The Annals of Statistics
1982, Vol. 10, No. 4, 1054–1067

THE FUNCTIONAL-MODEL BASIS OF FIDUCIAL INFERENCE

BY A. P. DAWID AND M. STONE

University College London

"... not to confute the genius but to perfect the conjecture"
(Ian Hacking, *Logic of Statistical Inference*)

The role of functional models and their associated fiducial analysis is explored in an attempt to uncover a general theory of fiducial inference.

1. Introduction. The chequered history of fiducial inference has been covered in recent reviews and books: Pedersen (1978), Seidenfeld (1979), Buehler (1982), Edwards (1982) and Stone (1982). Despite all this effort, we believe that the unifying and simplifying virtues of the *functional model* approach have been largely unrecognized. Although the term "functional model" appears to have been introduced by H. Bunke (1975), the concept is implicit in Fraser's structural inference (1961, 1968) and in earlier justifications of the "fiducial argument;" the more recent work of Plante (1979a, b) deals with functional models explicitly. The present paper sets out to integrate under a single mathematical structure the primary ideas that may be developed with functional models. However, we make no claim that this treatment reflects the whole spectrum of interpretations of the fiducial argument, or that it would have met with the approval of its originator, R. A. Fisher.

Sections 2 and 5 present the basic structure of functional models and some examples. In Section 3, we consider the problem of inference about a function of the parameter, and show agreement between two possible approaches. Section 4 considers the connections between fiducial probabilities and confidence coefficients. In Section 6, we come across an inconsistency that may arise when partial information becomes available about the parameter: again there are two approaches, but now they yield different answers in general. We show that this inconsistency vanishes for parametric information that indexes a pivotal submodel. In this case, the Fisher-Yates justification of a fiducial property involving confidence coverage probabilities may be demonstrated; this is the subject of Section 7. Finally, an Appendix shows the close connection between general functional models and those possessing certain properties of group-invariance. We believe that our approach through functional models unifies and streamlines the disjointed body of known results on fiducial inference.

2. Simple functional and fiducial models. The models we consider have three ingredients: *data* X, *parameter* Θ, and *error* E, whose possible values range over known spaces $\mathscr{X}$, Θ, $\mathscr{E}$ respectively. It is supposed that:

(i) X is a uniquely determined function of Θ and E (the value of X when $\Theta = \theta$ and $E = e$ being denoted simply by θe);
(ii) E has a known distribution P over $\mathscr{E}$, whatever the value of Θ (written $E \sim P$);
(iii) X alone is observed;
(iv) inference is required for Θ which is to be treated as "completely unknown."

Such models have been called *functional models* by H. Bunke (1975), in distinction to *distribution models* which merely specify the distributions $\{P_\theta\}$ of X given $\Theta = \theta$. While any functional model determines an associated distribution model in which P_θ is the

Received November 1981; revised February 1982.
AMS 1980 *subject classifications*. Primary 62A30; secondary 62A05, 62A25.
Key words and phrases. Fiducial inference, functional model, partitionability, reduction, marginalization and conditional consistency, fiducial-confidence, group-structural, pivotals.

distribution of θE where $E \sim P$, important structural distinctions are lost in such reduction. For example, the two distinct functional models with (a) $X_1 = E_1$, $X_2 = \Theta E_1 + (1 - \Theta^2)^{1/2}E_2$, and (b) $X_1 = (1 - \Theta^2)^{1/2}E_1 + \Theta E_2$, $X_2 = E_2$, in which E_1 and E_2 have independent standard normal distributions, both determine the same bivariate normal distribution model for (X_1, X_2).

For $x \in \mathscr{X}$, $e \in \mathscr{E}$, call x and e *compatible* if $x = \theta e$ for some $\theta \in \Theta$. We denote $\{e : x$ and e are compatible$\}$ by $\mathscr{E}_x$. Similarly, *mutatis mutandis*, we define the compatibility of $x \in \mathscr{X}$ and $\theta \in \Theta$, and the sets $\mathscr{X}_\theta$, Θ_x. We shall denote by $\mathscr{D}$ the set $\{(x, \theta) \in \mathscr{X} \times \Theta : x$ and θ are compatible$\}$.

Throughout this paper, we will assume *invertibility* in the sense that, for $x \in \mathscr{X}$ and $e \in \mathscr{E}_x$, the solution for θ of $x = \theta e$ is unique. For Sections 2, 3, 4, however, attention will be restricted to *simple* functional models defined as those for which $\mathscr{E}_x \equiv \mathscr{E}$. We will write $\langle X = \Theta E, E \sim P \rangle$ for the simple functional model (SFM for short) and $\theta = xe^{-1}$ for the unique solution of $x = \theta e$. After any observation $X = x$ and in the absence of definite information about Θ, no value of E is ruled out as incompatible with the data. In such circumstances, the present version of the fiducial argument is that the observation of X gives no cause to alter the original assignment of the distribution P to E. This motivates the following.

DEFINITION. The (*simple*) *fiducial model*, corresponding to the simple functional model $\langle X = \Theta E, E \sim P \rangle$, is given by

$$\langle \Theta = XE^{-1}, E \sim P \rangle. \tag{2.1}$$

In other words, the fiducial model is itself a simple functional model, with the roles of X and Θ interchanged: it is implicit in (2.1) that E is distributed independently of the value of X, in contrast to the assumption of the original functional model that E is independent of the value of Θ. *Fiducial inference* then consists of regarding Θ as a random variable generated by the fiducial model following the observation of X. This leads to an associated distribution model $\{\Pi_x\}$ for Θ, where Π_x denotes the distribution of xE^{-1} when $E \sim P$. We call Π_x the *fiducial distribution* for Θ based on data $X = x$. Then Π_x is concentrated on Θ_x.

EXAMPLE 2.1. *One parameter family of* cdfs. An interesting and theoretically useful class of specialized SFMs is defined as follows:

(i) $\mathscr{X} \subset R^1$, $\Theta \subset R^1$, $\mathscr{E} = (0, 1)$ and P is the uniform distribution on $\mathscr{E}$;

(ii) θe is a non-decreasing function of e for each θ;

(iii) if $\theta_1 \le \theta_2$ then $\theta_1 e \le \theta_2 e$ for all $e \in (0, 1)$;

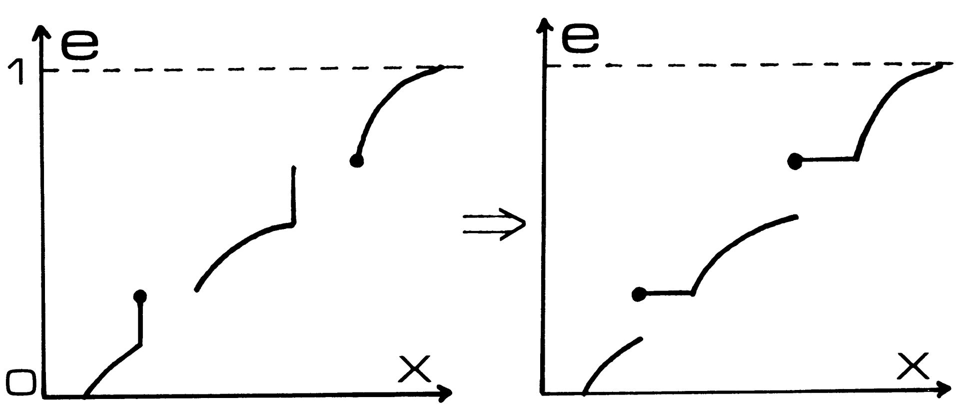

FIG. 1. *Generation of the* cdf *from the* SFM:
(a) *graph of* $x = \theta e$ (b) *graph of* $e = F_\theta(x)$.

(iv) for every pair (e, x) with $e \in (0, 1)$ and $x \in \mathscr{X}$, there is just one θ-value satisfying $x = \theta e$.

