22nd. November 1947.

My dear Irwin,

You wrote some time ago about the amount of information relative to the estimated value of the parameter α, measuring abundance of species. I remember that at the time the analysis was rather tricky, and I do not suppose that I can actually reproduce the paths that I then took. However, the starting point may be all that you really want.

If a species is observed, the probability that it has been observed r times is given by the truncated negative binomial distribution

$$\frac{1}{(1-x)^{-k}-1}\cdot\frac{(n+k-1)!}{(k-1)!\,n!}\,x^{n}$$

We may treat this as involving two parameters k and x, to be estimated from a sample of S individuals. The amount of information sought will be that relevant to k when estimated simultaneously with x. This is, indeed, equivalent to the simultaneous estimation of k and α, for samples of given size S, since

$$S = -\alpha \log(1-x) \quad .$$

We next, therefore, take out the values of

$\partial/\partial x \qquad S\cdot\dfrac{k(1-x)^{k-1}}{(1-x)^{-k}-1}+\dfrac{N}{x} \qquad \dfrac{S}{(1-x)\log(1-x)}+\dfrac{N}{x}$ when $k=0$.

$\partial/\partial k. \qquad \dfrac{S}{\log\frac{1}{1-x}}\sum \dfrac{1}{n}x^n\left(1+\frac{1}{2}+\cdots+\frac{1}{n-1}\right)+\dfrac{S}{2}\log(1-x).$

$-\partial^2/\partial x^2 \qquad -S\,\dfrac{x+\log(1-x)}{x(1-x)^2\log^2(1-x)}$

$-\dfrac{\partial^2}{\partial x\,\partial k}. \qquad \dfrac{S}{2(1-x)}$

$-\partial^2/\partial k^2 \qquad \dfrac{S}{12}\log^2(1-x)+\alpha_2+\alpha_3\left(1+\frac{1}{4}\right)+\alpha_4\left(1+\frac{1}{4}+\frac{1}{9}\right)+$

$\dfrac{S}{\log 1/_{1-x}}\sum_{n=2}^{\infty}\left\{1+\frac{1}{2^2}+\cdots+\frac{1}{(n-1)^2}\right\}\dfrac{x^n}{n}+\dfrac{S}{12}\log^2 x(1-x)$

by operating on the log likelihood for the limiting value $k = 0$.

To allow for the simultaneous estimation of x we must deduct the square of

$$\frac{\partial^2\alpha}{\partial x\,\partial k}.$$

from the value of $-\dfrac{\partial^2\alpha}{\partial x^2}$, so that the analytic expression for the amount of information sought comes out as follows

$$i=\frac{S}{\log 1/(1-x)}\sum_{n=2}^{\infty}\left\{1+\frac{1}{2^2}+\cdots+\frac{1}{(n-1)^2}\right\}\frac{x^n}{n}+\frac{S}{12}\log^2(1-x)+\frac{Sx\log^4(1-x)}{4\{x+\log(1-x)\}}$$

which is what I must have tabulated after removing the S.

Substituting $x=1-e^{-t}$ the three terms seem to be Bernoulli numbers?

$$\tfrac{1}{2}t-\tfrac{1}{12}t^2+\tfrac{1}{144}t^3+0\ t^4-\tfrac{1}{21600}t^5+0+\tfrac{1}{42\cdot 8}t^2$$

$$-\tfrac{1}{2}+\tfrac{1}{12}-\tfrac{1}{72}+\tfrac{1}{1580}+\tfrac{1}{12960}-\tfrac{1}{54432}+\tfrac{1}{1670\cdot 7!}$$

$$\tfrac{1}{12}t^2-\tfrac{1}{144}t^3+\tfrac{1}{1080}t^4+\tfrac{1}{32660}t^5-\tfrac{1}{443}$$

I suppose the first two are right, & the three may do

I do not at all clearly remember by what steps I obtained the numerical values given in the table. It is evidently possible, putting

$$x = 1 - e^{-t}$$

to obtain expansions in powers of t which are apparently good for the smaller values at the beginning of the table, but I do not think I can have used these expansions in the latter part of the table, where t may exceed 10. I rather fancy that no very tidy asymptotic formula exists for large values of t, or values of x very near to unity. So I feel pretty sure that I was using some effective analytic transformation other then those which have occurred to me since I received your letter.

Yours sincerely,

The argument of my table is

$$\log_{10} \frac{N}{s} = \log_{10} \left\{ \frac{-x}{(1-x)\log(1-x)} \right\} = \log_{10} \frac{e^t - 1}{t}.$$