Statistical Research Report
Institute of Mathematics
University of Oslo

No. 7
July 1969

An interpretation of fiducial probability that make sense to a Neyman-Pearson statisticien.

By

Else Sandved

1. Fiducial prediction or posterior fiducial inference.

R.A. Fisher has introduced a concept which he called "fiducial prediction" in [1] (p.113 etc.) and "posterior fiducial inferences" in [2]. In both these cases he considers the following example :

$x_1,\ldots x_n$ are independent and normally distributed with common unknown mean and common unknown variance. Then it is possible to make fiducial probability statements about a future observation x (drawn at random from the same population) by noticing that $t=\sqrt{\frac{n}{n+1}}\,\frac{x-\bar{x}}{s}$

(where $\bar{x}=\frac{1}{n}\sum_1^n x_i,\ s^2=\frac{1}{n-1}\sum_1^n (x_i-\bar{x})^2$) has a Student distribution

with n-1 d.f. In [2] (p.394) he writes,"... the known probability that t should exceed any assigned value t_1 is the probability that x should exceed the value

$\bar{x}+st_1\sqrt{\frac{n+1}{n}}$."Students" distribution, with the factor $\sqrt{\frac{n+1}{n}}$, therefore provides the fiducial distribution of

$\frac{x-\bar{x}}{s}$, in which the only unknown element is the future observation x."

Hence, the system of equations

$$_1\sqrt{\frac{n+1}{n}})=P(t\geq t_1);\quad -\infty<t_1<\infty$$

or equivalently

(1) $$P(x\geq\bar{x}+st_p\sqrt{\frac{n+1}{n}})=1-p;\quad o<p<1$$

(where t_p is the p-fractile of the Student distribution with n-1 d.f.) is obviously that Fisher meant by fiducial probability statements about a future observation in this example.

The probability statements in (1) are of course completely in accordance with "Neyman-Pearson probability", and would be of interest to a Neyman-Pearson statistician who wished to predict the value of x after having observed $x_1,\dots,x_n$. The system (1) also of course provides a sort of "distribution" for x, though not of the usual type, since the parenthesis in (1) contains other stochastic variables than x; $\bar{x}$ and s are also stochastic variables. (1) is not true conditionally given $x_1,\dots,x_n$.

2. Fiducial probability statements about parameters.

In [1], p. 114, Fisher writes,...."(fiducial) probability statements about the hypothetical parameters..." and further "... once their equivalence is understood to predictions in the form of (fiducial) probability statements about future observations, they are seen not to incur any logical vagueness by reason of the subjects of them being relatively unobservable."

To get probability statements about a parameter which are as far as possible eqivalent to the system(1), one may notice this : One thing characterizing (1) is that the parenthesis contains only a the future observation x considered as a stochastic variable, and b known functions of stochastic variables which are going to be observed at the present time.

Hence, in a statistical inference problem where the subject of interest is an unknown parameter, a system of equations similar to (1), where the parenthesis contains only a the parameter of interest considered as an unobservable stochastic variable and b known functions of stochastic variables which are going to be observed, could be interpreted as the fiducial probability distribution for the parameter.

To get a "distribution" of this type for the mean in the normal situation (treated by Fisher in [2], one may proceed in this way (using another notation than Fisher did) :

The parameter Ⓗ is an unobservable stochastic variable with an a priori completely unknown distribution function F. $X_1,\dots,X_n$, given Ⓗ $= \theta$, are conditionally independent and identically normally distributed with mean

θ and unknown variance σ^2. Hence, the conditional distribution of

$$T = \frac{\bar{X} - \textcircled{H}}{S} \sqrt{n}, \text{ (where } \bar{X} = \frac{1}{n}\sum_1^n X_i,$$

$S^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2$), given $\textcircled{H} = \theta$, is the Student distribution with $n-1$ d.f., and therefore we get the system of equations.

$$P(\textcircled{H} \geq \bar{X} + t_p \frac{S}{\sqrt{n}} \mid \textcircled{H} = \theta) = 1-p; \quad 0 < p < 1$$

(where t_p is the p-fractile of the Student distribution with $n-1$ d.f.), and since

$$\int_\theta P(\textcircled{H} \geq \bar{X} + t_p \frac{S}{\sqrt{n}} \mid \textcircled{H} = \theta) dF = \int_\theta (1-p) dF = 1-p$$

we finally get

$$(2) \qquad P(\textcircled{H} \geq \bar{X} + t_p \frac{S}{\sqrt{n}}) = 1-p; \quad 0 < p < 1.$$

The system (2) has a frequency or probability interpretation that make sense to and would be of interest to a Neyman-Pearson statistician in the case above, where the subject of interest is the realization θ of $\textcircled{H}$, and the system (2) also gives a sort of "distribution" for $\textcircled{H}$.