Figure 1 then illustrates the generation of the cdf F_θ that now expresses P_θ. Also, given $X = x$, $\Theta \geq \theta \Rightarrow E \leq F_\theta(x)$, while $E < F_\theta(x) \Rightarrow \Theta \geq \theta$. So $\Pi_x(\Theta \geq \theta) = F_\theta(x)$, since E is uniformly distributed.

If, additionally, $x = \theta e$ is a *strictly* increasing function of e for each θ, the distribution functions $\{F_\theta : \theta \in \Theta\}$ will be continuous, and conversely. The SFM will then be called *Fisherian*, since it would need only the supposition of differentiability of F_θ with respect to θ to obtain a fiducial density $-\partial F_\theta(x)/\partial\theta$ for the observation $X = x$, thereby matching the outcome of the original fiducial argument of Fisher (1930). A given one-parameter family of distribution functions $\mathscr{F} = \{F_\theta : \theta \in \Theta \subset R^1\}$ may then be naturally labeled *Fisherian* if and only if there is a Fisherian SFM for which, given $\Theta = \theta$, $X \sim F_\theta$ for all $\theta \in \Theta$.

We have not been able to establish necessary and sufficient conditions for a given $\mathscr{F}$ to be Fisherian. A set of sufficient conditions is: (i) each F_θ is continuous, (ii) for x_0, θ_0 such that $0 < F_{\theta_0}(x_0) < 1$, $\{F_\theta(x_0) : \theta \in \Theta\} \supset (0, 1)$, $F_{\theta_0}(x)$ is strictly increasing in x at $x = x_0$ and $F_\theta(x_0)$ is strictly decreasing in θ at $\theta = \theta_0$. An $\mathscr{F}$ satisfying these will be said to be *regular*.

EXAMPLE 2.2. *Correlation coefficient.* Suppose $\mathscr{X} = \Theta = R^1$, $\mathscr{E} = R^+ \times R^+ \times R$ and $\theta e = (\theta e_1 + e_3)/e_2$, with P such that the components E_1, E_2, E_3 of E are independently distributed $E_1 \sim \chi_{n-1}$, $E_2 \sim \chi_{n-2}$, $E_3 \sim N(0, 1)$. Defining $R = X/(1 + X^2)^{1/2}$ and $\Phi = \Theta/(1 + \Theta^2)^{1/2}$, it may be verified (Dempster, 1969, Equation (14.2.10); see also Section 3 below) that, given $\Theta = \theta$, R has the distribution of a sample correlation coefficient from a bivariate normal sample of size n with population correlation $\phi = \theta/(1 + \theta^2)^{1/2}$. For $X = x$, the fiducial distribution for Θ is that of

$$xE^{-1} = \frac{xE_2 - E_3}{E_1}$$

where $E \sim P$, and the corresponding fiducial distribution for Φ is found to have the *same* density $-\partial F_\phi(r)/\partial\phi$ as that produced, from a different logical standpoint, by Fisher (1930) working directly with the cdfs F_ϕ for R given ϕ. An explanation of this striking identity is given in Section 4.2.

EXAMPLE 2.3. *Coefficient of variation.* Suppose $\mathscr{X} = \Theta = R^1$, $\mathscr{E} = R \times R^+ = \{(e_1, e_2)\}$ and $\theta e = (\theta + e_1)/e_2$, with P such that E_1 and E_2 are independently $N(0, 1/n)$ and $\bar{\chi}_{n-1} = \chi_{n-1}/(n - 1)^{1/2}$ respectively. It may be verified that, when $\Theta = \theta$, $X = \theta E$ is distributed as the ratio of the sample mean to the sample standard deviation of a random sample of size n from a normal distribution for which θ is the corresponding population ratio. The fiducial distribution for Θ, given $X = x$, is that of $xE_2 - E_1$.

EXAMPLE 2.4. *Normal mean and variance.* Suppose $\mathscr{X} = \Theta = \mathscr{E} = R \times R^+$, with $x = (\bar{x}, s)$, $\theta = (\mu, \sigma)$, $e = (e_1, e_2)$ and $x = \theta e$ given by

$$\bar{x} = \mu + \sigma e_1, \quad s = \sigma e_2,$$

with P as in Example 2.3. It may be verified that, given $\Theta = (\mu, \sigma)$, $\bar{X}$ and S are distributed as the sample mean and standard deviation of a random sample of size n from an $N(\mu, \sigma^2)$ distribution. The fiducial model for $\Theta = (M, \Sigma)$ is then

$$M = \bar{X} - SE_1/E_2, \quad \Sigma = S/E_2, \tag{2.2}$$

giving the joint fiducial distribution of (M, Σ) first obtained by Fisher (1935).

EXAMPLE 2.5. *Random bias.* Suppose $\mathscr{X}$, Θ, x, θ and the distribution of E_1, E_2 are as in Example 2.4 but $E = (E_0, E_1, E_2)$ now includes independent random $N(0, 1)$ bias E_0 in

$\bar{X}$, so that $x = \theta e$ is given by

$$\bar{x} = \mu + e_0 + \sigma e_1, \quad s = \sigma e_2.$$

The fiducial model is then

$$M = \bar{X} - E_0 - SE_1/E_2, \quad \Sigma = S/E_2, \tag{2.3}$$

in which M now has a conditional $N(\bar{X}, 1 + \sigma^2)$ distribution, given $\Sigma = \sigma$ in the joint fiducial distribution.

There are two important special cases of SFMs:

I. *Pivotal model.* Suppose that, in a given SFM, for every $(x, \theta) \in \mathscr{D}$ it is possible to solve $x = \theta e$ for e uniquely. We will write $e = p(x, \theta)$, call $p: \mathscr{D} \to \mathscr{E}$ the *pivotal function* and say that we have a *simple pivotal model* $\langle \mathscr{D}, E = p(X, \theta), E \sim P \rangle$. Barnard (1977, 1981) takes such pivotal models as the starting point of his treatment of inference, although he goes on to drop the requirement that Θ and E should together determine X. In our development, the initial invertibility assumption of the SFM implies that, for every $x \in \mathscr{X}$ as well as for every $\theta \in \Theta$, the range of values of $p(x, \theta)$ is $\mathscr{E}$: the equation $p(x, \theta) = e$ has a unique solution for $x \in \mathscr{X}_\theta$, namely $x = \theta e$, when $e \in \mathscr{E}$ and $\theta \in \Theta$ are given, and for $\theta \in \Theta_x$, namely $\theta = xe^{-1}$, when $e \in \mathscr{E}$ and $x \in \mathscr{X}$ are given. The related simple functional and fiducial models may be crudely expressed in the forms

$$\langle p(X, \Theta) \sim P \quad \text{independently of} \quad \Theta \rangle$$

and

$$\langle p(X, \Theta) \sim P \quad \text{independently of} \quad X \rangle$$

respectively.

The Fisherian SFM in Example 2.1 is a simple pivotal model with $p(x, \theta) \equiv F_\theta(x)$ and, in general, with $\mathscr{D} \neq \mathscr{X} \times \Theta$. Example 2.4 is a simple pivotal model in which $\mathscr{D} = \mathscr{X} \times \Theta$ and $p(x, \theta) = ((\bar{x} - \mu)/\sigma, s/\sigma)$. However, Examples 2.2, 2.3 and 2.5 are not simple pivotal models since it is not possible to express their E's as functions of X and Θ.

II. *Simple group-structural model* (Fraser, 1961). Suppose $\mathscr{X}$, Θ and $\mathscr{E}$ can all be identified with a group G, and the operation θe is just group multiplication. This produces a simple pivotal model with $p = \theta^{-1}x$. Example 2.4 is an illustration of this structure.

