THE FIDUCIAL ARGUMENT IN STATISTICAL INFERENCE

By R. A. FISHER, Sc.D., F.R.S.

I. The nature of fiducial probability

In a series of papers from 1930, the author has called attention to a form of argument, which seems to have been entirely overlooked by the classical writers on probability, but which arises naturally from the exact tests of significance, when the variate is tabulated in terms of the probability. This form of argument leads in certain cases to rigorous probability statements about the unknown parameters of the population from which the observational data are a random sample, without the assumption of any knowledge respecting their probability distributions *a priori*. For such deductions we need to know the exact sampling distributions of statistical estimates, calculable from the observations only, of the unknown parameters, and these distributions must be continuous. It was probably these restrictions which stood in the way of the recognition, by the early writers on probability, of a form of argument having both theoretical interest and practical value; for the problems of distribution of which they possessed the exact solutions were nearly all discontinuous, being, like the binomial expansion, and the many similar generating functions given by Laplace, distributions of frequencies, rather than of continuously variable measurements, or functions calculated from these. The exact treatment of the mean of a normal sample was first given by "Student" in 1908, and since that time numerous exact solutions have become available, covering somewhat completely the problems connected with normally distributed variates, in addition to others of a more miscellaneous character.

The form of argument is extremely simple and may be illustrated by applying it to "Student's" solution. If a sample of n observations, $x_1, \ldots, x_n$, has been drawn from a normal population having a mean value μ, and if from the sample we calculate the two statistics

$$\bar{x} = \frac{1}{n} S(x)$$

and

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

where S stands for summation over the sample, "Student" has shown (1925) that the quantity t, defined by the equation

$$t = \frac{(\bar{x}-\mu)\sqrt{n}}{s},$$

is distributed in different samples in a distribution dependent only from the size of the sample, n. It is possible, therefore, to calculate, for each value of n, what value of t will be exceeded with any assigned frequency, P, such as 1 per cent. or 5 per cent. These values of t are, in fact, available in existing tables (Fisher, 1925–34).

It must now be noticed that t is a continuous function of the unknown parameter, the mean, together with observable values, $\bar{x}$, s and n, only. Consequently the inequality

$$t > t_1$$

is equivalent to the inequality $\mu < \bar{x} - st_1/\sqrt{n}$,

so that this last inequality must be satisfied with the same probability as the first. This probability is known for all values of t_1, and decreases continuously as t_1 is increased. Since, therefore, the right-hand side of the inequality takes, by varying t_1, all real values, we may state the probability that μ is less than any assigned value, or the probability that it lies between any assigned values, or, in short, its probability distribution, in the light of the sample observed.

It is of some importance to distinguish such probability statements about the value of μ, from those that would be derived by the method of inverse probability, from any postulated knowledge of the distribution of μ in the different populations which might have been sampled. It is only when the idea is totally set aside that we are seeking an inverse probability, that the meaning of fiducial probability is clearly apprehended. The inverse probability distribution would specify the frequency with which μ would lie in any assigned range $d\mu$, by an absolute statement, true of the aggregate of cases in which the observed sample yielded the particular statistics $\bar{x}$ and s. This can be found by Bayes' procedure, if the prior distribution of μ is known. The distribution which we have obtained is independent of all prior knowledge of the distribution of μ, and is true of the aggregate of all samples without selection. It involves $\bar{x}$ and s as parameters, but does not apply to any special selection of these quantities. To distinguish it from any of the inverse probability distributions derivable from the same data it has been termed the *fiducial* probability distribution, and the probability statements which it embraces are termed statements of fiducial probability. To attempt to define a prior distribution of μ which shall make the inverse statements coincide numerically with the fiducial statements is really to slur over this distinction between the meaning of statements of these two kinds.

