102

The Concepts of Inverse Probability and Fiducial Probability Referring to Unknown Parameters.

By R. A. Fisher, F.R.S.

(Received November 3, 1932.)

1. *Criticism of Dr. Jeffreys's Method.*

In a paper published in these 'Proceedings'* Jeffreys puts forward a form of reasoning purporting to resolve in a particular case the primitive difficulty which besets all attempts to derive valid results of practical application from the theory of Inverse Probability.

For a normally distributed variate, x, the frequency element may be written

$$df = \frac{h}{\sqrt{\pi}} e^{-h^2(x-\mu)^2}\, dx,$$

where μ is the mean of the distribution, and h the precision constant. For the convenience of the majority of statisticians who prefer to use the standard deviation, σ, of the distribution, in place of the precision constant, we may note that

$$h^2 = \frac{1}{2\sigma^2},$$

and that this substitution may be made at any stage of the argument.

Jeffreys considers the question: What distribution *a priori* should be assumed for the value of h, regarding it as a variate varying from population to population of the *ensemble* of populations which might have been sampled? He sets forth a proof to the effect that this distribution must be of the form

$$df \propto \frac{1}{h}\, dh \propto \frac{-1}{\sigma}\, d\sigma. \tag{1}$$

That there should be a method of evolving such a piece of information by mathematical reasoning only, without recourse to observational material, would be in all respects remarkable, especially since the same principle of reasoning should, presumably, be applicable to obtain the distribution *a priori* of other statistical parameters. The proof can, however, scarcely in any case establish all that is claimed, since there is nothing to prevent our setting up an

* 'Proc. Roy. Soc.,' A, vol. 138, p. 48 (1932).

artificially constructed series of populations having any chosen distribution of h, such as

$$df = ae^{-ah}\,dh,$$

in which case Jeffreys's reasoning would certainly lead to a false conclusion. Moreover, Jeffreys himself seems to feel some doubt as to the general validity of the distribution (1), for he says: "The solution must break down for very small h . . . and for large h . . ."; though he does not indicate in what way his mathematical proof fails for these parts of the range.

The principle he uses rests on the fact that, if we have three independent observations from the same population, the probability that the last of the three shall be intermediate between the first two must be exactly 1/3. This fact is sufficiently obvious if all three observations are made afresh for each test, but we may note at once that, for any particular population, the probability will generally be larger when the first two observations are far apart than when they are near together. This is important since, as will be seen, the fallacy of Jeffreys's argument consists just in assuming that the probability shall be 1/3, *independently of the distance apart of the first two observations.*

Since the property used is possessed by all distributions, and therefore amongst others by normal distributions having all possible values of h, it might be argued *a priori* that its existence could not possibly be used to throw light upon the frequency distribution of h. It is, in fact, only the illegitimate inference stressed above which makes such further inferences appear to be possible.

Jeffreys's argument proceeds in four steps:—(*a*) the probability of the first two observations having assigned values is expressed in terms of the two parameters, the mean and the precision constant, of the population; (*b*) introducing the probability *a priori* of the two parameters having assigned values, their posterior probability of having them is obtained; (*c*) the probability of the third observation having an assigned value is found and integrated over all possible values of the parameters; (*d*) the expression so obtained is equated to 1/3, *without averaging it for all possible pairs of initial observations*; had this essential step been taken, the equation would have degenerated to an identity for all possible distributions *a priori.*

The argument as developed involves the assumption of a particular distribution *a priori* of the mean; it will be advantageous, therefore, in order to make clear the exact point at which a fallacy is introduced to exhibit the analysis without this assumption.