3. Marginalization consistency. Suppose that, for the SFM

$$M_1 = \langle X = \Theta E, E \sim P \rangle,$$

(i) we are interested only in a subparameter $\Lambda = \lambda(\Theta)$; (ii) there is a function $Z = z(X)$ such that the relationship between Z, Λ and E itself constitutes an SFM

$$M_2 = \langle Z = \Lambda E, E \sim P \rangle.$$

We will call M_2 a *reduction* of M_1. (Note that the superficially similar notations ΘE and ΛE actually involve different parameter spaces and necessarily different binary operations.)

The *marginal* fiducial distribution of Λ for $X = x$ in M_1 is that of $\lambda(xE^{-1})$ where $E \sim P$. The *reduced* fiducial distribution of Λ, given by M_2, is that of zE^{-1} where $z = z(x)$. It is natural to inquire what relationship exists between these two distributions of Λ. The answer is an immediate and reassuring consequence of:

LEMMA 3.1. $$\lambda(XE^{-1}) = z(X)E^{-1}.$$

PROOF. $$\lambda(XE^{-1}) = \lambda(\Theta) = \Lambda = ZE^{-1} = z(X)E^{-1}.$$

Whether considered in the full or reduced model, Λ may be expressed as the same function of $z(X)$ and E, while $E \sim P$ in both cases. *A fortiori*, the marginal and reduced fiducial distributions coincide.

Wilkinson (1977) approaches fiducial inference from the standpoint of distribution models. Such an approach leads to marginalization inconsistency, which Wilkinson embraces under the umbrella of his "Noncoherence Principle." For example, if we start with the SFM

$$\langle X_i = \Theta_i + E_i,\ E_i \sim N(0, 1),\ i = 1, 2, \text{ independently} \rangle$$

and require inference for $\Lambda = \Theta_1^2 + \Theta_2^2$, the marginal fiducial distribution of Λ is non-central Chi squared with non-centrality parameter $Z = X_1^2 + X_2^2$. Moreover, the sampling distribution of Z depends only on Λ. Wilkinson applies his methods to the distribution of Z to yield a fiducial distribution for Λ differing from that induced by the joint fiducial distribution of (Θ_1, Θ_2). From the present point of view, there is no inconsistency, since reduction to Z does not produce an SFM.

Example 2.4 provides an illustration of Lemma 3.1. We have

$$\frac{\bar{X}}{S} = \frac{(M/\Sigma) + E_1}{E_2},$$

which implies the reduced SFM in which $Z = (\Lambda + E_1)/E_2$ for $Z = \bar{X}/S$ and $\Lambda = M/\Sigma$, already considered with different notation in Example 2.3.

Example 2.2 can likewise be derived as a reduction, M_2, of the simple group-structural model (the zero-mean *progression model*) $M_1 = \langle Y = \Gamma E \rangle$, where

$$Y = \begin{pmatrix} Y_1 & 0 \\ Y_3 & Y_2 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \Gamma_1 & 0 \\ \Gamma_3 & \Gamma_2 \end{pmatrix}, \quad E = \begin{pmatrix} E_1 & 0 \\ E_3 & E_2 \end{pmatrix},$$

the composition is provided by matrix multiplication, and the distribution of (E_1, E_2, E_3) is as in Example 2.2. The reduction is obtained on putting $X = Y_3/Y_2$, $\Theta = \Gamma_3/\Gamma_2$. As is well-known (Mauldon, 1955; Fraser, 1964), the distribution model implied by M_1 yields $S \sim W_2(n:\Sigma)$, where $S = YY'$, $\Sigma = \Gamma\Gamma'$. Thus $R = X/(1 + X^2)^{1/2} = S_{12}/(S_{11}S_{12})^{1/2}$, $\Phi = \Sigma_{12}/(\Sigma_{11}\Sigma_{22})^{1/2}$ are interpretable as genuine sample and population correlation coefficients. It has previously been noted (Sprott, reported by Fraser, 1964, Fisher, 1973, page 179) that the marginal fiducial distribution of Φ based on M_1 agrees with that obtained by Fisher's differentiation method. This is equivalent, by Lemma 3.1, to our assertion of this same result for M_2.

In the Appendix it is shown that every SFM can be considered as a reduction of a simple group-structural model. The possibilities for further reduction are also analysed by means of group-invariance arguments.

A reduction of a model involves (functional) contraction for X and Θ, but may or may not induce a contraction for E. No such contraction occurs in the reduction of Example 2.4 to 2.3 considered above. However, in the alternative reduction to $M_3 = \langle S = \Sigma E_2, (E_1, E_2) \sim P \rangle$, we may also contract (E_1, E_2) to E_2 yielding a now pivotal model $M_4 = \langle S = \Sigma E_2, E_2 \sim \chi_{n-1} \rangle$. Of course, both M_3 and M_4 yield the same sampling and fiducial distributions. By a trivial extension we can call M_4 a reduction of M_1.

The same reduction to M_3 and then M_4 is available in the non-pivotal model of Example 2.5.

In general, let F denote the maximal functional contraction of E available in a reduced model $M_2 = \langle Z = \Lambda E, E \sim P \rangle$, and let M_3 be the corresponding model $\langle Z = \Lambda F, F \sim P' \rangle$. We shall call M_3 the *minimal representation* of M_2.

4. Confidence properties.

4.1. *Pivotal confidence belts.* For the case of a pivotal SFM it is well-known that a confidence property can be found for fiducial probability. In the simple pivotal model $\langle \mathscr{D}, E = p(X, \Theta), E \sim P \rangle$, fix $A \subset \mathscr{E}$ with $P(A) = 1 - \alpha$ and write

$$xA^{-1} = \{xe^{-1} : e \in A\} = \{\theta \in \Theta_x : p(x, \theta) \in A\}.$$

Then

$$\Pi_x(xA^{-1}) = P(\{e : xe^{-1} \in xA^{-1}\}) = P(A) = 1 - \alpha$$

while

$$P_\theta(\theta \in XA^{-1}) = P_\theta(p(X, \theta) \in A) = P(A) = 1 - \alpha.$$

The class of regions $\{xA^{-1} : x \in \mathscr{X}\}$, which may be called a *pivotal confidence belt*, therefore has, simultaneously, fiducial probability $1 - \alpha$ for every x and the confidence property with confidence level $1 - \alpha$.

For the group-structural case of a pivotal model, the region xA^{-1} corresponds to left-multiplication of the fixed region A^{-1} by the data group-element x.

EXAMPLE 4.1. For the pivotal Example 2.4, take $A = \{(e_1, e_2) : e_1 + ce_2 > k\}$, where k is given and c is chosen to make $P(A) = 1 - \alpha$. Then $xA^{-1} = \{(\mu, \sigma) : \mu + k\sigma < \bar{x} + cs\}$, so that the construction in effect provides the level $1 - \alpha$ upper confidence or fiducial limit $\bar{X} + cS$ for the parameter $M + k\Sigma$, as given by Fisher (1973, Section V.5).

The Fisherian SFM of Example 2.1 is pivotal and generates standard pivotal confidence bands. For the choice $A = (\alpha, 1)$, the set xA^{-1} is $\{\theta \in \Theta_x : \alpha < F_\theta(x) < 1\}$, since the pivotal function $p(x, \theta)$ is just $F_\theta(x)$. The confidence belt may be expressed as $\{\{\theta < \theta_x\} \cap \Theta_x : x \in \mathscr{X}\}$ where $\theta_x = \sup\{\theta : F_\theta(x) > \alpha\}$. A Fisherian family $\mathscr{F}$ inherits the confidence-belt generating power of its associated (pivotal) Fisherian SFM.

EXAMPLE 4.2. The model of Example 2.5 is non-pivotal, but possesses the pivotal reduction $\langle S = \Sigma E_2 \rangle$. Both the sampling distributions of S and the fiducial distributions of Σ are the same in the full as in the reduced model. Thus sets of the form $\{\sigma : \sigma < ks\}$, which have fiducial probability $1 - \alpha$ for Σ under Π_x (where $\alpha = \Pr(\bar{\chi}_{n-1} < k^{-1})$), also form a $(1 - \alpha)$-level confidence-belt for Σ with respect to $\{P_\theta\}$.