It is necessary to emphasise also that statements similar to those of fiducial probability can only represent the true state of knowledge derivable from the sample, if the statistics used contain the whole of the relevant information which the sample provides. If, for example, an estimate s', derived from the mean error, had been used in place of one derived from the mean square error, and a quantity t' had been defined by the equation

$$t' = \frac{(\bar{x} - \mu)\sqrt{n}}{s'},$$

the distribution of t', like that of t, would depend only on the size of the sample; and probability statements accurate for t' could be expressed in terms of μ. The probability distribution for μ obtained in this way would, of course, differ from that obtained from t, and the probability statements derived from the two distributions would be discrepant. There is, however, in the light of the theory of estimation, no difficulty in choosing between such inconsistent results, for it has been proved that, whereas s' uses only a portion of the information utilised by s, on the contrary, s utilises the whole of the information used by s', or indeed by any alternative estimate. To use s', therefore, in place of s would be logically equivalent to rejecting arbitrarily a portion of the observational data, and basing probability statements upon the remainder as though it had been the whole. Dr J. Neyman has unfortunately attempted to develop the argument of fiducial probability in a way which ignores the results from the theory of estimation, in the light of which it was originally put forward. His proofs, therefore, purport to establish the validity of a host of probability statements many of which are mutually inconsistent.

When, as with inferences respecting a single parameter based on the use of sufficient statistics only, we obtain a unique probability distribution for that parameter, all possibility of admitting mutually conflicting inferences is excluded, and the resulting distribution may be properly termed the fiducial distribution of the parameter. The same would be true of inferences concerning the simultaneous values of two or more parameters, if such a unique simultaneous distribution could be obtained. Such a simultaneous distribution does not in general follow by any simple generalisation of the argument from the probability integral by which the distribution of a single parameter may be obtained, such as that recently developed by Dr Neyman. It is the purpose of the present note to demonstrate the possibility of thus completing the solution, in the simple class of problems which arise from the normal distribution.

II. Posterior fiducial inferences

As a preliminary we may consider the problem: Given the value of n observations $x_1, \ldots, x_n$ drawn from a normal population, to find the fiducial distribution of an additional observation, x, not yet made.

Two points may be noted. First, that, in the theory of inverse probability, problems of this kind are only to be solved after the simultaneous distribution of the population parameters has been obtained, and by means of this simultaneous distribution; whereas by the fiducial argument they are solved directly. Second, that the concept of a fiducial distribution is now being applied not to a parameter of the population, but to an observation, the distribution of which is given in terms of such parameters. With these comments we may now suppose that the observed sample has yielded statistics

$$\bar{x} = \frac{1}{n} S(x)$$

and

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

Let, now,

$$t = \sqrt{\frac{n}{n+1}} \frac{x-\bar{x}}{s}.$$

That t, so defined, is distributed, independently of the parameters of the population, in "Student's" distribution for $n-1$ degrees of freedom, follows (Fisher, 1925) from the fact that it is the ratio of a quantity $x-\bar{x}$, normally distributed about zero, to an estimate of its standard error, statistically independent of $x-\bar{x}$, based on the mean square deviation among the different members of the original sample.

It follows that the known probability that t should exceed any assigned value t_1 is the probability that x should exceed the value

$$\bar{x}+st_1\sqrt{\frac{n+1}{n}}.$$

"Student's" distribution, with the factor $\sqrt{1+\frac{1}{n}}$, therefore provides the fiducial distribution of

$$\frac{x-\bar{x}}{s},$$

in which the only unknown element is the future observation, x.

We may next consider the analogous, more general problem: Given a sample of n observations yielding the statistics $\bar{x}$ and s, to find the fiducial frequency distribution of the statistics $\bar{x}'$ and s' derived from a subsequent sample of n' observations.

In proceeding to the solution, we again avoid all reference to unknown and hypothetical parameters, and develop the solution in terms of directly observable quantities only, by obtaining two quantities t and z, the simultaneous distribution of which is known with exactitude, and which are expressible in terms of the observable features of the two samples.