Let the population sampled have mean μ, and precision constant h, then the probability that the first two observations should lie in the ranges dx_1, dx_2, is

$$\frac{h^2}{\pi} e^{-h^2[(x_1-\mu)^2+(x_2-\mu)^2]} \, dx_1 \, dx_2.$$

Let the larger of these observations be $u + v$ and the smaller $u - v$, then the frequency element may be re-written

$$\frac{2h^2}{\pi} e^{-2h^2[(u-\mu)^2+v^2]} \, du \, dv.$$

If now $f(h)\,dh$ is the prior probability that h lies in the range dh, the probability of assigned values for u, v and h will be

$$\frac{2h^2 f(h)}{\pi} e^{-2h^2[(u-\mu)^2+v^2]} \, du \, dv \, dh,$$

and that for assigned values of u, v, h, and the third observation, x_3, will be

$$\frac{2h^3 f(h)}{\pi^{3/2}} e^{-2h^2[(u-\mu)^2+v^2]-h^2(x_3-\mu)^2} \, du \, dv \, dh \, dx_3.$$

Writing $x_3 = u + c$, since the magnitude of c determines whether or not the third observation lies between the others, we may now average over all values of u, by integrating with respect to that variate from $-\infty$ to ∞. This gives

$$\frac{2h^2 f(h)}{\pi\sqrt{3}} e^{-2h^2v^2-\frac{2}{3}h^2c^2} \, dv \, dh \, dc, \tag{2}$$

in which μ, the mean of the population, has disappeared, showing that its value, and therefore its distribution *a priori*, is irrelevant. Equation (2) corresponds with Jeffreys's equation (3) (p. 49), save that the differential elements, dv and dc, have been retained.

For any given value of v, therefore, the probability of the third observation lying between the first two will be found by integrating (2) with respect to c and h, and evaluating the frequency with which c is less than v.

Writing

$$\alpha(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} \, dt,$$

the integral with respect to c from $-v$ to v is

$$\frac{2hf(h)}{\sqrt{2\pi}} \cdot \alpha\left(hv\sqrt{\tfrac{2}{3}}\right) e^{-2h^2v^2} \, dv \, dh, \tag{3}$$

while the integral over all values is

$$\frac{2hf(h)}{\sqrt{2\pi}} e^{-2h^2v^2}\, dv\, dh\,; \tag{4}$$

it is by equating the integral of (4) with respect to h, to three times that of (3) *for every possible value of* v, that Jeffreys obtains a unique form for $f(h)$.

All that we really know, however, is that *on the average of all values of* v, the probability is exactly one-third. We should, therefore, integrate (3) and (4) with respect to v, over all values from 0 to ∞. For (3) we have, as is also shown by Jeffreys (p. 50),

$$\tfrac{1}{3} f(h)\, dh,$$

and for (4),

$$f(h)\, dh,$$

so that the fact that the probability is just one-third is assured, irrespective of h, and therefore for every frequency element of that variate independently. It is merely because the substitution $f(h) = 1/h$ makes integration with respect to h equivalent to integration with respect to v, that this special distribution *a priori* satisfies Jeffreys's condition.

2. *The Method of the Fiducial Distribution.*

An altogether different approach may perhaps make clear why the consideration of proportional rather than absolute increments in the variables h and σ should lead to simpler mathematical consequences. If from a series of n' observations, x, drawn from a normal population with standard deviation σ, or variance σ^2, we make an estimate of this variance, based on the sum of the squares of the deviations of the observations from their mean, in the form

$$s^2 = \frac{1}{n'-1}\, \mathrm{S}\,(x - \bar{x})^2,$$

then the estimate s^2 is known* to be distributed, in random samples, in a manner, which depends only on the unknown parameter σ of the sampled population, and is specified by the formula

$$df = \frac{1}{\frac{n'-1}{2}\,!} \left\{\frac{(n'-1)\,s^2}{2\sigma^2}\right\}^{\frac{1}{2}(n'-1)} e^{-\frac{(n'-1)\,s^2}{2\sigma^2}}\, d\left\{\frac{(n'-1)\,s^2}{2\sigma^2}\right\}$$

The distribution of the ratio s/σ is thus independent of all unknown parameters, and is calculable solely from the number of observations in the sample, or, to

* Fisher, 'Metron.,' vol. 5, p. 90 (1926).

