J. Austral. Math. Soc. (*Series A*) 32 (1982), 339–348

IDENTIFIABILITY OF MIXTURES

G. M. TALLIS and P. CHESSON

(Received 13 August 1980; revised 5 February 1981)

Communicated by R. L. Tweedie

Abstract

Let $F(x, \theta)$ be a family of distribution functions indexed by $\theta \in \Omega$. If $G(\theta)$ is a distribution function on Ω, $H(x) = \int_\Omega F(x, \theta)\, dG(\theta)$ is a mixture with respect to G. If there is a unique G yielding H, the mixture is said to be identifiable. This paper summarises some known results related to identifiability of special types of mixtures and then discusses the general problem of identifiability in terms of mappings. Some new results follow for mappings with special features.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 62 E 10.

Keywords: mixture, identifiability, kernel, linear independence, strong independence, bounded inverse, unbounded inverse.

1. Introduction and review

The problem of identifiability of mixtures of distribution functions concerns the transformation

$$H(x) = \int_\Omega F(x, \theta)\, dG(\theta)$$

where $F(x, \theta)$ is a distribution function for all $\theta \in \Omega$ and G is a distribution function defined on Ω. Standard measurability conditions imposed on $F(x, \theta)$ ensures the integral makes sense, Robbins (1948). It is easy to see that $H(x)$ is a distribution function and it is called a mixture. The family $F(x, \theta)$, $\theta \in \Omega$, is referred to as the kernel of the mixture and G as the mixing distribution function. The mixture H is said to be identifiable if and only if there is a unique G yielding H.

https://doi.org/10.1017/S1446788700024897 Published online by Cambridge University Press

In general x and θ can be vector quantities but for notational simplicity the results are often written for the scalar case.

Various approaches to constructing conditions which will guarantee identifiability have been tried. These have invariably involved imposing conditions on $F(x, \theta)$ and Ω to simplify the problem to the stage of mathematical analysis. For example, if $F(x, \theta) = F(x - \theta)$, $\Omega = (-\infty, \infty)$ and $\phi_H(t) = \int_{-\infty}^{\infty} e^{itx}\, dH(x)$ has no intervals where it vanishes, then H is an identifiable mixture. Other results using special kernels and transform theory are available; for example, Barndorff-Nielsen (1965) discusses mixtures of exponential families and Maritz (1970) gives a summary of practical results related to indentifiability.

The obvious first restriction to impose on Ω is that it consists of a finite number of elements, that is, $\Omega = \{\theta_1, \theta_2, \ldots, \theta_n\}$. Then (1) becomes

$$H(x) = \sum_1^n \beta_i F_i(x) \tag{2}$$

where, $F_i(x) = F(x, \theta_i)$, β_i real, $\Sigma_1^n \beta_i = 1$. Identifiability of (2) revolves around the linear independence of the set of functions $\mathcal{F}_n = \{F_1(x), \ldots, F_n(x)\}$ in the following way.

DEFINITION. $\mathcal{F}_n$ is said to be linearly independent if for real constants c_i, $\Sigma_1^n c_i F_i(x) \equiv 0$ implies $c_i = 0$, $i = 1, 2, \ldots, n$.

THEOREM 1. *The mixture* (2) *is identifiable if and only if* $\mathcal{F}_n$ *is linearly independent.*

Thus testing for identifiability is reduced to testing $\mathcal{F}_n$ for linear dependence. At this stage we note that the sample space can always be transformed so that $F_i(-1) \equiv 0$ and $F_i(1) \equiv 1$ without altering identifiability. Thus, to test $\mathcal{F}_n$ for dependence the Gram determinant $\mathcal{G} = | \int_{-1}^{1} F_i(x) F_j(x)\, dx |$ can be evaluated, non-vanishing $\mathcal{G}$ corresponding to independence. On the other hand, by the very definition of linear dependence, if $\mathcal{F}_n$ is independent there must exist values x_1, $x_2, \ldots, x_n$ such that $| F_i(x_j) | \neq 0$. This is the condition of Teicher (1963), Theorem 1.

