

October 9, 1940

Dear Jeffreys

I thought the paper on scoring and enumeration might interest you, and I can see that it has done so. Section 2 is the justification of the statement in Section 1: "We shall show that the familiar process, etc. ~~is~~ a special case of the summation in the numerator of variable scores from different independent bodies of data, and in the denominator of the corresponding quantities of information". This general procedure is then applied in Sections 3 and 4 to cases where a simple enumeration is clearly inadequate, as it is in the final example on the chimpanzees.

To say that the whole population is interrelated is - is it not? - merely to say that it differs from other populations, real or imaginary, with which it is less related. The distinction I am using is the quite simple one between known close relationship and unknown, presumably remote, relationship, as is implied in such a phrase as consanguineous or non-consanguineous marriages. The fact that the close relationships I ~~investigate~~ do not give results, or only seldom give results greatly differing ^{ent} from a simple enumeration ignoring relationship is good evidence that remote relationships can be safely ignored.

I should guess that the negative binomial derived from the

distribution of the Poisson parameter \underline{m} being of the type $m^p e^{-m}$ is probably the only case worth detailed investigation among distributions where some continuous variation of \underline{m} is suspected, i.e. it is simple and exact, and any degree of variability can be imposed on \underline{m} . Of course other suitable forms can be given, such as $m^p e^{-m^2}$. These will give very similar results involving more complicated functions. A discontinuous distribution with a finite ~~possibility~~^{probability} for $n = \text{zero}$ must fairly often be helpful in practice.

Years ago I tried fitting some of these fertility distributions, and came to the conclusion that the deviations from reasonable theoretical forms are, when we use data large enough to give evidence worth having, probably due to errors of statement ~~&~~^{and to} heterogeneity, these being common defects in extensive data.

I am sending an offprint of Stevens' paper and my letter note under separate cover.

Yours sincerely,