The Annals of Statistics
1973, Vol. 1, No. 3, 478–485

EXPECTATION CONSISTENCY AND GENERALIZED BAYES INFERENCE

By A. P. Dawid and M. Stone

University College London

In a previous paper, the authors introduced a new criterion of expectation consistency between probability distributions of discrete data given discrete parameter values and arbitrary posterior probability distributions for the parameter. It is here shown, under very weak assumptions, that expectation consistency implies that the posterior distributions are generalized Bayes. However when the posterior distributions are generalized Bayes, the implied prior distribution need not be unique. The class of implied distributions is characterized in terms of a partition of parameter space.

1. Introduction and summary. Consider an inference problem with parameter space Θ and data space X, each finite or countable. The fixed set $p = \{p_\theta(x) : x \in X, \theta \in \Theta, \sum_{x \in X} p_\theta(x) = 1\}$ gives the conditional probabilities of the data given the parameter. We postulate a statistician S who is willing to express his inference about θ in the form of a set $\pi = \{\pi_x(\theta) : x \in X, \theta \in \Theta, \sum_{\theta \in \Theta} \pi_x(\theta) = 1\}$ of "inverse probabilities" for the parameter given the data. For an arbitrary decision problem with utility function $u(d, \theta)$, S, having observed x, is prepared to use π to evaluate any decision $d(x)$ about θ by means of the expectation

$$\sum_{\theta \in \Theta} \pi_x(\theta) u(d(x), \theta) .$$

S does not necessarily accept that π must be constructed by Bayes' theorem with some prior on Θ. However he claims he would be averse to the possibility that, in comparing two decision rules $d_1(\cdot)$ and $d_2(\cdot)$ in a decision problem,

$$E_{\pi_x}\{u(d_1(x), \tilde{\theta}) - u(d_2(x), \tilde{\theta})\} \leqq 0$$

for all $x \in X$ while

$$E_{p_\theta}\{u(d_1(\tilde{x}), \theta) - u(d_2(\tilde{x}), \theta)\} > 0$$

for all $\theta \in \Theta$. Writing $t(x, \theta) = u(d_1(x), \theta) - u(d_2(x), \theta) - E_{\pi_x}[u(d_1(x), \tilde{\theta}) - u(d_2(x), \tilde{\theta})]$, this would imply

$$E_{\pi_x} t(x, \tilde{\theta}) = 0 \qquad \forall x \tag{1.1}$$

while

$$E_{p_\theta} t(\tilde{x}, \theta) > 0 \qquad \forall \theta . \tag{1.2}$$

Conversely, given any function $t(x, \theta)$ for which $E_{\pi_x}[t(x, \tilde{\theta})] \equiv 0$, it is straight-

Received January 1972; revised July 1972.

AMS 1970 *subject classifications.* Primary 62A15; Secondary 90D30.

Key words and phrases. Inverse inference, expectation consistency, generalized Bayes, countable parameter space.

forward to construct some utility function $u(d, \theta)$ and decision rules d_1 and d_2 for which $u(d_1(x), \theta) - u(d_2(x), \theta) = t(x, \theta)$, and hence S must be averse to the possibility that (1.1) and (1.2) could hold simultaneously.

However, as we will indicate in Section 3, this aversion is generally too strong, when Θ and X are infinite, in the sense that no π will exist that avoids this possibility—unless the class of *test functions* t is restricted in some way. This difficulty is associated with the breakdown of the condition for interchanging the order of a doubly infinite summation.

In Section 2 we propose, following [1], a reasonable weakening of S's aversion criterion and show that, under simple conditions on p and π, π is then necessarily equivalent to a Bayes posterior distribution. This result extends previous work in [1].