The next case is that of countably infinite mixtures defined by

$$H(x) = \sum_1^\infty \beta_i F_i(x), \qquad \sum_1^\infty |\beta_i| < \infty, \qquad \sum_1^\infty \beta_i = 1. \tag{3}$$

Again, H is identifiable if there exists a unique set $\{\beta_i\}$ which satisfies (3). The counterpart of linear independence in this situation is

DEFINITION 2. The set $\mathcal{F}_\infty = \{F_1(x), F_2(x), \ldots\}$ is strongly independent if $\Sigma_1^\infty c_i F_i(x) \equiv 0$ implies $c_i \equiv 0$, $\Sigma_1^\infty |c_i| < \infty$.

THEOREM 2. *The mixture* (3) *is identifiable if and only if* $\mathcal{F}_\infty$ *is strongly independent.*

Again, the proof is easy and the problem of identifiability reduces to testing $\mathcal{F}_\infty$ for strong independence. Set $d_{ij} = \int_{-1}^{1} F_i(x) F_j(x)\, dx$ and $D = [d_{ij}]$, $i, j = 1, 2, \ldots$, then we have

LEMMA 1. $\mathcal{F}_\infty$ *is strongly independent if and only if* D^{-1} *exists.*

PROOF. We show that $\Sigma_1^\infty c_i F_i(x) \equiv 0$ if and only if $Dc = 0$. Suppose that $\Sigma_1^\infty c_i F_i(x) \equiv 0$, then $\Sigma_1^\infty c_i d_{ij} = 0$ for all j and hence $Dc = 0$.

On the other hand if $Dc = 0$, then $c'Dc = 0 = \lim_{n\to\infty} \int_{-1}^{1} (\Sigma_1^n c_i F_i(x))^2\, dx = \int_{-1}^{1} (\Sigma_1^\infty c_i F_i(x))^2\, dx$ and $\Sigma_1^\infty c_i F_i(x) = 0$ almost everywhere.

Various related results, examples and tests for the existence of D^{-1} are given in Patil and Bildekar (1966) and Tallis (1969).

In the above two theorems it appears that a slightly more general question has been formulated than is needed in that the condition $\beta_i > 0$ was not required, but only $\Sigma_1^\infty |\beta_i| < \infty$. However, Lemma 2 of the following section establishes that the two different ways of formulating the identifiability problem are equivalent.

When $F(x, \theta)$ is continuous in θ a different approach is called for. Without loss of generality we assume $\Omega = [-1, 1]$, then if $T(x, \theta) = \partial F(x, \theta)/\partial\theta$ is continuous in θ and square integrable over $[-1, 1] \times [-1, 1]$, (1) can be put in the form

$$L(x) = \int_{-1}^{1} T(x, \theta) G(\theta)\, d\theta \tag{4}$$

where $L(x) = F(x, 1) - H(x)$. Thus (1) is said to be identifiable if there is a unique square integrable solution to (4). Defining $K(x, y) = \int_{-1}^{1} T(x, z) T(y, z)\, dz$, then we have

THEOREM 3. *The mixture* (1) *is identifiable if and only if the set of eigenfunctions for nonzero eigenvalues of* K *is complete.*

The result follows from the Hilbert-Schmidt theorem for symmetric kernels.

The purpose of this paper is to provide more general settings for the solution of various classes of identifiability problems. Section II formulates the problem in the context of probability measures on spaces of probability measures and establishes some basic results. Section III looks at the situaton where the event space is finite and obtains a complete solution to the identifiability problem. In Section IV we use the structure of II to give a general treatment of identifiability when there is a bounded inverse. Theorem 2 is a special case of this treatment. Finally in Section V we look at some general problems involving unbounded inverses and relate them to the results of Theorem 3.

II. General formulation and basic results

Let $(X, \mathcal{C})$ be a measurable space and $\{P_\theta, \theta \in \Omega\}$ a family of probability measures on Ω. Let $\mathcal{B}$ be a σ-field of subsets of Ω such that $P_\theta(A)$ is $\mathcal{B}$ measurable for every $A \in \mathcal{C}$. In the case where P_θ is defined in terms of a distribution function, as in the introduction, measurability of $F(x, \theta)$ as a function of θ, for each x, is necessary and sufficient for measurability of $P_\theta(A)$ for all $A \in \mathcal{C}$. A mixture of P_θ is a measure $Q(A) = \int_\Omega P_\theta(A)\, d\mu(\theta)$, $A \in \mathcal{C}$, for some probability measure μ on $(\Omega, \mathcal{B})$. We shall simply use the notation $Q = \int_\Omega P_\theta\, d\mu(\theta)$ to denote a mixture of P_θ. The problem of identifiability of mixtures can now be stated in general parametric form: Mixtures of P_θ are said to be identifiable if the mapping $\mu \mapsto \int_\Omega P_\theta\, d\mu(\theta)$, for probability measures μ on $(\Omega, \mathcal{B})$, is 1-1.