For this purpose we have

$$t=\frac{(\bar{x}-\bar{x}')\sqrt{nn'(n+n'-2)}}{\sqrt{n+n'}\sqrt{(n-1)s^2+(n'-1)s'^2}}$$

and

$$z=\log s-\log s'.$$

Then it is well known that, in the aggregate of pairs of samples drawn from such a population, t is distributed in "Student's" distribution for $n+n'-2$ degrees of freedom, while z is distributed in its own characteristic distribution, determined by the two numbers of degrees of freedom,

$$n_1=n-1,\quad n_2=n'-1.$$

The values of $\bar{x}'$ and s' are, of course, determined by the values of t and z to which they lead, so that, substituting for t and z in their known distributions, we have for the simultaneous fiducial distribution of $\bar{x}'$ and s' the following expression

$$df=\frac{\frac{1}{2}(n+n'-3)!}{\frac{1}{2}(n-3)!\,\frac{1}{2}(n'-3)!\sqrt{\pi}}\,\frac{2\,(n-1)^{\frac{1}{2}(n-1)}(n'-1)^{\frac{1}{2}(n'-1)}s^{n-1}s'^{\,n'-2}\,dx'\,ds'}{\sqrt{\frac{1}{n}+\frac{1}{n'}}\left\{(n-1)s^2+(n'-1)s'^2+\frac{(\bar{x}-\bar{x}')^2}{\sqrt{\frac{1}{n}+\frac{1}{n'}}}\right\}^{\frac{1}{2}(n+n'-1)}}.$$

The general distribution has been established merely from the known simultaneous distribution of t and z, and in itself gives the essential information respecting a subsequent sample for the sake of which knowledge of the population parameters would, in the traditional procedure, be sought. We may, however, now, as a particular case, allow n' to tend to infinity, and in consequence, the statistics $\bar{x}'$ and s' to tend to the parametric values, μ and σ. The simultaneous fiducial distribution of μ and σ found in this way is

$$df = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2}(\mu-\bar{x})^2} d\mu . \frac{1}{\frac{1}{2}(n-3)!} \left\{\frac{(n-1)s^2}{2\sigma^2}\right\}^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2d\sigma}{\sigma}.$$

The two parameters μ and σ are not independently distributed, that of μ being distributed for any given σ with variance σ^2/n, but the marginal distribution of μ found by integrating, with respect to σ, from 0 to infinity is, as might be expected,

$$df = \frac{\frac{1}{2}(n-2)!}{\frac{1}{2}(n-3)!\sqrt{\pi(n-1)}} \frac{1}{\left\{1+\frac{n(\mu-\bar{x})^2}{(n-1)s^2}\right\}^{\frac{1}{2}n}} \frac{\sqrt{n}\,d\mu}{s},$$

in accordance with the fact that $$\frac{(\mu-\bar{x})\sqrt{n}}{s}$$

is distributed as is t for $(n-1)$ degrees of freedom. Similarly, the marginal distribution of σ is

$$df = \frac{1}{\frac{1}{2}(n-3)!} \left\{\frac{(n-1)s^2}{2\sigma^2}\right\}^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2d\sigma}{\sigma},$$

in accordance with the fact that $$\frac{(n-1)s^2}{\sigma^2}$$

is distributed as is χ^2 for $(n-1)$ degrees of freedom.

It thus appears that, for the special case of the mean and variance of the normal distribution, there is no difficulty in extending the notion of fiducial probability unambiguously to the simultaneous distribution of two parameters.

In general, it appears that if statistics $T_1, T_2, T_3, \ldots$ contain jointly the whole of the information available respecting parameters $\theta_1, \theta_2, \theta_3, \ldots$, and if functions $t_1, t_2, t_3, \ldots$ of the T's and θ's can be found, the simultaneous distribution of which is independent of $\theta_1, \theta_2, \theta_3, \ldots$, then the fiducial distribution of $\theta_1, \theta_2, \theta_3, \ldots$ simultaneously may be found by substitution. For a single statistic having a distribution dependent only on a single parameter, of which it is a sufficient estimate, such a function is always provided by the probability integral. Hence the generality of the method in univariate cases.