For the case Θ and X finite, Freedman and Purves [2] develop analogous arguments and implications, using betting terminology and starting at a more primitive level. At least in the finite case, our π and t can be roughly identified with Freedman and Purves's notation as follows (although their technical conditions differ from ours):

$$\sum_{\theta \in A} \pi_x(\theta) = (1 + \lambda(x, A))^{-1}$$
$$t(x, \theta) = \sum_A w_A(\theta, x) .$$

2. Definitions and background. Our weakening of S's aversion to the possibility expressed in (1.1) and (1.2) is that he should avoid *expectation inconsistency* which, following [1], is defined by:

DEFINITION 2.1. We say that π is *expectation inconsistent* with p if, for each finite subset Ψ of Θ, there is a t and finite $\Phi \supset \Psi$ such that

(i) $t(x, \theta) \equiv 0$ if $\theta \notin \Phi$
(ii) $E_{\pi_x} t(x, \tilde{\theta}) = 0$ for all $x \in X$
(iii) $E_{p\theta} t(\tilde{x}, \theta) > 0$ for all $\theta \in \Phi$.

π is *expectation consistent* with p if it is not expectation inconsistent, that is, if there is a finite set Ψ such that, for each $\Phi \supset \Psi$, there does not exist t satisfying (i), (ii) and (iii).

The restriction (i) is clearly designed to deal with the problems created by doubly infinite summation. However the force of this restriction is weakened by the arbitrariness of Ψ, which effectively means that the positivity in (iii) has force for the whole of Θ.

Our next definition is also taken from [1]:

DEFINITION 2.2. We say that π is *Bayes* if

$$\omega(\theta)p_\theta(x) \equiv \pi_x(\theta) \sum_{\phi \in \Theta} \omega(\phi)p_\phi(x) \tag{2.1a}$$

holds for some non-null nonnegative function $\omega(\cdot)$ on Θ. Equivalently, π is

Bayes if there exist nonnegative non-null functions $\omega(\cdot)$ on Θ and $g(\cdot)$ on X such that

$$\omega(\theta)p_\theta(x) \equiv g(x)\pi_x(\theta) . \tag{2.1b}$$

The function ω is a *Bayes prior distribution.*

Note that this definition does not put any restrictions on $\pi_x(\cdot)$ for any x such that $\sum_{\phi \in \Theta} \omega(\phi)p_\phi(x) = 0$. Interpreting this sum as a 'marginal probability' for x, the arbitrariness of $\pi_x(\cdot)$ for such x appears reasonable.

The following results were established in [1]:

THEOREM 2.1. *If π is Bayes then π is expectation consistent.*

THEOREM 2.2. *If π is expectation consistent with p and $p_\theta(x) > 0$ for all $x \in X$, $\theta \in \Theta$, then π is Bayes. Moreover any ω satisfying* (2.1a) *is unique up to multiplication by a positive quantity.*

The next section shows how the second condition of Theorem 2.2 may be considerably relaxed.

3. The necessity of Bayes. If (2.1b) is to hold, it is clear that we must have $\omega(\theta) = 0$ unless $\pi_x(\theta) > 0$ whenever $p_\theta(x) > 0$. Hence we shall be particularly concerned with the set $\Delta = \{\theta : \pi_x(\theta) > 0$ whenever $p_\theta(x) > 0\}$.

The following lemma is a trivial extension of the corollary to Lemma 1 of [1]:

LEMMA 3.1. *If π is expectation consistent with p, there exists a finite set Ψ and, for every finite $\Phi \supset \Psi$, non-null, nonnegative functions ω_Φ on Φ and g_Φ on X such that*

(i) $\omega_\Phi(\theta)p_\theta(x) \equiv g_\Phi(x)\pi_x(\theta) \quad (\theta \in \Phi, x \in X)$
(ii) ω_Φ *is null in the complement of* Δ.

Although Lemma 3.1 appears to come very close to our principal objective in this section, there are considerable difficulties yet to be surmounted, and a further condition, (C), will be needed. In order to state this condition we introduce the following analysis. Following Freedman and Purves [2] we define a relation R on Θ by: $\theta_1 R \theta_2$ if there exists $x \in X$ such that both $p_{\theta_1}(x) > 0$ and $p_{\theta_2}(x) > 0$. The relation R is reflexive and symmetric, but need not be transitive. Let Γ be a nonempty subset of Θ. We define an equivalence relation $\sim_\Gamma$ on Γ by: $\theta \sim_\Gamma \tau$ if there exists a finite integer $n \geqq 0$ and a sequence $(\theta_0, \theta_1, \cdots, \theta_{n+1})$ such that $\theta_i \in \Gamma$, $i = 0, \cdots, n+1$, $\theta_0 = \theta$, $\theta_{n+1} = \tau$, and $\theta_i R \theta_{i+1}$, $i = 0, \cdots, n$. Any such sequence will be called a Γ-*ladder* for (θ, τ).