One may also notice that (2) is not true conditionally given $X_1 = x_1, \ldots, X_n = x_n$.

Fraser [3] (p.662) gives a different frequency interpretation of fiducial probability in the following example :

Consider a very large number of normal independent variables with variance equal to 1 and with mean μ, and let μ^* designate the values for the parameters relative to the observed variables. Then the values μ^* has a normal frequency distribution with center in the observed variable and with variance equal to 1.

The arguments in sections 1 and 2 of the present paper may indicate that Fisher by fiducial probability distribution for a parameter in the example above could have meant something like the system (2). However, this indication is somewhat weakened in the next section.

3. The Behrens-Fisher problem.

If in the example in section 2

$P(\Theta \geq \bar{X}+t_p \frac{S}{\sqrt{n}}) = 1-p$ for all a priori possible F, then it

is also true for the F assigning probability 1 to a fixed θ. Hence we get

$$(3) \qquad P(\theta \geq \bar{X} + t_p \frac{S}{\sqrt{n}}) = 1-p \quad ; \qquad 0 < p < 1$$

for every θ. That (2) follows from(3) were shown in section 2. This shows that there is a numerical agreement between (one-sided) confidence intervals and fiducial (in the interpretation above) intervals (or between confidence coefficients and fiducial probability). But the logical content of the two concepts differs of course, since in the argument leading to a confidence interval the parameter is not considered as an unobservable random variable, but as an unknown constant.

Since in the Behrens-Fisher problem there is a numerical difference between the two sorts of probabilities, as shown by Neyman [4], two solutions are possible :

1 Either Fisher has meant by fiducial probability distribution something more than or something different from a system of the type (2),

2 or Fisher has made a logical mistake when he deduced his result in the Behrens-Fisher problem.

To try to judge the second possibility, we proceed in the following way, which has some resemblance with the procedure of Fisher in [2] p. 396.

Let M_1 and M_2 be unobservable stochastic variables. Given $M_1 = \mu_1$ and $M_2 = \mu_2$ the variables $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ are conditonally independent and normally distributed with $EX_i = \mu_1$, $EY_i = \mu_2$, Var $X_i = \sigma_1^2$, Var $Y_i = \sigma_2^2$ for all i, where σ_1^2 and σ_2^2 are unknown. (One could also let the variances be unobservable stochastic variables with completely unknown distribution, and let the above be true conditionally, given the variances equal to σ_1^2 and σ_2^2, respectively.) The subject of interest is the realization $\theta = \mu_1 - \mu_2$ of the stochastic variable $\Theta = M_1 - M_2$.

Let $\bar{X} = \frac{1}{m}\sum_1^m X_i$, $\bar{Y} = \frac{1}{n}\sum_1^n Y_i$ $S_1^2 = \frac{1}{m(m-1)}\sum_1^m (X_i - \bar{X})^2$,

$S_2^2 = \frac{1}{n(n-1)}\sum_1^n (Y_i - \bar{Y})^2$.

Then $T_1 = \frac{\bar{X} - \mu_1}{S_1}$ and $T_2 = \frac{\bar{Y} - \mu_2}{S_2}$

are conditionally independent and Student distributed with $m-1$ and $n-1$ d.f. respectively, given $M_1 = \mu_1$ and $M_2 = \mu_2$

Consider the cartesian coordinate system with axes t_1 and t_2. Let r, $0 < r < \frac{\pi}{2}$ be a given angle and let a line through origin make the angle r with the t_1-axis. The distance from a point (t_1, t_2) to the line equals $t_2 \cos r - t_1 \sin r$. Since the distribution of (T_1, T_2) is completely specified, it is possible to find, for every p such that $0 < p < 1$, a constant $c(p,r)$ such that

(4) $$P(T_2 \cos r - T_1 \sin r \geq c(p,r)) = p .$$

Now let $\sin R = \frac{S_1}{\sqrt{S_1^2 + S_2^2}}$ and $\cos R = \frac{S_2}{\sqrt{S_1^2 + S_2^2}}$.