4.2. *Monotone* SFMs. When an SFM is not pivotal, a confidence property for fiducial probability is sometimes obtainable by an entirely different construction.

Suppose there are real functions $z(x)$ (from $\mathscr{X}$ onto $\mathscr{Z} \subset R^1$) and $\lambda(\theta)$ (from Θ onto $\mathscr{L} \subset R^1$) such that, for each $e \in \mathscr{E}$, $\lambda(\theta_1) \le \lambda(\theta_2) \Leftrightarrow z(\theta_1 e) \le z(\theta_2 e)$. Then it may be seen that (a) $z(\theta e)$ is a function of $\lambda(\theta)$ (and e), (b) $\lambda(xe^{-1})$ is a function of $z(x)$ (and e); that is, $\langle Z = \Lambda E, E \sim P \rangle$ is a reduced SFM with, however, the additional monotonicity property

$$\forall e \in \mathscr{E}, \qquad \lambda_1 \le \lambda_2 \Leftrightarrow \lambda_1 e \le \lambda_2 e. \tag{4.1}$$

(Note the special meaning of λe even when λ is real!) An SFM with property (4.1) will be called *monotone* whether or not it has been obtained by reduction. By Section 3, we know that fiducial inference for $\lambda(\Theta)$ from the observation $X = x$ in the full model coincides with that for Λ from the observation $Z = z(x)$ in the reduced model. Therefore, we may confine our interest to the reduced model as far as Λ is concerned and state the following lemma:

LEMMA 4.1.

$$\Pi_z(\Lambda \le \lambda) = P_\lambda(Z \ge z).$$

PROOF.

$$\begin{aligned} \Pi_z(\Lambda \le \lambda) &= P(zE^{-1} \le \lambda), && \text{by definition,} \\ &= P(z \le \lambda E), && \text{by (4.1),} \\ &= P_\lambda(z \le Z), && \text{by definition.} \end{aligned}$$

COROLLARY 4.1. *The fiducial distributions for a monotone functional model depend only on the sampling distributions of the data and not further on the transformation structure of the model.*

Examples 2.2 and 2.3 are monotone functional models (without reduction). So is Example 2.1, (i) through (iv), a special case of which is the Fisherian SFM corresponding to the regular Fisherian family of cdfs $\{F_\phi : -1 < \phi < 1\}$ of Example 2.2. Corollary 4.1 then establishes the identity asserted at the end of Example 2.2. It also explains (in conjunction with Lemma 3.1) the identity exhibited by Dempster (1963) between the marginal fiducial distribution of M/Σ obtained from Example 2.4, and that obtained by differentiation of the distribution function of $\bar{X}/S$.

COROLLARY 4.2. *If λ_z is a given strictly-increasing function of z with values in $\mathscr{L}$,*

$$\Pi_z(\Lambda \le \lambda_z) = P_{\lambda_z}(\lambda_z \le \lambda_Z).$$

To establish the confidence property of a monotone SFM, we introduce a regularity condition:

DEFINITION. The monotone SFM $\langle Z = \Lambda E, E \sim P \rangle$ is *regular* if
(i) $\mathscr{Z}$ and $\mathscr{L}$ are open intervals;
(ii) for each $z \in \mathscr{Z}$, $P(z \le \lambda E)$ is a continuous, *strictly* increasing function of λ, taking all values in (0, 1);
(iii) for each $\lambda \in \mathscr{L}$, $P(z \le \lambda E)$ is a continuous, *strictly* decreasing function of z, taking all values in (0, 1).

THEOREM 4.1. *Let $\langle Z = \Lambda E, E \sim P \rangle$ be a regular, monotone* SFM. *Then, for any fixed $\alpha \in (0, 1)$, there exists a function, $z \rightsquigarrow \lambda_z$ say, with values in $\mathscr{L}$, such that*

(a) $$\Pi_z(\Lambda \le \lambda_z) \equiv 1 - \alpha \qquad (\text{all } z \in \mathscr{Z}), \qquad \text{and}$$

(b) $$P_\lambda(\lambda \le \lambda_Z) \equiv 1 - \alpha \qquad (\text{all } \lambda \in \mathscr{L}).$$

PROOF. Define λ_z by $P(z \le \lambda_z E) = 1 - \alpha$; this is unique, by regularity. Lemma 4.1 then shows that, for the observation $Z = z$, the interval $\Lambda \le \lambda_z$ has fiducial probability $1 - \alpha$. Moreover, with regularity, λ_z is strictly increasing in z and $\{\lambda_z : z \in \mathscr{Z}\} = \mathscr{L}$. So, by Corollary 4.2, $P_\lambda(\lambda \le \lambda_Z) = 1 - \alpha$ for all $\lambda \in \mathscr{L}$ and we conclude that $\{\Lambda \le \lambda_z : z \in \mathscr{Z}\}$ is also a confidence belt with confidence level $1 - \alpha$.

In fact, it is not necessary to use the full force of Lemma 4.1 to derive this result. The assumption of regularity for a monotone SFM $M = \langle Z = \Lambda E, E \sim P \rangle$ implies regularity, as defined in Section 2, of the family of cdfs, $\mathscr{F}$, associated with its sampling distributions $\{P_\lambda : \lambda \in \mathscr{L}\}$. This then implies that $\mathscr{F}$ is Fisherian and hence that there is a (monotone, pivotal) Fisherian SFM, M^* say, having the same sampling distributions as M and consequently, by Corollary 4.1, the same fiducial distributions. The confidence property now follows from Section 4.1.

Hora and Buehler (1966, Theorem 6.1) proved a result similar to our Theorem 4.1, in a group-invariant setting. However, they conjectured, wrongly, that it only applied to cases which already had the pivotal confidence property. Pierce and Bogdanoff (1971) showed the result to hold for a special non-pivotal case, essentially identical with our Example 2.3.

5. Partitionable functional models and associated fiducial inference. We now relax the assumption $\mathscr{E}_x \equiv \mathscr{E}$, characteristic of an SFM. Conditions supporting inference in the resulting general functional models have been considered by Fraser (1971), Brenner and Fraser (1979), H. Bunke (1975) and O. Bunke (1976), to which we refer the reader for further motivation.

A *partitionable functional model* is defined as one for which $\{\mathscr{E}_x : x \in \mathscr{X}\}$ constitutes a partition of $\mathscr{E}$. It is easily seen that partitionability holds when the condition that x and e be compatible is expressible in the form $a(x) = u(e)$ for some functions a on $\mathscr{X}$ and u on $\mathscr{E}$. Conversely, assuming partitionability, we may label the sets in this partition and let

$a(x)$ and $u(e)$ denote the values of the label attached to $\mathscr{E}_x$ and to the set of the partition that contains e, respectively. Then $a(X) = u(E)$.

The quantity $a(X)$, being a function of E alone, is distributed independently of Θ and will be called the *functional ancillary*: it is always uniquely determined (up to equivalence), unlike ancillary statistics for distribution models. The observation $X = x$ yields direct information about E of the form $u(E) = a(x) = a$, say, making it appropriate to analyze the *conditional* SFM $\langle X = \Theta E, E \sim P_a \rangle$ where P_a is the distribution of E conditional on $u(E) = a$, yielding the fiducial model $\langle \Theta = xE^{-1}, E \sim P_a \rangle$.

EXAMPLE 5.1. *Location-scale model.* Suppose $\mathscr{X} = \mathscr{E} = R^n$ ($n \geq 2$) and $\Theta = R \times R^+$. For $x = (x_1, \cdots, x_n)$, $e = (e_1, \cdots, e_n)$, $\theta = (\mu, \sigma)$, suppose $x = \theta e$ is given by $x_i = \mu + \sigma e_i$ ($i = 1, \cdots, n$). The distribution P of E is arbitrary. We may take $u(e) = ((e_1 - \bar{e})/s_e, \cdots, (e_n - \bar{e})/s_e)$ where s_e is the sample standard deviation of $e_1, \cdots, e_n$, and hence $a(x) = ((x_1 - \bar{x})/s_x, \cdots, (x_n - \bar{x})/s_x)$. For compatible x and e, $\theta = xe^{-1}$ may be expressed as, for example, $(\bar{x} - \bar{e}s_x/s_e, s_x/s_e)$ and the resulting fiducial model is $\langle \Theta = (\bar{x} - s_x\bar{E}/s_E, s_x/s_E), E \sim P_a \rangle$.