For any subset U of Θ, we define its *associate* $\bar{U} \subset X$ to be the set

$$\bar{U} = \{x : p_\theta(x) > 0 \text{ for some } \theta \in U\} .$$

Taking $\Gamma = \Theta$, consider the equivalence classes under $\sim_\Theta$, and their respective associates which are readily seen to be pairwise disjoint. Each such pair could be regarded as the parameter space and sample space of an isolated subproblem.

It is tempting to suppose that we might simplify matters by investigating each subproblem separately. However this does not appear to be possible since the subproblems are defined without relation to π.

A more relevant partition of Θ is into the family $\mathscr{T}$ of equivalence classes in Δ under $\sim_\Delta$ together with $\Theta - \Delta$. Again, the associates $\{\bar{T}: T \in \mathscr{T}\}$ are pairwise disjoint.

In Lemma 3.6 and Theorem 3.1, we shall impose the condition:

(C) $\mathscr{T}$ *is finite.*

We need some lemmas.

LEMMA 3.2. *Let G be any finite subset of Δ. Then there exists a finite set F such that $G \subseteq F \subseteq \Delta$, and*

$$\theta, \theta' \in F \quad \text{and} \quad \theta \sim_\Delta \theta' \Rightarrow \theta \sim_F \theta' . \tag{3.1}$$

PROOF. For each pair $(\theta, \tau) \in G$ such that $\theta \sim_\Delta \tau$, choose a Δ-ladder, $G(\theta, \tau)$ say, for (θ, τ). Then we can take F to be the union of G and all the $G(\theta, \tau)$'s. This set is easily seen to satisfy (3.1).

Now let S denote a property which elements of Θ may or may not possess (examples will appear presently). Write $S(\theta)$ for "θ has S." The importance of the equivalence relations introduced is that frequently a property S will carry through these relations in the following obvious way.

LEMMA 3.3. *Suppose that $S(\theta_1) \Rightarrow S(\theta_2)$ whenever $\theta_1, \theta_2 \in \Gamma$ and $\theta_1 R \theta_2$. Then $S(\theta_1) \Rightarrow S(\theta_2)$ whenever $\theta_1 \sim_\Gamma \theta_2$.*

LEMMA 3.4. *Let F be any subset of Δ for which* (3.1) *holds. Suppose further that there exist nonnegative functions ω_F on F and g_F on X such that*

$$\omega_F(\theta)p_\theta(x) \equiv g_F(x)\pi_x(\theta) \qquad (\theta \in F, x \in X) . \tag{3.2}$$

Then, if $\theta_1, \theta_2 \in F$ and $\theta_1 \sim_\Delta \theta_2$, $\omega_F(\theta_1)$ and $\omega_F(\theta_2)$ are either both zero or both positive.

PROOF. Let $S(\theta)$ be "$\theta \in F$ and $\omega_F(\theta) = 0$." If $\theta_1, \theta_2 \in F$ and $\theta_1 R \theta_2$ then for some $x_0 \in X$, $p_{\theta_1}(x_0) > 0$ and $p_{\theta_2}(x_0) > 0$. If $S(\theta_1)$ holds, (3.2) implies that

$$0 = \omega_F(\theta_1)p_{\theta_1}(x_0) = g_F(x_0)\pi_{x_0}(\theta_1) .$$

But since $p_{\theta_1}(x_0) > 0$ and $\theta_1 \in \Delta$, $\pi_{x_0}(\theta_1) > 0$. Hence $g_F(x_0) = 0$. So

$$\omega_F(\theta_2)p_{\theta_2}(x_0) = g_F(x_0)\pi_{x_0}(\theta_2) = 0 ,$$

whence $\omega_F(\theta_2) = 0$ and so $S(\theta_2)$ holds. The result now follows from Lemma 3.3.

