March 17, 1942

Dear Jackson,

I was glad to get your letter and to see that you are still concerned with what I have always thought an exceedingly important type of data, i.e., your recapture frequencies.

I entirely agree with the formulae you suggest for R_-, where $1-R_-$ is the combined death and emigration rate, and for R_+; SIMILARLY related to the combined rates of emergence and immigration over the same time interval. It was playing about with Ford's material from the Scilly Isles that made me do this. I expect you saw the paper in the Annals.

With respect to exterpolation, or properly to the estimation of the effective number of the population sampled, I should now be inclined to work out R_+ and R_- for each interval of the series, and to make two estimates based respectively on forward and backward recaptures i.e.

$$\frac{10,000}{N} \qquad \frac{y+m+n+o+p+q}{R+(1+R_{+1}(1+R_{+2}(1+R_{+3}(1+R_{+4}(1+R_{+5})))))}$$

where R_{+s} stands for R_+ taken s intervals later. Equally, and practical independently, one could write:

$$\frac{10000}{N} \qquad \frac{b+c+d+e+f+g}{R-(1+R_{-1}(1+R_{-2}(1+R_{-3}(1+R_{-4}(1+R_{-5})))))}$$

The only point I can see in duplicating the estimates lies in using the mean square discrepancy between them as an estimate of (four times) the sampling variance of their means, or of the combined estimate formed by adding their numerators and denominators to make a new fraction.

With data such as yours, I feel sure that such an empyrical estimate of precision is better worth having than anything which depends on the approximate geometric progression of the y values.

With respect to the significance of population increase or decrease, this depends essentially on whether R_+ ~~and~~ R_- consistently retain the same sign over a series of consecutive intervals. As a test of significance, the kind of thing I should do would be, supposing you think you know enough to predict that over a period of nine weeks chosen in advance the population should be decreasing, to take crude differences ~~and~~ $R_+ - R_-$ and apply a t test to what you get. Equally, of course, the same could be done with the aggregate apparent increase over a chosen period, testing the consistency over a number of consecutive years.

I do not know whether sunspots are really associated with weather in the Tropics. They have, I think, no appreciable association in these latitudes, but perhaps someone has shown that African rainfall does follow the same cycle.

Stevens left here some months ago for Portugal, but I have not heard how he is getting on there. He was offered a post at Coimbra University.

P.T.O.

About your probability conundrum, I think the discrepancy arises from the fact that you have considered in one case the chance that the product of four probabilities is less than one in 10,000, and in the other the chance that the weighted mean of $\underline{z}$ exceeds a certain value for which, in one particular configuration, the product of the probabilities has this value; if one were to take out other sets of populations having the same product, the mean $\underline{z}$ would, of course, vary somewhat from set to set.

I am very glad that the Tsetse work is going on so well; perhaps you will be sending me something more for the Annals one of these days.

Yours sincerely,