SOME EXAMPLES WITH FIDUCIAL RELEVANCE

BY JOHN W. TUKEY

Princeton University

1. Summary. It has been believed by some ([17], p. 204, and perhaps, by implication [21], p. 2, near line 10) that—and R. A. Fisher [e.g. at the Lake Junaluska conference in 1946] has urged the desirability of determining whether—the distribution induced by a pivotal, sufficient and smoothly invertible set of quantities is unique (that is to say, the induced distribution is independent of the choice of a particular set of pivotal quantities among those sets with these properties). If true, such uniqueness would be important in connection with the theory of fiducial probability.

It is the purpose of this paper to present certain examples of particular interest showing that these conditions do not provide uniqueness. The first example applies to any family of two-dimensional normal distributions with fixed and known variances and covariances. A one-parameter family of pivotal pairs of quantities are provided, such that no two of the induced distributions are the same. Each pair is sufficient, and consists of two independent quantities, each distributed according to a unit normal distribution. Each pair is shown to be smoothly invertible of every finite order. This example can be extended to the Behrens-Fisher situation. The second example is due to L. J. Savage, and exhibits a two-parameter situation where the two alternative pairs of pivotal quantities constructed according to the prescription of Segal [24] give rise to different distributions.

Mauldon [19] has recently published a quite different example of nonuniqueness which is also based on the bivariate normal distribution. In his example, the means are known and the second moments are to be estimated, so that there are 3 essential parameters.

The paper concludes with a reasonably complete bibliography of papers on fiducial probability.

2. Introduction. The history of fiducial inference has been clouded with dispute and failures of understanding—possibly, however, to no greater extent than is reasonably to be expected when basic new concepts are being forged between the hammer of mathematics and the anvil of concrete applications. This is not the place to review this history, to try to describe fiducial inference as it appears today, or to compare it with other schemes of inference (even as seen by one person). [The writer hopes to do the latter two of these elsewhere (cf. [25]).]

The uniqueness of the result of the fiducial argument has been held by Fisher to be of central importance, and conditions which ensure, or might ensure, unique-

Received April 9, 1956; revised February 1, 1957.

Prepared in connection with research sponsored by the Office of Naval Research.
ness have been important to him and his colleagues (e.g., [21]). The uniqueness problem does not seem to have received the attention which it deserved, even though it was a relatively completely formulated mathematical problem, and could be discussed without touching on any of the relatively sensitive issues of philosophy or principle associated with fiducial inference.

The main example given here appeared first in a paper on "The Purposes of Fiducial Inference" which was presented, as part of a symposium on "Probability and Statistical Inference", to the Econometric Society and the Institute of Mathematical Statistics in Minneapolis on September 6, 1951. The example, as presented at that time, was formally the same, but no detailed proof of smooth invertibility was given, and the need for one was not adequately recognized. The present improvements both in scope and simplicity of approach are the results of comments and stimulation by L. J. Savage, to whom go the author's best thanks.

The example showing that the Segal construction can lead to nonuniqueness is due entirely to Savage, and dates from about the same time (summer, 1952). It was first communicated to the writer in November, 1952.

In so far as examples of nonuniqueness are concerned, the earliest seems to be an unpublished example of P. H. Diananda. Savage informed the writer (in November, 1952) that this is referred to in a Cambridge thesis [29], also unpublished, of R. M. Williams. M. S. Bartlett says that this example is based on the Wishart distribution, and obtains different results in a symmetrical way by starting with the variance of first one variable and then the other. It thus presumably belongs to the same class as Savage's example, and is probably somewhat more complicated to discuss.

More recently, Mauldon has published an explicit example, and has indicated the existence of a family of examples based on the second moments of a family of normal distributions with known, fixed first moments. His examples involve a minimum of 3 parameters, and are thus somewhat more complex. His explicit example involves the permutation of a pair of observations and parameters, as does Savage's example, but in a somewhat less informative situation, since the set of pivotal quantities used were not constructed according to a general method of independent interest, as was the case with Savage's example. The indicated extensions include continuous-parameter families of alternatives, but seem likely to be less understandable than the continuous-parameter cases described here. (Charles Stein has pointed out a way of looking at Mauldon's example in terms of the subgroup of triangular matrices and its conjugate subgroups which makes this example more interesting.)

So far as I know, these are the available examples concerning the nonuniqueness of the result of the fiducial argument (as described in [33]) under varying assumptions. (Conditions under which uniqueness clearly exists can be given, but we shall not discuss them here.)