LEMMA 3.5. *If π is expectation consistent with p, there exists an increasing sequence of finite subsets of Θ, $\Phi_1, \Phi_2, \cdots$ say, with $\lim_{i\to\infty} \Phi_i = \Theta$ such that, for $i = 1, 2, \cdots$,*

(i) *the implication* (3.1) *holds with $F = F_i =_{\text{def}} \Delta \cap \Phi_i$*

(ii) *there exist nonnegative functions ω_i on Φ_i and g_i on X such that*

$$\omega_i(\theta)p_\theta(x) \equiv g_i(x)\pi_x(\theta) \qquad (x \in X, \theta \in \Phi_i) \tag{3.3}$$

with ω_i non-null when restricted to F_i.

PROOF. Write $\Theta = \{\theta_1, \theta_2, \cdots\}$. After the ith stage of construction we have Φ_i. In case Θ is finite and $\Phi_i = \Theta$, set $\Phi_{i+1} = \Theta$. Otherwise let $j = \min\{k : \theta_k \notin \Phi_i\}$. Taking $G = \Delta \cap (\Phi_i \cup \{\theta_j\})$, construct F from Lemma 3.2 and set $\Phi_{i+1} = F \cup \Phi_i \cup \{\theta_j\}$. We start with arbitrary finite $\Phi_0 \supseteqq \Psi$, where Ψ is given by Lemma 3.1.

It is clear that (i) holds for $i = 1, 2, \cdots$, and, by Lemma 3.1, (3.3) holds for $\omega_i = \omega_{\Phi_i}$, $g_i = g_{\Phi_i}$; and ω_i restricted to Δ is non-null.

LEMMA 3.6. *Suppose π is expectation consistent with p and* (C) *holds. Then there is an equivalence class $T^* \in \mathscr{T}$ and sequences $\{\Phi_i\}$, $\{\omega_i\}$, $\{g_i\}$, having the properties given by Lemma* 3.5, *and functions μ on T^* and λ on $\bar{T}^*$, such that*

$$\omega_i(\theta) = c_i\,\mu(\theta) \qquad (\theta \in \Phi_i \cap T^*)$$
$$g_i(x) = c_i\,\lambda(x) \qquad (x \in \bar{T}^*)$$

where $c_i > 0$. Moreover $\mu(\theta) > 0$ ($\theta \in T^*$) and $\lambda(x) > 0$ ($x \in \bar{T}^*$).

PROOF. Select a point τ from each equivalence class $T \in \mathscr{T}$, yielding a finite set $A = \{\tau_1, \cdots, \tau_k\}$ say. Take $\Phi_0 = \Psi \cup A$ and construct $\{\Phi_i\}$, $\{F_i\}$, $\{\omega_i\}$ as in Lemma 3.5. Then $A \subseteqq \Phi_i$ for each i. For any T, Lemma 3.4 shows that $\omega_i(\theta) = 0$ or $\omega_i(\theta) > 0$, simultaneously for all $\theta \in \Phi_i \cap T$. Since ω_i is non-null on Δ there exists $\tau \in A$ with $\omega_i(\tau) > 0$. Then since A is finite, there exists $\tau^* \in A$ for which $\omega_i(\tau^*) > 0$ for infinitely many i values. Choosing a subsequence if necessary, we may suppose $\omega_i(\tau^*) > 0$ for every i. Let T^* be the equivalence class containing τ^*.