EXAMPLE 5.2. Suppose $\mathscr{E}$ is as in Example 5.1, but $\mathscr{X} = \{x \in R^n : s_x = 1\}$ and $\Theta = R^1$. Suppose $x = \theta e$ is given by $x_i = (\theta + e_i)/s_e$ ($i = 1, \cdots, n$). We may take $u(e)$ as in Example 5.1, whence $a = (x_1 - \bar{x}, \cdots, x_n - \bar{x})$. Then $\theta = xe^{-1}$ may be expressed as $\bar{x}s_e - \bar{e}$ and the fiducial model is $\langle \Theta = \bar{x}s_E - \bar{E}, E \sim P_a \rangle$.

Clearly Examples 5.1 and 5.2 generalize Examples 2.4 and 2.3, respectively, to which they effectively reduce when, under P, the components of E are independent standard normal, so that $(\bar{E}, s_E)$, being independent of $u(E)$, has the same distribution whether under P or P_a.

A simple example of a functional model that is not partitionable is given by $\mathscr{X} = \mathscr{E} = R^1$, $\Theta = R^+$ and $\theta e = \theta + e$. In this case $\mathscr{E}_x = \{e : e < x\}$. A more interesting case is the *variance component model*, in which $\mathscr{X}$ is the space of real $I \times J$ matrices ($I, J \geq 2$), $\mathscr{E} = R^I \times \mathscr{X}$, $\Theta = R^2 \times R^+$, and where, with $x = (x_{ij})$, $\theta = (\mu, \sigma, \tau)$, $e = ((e_i), (f_{ij}))$, $x = \theta e$ is given by $x_{ij} = \mu + \sigma e_i + \tau f_{ij}$. It turns out that this model is partitionable for $I = 2$, but not otherwise.

Fraser's general group-structural models (Fraser, 1961, 1968) are partitionable, as are their reductions of the type of Example 5.2. However, a general necessary and sufficient criterion for partitionability does not appear to be currently available.

6. Conditional consistency. Suppose that, for the set-up of Section 3, the value λ of Λ becomes *known* and we are interested in "fiducial" inference about Θ, conditional on this information and the observation $X = x$.

Two distinct ways of taking account of the information are:

(I) to condition the overall fiducial distribution, Π_x, of Θ on $\Lambda = \lambda$ in the usual manner to give the *conditional fiducial distribution* $\Pi_x(\cdot \mid \lambda)$;

(II) to use the fact that the restriction of possible values of Θ to $\Theta^\lambda = \{\theta : \lambda(\theta) = \lambda\}$ yields, for the set-up considered, a partitionable functional model with $a(X) = Z$ and $u(E) = \lambda E$. (This is because $x = \theta e$ for some $\theta \in \Theta^\lambda$ if and only if $z(x) = \lambda e$.) The method of Section 5 then gives a *restricted fiducial distribution* on Θ^λ, Π_x^λ say, for Θ based on $X = x$ from the fiducial model $\langle \Theta = xE^{-1}, E \sim P_z \rangle$, where P_z is P conditional on $\lambda E = z$ $(= z(x))$.

Dempster (1963) has investigated these two ways for our Example 2.4 with $\Lambda = M/\Sigma$, $Z = \bar{X}/S$ and has shown that they yield different answers. This breakdown in the self-consistency which has so far characterized our development of the functional-model approach to fiducial inference stems from different conditionings of the distribution of E: in method (I), on the value λ for the function zE^{-1} [$= zE_2 - E_1$ in Example 2.4] but, in method (II), on the value z for the function λE [$= (\lambda + E_1)/E_2$ in Example 2.4]. Note that the logical restriction on E expressed by either condition is the same, but it is embedded in different partitions of $\mathscr{E}$ in the two cases, leading to different conditional distributions.

The same inconsistency may lead to discrepancy between (a) the fiducial distribution for Θ for the SFM $\langle X = \Theta E, E \sim P \rangle$ and (b) the "step-by-step" distribution for Θ obtained by compounding the distribution for Λ from the reduced SFM $\langle Z = \Lambda E, E \sim P \rangle$ and the restricted fiducial distribution Π_x^λ for Θ given $\Lambda = \lambda$ and $X = x$.

As Pedersen (1978) has pointed out, Example 5.1 shows that discrepancy may likewise occur between two alternative step-by-step decompositions: the one above and the other in which $\Lambda = \Sigma$ and $Z = S$ define the reduced SFM.

The fact that the latter decomposition happens to be consistent with the overall fiducial distribution is an illustration of consistency due to pivotality of $\langle Z = \Lambda F, F \sim P' \rangle$, the minimal representation of the reduced model. In this case, $Z = \Lambda F$ can be re-expressed in the form $q(Z, \Lambda) = F$. For given λ, the equations $z = \lambda f$, $f = q(z, \lambda)$ set up a one-one correspondence between z and f. Consequently, conditioning on $\lambda F = z$, as in (II), is equivalent to conditioning on $F = q(z, \lambda)$. Similarly, conditioning on $zF^{-1} = \lambda$, as in (I), is also equivalent to conditioning on $F = q(z, \lambda)$, and we get a unique conditional inference.

If we start with a partitionable functional model having functional ancillary $b(X) = w(E)$ say, and such that its induced conditional SFM's all admit a pivotal reduction as above, then we similarly obtain identical results from routes (I) and (II). (In (II), further conditioning on $w(E) = b(x)$ is needed throughout.)

Although it seems plausible that (I) and (II) will yield identical answers *only* when the reduced model used for conditioning is pivotal, we do not have a proof of this.

7. Fiducial-confidence. Yates (1939) stated, and Fisher (1945) reiterated, essentially the following "coverage probability" property of the probability-$(1 - \alpha)$ fiducial interval $(-\infty, \bar{x} + st_{n-1}(\alpha)/\sqrt{n})$ for a normal mean. Fixing s at the observed value, give σ its fiducial distribution; given the resulting σ and the true value μ, assign to $\bar{x}$ its conditional sampling distribution given s. The probability of coverage, that is, of $\mu < \bar{x} + st_{n-1}(\alpha)/\sqrt{n}$, in the resulting distribution (indexed by (μ, s)) of $\bar{x}$ equals the fiducial probability $1 - \alpha$.

The validity of this interpretation of a fiducial probability, which may be called the *fiducial-confidence* property, does not depend on normality but holds for a general location-scale distribution model conditioned on the (ancillary) maximal invariant statistic.

Note that the Neyman confidence property holds for any value (or distribution) of (μ, σ) but does not condition on s. Note also the *double* conditionality on s in the statement of the fiducial-confidence property. Given the observed s, attention is confined to a population of σ-values plausibly representative of the true σ-value, rather than being extended to all possible values of σ. Then, given that population, attention is confined to the reference set of samples with the observed value of s.

A general treatment of fiducial-confidence may be given in terms of

(a) an SFM $M = \langle X = \Theta E, E \sim P \rangle$ (which may have been derived as a conditional SFM in a partitionable functional model),

(b) a reduction of M yielding a *pivotal* SFM $\langle Z = \Lambda F, F \sim P' \rangle$ with $F = q(Z, \Lambda)$ as in Section 6.