3. Formalities. A specification determines the probability distribution as a function of parameters. A quantity is a function of observations and parameters,
Examples with fiducial relevance

defined for all possible combinations of admissible values of its arguments. (Values of the parameters are appropriate as arguments whether or not they determine the distribution of the observations concerned.) A set of quantities is pivotal (with respect to a specification) if, whatever admissible values of parameters are inserted in the quantities and are, at the same time, the parameters determining the distribution of the observations appearing in the quantities, the resulting set of statistics has the same distribution. A set of quantities is sufficient if, when any arbitrary fixed values of the parameters are inserted as arguments, the resulting set of statistics is sufficient for the parameters determining the distribution of the observations. A set of quantities is smoothly invertible (of class $\alpha$), if, when any possible set of observations is inserted as arguments, the mapping from parameters to quantities:

1. has the same range for any possible set of observations,
2. is 1 to 1, and hence has a single-valued inverse, and
3. this inverse is continuous (and has continuous derivatives of all orders up to $\alpha$).

If a set of quantities are pivotal and smoothly invertible, then each set of possible observations induces a distribution on parameter space. [A distribution is uniquely associated with the pivotal quantities by their pivotal property. Fixing a set of possible observations fixes a 1 to 1 bicontinuous (i.e., continuous in both directions) relation between quantities and parameters which transfers this distribution to parameter space.]

4. Twisted two-dimensional normals. We now start from a family of two-dimensional normal distributions in which all second-degree moments (about the mean) are fixed, but where all locations in the plane are possible. If we introduce an appropriate coordinate system, the specification becomes the following:

$$x \text{ and } y \text{ are normally and independently distributed with averages } \mu \text{ and } \nu \text{ and unit variances.}$$

[We shall abbreviate such statements in the form "$x \text{ and } y \text{ are NID}(\mu, \nu; I)$."]

We deal with one bivariate observation drawn from some distribution of this family. (This observation may, of course, be the vector of means from a sample of $n$ from a population NID$(\mu, \nu; \sqrt{n} I)$.) Then the quantities

$$w_1 = x - \mu,$$
$$w_2 = y - \nu,$$

are immediately seen to be NID$(0, 0; I)$ for any $\mu$ and $\nu$, and hence to be pivotal. Fixing $\mu$ and $\nu$, the values of $w_1$ and $w_2$ determine $x$ and $y$. The latter are surely sufficient, since they specify the sample (of one) completely. Thus $w_1, w_2$ are sufficient and pivotal quantities, and since they are obviously smoothly invertible of all orders, we have one smoothly induced distribution.

But the fact that the probability density for $(w_1, w_2)$ is constant on circles
about the origin enables us to modify these quantities without disturbing their distribution.

Let \( f(\mu, \nu, r) \) be a sufficiently smooth, but otherwise arbitrary, function of three variables. Put

\[
r = \sqrt{w_1^2 + w_2^2} = \sqrt{(x - \mu)^2 + (y - \nu)^2},
\]

and

\[
w_{1f} = (x - \mu) \cos f + (y - \nu) \sin f = w_1 \cos f + w_2 \sin f,
\]

\[
w_{2f} = -(x - \mu) \sin f + (y - \nu) \cos f = -w_1 \sin f + w_2 \cos f,
\]

where \( f = f(\mu, \nu, r) \). Then \( w_{1f} \) and \( w_{2f} \) are, for each fixed \( \mu \) and \( \nu \), also NID(0, 0; \( I \)) and, moreover, we have \( r^2 = w_{1f}^2 + w_{2f}^2 \). For we have merely twisted each circle of radius \( r \) through the angle \( f(\mu, \nu, r) \). Each of these pairs \((w_{1f}, w_{2f})\) is pivotal, and sufficient (since \( r^2 = w_{1f}^2 + w_{2f}^2 \) is known as soon as \( w_{1f} \) and \( w_{2f} \) are known, so that fixing \( \mu \) and \( \nu \) fixes \( r \) and hence make \( w_1 \) and \( w_2 \) available, and thus makes \( x \) and \( y \) available!), and, if they are smoothly invertible, are candidates for the construction of our counterexample.