III. Application of the method to special problems

The process of solution of the fiducial distributions of parameters, by the recognition of quantities of known distribution functionally related to them, is a powerful tool for the solution of a variety of problems which offer difficulties to other methods of approach. This may be illustrated in two such problems, which are occasionally of practical interest.

(i) *The difference between the means of two normally distributed populations*

Let us suppose that a sample of n observations has yielded a mean, $\bar{x}$, and an estimated variance of the *mean*, s^2, so that

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

then we know that if μ is the mean of the population,

$$\mu = \bar{x} + st,$$

where t is distributed in "Student's" distribution. Similarly, for the mean of a second population, of which we have n' observations, we may write

$$\mu' = \bar{x}' + s't',$$

where t is distributed in "Student's" distribution for $n'-1$ degrees of freedom, independently of t. If now

$$\mu' - \mu = \delta, \quad \bar{x}' - \bar{x} = d,$$

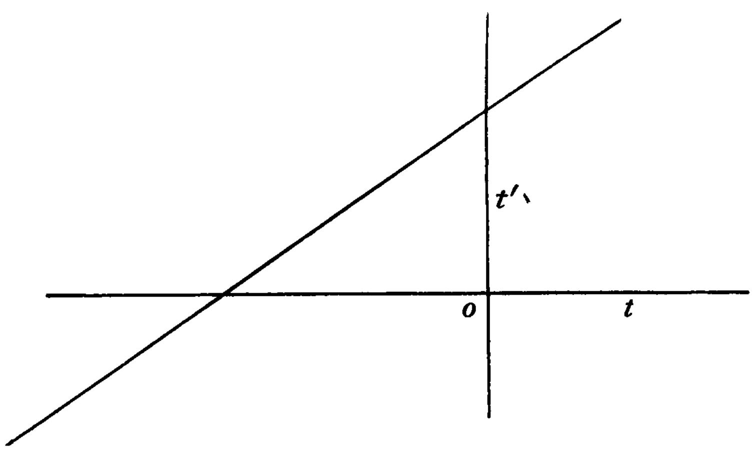
we find that

$$\epsilon = \delta - d = s't' - st;$$

and since s' and s are known, the quantity represented on the right has a known distribution, though not one which has been fully tabulated. The equation may be written

$$\epsilon = \sqrt{s^2 + s'^2}\,(t' \cos R - t \sin R),$$

where $\tan R = s/s'$, so that R is a known angle.



If t and t' be taken as the co-ordinates of a point on a plane, the frequency of the observations falling within any area of the plane is calculable. The points for which ϵ has any given value lie on a straight line, at a distance from the origin $\pm\,\epsilon/\sqrt{s^2+s'^2}$, and making an angle R with the axis of t. The fiducial probability that ϵ exceeds any given value is the frequency in the area above this line. If n and n' are both increased, the distribution of ϵ tends to be normal and independent of R; when R is 0° or 90° the distribution is of "Student's" form. In general it involves n, n' and R, and for any chosen probability, therefore, requires a table of triple entry.

Any fiducial distribution supplies a series of possible tests of significance. In this case, since d is known, we may use $d/\sqrt{s^2+s'^2}$ to test the hypothesis that δ has the chosen value zero. This is, in fact, the exact test for the significance of the difference, d, between the observed means, equivalent to that given in 1929 by W.-V. Behrens.