There is a degree of arbitrariness in defining the parameters of a probability distribution, but identifiability should not depend on the particular parameterisation chosen; therefore, for a general discussion of the mixture problem, we seek a standard parameterisation. The most general parameterisation is parameterisation of P_θ by itself. For this the parameter space is some set S of probability measures on $(X, \mathcal{C})$ and a mixture is defined as an integral $\int_S P\, d\mu(P)$ where μ is a probability measure on some σ-field $\mathbb{S}$ of subsets of S. We shall say that mixtures of S are identifiable if and only if the mapping $\mu \mapsto \int_S P\, d\mu(P)$ is 1-1 for probability measures μ on $(S, \mathbb{S})$.

The above formulations of the mixture problem in terms of θ and in terms of S appear equivalent when $S = \{P_\theta \mid \theta \in \Omega\}$ and indeed are equivalent whenever $\mathcal{B}$ and $\mathbb{S}$ are respectively the smallest σ-fields for which the mappings $\theta \mapsto P_\theta(A)$ and $P \mapsto P(A)$ are measurable for all $A \in \mathcal{C}$. To see this define $f(\theta) = P_\theta$ then, for all real numbers r and all $A \in \mathcal{C}$

$$f^{-1}\{P \in S \mid P(A) < r\} = \{\theta \in \Omega \mid P_\theta(A) < r\}.$$

Since $\mathbb{S}$ is generated by the sets of the form $\{P \in S \mid P(A) < r\}$ and $\mathcal{B}$ is generated by the sets $\{\theta \in \Omega \mid P_\theta(A) < r\}$ it follows that $f^{-1}(\mathbb{S}) = \mathcal{B}$. Since S is the range of f it follows from Breiman (1968, Proposition 2.12) that the relation

$$\nu(f^{-1}(T)) = \mu(T), \qquad T \in \mathbb{S}$$

establishes a 1-1 correspondence between probability measures ν on $(\Omega, \mathfrak{B})$ and probability measures μ on $(S, \mathfrak{S})$. For measures related in this way $\int_\Omega P_\theta \, d\nu(\theta) = \int_S P \, d\mu(P)$ and it is clear that the mixtures of P_θ are identifiable if and only if mixtures of S are identifiable.

In the above discussion μ has always been a probability measure. However sometimes (Teicher (1963), Tallis (1969) and the introducton to this paper) the measure μ defining the mixture is allowed to be any signed measure with $\mu(S) = 1$. In this formulation a necessary and sufficient condition that mixtures of S are identifiable is that there exists no non zero signed measure μ with the property

$$(5) \qquad \int_S P \, d\mu(P) = 0.$$

Clearly this condition is also sufficient for identifiability as formulated here. Lemma 2 below shows that it is also a necessary condition. It follows that the two different formulations of the mixture problem are equivalent.

LEMMA 2. *If μ is a non zero measure on $(S, \mathfrak{S})$ such that* (5) *holds then there are probability measures μ_1 and μ_2 on $(S, \mathfrak{S})$ such that*

$$(6) \qquad \int_S P \, d\mu_1(P) = \int_S P \, d\mu_2(P)$$

and hence mixtures of S are not identifiable.

PROOF. Since μ is a signed measure $\mu = \mu_1 - \mu_2$ with μ_1 and μ_2 non negative measures. Also one of μ_1 and μ_2 must be finite. Now $P(X) = 1$ for all $P \in S$ so (5) entails $\int_S d\mu(P) = 0$, that is, $\mu_1(S) = \mu_2(S)$, so that both μ_1 and μ_2 are finite with the same maximum. Rescaling the μ_i if necessary so that $\mu_1(S) = \mu_2(S) = 1$ we see that (6) is satisfied.