cover a more general class of cases, from the number of degrees of freedom of the residuals, from which the variance is estimated. From this distribution, which has been sufficiently tabulated, we can assert, without reference to any unknown quantities, or to their unknown probabilities *a priori*, with what frequency any particular value of the ratio s/σ will be exceeded in random samples ; or, what is often more convenient, for any given probability, such as 0·99 or 0·01, what is the value of the ratio which will be exceeded with this probability. Thus for 10 degrees of freedom, such as we should have from a sample of 11 observations of a normally distributed variate, it is known that the ratio will exceed 2·3209 in 1 per cent. of cases,* and will fall short of 0·2558 in another 1 per cent. If, therefore, we designate by $s_{0\cdot01}(\sigma)$ that value of s which, for a given σ, will be exceeded in exactly 1 per cent. of trials, we have the simple relationship

$$s_{0\cdot01}(\sigma) = \sigma \times 2\cdot3209,$$

and this value of s we may term the 1 per cent. value of s for the value of σ considered. If now we designate by $\sigma_{0\cdot99}(s)$ that value of σ for which s is the 1 per cent. value, then evidently

$$\sigma_{0\cdot99}(s) = s/2\cdot3209,$$

and we may term this value of σ the 99 per cent. value of σ for the given value of s. Evidently where s is known this value of σ is also known. Moreover, the inequality

$$s > s_{0\cdot01}(\sigma), \tag{5}$$

is equivalent to the inequality,

$$\sigma < \sigma_{0\cdot99}(s), \tag{6}$$

since for any value of the probability chosen the corresponding values of s and σ increase together from 0 to ∞.

Now we know that the inequality (5) will be satisfied in just 1 per cent. of random trials, whence we may infer that the inequality (6) will also be satisfied with the same frequency. Now this is a probability statement about the unknown parameter σ, easily translatable into an equivalent statement about the unknown parameter h, in terms of known quantities only. For example, if s is an estimate derived from 10 degrees of freedom we know that σ has a probability 0·01 of being less than $s/2\cdot3209$, and in like manner we know its probability of falling between any other assigned limits. Probability statements of this type are logically entirely distinct from inverse probability

* Fisher, "Statistical Methods for Research Workers," Table III, 4th ed. (1932.)

statements, and remain true whatever the distribution *a priori* of σ may actually be. To distinguish them from statements of inverse probability I have called them statements of fiducial probability. This distinction is necessary since the assumption of a given frequency distribution *a priori*, though in practice always precarious, might conceivably be true, in which case we should have two possible probability statements differing numerically, and expressible in a similar verbal form, though necessarily differing in their logical content. The probabilities differ in referring to different populations; that of the fiducial probability is the population of all possible random samples, that of the inverse probability is a group of samples selected to resemble that actually observed.

It is the lack of this distinction that gives a deceptive plausibility to the frequency distribution *a priori*

$$df = d\sigma/\sigma = d(\log \sigma).$$

For this particular distribution *a priori* makes the statements of inverse and of fiducial probability numerically the same, and so allows their logical distinctness to be slurred over. It is, moreover, as Jeffreys, by his references to large and small values of h, clearly perceives, an impossible distribution *a priori*, since it gives a zero probability *a priori* for h lying between any finite limits, however far apart. In the fiducial form of statement this difficulty does not occur.

3. *Summary.*

(1) The argument of Jeffreys in favour of a particular frequency distribution *a priori* for the precision constant of a normally distributed variate rests on the fallacy that the probability of the last of three observations, lying between the previous two, should be one-third, *irrespective of the distance apart of the two previous observations.*

(2) The apparent simplicity of the results of assuming this particular distribution *a priori* rests on the fact that the *inverse* and the *fiducial* probability statements about the unknown parameter are thereby made to coincide, though logically they are entirely distinct. This particular distribution *a priori* is, however, not only hypothetical but unacceptable as such, since it implies that all ranges of values of the parameter covering finite ratios, however great, are infinitely improbable.