \cdot \frac{1}{p} \cdot \frac{1}{q} \). The other chances being given by the terms of the expansion

\[
\left( \frac{1 + (n-1)}{n} \right) \left( \frac{1 + (m-1)}{m} \right) \left( \frac{1 + (p-1)}{p} \right) \left( \frac{1 + (q-1)}{q} \right)
\]

Is this correct, and ought some allowance to be made for it? It would seem to me that a good with criterion would be that the chosen odds or half or more should be significant. Is this sound? It seems to me just as logical to say "What is the probability that a given distribution would yield this set of samples?" as ditto with sample instead of set.

I have been thinking about the degrees of freedom absorbed by curves fitted by my spring instrument which makes the tension on the
springs at each point proportional to the distance between point & curve.
I am assuming tension to be parallel to y-axis (it is quite easy to get this).
I believe that with any given flexibility of ruler, that the proportion of a degree of freedom absorbed by any point may be found as follows: join up all the points except the one to be tested. Measure the distance between the ruler & (a). This distance represents a whole 'degree of freedom.' Join up (a) & find the proportion 'absorbed' of the 'degree of freedom.'
This is perfectly correct for a straight ruler & three points in an equilateral triangle.

The proportions absorbed are \( \frac{5}{6} + \frac{1}{3} + \frac{5}{6} = 2 \).
Is this O.K.? If so, I will go ahead with my instrument. Please keep it secret.
I have finished my Snakebeck criticism a long time ago, but it has been typed in spottel form of the typewriter. It is almost ready. When it is I will send you a copy to be suitably criticized. Please do it suitably. You were so nice about seeing the manuscript of my work that you deserve I should be off-hand but shall welcome criticism. Would you please thank Yates for his letter & paper which I found most useful & interesting. I will send back the copy very shortly. Please give my kind regards to Mr. Fisher.

Yours sincerely,

[Signature]

[Handwritten name]