It is often convenient to make the assumption that the σ-field $\mathfrak{S}$ contains all singletons $\{P\}$, $P \in S$ for then the set of mixtures of S contains the convex hull of S, that is, S contains the set

$$\left\{ \sum_1^n \lambda_i P_i \mid \sum_1^n \lambda_i = 1, \lambda_i \geq 0, P_i \in \mathfrak{S}, i = 1, \ldots, n, n \geq 1 \right\}.$$

This assumption is made in the introduction and in later sections. The assumption is not very restrictive as it is satisfied whenever X is a separable metric space with Borel sets $\mathfrak{A}$. To see this we note that $P \in \mathfrak{S}$ is determined by its values on the open sets. There are open sets $U_1, U_2, \ldots$ such that every open set is a union of members of this class. Writing $V_1 = U_1$, $V_n = U_n - \bigcup_1^{n-1} U_i$ we see that every open set is a union of members of the disjoint class $V_1, V_2, \ldots$. It follows that any

https://doi.org/10.1017/S1446788700024897 Published online by Cambridge University Press

$P \in S$ is determined by its values on the countable collection $V_1, V_2, \ldots$. Clearly $\{P \in S \mid P(V_i) = P_0(V_i)\} \in \mathbb{S}$ for any $P_0 \in S$ and $\bigcap_{i=1}^{\infty} \{P \in S \mid P(V_i) = P_0(V_i)\} = \{P_0\}$. Hence $\mathbb{S}$ does contain all singletons.

III. The mixture problem for finite σ-fields

If the σ-field $\mathcal{Q}$ is finite the mixture problem can be reduced to a problem of linear dependence of vectors in R^n (Theorem 4). In general $\mathcal{Q}$ is generated by a unique partition $A_1, \ldots, A_n$ of non empty subsets of X. The case of a finite σ-field is equivalent to the situation where X is finite and $\mathcal{Q}$ is the power set of X or the case where S is a set of distribution functions having a common finite set of points of increase. The mixture problem for finite σ-fields is closely related to the finite mixture problem (Theorem 5).

THEOREM 4. *Mixtures of S are identifiable if and only if*

$$\{(P(A_1), \ldots, P(A_n)) \mid P \in S\} \tag{6}$$

is a linearly independent subset of R^n.

Note that the vector $(P(A_1), \ldots, P(A_n))$ uniquely determines an element of S and furthermore the mixture can only be identifiable if S has no more than n elements.

PROOF. Suppose (6) is not linearly independent. There are numbers $\lambda_1, \ldots, \lambda_n$, not all zero, and members of S, $P_1, \ldots, P_n$, such that $\Sigma \lambda_i P_i(A_j) = 0$ for $j = 1, \ldots, n$. The equation $\mu(\{P_i\}) = \lambda_i$ defines a signed measure μ such that

$$\int_S P \, d\mu(P) = 0$$

and from Lemma 2 we see that the mixture is not identifiable.

If (6) is linearly independent then $S = \{P_1, \ldots, P_r\}$ for $r \leqslant n$ and furthermore there are no $\lambda_1, \ldots, \lambda_r$, not all zero, such that $\Sigma_1^r \lambda_i P_i(A_j) = 0$ for $j = 1, \ldots, n$. Hence the mixture is identifiable.

Now suppose that we have a finite mixture problem, that is S has just $n < \infty$ elements.

THEOREM 5. *Mixtures of S are identifiable if and only if there is a finite σ-subfield* $\mathcal{Q}^*$ *of* $\mathcal{Q}$ *such that the mixtures of* S^* *are identifiable where* S^* *is the set of restrictions of members of S to* $\mathcal{Q}^*$.