We shall have use for the Jacobian of the transformation from \( \mu, \nu \) to \( w_{1f}, w_{2f} \) (with \( x, y \) held fixed). The details of calculation are presented in a later section, but the result is

\[
J = \frac{\partial (w_{1f}, w_{2f})}{\partial (\mu, \nu)} = 1 + (x - \mu) \frac{\partial f}{\partial \nu} - (y - \nu) \frac{\partial f}{\partial \mu} = 1 + \begin{vmatrix} x - \mu & f_{\mu} \\ y - \nu & f_{\nu} \end{vmatrix},
\]

where the partial derivatives of \( f \) are taken considering \( f \) as a function of the three independent variables \( \mu, \nu \) and \( r \). It is shown in the same section that solutions \( \mu, \nu \) of

\[
w_{1f}(x, y, \mu, \nu) = a,
\]

\[
w_{2f}(x, y, \mu, \nu) = b,
\]

exist for any chosen \( x, y, a \) and \( b \) for any continuous \( f(\mu, \nu, r) \) and that, if the Jacobian above is always positive, then these solutions are unique, and as many times continuously differentiable as \( f_\mu \) and \( J \) themselves.

5. Detailed example. In order to give a detailed counterexample, we specialize \( f(\mu, \nu, r) \) to the form

\[
f(\mu, \nu, r) = \frac{\beta \mu}{1 + r^2},
\]

when our pivotal quantities become, when written out explicitly,

\[
w_{1n} = (x - \mu) \cos \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2} + (y - \nu) \sin \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2},
\]

\[
w_{2n} = -(x - \mu) \sin \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2} + (y - \nu) \cos \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2}.
\]
$w_{2\alpha} = -(x - \mu) \sin \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2} + (y - \nu) \cos \frac{\beta \mu}{1 + (x - \mu)^2 + (y - \nu)^2}$.

Since $f_\nu = 0$ and $f_\mu = \beta/(1 + r^2)$ we find that

$$\frac{\partial (w_{1\alpha}, w_{2\alpha})}{\partial (\mu, \nu)} = 1 - \beta \frac{y - \nu}{1 + (x - \mu)^2 + (y - \nu)^2},$$

which is surely positive for $|\beta| < 2$. To complete the example, we have only to show that different values of $\beta$ lead, for one and the same set of observations $x, y$, to different induced distributions for $(\mu, \nu)$.

In view of the existence and continuity of all derivatives the induced density in the $(\mu, \nu)$ plane is

$$\frac{\partial (w_{1\beta}, w_{2\beta})}{\partial (\mu, \nu)} \text{ (density for } w_{1\beta}, w_{2\beta}) = \left(1 - \beta \frac{y - \nu}{1 + r^2} \right) \frac{1}{2\pi} e^{-r^2},$$

where $r^2 = (x - \mu)^2 + (y - \nu)^2$. This induced density is clearly different for different values of $\beta$. Subject, then to the verification of (i) the general value of the Jacobian, (ii) the general existence of solutions, and (iii) the uniqueness of solutions when the Jacobian is positive, the announced example is complete.

6. The existence and uniqueness of solutions, and the value of the Jacobian.

We now investigate the existence and uniqueness of solutions, $\mu, \nu$ of

$$w_{1f}(x, y, \mu, \nu) = a,$$

$$w_{2f}(x, y, \mu, \nu) = b,$$

for any given $x, y, a$ and $b$. We shall use direct methods, noting that smooth invertibility, under the conditions we use is also a consequence of the following result:

Any $\alpha$ times continuously differentiable mapping from a connected open domain to a simply connected range whose Jacobian determinant is continuous and of constant sign, and whose inverse carries compact sets into compact sets, is smoothly invertible of order $\alpha$. This invertibility theorem is proved elsewhere [26], and examples are given there to show that no one of its topological conditions can be omitted.

Introduce polar coordinates for $(a, b)$ through

$$a = \rho \cos \theta, \quad b = \rho \sin \theta$$

and polar-like coordinates for $\mu, \nu$ by

$$\mu = x - r \cos A,$$

$$\nu = y - r \sin A,$$

then the equations to be solved for $r$ and $A$ are, when converted to polar co-
ordinate form, \( r = \rho \), and
\[
A - f(\mu, \nu, r) = A - f(x - r \cos A, y - r \sin A, r) = \theta + 2k\pi, \quad (k \text{ an integer}).
\]

Clearly we need only study the second equation
\[
\varphi(A) = A - f(x - \rho \cos A, y - \rho \sin A, \rho) = \theta + 2k\pi, \quad (k \text{ an integer})
\]
for given \( x, y, \rho, \theta \) and see when it always has a unique solution.