(ii) *The variance of a normally distributed set of means*

Suppose we have k samples of n observations each from equally variable populations, or other such material suitable for the analysis of variance. Let the analysis obtained, apart from excluded items, be as follows:

	Degrees of freedom	Sum of squares	Mean square
Among samples	n_1	A	a
Error	n_2	B	b

If a is significantly greater than b, as shown by the z test, where

$$2z = \log a - \log b,$$

there is reason to suppose that the means of the k population sampled were not all equal. In some such cases, though not in all, it is appropriate to suppose that these means themselves constitute a sample from a normal population with unknown variance. Then the test of significance will have indicated that this variance is significantly greater than zero. Now if θ is the variance within samples, and ϕ the variance of the population of means,

$$B = \theta\chi_2^2$$

or

$$b = \frac{\chi_2^2}{n_2}\theta,$$

where χ_2^2 is distributed as is the sum of the squares of n_2 independent normal deviates all having unit variance. Similarly

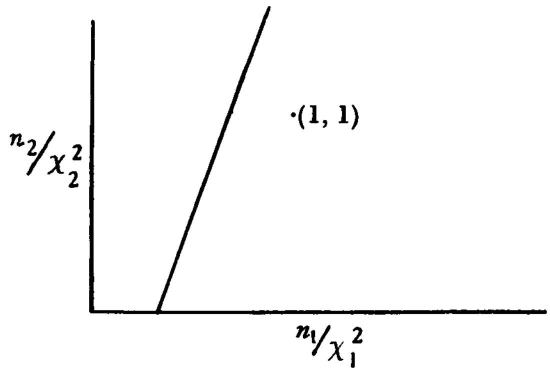
$$a = \frac{\chi_1^2}{n_1}(\theta + k\phi).$$

Hence

$$a\frac{n_1}{\chi_1^2} - b\frac{n_2}{\chi_2^2} = k\phi.$$

But χ_1^2 and χ_2^2 are distributed independently in distributions of known form, hence the distribution of ϕ may be calculated from their simultaneous distribution.

Let n_1/χ_1^2 and n_2/χ_2^2 be taken as the coordinates of a point, then the set of points consistent with any given value of ϕ lie on a straight line, making with the axis of n_1/χ_1^2 an angle, the tangent of which is a/b, or e^{2z}. The fiducial probability of ϕ exceeding any chosen value is the total frequency to the right of the corresponding line. The observed



value of z is significant if the parallel line through the origin has to the right of it some 95 or 99 per cent. of the total, according to the level of significance chosen. Lines which do not strike the axis of abscissae to the right of the origin, correspond to negative values of ϕ, and are without interest for the problem stated. For large values of n_1 and n_2, the frequency distribution tends to a normal form, without correlation, centred on the point (1, 1).

Tabulation of the exact solution encounters in both problems the difficulty of triple entry, which may be largely mitigated by the use of a harmonic series of values of n_1 and n_2 appropriate for asymptotic interpolation, as in the table of z. It is doubtful, too, whether the exact tests will often differ materially from the simple approximate tests commonly used. Nevertheless, a table of either kind, once constructed, would be probably of sufficiently frequent use in enabling experimenters to form a rapid opinion as to what interpretations of their data should be regarded as acceptable. They are here principally of interest as illustrating the simplicity with which the fiducial argument may be applied to this type of problem.

REFERENCES

Bayes, T. (1763). "An essay towards solving a problem in the doctrine of chances." *Philos. Trans.* **53**, 370–418.

Behrens, W.-V. (1929). "Ein Betrag zur Fehlenberechnung bei wenigen Beobachtungen." *Landw. Jb.* **68**, 807–37.

Fisher, R. A. (1925–34). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

—— (1925). Applications of "Student's" Distribution. *Metron.* **3**, 90–104.

—— (1930). "Inverse probability." *Proc. Camb. phil. Soc.* **26**, 528–35.

—— (1933). "The concepts of inverse probability and fiducial probability referring to unknown parameters." *Proc. roy. Soc.* A, **139**, 343–8.

—— (1934). "Two new properties of mathematical likelihood." *Proc. roy. Soc.* A, **144**, 285–307.

Neyman, J. (1934). "On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection." *J. R. Statist. Soc.* **97**, 558–625.

"Student" (1925). New tables for testing the significance of observations. *Metron.* **3**, 105–8.