PROOF. If such an $\mathcal{Q}^*$ can be found the mixture is clearly identifiable. Suppose now that mixtures of $S = \{P_1, \ldots, P_n\}$ are identifiable. We shall prove by induction that there exist $B_1, \ldots, B_n \in \mathcal{Q}$ such that the matrix $(P_i(B_j))$, $i = 1, \ldots, n$, $j = 1, \ldots, n$, has rank n. Let $r < n$ and $B_1, \ldots, B_r$ be such that $(P_i(B_j))$, $i = 1, \ldots, r$, $j = 1, \ldots, r$, has rank r. There are unique real numbers $\alpha_1, \ldots, \alpha_r$ such that $P_{r+1}(B_j) = \sum_{i=1}^r \alpha_i P_i(B_j)$. However, since mixtures are identifiable there is a $B_{r+1} \in \mathcal{Q}$ such that $P_{r+1}(B_{r+1}) \neq \sum_{i=1}^r \alpha_i P_i(B_{r+1})$ and the matrix, $(P_i(B_j))$, $i = 1, \ldots, r+1$, $j = 1, \ldots, r+1$, has rank $r+1$. Thus the required set $B_1, \ldots, B_n$ exist. Let $\mathcal{Q}^*$ be the σ-field generated by $B_1, \ldots, B_n$ and since $(P_i(B_j))$ has full rank it is clear that mixtures of S^* are identifiable.

IV. Mixtures with bounded inverses

Some simplifications of the mixture problem are achieved by imposing restrictions on the kind of acceptable 1-1 relationships between the mixture and the mixing distribution. This leads us to consider bounded inverses.

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be sets with functions ρ_i defined on $\mathfrak{M}_i \times \mathfrak{M}_i$, $i = 1, 2$, such that

(1) ρ_i: $\mathfrak{M}_i \times \mathfrak{M}_i \to E_1$ (Euclidean 1-space),

(2) $\rho_i(x, x) = 0 \ \forall x \in \mathfrak{M}_i$,

(3) $\rho_i(x, y) > 0 \ \forall x \neq y \in \mathfrak{M}_i$.

If f is a mapping f: $\mathfrak{M}_1 \to \mathfrak{M}_2$ with domain $\mathfrak{D} \subset \mathfrak{M}_1$ and range $\mathfrak{R} \subset \mathfrak{M}_2$, we define the norm of f by

$$\|f\| = \inf_{\alpha} \{\rho_2(f(x), f(y)) \leqslant \alpha\rho_1(x, y), \forall x, y \in \mathfrak{D}\}$$

and if f^{-1} exists

$$\|f^{-1}\| = \inf_{\alpha} \{\rho_1(f^{-1}(u), f^{-1}(v)) \leqslant \alpha\rho_2(u, v), \forall u, v \in \mathfrak{R}\}.$$

THEOREM 6. *f^{-1} exists and $\|f^{-1}\| < \infty$ if and only if $\rho_2(f(x), f(y)) \geqslant \alpha\rho_1(x, y)$ $\forall x, y \in \mathfrak{D}$ and some $\alpha > 0$.*

The lemma asserts essentially that f has a bounded inverse provided f separates points sufficiently.

PROOF. If $f(x) = f(y)$ then $0 = \rho_2(f(x), f(y)) \geqslant \alpha\rho_1(x, y)$ which implies that $\rho_1(x, y) = 0$, that $x = y$ and that f^{-1} exists. Moreover,

$$\rho_2(f(x), f(y)) \geqslant \alpha\rho_1(f^{-1}(f(x)), f^{-1}(f(y)))$$

or

$$\rho_1(f^{-1}(u), f^{-1}(v)) \leqslant \alpha^{-1}\rho_2(u, v) \qquad \forall u, v \in \mathfrak{R}$$

and hence $\|f^{-1}\| \leqslant \alpha^{-1} < \infty$.

On the other hand if f^{-1} exists and $\|f^{-1}\| < \infty$, then

$$\rho_1(x, y) = \rho_1(f^{-1}(f(x)), f^{-1}(f(y))) \leqslant \|f^{-1}\|\rho_2(f(x), f(y))$$

and hence $\rho_2(f(x), f(y)) \geqslant \alpha\rho_1(x, y)$ for $\alpha^{-1} = \|f^{-1}\|$ and all $x, y \in \mathfrak{D}$.

A standard result is retrieved by letting $\mathfrak{M}_1 = \mathfrak{M}_2 = \mathfrak{M}$ be a linear space with inner product $p(x, y)$. Then $\rho_1 = \rho_2 = \rho$ is the metric induced by the norm $[p(x, x)]^{1/2} = \|x\|$ that is $\rho(x, y) = \|x - y\|$ and we assume that T is a linear transformation of $\mathfrak{M}$ on to $\mathfrak{R} = \mathfrak{M}$. Theorem 6 shows that T^{-1} exists and $\|T^{-1}\| < \infty$ if and only if $\|Tx\| \geqslant \alpha\|x\| \; \forall x \in \mathfrak{M}, \alpha > 0$.