Let the belt $\mathscr{A} \subset \mathscr{X} \times \Theta$ be such that the condition "$(X, \Theta) \in \mathscr{A}$" is re-expressible as a condition "$(E, Z) \in \mathscr{B}$" for some $\mathscr{B} \subset \mathscr{E} \times \mathscr{Z}$. Then

$$\Pi_x((x, \Theta) \in \mathscr{A}) = \int \Pi_x((x, \Theta) \in \mathscr{A} \mid \Lambda = \lambda)\, d\Pi_x(\lambda)$$

$$= \int \Pi_x^\lambda((x, \Theta) \in \mathscr{A})\, d\Pi_z(\lambda) \qquad (z = z(x)),$$

by Sections 3 and 6,

$$= \int P((E, z) \in \mathscr{B} \mid F = q(z, \lambda))\, d\Pi_z(\lambda).$$

Now, for any θ such that $\lambda(\theta) = \lambda$,

$$P_\theta((X, \theta) \in \mathscr{A} \mid Z = z) = P((E, z) \in \mathscr{B} \mid F = q(z, \lambda)).$$

So we can write

$$\Pi_x((x, \Theta) \in \mathscr{A}) = \int P_\theta((X, \theta) \in \mathscr{A} \mid Z = z)\, d\Pi_z(\lambda), \tag{7.1}$$

where the integrand is the coverage probability, conditional on z, of the *fiducial confidence belt* $\mathscr{A}$, which probability depends on θ only through $\lambda = \lambda(\theta)$. Thus we see that the unconditional fiducial probability of the belt depends only on z, and is realized as the expectation, with respect to the fiducial distribution for Λ given $Z = z$, of its sampling coverage probability conditional on $Z = z$ (dependent on Λ).

The above argument fails if the reduction used is nonpivotal. In that case we cannot generally expect the fiducial-confidence property for $\mathscr{A}$.

For the normal-mean application (using the notation of Example 2.4), we merely have to take $Z = S$, $\Lambda = \Sigma$, $F = E_2$ and the belt $\mathscr{A}$ defined by: $\bar{X} - k_1(S) < M < \bar{X} + k_2(S)$, equivalent to $-k_2(S) < S(E_1/E_2) < k_1(S)$. The conditional coverage probability, the integrand of (7.1), is $\Phi(k_2(s)\sqrt{n}/\sigma) - \Phi(-k_1(s)\sqrt{n}/\sigma)$, whose fiducial expectation, for $\Sigma \sim s/\bar{\chi}_{n-1}$, is the overall fiducial probability $\Pr(-k_1(s)\sqrt{n}/s < t_{n-1} < k_2(s)\sqrt{n}/s)$. The usual choice $k_i(S) = k_i S/\sqrt{n}$, with k_i a percentage point of the t_{n-1} distribution, gives a constant prespecifiable value of the fiducial probability. However, it is of interest that the choice $k_i(S) = w$ yields symmetric intervals of fixed width $2w$, but with fiducial probability $\Pr(|t_{n-1}| < w\sqrt{n}/s)$, a function of the observed value s of S.

If we start with a partitionable functional model, with functional ancillary $a(X) = u(E)$, we can still attach a fiducial-confidence interpretation to belts of the form of $\mathscr{A}$. In this case the integrand in (7.1) becomes

$$P_\theta((X, \theta) \in \mathscr{A} \mid Z = z, a(X) = a(x)),$$

for θ such that $\lambda(\theta) = \lambda$; this is the coverage probability, conditional on z, in the frame of reference which fixes $a(X)$ at its observed value. The appropriateness of such conditioning on $a(X)$ has been argued by Fisher (1934), Barnard (1976). For Example 5.1, we can thus justify belts of the form $\bar{X} - k_1(a(X), S) < M < \bar{X} + k_2(a(X), S)$. It is then always possible (but by no means necessary) to choose the functions k_1 and k_2 to give a fixed fiducial probability, irrespective of the data.

The fiducial-confidence property was used by Yates (1939) and Fisher (1939) in defense of the Behrens 'confidence interval' for the difference of two normal means. This application of our general treatment is a special case of:

$$X_i = M_i + \Sigma_i E_i, \quad S_i = \Sigma_i U_i, \quad i = 1, 2,$$

relating data $X = (X_1, S_1; X_2, S_2)$, parameter $\Theta = (M_1, \Sigma_1; M_2, \Sigma_2)$ and error $E = (E_1, U_1; E_2, U_2)$ with $E \sim P$ (unspecialized), for which we take $Z = S_2/S_1$, $\Lambda = \Sigma_2/\Sigma_1$, $F = U_2/U_1$, and a belt of the form

$$(X_1 - X_2) - S_1 \cdot k_1(S_2/S_1) < M_1 - M_2 < (X_1 - X_2) + S_1 \cdot k_2(S_2/S_1),$$

equivalent to $-k_2(Z) < (E_1/U_1) - Z(E_2/U_2) < k_1(Z)$. Suitable choices for k_1 and k_2 will ensure constant fiducial probability.

A variation on the Yates-Fisher analysis of Behrens's problem is obtained on taking $Z = (S_1, S_2)$, $\Lambda = (\Sigma_1, \Sigma_2)$, $F = (U_1, U_2)$, and a belt of the form

$$(X_1 - X_2) - b_1(S_1, S_2) < M_1 - M_2 < (X_1 - X_2) + b_2(S_1, S_2)$$

equivalent to $-b_2(S_1, S_2) < S_1(E_1/U_1) - S_2(E_2/U_2) < b_1(S_1, S_2)$. This allows the possibility of fixed-length intervals.

In these nearly standard examples, the basic SFM $\langle X = \Theta E, E \sim P\rangle$ is itself pivotal. For Example 2.5 this is not the case: only E_2, and no other function of E, is a function of $X = (\bar{X}, S)$ and $\Theta = (M, \Sigma)$. However, we may apply our general treatment of fiducial-confidence with $Z = S$, $\Lambda = \Sigma$ and $F = E_2$, and a belt $\mathscr{A}$ of the form $\bar{X} - k_1(S) < M < \bar{X} + k_2(S)$, equivalent to $-k_2(S) < E_0 + S(E_1/E_2) < k_1(S)$. We are thus assured that the fiducial probability that $\bar{x} - k_1(s) < M < \bar{x} + k_2(s)$, namely

$$\Pr(-k_1(s) < N(0, 1) \cdot (1 + (s^2/\bar{\chi}_{n-1}^2))^{1/2} < k_2(s))$$

where $N(0, 1)$ and $\bar{\chi}_{n-1}$ are independent, is the expectation of the conditional (given s, σ) sampling coverage probability

$$\Phi(k_2(s)/(1 + \sigma^2)^{1/2}) - \Phi(-k_1(s)/(1 + \sigma^2)^{1/2})$$

with respect to the fiducial distribution of Σ.

8. Discussion. Our intention in this work has been, not so much to advocate the use of fiducial inference, as to investigate how far a theory of fiducial inference can be developed in a self-consistent way. In a functional model setting, this is further, perhaps, than might be expected. However, some signs of breakdown have appeared in Sections 6 and 7, in relation to non-pivotal reductions. The theory as presented here becomes completely self-consistent, but much more restricted, if non-pivotal functional models are excluded.

We have not considered problems related to the combination of functional models, for example for independent experiments. Generally this does not yield a partitionable functional model overall. It may well be that criteria of self-consistency at this level would effectively exclude all but group-structural models (Fraser, 1962; Brenner and Fraser, 1979). A related unexplored topic is fiducial prediction.

Marginal sufficiency. In Section 3, we considered the case where we are interested only in a subparameter Λ, and we have a reduced model relating Λ to a function Z of the data. This is a problem of inference in the presence of nuisance parameters, and the decision to work only with the reduced model, thus discarding some (possibly relevant) data, requires justification. Our analysis has provided such a justification when inference is to be made in fiducial terms. From a broader viewpoint, Barnard's criterion of *G-sufficiency* (Barnard, 1963a; Barndorff-Nielsen, 1978) allows use of the reduced model for any kind of inference about Λ, when the full model is group-structural and (as is guaranteed by our Appendix, Section A2), Λ and Z are maximal invariants under a subgroup. For an arbitrary SFM, we may use the construction of the Appendix to represent the full and the reduced model as successive reductions from a group-structural model; two applications of the criterion of G-sufficiency now validate inference about Λ in the reduced model, as exemplified in Dawid (1975).