If $\theta \in \Phi_i \cap T^*$ then $\theta \in F_i$ and $\theta \sim_\Delta \tau^*$. By (3.1), $\theta \sim_{F_i} \tau^*$. Let $(\theta_0, \theta_1, \cdots, \theta_{n+1})$ be a F_i-ladder for (θ, τ^*). Then clearly, for $r = 0, 1, \cdots, n+1$, $\theta_r \sim_\Delta \tau^*$ so that $\theta_r \in \Phi_i \cap T^*$. Hence $\omega_i(\theta_r) > 0$, $r = 0, 1, \cdots, n+1$, and

$$\frac{\omega_i(\theta)}{\omega_i(\tau)} = \prod_{r=0}^{n} \frac{\omega_i(\theta_r)}{\omega_i(\theta_{r+1})}\,.$$

Since $\theta_r\, R\, \theta_{r+1}$, there exists x such that both $p_{\theta_r}(x) > 0$ and $p_{\theta_{r+1}}(x) > 0$ (and so too $\pi_x(\theta_r) > 0$ and $\pi_x(\theta_{r+1}) > 0$). Then from (3.3),

$$\frac{\omega_i(\theta_r)}{\omega_i(\theta_{r+1})} = \frac{\pi_x(\theta_r)}{\pi_x(\theta_{r+1})} \cdot \frac{p_{\theta_{r+1}}(x)}{p_{\theta_r}(x)}$$

and this is independent of i; hence so is $\omega_i(\theta)/\omega_i(\tau)$.

Defining μ on T^* by $\mu(\theta) = \omega_i(\theta)/\omega_i(\tau^*)$ if $\theta \in \Phi_i \cap T^*$ we see that μ is well defined and $\mu(\theta) > 0$.

If $x \in \bar{T}^*$, there exists $\theta^* \in T^*$ with $p_{\theta^*}(x) > 0$. Then as soon as $\theta^* \in \Phi_i$, (3.3) yields $\mu(\theta)p_\theta(x) \equiv \lambda(x)\pi_x(\theta)$ ($\theta \in \Phi_i \cap T^*$) (where $\lambda(x) = g_i(x)/\omega_i(\tau^*)$) and this equation defines $\lambda(x)$ uniquely since all terms are positive when $\theta = \theta^*$. In particular $\lambda(x) > 0$, and

$$\mu(\theta)p_\theta(x) \equiv \lambda(x)\pi_x(\theta) \qquad (\theta \in T^*)\,. \tag{3.4}$$

It is clear that the lemma is satisfied with $c_i = \omega_i(\tau^*)$. We can now prove the main result.

THEOREM 3.1. *Suppose that π is expectation consistent with p, and that Δ has a* finite *number of equivalence classes under $\sim_\Delta$. Then π is Bayes, and* (2.1b) *holds for*

$$\omega(\theta) = \mu(\theta) \qquad (\theta \in T^*)$$
$$= 0 \qquad (\theta \notin T^*)$$

and

$$g(x) = \lambda(x) \qquad (x \in \bar{T}^*)$$
$$= 0 \qquad (x \notin \bar{T}^*)$$

where T^*, μ and λ are provided by Lemma 3.6.

PROOF. If $x \in \bar{T}^*$ and $\theta \in T^*$ then (3.4) shows that (2.1b) holds. If $x \notin \bar{T}^*$ then $g(x) = 0$; also $p_\theta(x) = 0$ if $\theta \in T^*$ and $\omega(\theta) = 0$ if $\theta \notin T^*$. The remaining case to consider is when $x \in \bar{T}^*$ and $\theta \notin T^*$, when $\omega(\theta) = 0$. But $\pi_x(\theta) = 0$ also. For suppose $\pi_x(\theta) > 0$. If $\theta \in \Delta - T^*$ then $p_\theta(x) = 0$, since otherwise $\theta \in T^*$; and if $\theta \notin \Delta$ then $\omega_i(\theta) = 0$ whenever $\theta \in \Phi_i$, by Lemma 3.1 (ii). In either case (3.3) implies that $g_i(x) = 0$ if i is large enough. On the other hand, there exists $\theta^* \in T^*$ with $p_{\theta^*}(x) > 0$, and $\omega_i(\theta^*) > 0$ whenever $\theta^* \in \Phi_i$. Since $g_i(x)\pi_x(\theta^*) = p_{\theta^*}(x)\omega_i(\theta^*)$ for $\theta^* \in \Phi_i$, it follows that $g_i(x) > 0$ if i is large enough. This contradiction shows that $\pi_x(\theta) = 0$ and completes the proof.

The condition (C) used in this theorem is a wide one, but a trivial example shows that it is not necessary.