Let $R=r$ be given, and write the system (4) as a system of conditional probabilities in this way

(5) $$P(T_2 \cos R - T_1 \sin R \geq c(p,R) \mid R = r) = p \; ; \; 0 < p < 1 .$$

By integrating (5) with respect to the marginal distribution for R we get

(6) $P(T_2 \cos R - T_1 \sin R \geq c(p,R)) = p; \quad 0 < p < 1.$

The system (6) is equivalent to

(7) $P(\theta - (\bar{X} - \bar{Y}) \geq \sqrt{S_1^2 + S_2^2}\, c(p,R)) = p; \quad 0 < p < 1.$

The probabilities in (4) - (7) are all conditional, given $\textcircled{H} = \theta$. In the same way as in sec tion 2 we get by integrating (7) with respect to the marginal distribution for $\textcircled{H}$

(8) $P(\textcircled{H} \geq (\bar{X} - \bar{Y}) + \sqrt{S_1^2 + S_2^2}\, c(p,R)) = p; \quad 0 < p < 1,$

which could have been interpre ted as the fiducial distribution for $\textcircled{H}$, if it had been true, but it is not because of the erranous step from (4) to (5) . (All other steps are valid.) Is it possible that Fisher did an error similar to the error mentioned, i.e. that he thought he worked with conditional probabilities when he did not ? If this is the case, then that could be the reason why the numerical results of Fisher and Neyman [4] in this problem do not agree; this disagreement need not be explained by interpreting the fiducial probability as a special "sort" of probability.

In any case, whatever Fisher has meant by fiducial probability, it is clear that he meant it to have a frequency interpretation (perhaps in addition to something else) because of his words in [2], p.396:

"The fiducial probability.....is the frequency in the area...".

With an interpretation such as (2) of fiducial probability it is possible to get other fiducial probability statements for $\textcircled{H}$ in the Behrens-Fisher case, instead of the wrong ones (8), by using the Welch method. Let the notation be as above. Then (see for instance Sverdrup [5] p.166 etc.), given all the parameters, the conditional distribution of

$$\dot{T} = \frac{\bar{X} - \bar{Y} - \theta}{\sqrt{S_1^2 + S_2^2}} \quad \text{is}$$

(approximately) Student distributed with

ν d.f., where $\frac{1}{\nu} = \frac{\alpha^2}{m-1} + \frac{(1-\alpha)^2}{n-1}$

and where $\alpha = \frac{1}{1 + \frac{m}{n}\frac{\sigma_2^2}{\sigma_1^2}}$. Hence ν

lies between $\min(m-1, n-1)$ and $m+n-2$,

and

(9) $p(t,\nu), = P(T \leq t) = P(\theta \geq \bar{X} - \bar{Y} - t\sqrt{S_1^2 + S_2^2})$; $-\infty < t < +\infty$

lies between certain values, say $p_1(t) \leq p(t,\nu) \leq p_2(t)$, which can be found from the tables of the Student distribution. As (9) is true conditionally, given $\textcircled{H} = \theta$, we get by integration with respect to the marginal distribution for $\textcircled{H}$, the system of equations

(10) $p_1(t) \leq P(\textcircled{H} \geq \bar{X} - \bar{Y} - t\sqrt{S_1^2 + S_2^2}) \leq p_2(t)$; $-\infty < t < +\infty$.

One can hardly call (10) a "distribution" for $\textcircled{H}$, but (10) gives probability statements of interest in the Behrens-Fisher case, especially when $p_1(t)$ and $p_2(t)$ are close together, which is often the case.

4. Which "distribution" should one choose ?

A "distribution" for $\textcircled{H}$ of the type (2) is of course not unique in the same way as the usual probability distribution for a variable is unique. If for instance the a priori distribution for all the parameters is known, then the conditional distribution for the parameters, given the observations, is unique and known.

Fisher has discussed the uniqueness of "fiducial probability distributions" in [2] pp. 392-393. Correspondingly, one could construct another "distribution" for $\textcircled{H}$, different from (2), by means of the known conditional distribution for

$T' = \frac{\bar{X} - \textcircled{H}}{S'}\sqrt{n}$, given $\textcircled{H} = \theta$, where S_1^2

is an estimator for σ^2 derived from the mean error of the observations. That (2) is preferable to the "distribution" for (H) derived from T^1 could possibly be established by means of reasonable criterions for adequate or useful "distributions" of the type (2).

Refernces.

[1] R.A. Fisher, "Statistical Methods and Scientific Inference", Oliver and Boyd, 1956, London.

[2] R.A. Fisher, "The fiducial argument in statistical inference", Annals of Eugenics, Vol. VI Pt IV pp 391-398, 1935.

[3] D.A.S. Fraser, "On fiducial inference", Annals of Math. Statist. Vol. 32 pp 661-676.

[4] J. Neyman, "Fiducial argument and the theory of confidence intervals", Biometrika vol. 32, pp 128-150.

[5] E. Sverdrup, "Laws and chance variations", vol. 2, 1967 North-Holland Publ. Comp., Amsterdam.

ESa/elr July 1969.