Moreover, we may use these auxiliary variables to evaluate the Jacobian \( J \) of \( \omega_{1/2} \) and \( \omega_{2/1} \), or, equivalently, of \( a \) and \( b \), with respect to \( \mu \) and \( \nu \). We have
\[
\frac{\partial(a, b)}{\partial(\mu, \nu)} = \frac{\partial(a, b)}{\partial(\rho, \theta)} \frac{\partial(\rho, \theta)}{\partial(r, A)} / \frac{\partial(\mu, \nu)}{\partial(r, A)}
\]
\[
= \rho \frac{\partial(\rho, \theta)}{\partial(r, A)} / \rho = \frac{\partial(\rho, \theta)}{\partial(r, A)}
\]
\[
= \begin{bmatrix} 1 & 0 \\ \partial \theta & \partial A \end{bmatrix} = \frac{\partial \theta}{\partial r} = \varphi(A)
\]
where we have used \( \rho = r \) as necessary. Thus if we evaluate \( \varphi(A) \) we will also evaluate the desired Jacobian \( J \).

As \( A \) increases from \( 0 \) to \( 2\pi \), the value of \( \varphi(\nu) \) varies from
\[
\varphi(0) = -f(x - \rho, y, \rho) \to \varphi(2\pi) = 2\pi - f(x - \rho, y, \rho) = \varphi(0) + 2\pi
\]
and since it is continuous it must pass through the value \( \theta + 2k\pi \) for some integer \( i \). Thus we can always solve the initial pair of equations for any \( f(\mu, \nu, r) \) which is continuous in its arguments (at least in its first and third arguments together.)

If now we show that \( \varphi(A) \) is strictly increasing, it will follow that it takes every value only once, and that solutions not only exist but are unique. Clearly,
\[
J = \varphi(A) = 1 - (\rho \sin A)f_\mu + \rho \cos A f_\nu,
\]
\[
= 1 - (r \sin A)f_\mu + r \cos A f_\nu,
\]
\[
= 1 - (y - \nu)f_\mu + (x - \mu)f_\nu,
\]
and if the Jacobian is always positive, \( \varphi(A) \) is monotone, and solutions not only exist but are unique.

Moreover, since
\[
\frac{\partial \theta}{\partial A} = \varphi(A) = J, \quad \frac{\partial \theta}{\partial r} = f_r,
\]
\[
\frac{\partial \rho}{\partial r} = 1, \quad \frac{\partial \rho}{\partial A} = 0,
\]
the unique inverse is clearly as many times continuously differentiable as are \( f_r \) and the Jacobian.
7. The Behrens-Fisher problem. The best-known example to which the fiducial argument has been applied is the so-called Behrens-Fisher problem, first treated by W. V. Behrens [31] and afterwards discussed by many writers (see Breny [32] for a recent review and [33] for Fisher's most recent statement). The problem arises from two samples, one of \( n_1 \) observations from a normal distribution with average \( \mu_1 \) and the variance \( n_1 \sigma_1^2 \), and other of \( n_2 \) observations from a normal distribution with average \( \mu_2 \) and variance \( n_2 \sigma_2^2 \), when it is desired to make an inference about \( \mu_1 - \mu_2 \). (For convenience we have defined \( \sigma_1 \) and \( \sigma_2 \) in an unusual way.)

If we take \( x_1 \) and \( x_2 \) as the means of the two samples, and \( s_1^2 \) and \( s_2^2 \) as the conventional estimates of the variances of these means, then

\[
\begin{align*}
w_1 &= \frac{x_1 - \mu_1}{\sigma_1}, \\
w_2 &= \frac{x_2 - \mu_2}{\sigma_2}, \\
w_3 &= \frac{s_1}{\sigma_1}, \\
w_4 &= \frac{s_2}{\sigma_2},
\end{align*}
\]

are independently distributed sufficient pivotal quantities. The distribution of \( w_1 \) and \( w_2 \) is unit bivariate normal, so that \( w_1 \) and \( w_2 \) may be introduced as before, and (because \( \mu_1 \) and \( \mu_2 \) do not appear in \( w_3 \) or \( w_4 \), etc.)

\[
\frac{\partial(w_1, w_2, w_3, w_4)}{\partial(\mu_1, \mu_2, \sigma_1, \sigma_2)} = \frac{\partial(w_1, w_2)}{\partial(\mu_1, \mu_2)} \left( \frac{\partial w_3}{\partial \sigma_1} \right) \left( \frac{\partial w_4}{\partial \sigma_2} \right)
\]

and if the Jacobian of \( w_1, w_2 \), with respect to \( \mu_1, \mu_2 \), depends on \( \alpha \) (as it does) so too does the Jacobian of all four pivotal quantities with respect to all four parameters. Consequently the induced distribution on parameter space is not unique for the Behrens-Fisher specification. (It might be interesting to examine other distributions than the conventional one.)