EXAMPLE. Let $X = \Theta = \{\text{natural integers}\}$, and suppose $p_\theta(\theta) = 1$ for $\theta \in \Theta$, $\pi_x(x) = 1$ for $x \in X$. Then π is expectation consistent with p and is Bayes; but Δ has an infinite number of equivalence classes.

We are now in a position to support the remark made in Section 1 concerning the necessity for restricting the class of test functions.

Suppose that (1.1) and (1.2) could be simultaneously avoided for arbitrary t. Then π is *a fortiori* expectation consistent with p. If moreover, condition (C) holds then π is Bayes. However Appendix 1 of [1] indicates how one may then be able to construct a test-function for which (1.1) and (1.2) both hold, contradicting the supposition.

4. Characterization of the Bayes prior distributions. In this section we suppose that π is expectation consistent with p, and characterize the class of functions ω for which (2.1a) can hold. We start with a special case.

THEOREM 4.1. *Suppose that $\pi_x(\theta) > 0$ whenever $p_\theta(x) > 0$ and that $\theta \sim_\Theta \tau$ for all $\theta, \tau \in \Theta$. Then π is Bayes and ω, satisfying* (2.1a), *is everywhere positive, and is unique up to a positive multiple.*

PROOF. In this case, Δ is Θ, and there is only one equivalence class under

$\sim_\Delta$. So, by Theorem 3.1, π is Bayes. Positivity of ω follows by Lemma 3.4 taking $F = \Theta$, $\omega_F \equiv \omega$ and $g_F(x) = \sum_{\phi \in \Theta} \omega(\phi)p_\phi(x) = g(x)$ say. Finally suppose (2.1b) holds for ω and g, and also for ω' and g'. Take $\tau \in \Theta$ and let $S(\theta)$ denote "$\omega(\theta)/\omega'(\theta) = \omega(\tau)/\omega'(\tau)$." If $\theta_1 \, R \, \theta_2$, there exists x^* such that $p_{\theta_i}(x^*) > 0$ and $\pi_{x^*}(\theta_i) > 0$ for $i = 1, 2$. Then $g(x^*)$, $g'(x^*) > 0$, and (2.1b) implies that $\omega(\theta_1)/\omega'(\theta_1) = g(x^*)/g'(x^*) = \omega(\theta_2)/\omega'(\theta_2)$. So $S(\theta_1) \Rightarrow S(\theta_2)$; hence by Lemma 3.3, $\omega(\theta)/\omega'(\theta)$ is the same for all $\theta \in \Theta$.

For the general case, consider the (finite or countable) set $\mathscr{F} = \{T_1, T_2, \cdots\}$ of equivalence classes of Δ under $\sim_\Delta$. Suppose that, for some $T \in \mathscr{F}$,

$$(4.1) \qquad \sum_{\theta \in T} \pi_x(\theta) = 1 \qquad \text{for all } x \in \bar{T}\,.$$

Then p and π may be restricted to $(x, \theta) \in \bar{T} \times T$ while retaining the structure of Section 1, and we can ask whether, in this *restricted* problem, π is expectation consistent with p. Define $\mathscr{T}$ to be the set of equivalence classes $T \in \mathscr{F}$ such that (4.1) holds, and also

(4.2) π is expectation consistent with p when restricted to $(x, \theta) \in \bar{T} \times T$.

If $T \in \mathscr{T}$ then Theorem 4.1 applies to the restricted problem; hence there exists ω^T on T, strictly positive and unique up to a positive multiple, and g^T on $\bar{T}$ such that

$$(4.3) \qquad \omega^T(\theta)p_\theta(x) \equiv \pi_x(\theta)g^T(x) \qquad (x \in \bar{T}, \theta \in T)\,.$$

Clearly $g^T(x) > 0$ for $x \in \bar{T}$.

We now show that ω^T and g^T may be extended to the whole of Θ and X respectively whilst preserving (4.3).