Bayesian connections. The marginalization consistency established in Section 3 shows that fiducial inference for functional models satisfies the *Reduction Principle* of Dawid (1977). In contrast, the marginalization paradox of Dawid, Stone and Zidek (1973) deals with violations of this Principle in problems with exactly the same structure considered here. However, the central inconsistency of that paradox related only to attempted interpretations of the fiducial distributions as formal posteriors, derived from an improper prior distribution by the formal application of Bayes' theorem. And, with the exception of group-structural models (for which a right-invariant prior will serve) fiducial distributions are usually *not* formal posteriors (Lindley, 1958).

Nevertheless, it follows from Section 3 and the Appendix that a fiducial distribution for an SFM can always be regarded as produced by a suitable marginalization from the full fiducial distribution in a structural model, based on an induced transformation group G. If

G is *amenable* (Bondar and Milnes, 1981), this fiducial distribution, being formal Bayes for a right invariant prior, is also interpretable as a true Bayes posterior for a suitable *finitely additive* prior; and this property extends, through marginalization, to the required fiducial distribution of the SFM (Heath and Sudderth, 1978; Sudderth, 1980). (In this case, the fiducial distributions possess a certain weak coherence property.) For finitely additive priors, posteriors need not look "Bayesian" nor are they uniquely determined (Stone, 1980), but the above considerations do lend a little Bayesian support to our general fiducial argument.

Problems of the Nile. In applications of the confidence theory of Section 4.1, it is of interest to solve problems of the type "Find the maximal function of E which is expressible as a function of X and Δ," where Δ is a given function of Θ. In general the solution to this problem governs the possibility of making pivotal confidence statements about Δ. The case $\Delta = \Theta$ corresponds to a maximal pivotal reduction, while Δ trivial yields the functional ancillary.

Likewise, in Section 7, we should wish to find a maximal function of E and Z expressible as a function of X and Δ. Such problems have been partially investigated by Barnard and Sprott (1982), under the name "generalized problems of the Nile". Some abstract theory is given by Plante (1979b).

The fiducial-confidence justification of fiducial probability of Section 7 is only convincing to the extent that one is prepared to accept, unjustified, the fiducial distribution of Λ used to average out sampling probabilities. Now if Λ_1 and Λ_2 are the parameters of two pivotal reductions, with Λ_1 a contraction of Λ_2, then any fiducial-confidence property holding with respect to Λ_1 also holds with respect to Λ_2. So (if there are degrees of acceptability) one would presumably prefer, in interpreting fiducial probabilities about a parameter Δ in these terms, to use the maximal contraction Λ of Θ which will allow a fiducial-confidence property. (The case Λ absent returns us to the full confidence interpretation of Section 4.1.)

On the other hand, use of the larger reduction, with parameter Λ_2, will allow a greater richness to the admissible fiducial-confidence belts, as seen in the Student's t and Behrens-Fisher examples. In the limit, with $\Lambda = \Theta$, any belt is allowable, but there can no longer be any confidence in the fiducial-confidence justification.

APPENDIX

A1. Coset models. Let H be a subgroup of a group G. Then any $h \in H$ can be identified with the one-one transformation $g \rightsquigarrow hg$ on G, and the orbit of any $g \in G$ under this transformation group is the *right coset* $Hg = \{hg : h \in H\}$. Thus a maximal invariant function of G under the action of H is the projection $g \rightsquigarrow Hg$ of G onto the quotient set $G/H = \{Hg : g \in G\}$. Moreover, identifying any $g \in G$ with the (well-defined) transformation on G/H given by $Hk \rightsquigarrow (Hk)g = H(kg)$ exhibits G as a group of transformations (written on the right) acting on G/H. This action is transitive, but not generally exact, since $(Hk)g = Hk$ whenever $kgk^{-1} \in H$.

Now consider an SFM $\langle X = \Theta E, E \sim P \rangle$ having $\Theta = \mathscr{X} = G/H$, $\mathscr{E} = G$, with the above action. We call this a *simple coset model* $[G, H; P]$. It arises from a simple group-structural model in which G serves as parameter-, data- and error-space, by reduction of the parameter- and data-spaces, as in Section 3, to the maximal invariants under the action of H. A number of our examples clearly illustrate this structure, for instance Example 2.3 with G the location-scale group of Example 2.4 and H its scale-subgroup.

In fact, any SFM $\langle X = \Theta E, E \sim P \rangle$ may be regarded as a simple coset model. For fix $e_0 \in \mathscr{E}$. As e_0 determines a one-one correspondence between Θ and $\mathscr{X}$, we can, without any essential change, work with the recoded data $X^0 = Xe_0^{-1}$, taking values in $\mathscr{X}^0 = \Theta$. This recoding induces an action of any $e \in \mathscr{E}$ as a one-one transformation from Θ onto $\mathscr{X}^0 =$

Θ. Changing notation so that θe now denotes the result of this transformation on θ, we have an equivalent SFM $\langle X^0 = \Theta E, E \sim P \rangle$. Note that $\mathscr{E}$ now determines a set $\mathscr{E}^0$ of invertible transformations of Θ onto itself, with $e_0 \in \mathscr{E}$ acting as the identity. Let G be the group of transformations generated by $\mathscr{E}^0$. Then the distribution P can be regarded as defined over $\mathscr{E}^0$, and thus as a distribution over G, confined to $\mathscr{E}^0$. Thus, in our SFM, we may, without loss of generality, regard E as taking values in a group G of transformations of Θ.

Suppose first that G is transitive on Θ. Fix $\theta_0 \in \Theta$, and let H be the subgroup $\{g \in G : \theta_0 g = \theta_0\}$. Then $\{g : \theta_0 g = \theta\} = Hk$ for any k such that $\theta_0 k = \theta$. We thus obtain an isomorphism between Θ and G/H: note that, if θ and Hk correspond, then so do θe and $(Hk)e$ for $e \in G$. Consequently, we may identify Θ with G/H, and the action of $e \in G$ with the standard action of G on the right of G/H, so obtaining, finally, a simple coset model equivalent to the original SFM.

If G is not transitive on Θ, the above construction applies to each orbit of Θ under G, with H possibly dependent on the orbit. However, the observation $X^0 = x^0$ tells us that Θ belongs to the same orbit as x^0 and the whole problem may be confined to this orbit, again yielding a simple coset model.

A2. Reduction. Let $[G, H; P]$ be a simple coset model, and K a subgroup of G containing H. Define the projection $\pi : G/H \to G/K$ by $\pi(Hg) = Kg$. This is well-defined, since $Hg_1 = Hg_2 \Rightarrow g_1 g_2^{-1} \in H \Rightarrow g_1 g_2^{-1} \in K \Rightarrow Kg_1 = Kg_2$. Moreover, $\pi((Hg_1)g) = \pi(H(g_1 g)) = K(g_1 g) = (Kg_1)g$, so that π effects a reduction in the sense of Section 3, with π playing the role of both $z(\cdot)$ and $\lambda(\cdot)$.