8. Savage’s Example. We now set forth the example due to L. J. Savage, which shows how uniqueness can escape us in a different way.

Let \( x \) and \( y \) be distributed according to

\[
\psi(x, y \mid \alpha, \beta) \, dx \, dy = \frac{\alpha^\beta \gamma}{\alpha \beta} (x + y)^{-\alpha \beta} \, dx \, dy,
\]

where \( \alpha \) and \( \beta \) are positive and \( 0 \leq x, y < \infty \). If the cumulative distribution of \( x \) is \( 1 - S \), and the cumulative conditional distribution of \( y \) given \( x \) is \( 1 - T \) then

\[
S = \left( 1 + \frac{\alpha \beta x}{\alpha + \beta} \right) e^{-\alpha x},
\]

\[
T = \left( 1 + \frac{\beta y}{1 + \beta x} \right) e^{-\beta y},
\]

as is easily verified by integration. From their definition these quantities are uniformly distributed on \( 0 \leq S, T \leq 1 \) and are pivotal. (They are essentially examples of the pivotal quantities pointed out by Segal [24].) Moreover, for \( x \) and \( y \) fixed, \( T \) takes any values between 0 and 1 for suitably chosen \( \beta \)—and for
$x, y$ and $\beta$ fixed, $S$ takes all values between 0 and 1 for suitably chosen $\alpha$—thus the open square is covered for each choice of $(x, y)$.

Now let us calculate the Jacobian from $(\alpha, \beta)$ to $(S, T)$. Clearly $\partial T/\partial \alpha$ vanishes, so that $\partial S/\partial \beta$ is not involved. The other two derivatives are

$$
\frac{\partial S}{\partial \alpha} = - \frac{\alpha x}{(\alpha + \beta)^2} [(\alpha + 2\beta) + \beta(\alpha + \beta)x]e^{-\alpha x},
$$

$$
\frac{\partial T}{\partial \beta} = - \frac{y}{(1 + \beta)^2} [(1 + \beta x)(1 + \beta x + \beta y) - 1]e^{-\beta y}
$$

and are each clearly negative for all positive $\alpha, \beta, x, y$. The Jacobian is their product and is clearly positive.

The induced density on $0 \leq \alpha, \beta < +\infty$ is given by the pivotal density (identically unity) multiplied by the Jacobian, namely

$$
\frac{\alpha xy}{(\alpha + \beta)^2} e^{-\alpha x - \beta y} \frac{[\alpha + 2\beta + \beta(\alpha + \beta)x][(1 + \beta x)(1 + \beta x + \beta y) - 1]}{(1 + \beta x)^2} d\alpha d\beta,
$$

where we have separated a factor symmetric in $(\alpha, x)$ and $(\beta, y)$ from a factor manifestly not so symmetric.

The specification at the beginning of this section was symmetric in $(\alpha, x)$ and $(\beta, y)$ so that if we take $1 - S$ as the cumulative distribution of $y$, and $1 - T$ as the cumulative conditional distribution of $x$ given $y$ and go through the identical argument, we will find that the induced distribution for $\alpha, \beta$ is obtained from the above by symmetry—by interchanging $\alpha$ with $\beta$ and $x$ with $y$. The result will clearly not be the same!

Thus the two applications of Segal’s process to this specification do not lead to the same induced distribution. As a consequence we see that conditions of monotony of pivotal quantities, though they enforce smooth invertibility, cannot enforce uniqueness of induced distribution. For $S$ and $T$ are monotone in all of $x, y, \alpha, \beta$, as follows when we note

$$
\frac{\partial S}{\partial \beta} = \frac{\alpha^2 x}{(\alpha + \beta)^2} e^{-\beta x} > 0
$$

and use our earlier results.

REFERENCES