THEOREM 4.2. *If $T \in \mathscr{T}$ then π is Bayes; and* (2.1a) *holds with $\omega(\theta) = \omega^T(\theta)$ if $\theta \in T$, $\omega(\theta) = 0$ otherwise.*

PROOF. We show that (2.1b) holds for this ω and for $g(x) = g^T(x)$ $(x \in \bar{T})$, $g(x) = 0$ otherwise. If $\theta \in T$ and $x \in \bar{T}$, (2.1b) follows from (4.3). If $\theta \in T$ and $x \notin \bar{T}$, then $p_\theta(x) = 0$ and $g(x) = 0$. If $\theta \notin T$ and $x \in \bar{T}$, $\omega(\theta) = 0$ and $\pi_x(\theta) = 0$ by (4.1). Finally if $\theta \notin T$ $x \notin \bar{T}$, $\omega(\theta) = 0$ and $g(x) = 0$. Hence (2.1b) holds in all cases.

Theorem 4.2 shows that if $\mathscr{T}$ is nonempty then π is Bayes. The next theorem gives a converse result.

THEOREM 4.3. *Suppose π is Bayes. Then $\mathscr{T}$ is nonempty, and the general solution to* (2.1a) *is given by*:

$$(4.4) \qquad \begin{array}{l} \omega(\theta) = \lambda_T \omega^T(\theta) \quad \textit{if} \quad \theta \in T \quad \textit{and} \quad T \in \mathscr{T} \\ \omega(\theta) = 0 \quad \textit{otherwise; where} \quad \lambda_T \geqq 0 \quad \textit{and} \quad \sum_{T \in \mathscr{T}} \lambda_T > 0\,. \end{array}$$

PROOF. Suppose ω satisfies (2.1a). Then (2.1b) holds with

$$g(x) = \sum_{\phi \in \Theta} \omega(\phi)p_\phi(x)\,.$$

Now $\omega(\theta) = 0$ if $\theta \notin \Delta$, since then, for some x, $p_\theta(x) > 0$ and $\pi_x(\theta) = 0$. Also, if $T \in \mathscr{F}$, (3.1) holds with $F = T$. Hence, by Lemma 3.4, $\omega(\theta)$ restricted to T is either null or everywhere positive. Since ω is non-null, there exists T such that $\omega(\theta) > 0$ for all $\theta \in T$. Let $x \in \bar{T}$. Then $p_\theta(x) > 0$ for some $\theta \in T$, so that (2.1b) implies $g(x) > 0$. If $\phi \notin \Delta$, $\omega(\phi) = 0$; and if $\phi \in \Delta - T$, $p_\phi(x) = 0$. In either case (2.1b) shows that $\pi_x(\phi) = 0$, and hence (4.1) holds. Also,

$$\omega(\theta)p_\theta(x) \equiv g(x)\pi_x(\theta) \qquad (\theta \in T,\ x \in \bar{T}) \tag{4.5}$$

and ω is non-null on T, so that π is Bayes in the problem restricted to $(x, \theta) \in \bar{T} \times T$. Then by Theorem 2.1, π is expectation consistent with p in the restricted problem; i.e. (4.2) holds. Hence $T \in \mathscr{T}$ and $\mathscr{T}$ is nonempty. We see that $\mathscr{T}$ contains all those equivalence classes T for which ω is non-null on T.

Hence if $\theta \notin \bigcup \{T\colon T \in \mathscr{T}\}$, $\omega(\theta) = 0$; while if $T \in \mathscr{T}$, either ω is null on T, or else, comparing (4.3) and (4.5), ω is a positive multiple of ω^T on T. This shows that any solution to (2.1a) is of the form (4.4). Conversely, if ω is of this form, it is easily seen, using Theorem 4.2, that ω satisfies (2.1a). This completes the proof.

In particular, this characterization of the solution to (2.1a) holds when the conditions of Theorem 3.1 hold.

REFERENCES

[1] Dawid, A.P. and Stone, M. (1972). Expectation consistency of inverse probability distributions. *Biometrika* **59** 486–489.

[2] Freedman, D.A. and Purves, R.A. (1969). Bayes' method for bookies. *Ann. Math. Statist.* **40** 1177–1186.

Department of Statistics
University College London
Gower Street
London WC1E 6BT
England