Conversely, suppose $\langle Z = \Lambda E, E \sim P \rangle$ is a reduction of an SFM $\langle X = \Theta E, E \sim P \rangle$. Representing the latter as a simple coset model $[G, H; P]$, with $\mathscr{E} \subset G$ (as in A1), relabel $\mathscr{Z}$ so that $z(Hg) \equiv \lambda(Hg) \equiv f(g)$ say. (There is a one-one correspondence between $\mathscr{Z}$ and $\mathscr{L}$ by the invertibility of $z(Hg) = \lambda(Hg)i$, where i is the identity in G which is also in $\mathscr{E}$.) Now, for $e \in \mathscr{E}$, $\lambda(Hg_1) = \lambda(Hg_2) \Leftrightarrow z(Hg_1 e) = z(Hg_2 e)$, by reduction, so that $f(g_1) = f(g_2) \Leftrightarrow f(g_1 e) = f(g_2 e) (e \in \mathscr{E})$. Since $\mathscr{E}$ generates G, it follows that $f(g_1) = f(g_2) \Leftrightarrow f(g_1 g_2^{-1}) = f(i)$. Defining $K = \{k \in G : f(k) = f(i)\}$, we see that K is a subgroup of G (containing H) and that $z(Hg_1) = z(Hg_2)$ (and $\lambda(Hg_1) = \lambda(Hg_2)$) iff $g_1 g_2^{-1} \in K$ i.e. $Kg_1 = Kg_2$. That is, z is a projection from G/H *to* G/K, as is the associated λ, i.e. any reduction is such a projection.

If K happens to be a normal subgroup of G, then G/K is itself a group, and the minimal representation of the reduced model is a simple group-structural model and therefore pivotal. This, then, is a sufficient condition for conditional consistency as defined in Section 5.

REFERENCES

BARNARD, G. A. (1963a). Some logical aspects of the fiducial argument. *J. Roy. Statist. Soc.*, B **25** 111–114.

BARNARD, G. A. (1963b). Logical aspects of the fiducial argument. *Bull. Int. Statist. Inst.* **40** 870–883.

BARNARD, G. A. (1976). Conditional inference is not inefficient. *Scand. J. Statist.* **3** 132–134.

BARNARD, G. A. (1977). Pivotal inference and the Bayesian controversy. *Bull. Int. Statist. Inst.* **47** (1) 543-551.

BARNARD, G. A. (1981). A coherent view of statistical inference for public science. Paper presented at *Symposium on Statistical Inference and Applications*, University of Waterloo, August 1981.

BARNARD, G. A. and SPROTT, D. A. (1983). The generalized problem of the Nile: robust confidence sets for parametric functions. *Ann. Statist.* **11** to appear.

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory.* Wiley, New York.

BONDAR, J. V. and MILNES, P. (1981). Amenability: a survey for statistical applications of Hunt-Stein and related conditions on groups. *Z. Wahrsch. verw. Gebiete* **57** 103–128.

BRENNER, D. and FRASER, D. A. S. (1979). On foundations for conditional probability with statistical models—when is a class of functions a function? *Statist. Hefte* (N.F.) **20** 148–159.

BUEHLER, R. J. (1982). Fiducial inference I. Entry in *Encyclopedia of Statistical Sciences*, ed. by S. Kotz and N. L. Johnson. Wiley, New York.

BUNKE, H. (1975). Statistical inference: Fiducial and structural vs. likelihood. *Math. Operationsforsch. u. Statist.* **6** 667–676.

BUNKE, O. (1976). Conditional probability in incompletely specified stochastic equations and statistical inference. *Math. Operationsforsch. u. Statist.* **7** 673–678.

DAWID, A. P. (1975). On the concepts of sufficiency and ancillarity in the presence of nuisance parameters. *J. Roy. Statist. Soc.* B **37** 248–258.

DAWID, A. P. (1977). Conformity of inference patterns. In *Recent Developments in Statistics*, eds. J. R. Barra, B. van Cutsen, F. Brodeau and G. Romier. North Holland, Amsterdam. 245–256.

DAWID, A. P., STONE, M., and ZIDEK, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference (with Discussion). *J. Roy. Statist. Soc.* B **35** 189–233.

DEMPSTER, A. P. (1963). Further examples of inconsistencies in the fiducial argument. *Ann. Math. Statist.* **34** 884–891.

DEMPSTER, A. P. (1969). *Elements of Continuous Multivariate Analysis.* Addison-Wesley, Reading, Mass.

EDWARDS, A. W. F. (1982). Fiducial Inference II. Entry in *Encyclopedia of Statistical Sciences*, eds. S. Kotz and N. L. Johnson, Wiley, New York.

FISHER, R. A. (1930). Inverse probability. *Proc. Camb. Phil. Soc.* **26** 528–535.

FISHER, R. A. (1934). Two new properties of mathematical likelihood. *Proc. Roy. Soc.* Ser. A **144** 285–307.

FISHER, R. A. (1935). The fiducial argument in statistical inference. *Ann. Eugenics* **6** 391–398.

FISHER, R. A. (1939). The comparison of samples with possibly unequal variances. *Ann. Eugenics* **9** 174–180.

FISHER, R. A. (1945). The logical inversion of the notion of the random variable. *Sankhyā* **7** 129–132.

FISHER, R. A. (1973). *Statistical Methods and Scientific Inference.* (Third Edition.) Hafner, New York.

FRASER, D. A. S. (1961). The fiducial method and invariance. *Biometrika* **48** 261–280.

FRASER, D. A. S. (1962). On the consistency of the fiducial method. *J. Roy. Statist. Soc.* B **24** 425–434.

FRASER, D. A. S. (1964). On the definition of fiducial probability. *Bull. Int. Statist. Inst.* **40** 842–856.

FRASER, D. A. S. (1968). *The Structure of Inference.* Wiley, New York.

FRASER, D. A. S. (1971). Events, information processing, and the structural model. In *Foundations of Statistical Inference*, ed. by V. P. Godambe and D. A. Sprott. (Proceedings of the Symposium on the Foundations of Statistical Inference, University of Waterloo, 1976), Holt, Rinehart & Winston, Toronto.

HEATH, D. and SUDDERTH, W. (1978). On finitely additive priors, coherence and extended admissibility. *Ann. Statist.* **6** 333–345.

HORA, R. B. and BUEHLER, R. J. (1966). Fiducial theory and invariant estimation. *Ann. Math. Statist.* **37** 643–656.

LINDLEY, D. V. (1958). Fiducial distributions and Bayes' theorem. *J. Roy. Statist. Soc.* B **20** 102–107.

MAULDON, J. G. (1955). Pivotal quantities for Wishart's and related distributions, and a paradox in fiducial theory. *J. Roy. Statist. Soc.* B **17** 79–85.

PEDERSEN, J. G. (1978). Fiducial inference. *Int. Statist. Review* **46** 147–170.

PIERCE, D. A. and BOGDANOFF, D. A. (1971). Note on Bayes-Fiducial intervals for problems of location and scale. *Ann. Math. Statist.* **42** 833–836.

PLANTE, A. (1979a). On the validation of fiducial techniques. *Can. J. Statist.* **7** 217–226.

PLANTE, A. (1979b). Structured probability statements. *Can. J. Statist.* **7** 227–232.

SEIDENFELD, T. (1979). *Philosophical Problems of Statistical Inference.* Reidel, Boston.

STONE, M. (1980). Review and analysis of some inconsistencies related to improper priors and finite additivity. *Logic, Methodology and Philosophy of Science VI*, ed. by L. J. Cohen, J. Løs, H. Pfeiffer and K.-P. Podewski. (Proc. 6th Int. Congress). North-Holland, Amsterdam.

STONE, M. (1982). Fiducial probability. Entry in the *Encyclopedia of Statistical Sciences*, ed. by S. Kotz and N. L. Johnson, Wiley, New York.

SUDDERTH, W. (1980). Finitely additive priors, coherence and the marginalization paradox. *J. Roy. Statist. Soc.* B **42** 339–341.

WILKINSON, G. N. (1977). On resolving the controversy in statistical inference (with Discussion). *J. Roy. Statist. Soc.* B **39** 119–171.

YATES, F. (1939). An apparent inconsistency arising from tests of significance based on fiducial distributions of unknown parameters. *Proc. Camb. Phil. Soc.* **35** 579–591.

DEPARTMENT OF STATISTICAL SCIENCE
UNIVERSITY COLLEGE LONDON
GOWER STREET
LONDON WC1E 6BT, ENGLAND