We apply these results to the mixture problem. As in Section II we consider the event space $\{X, \mathfrak{A}\}$ and the probability space $\{S, \mathbb{S}\}$. Now if $\mathfrak{M}_1$ is the set of all P-measures defined on $\{X, \mathfrak{A}\}$ then $S \subset \mathfrak{M}_1$ and, since X is a separable metric space, by Parthasarathy (1967) Theorem 6.2, page 43, there exists a metric ρ_1^* such that $\mathfrak{M}_1$ can be metrized as a separable metric space.

Consider now the set $\mathfrak{M}_2$ of all measures defined on $\{S, \mathbb{S}\}$. Since $S \subset \mathfrak{M}_1$, S is also a separable metric space with metric ρ_1^*. Thus, applying the above result again, there exists a metric ρ_2^* such that $\mathfrak{M}_2$ is metrized as a separable metric space.

Define the function f: $\mathfrak{M}_2 \to \mathfrak{M}_1$ by

$$f(\mu) = m = \int P \, d\mu(P), \qquad \mu \in \mathfrak{M}_2,$$

then we have the following corollary to Theorem 6.

COROLLARY. *The mixture m is identifiable with bounded inverse if and only if*

$$\rho_1(m_1, m_2) = \rho_1\left(\int_S P \, d\mu_1(P), \int_S P \, d\mu_2(P)\right) \geqslant \alpha\rho_2(\mu_1, \mu_2)$$

for some $\alpha > 0$ *and all* $\mu_1, \mu_2 \in \mathfrak{M}_2$, *where* ρ_1 *and* ρ_2 *are metrics on* $\mathfrak{M}_1$ *and* $\mathfrak{M}_2$ *respectively.*

It is convenient to illustrate the corollary in the countably infinite case. Choose

$$\rho_1^2(F, G) = \int_{-1}^{1} [F(x) - G(x)]^2 \, dx$$

and

$$\rho_2^2(\beta^{(1)}, \beta^{(2)}) = \sum_1^\infty (\beta_i^{(1)} - \beta_i^{(2)})^2$$

then the condition becomes

$$\int_{-1}^1 \left[\sum_1^\infty F_i(x)\gamma_i\right]^2 dx \geqslant \alpha\|\gamma\|^2, \qquad \gamma = \beta^{(1)} - \beta^{(2)},$$

that is $\gamma' D\gamma \geqslant \alpha\|\gamma\|^2$, $d_{ij} = \int_{-1}^1 F_i(x)F_j(x)\,dx$. But $|\gamma' D\gamma| \leqslant \|D\gamma\|\,\|\gamma\|$, and hence the above implies $\|D\gamma\| \geqslant \alpha\|\gamma\|$, the condition that D have a bounded inverse. In practice, sometimes this test is easily applied and may save the check as to whether or not zero is a regular point of the linear operator D.

V. Mixtures with unbounded inverses

In the notation of the previous section, we consider here briefly the more general situation of $\rho_2(\mu_1, \mu_2) > 0 \Rightarrow \rho_1(m_1, m_2) > 0$. As first example we re-examine Theorem 3.

We have

$$L(x) = \int_{-1}^1 T(x, \theta)G(\theta)\,d\theta$$

and choose $\rho_1 = \rho_2 = \rho$ defined on $[-1, 1]$ by

$$\rho(F, G) = \int_{-1}^1 [F(x) - G(x)]^2\,dx.$$

The above condition becomes

$$\int_{-1}^1 [G_1(\theta) - G_2(\theta)]^2\,d\theta$$

$$> 0 \Rightarrow \int_{-1}^1 \left(\int_{-1}^1 T(x, \theta)[G_1(\theta) - G_2(\theta)]\,d\theta\right)^2 dx > 0.$$

Defining $K(x, y)$ as in Theorem 3 we get, putting $\chi(\theta) = G_1(\theta) - G_2(\theta)$,

$$\int_{-1}^1 \chi^2(\theta)\,d\theta > 0 \Rightarrow \int_{-1}^1 \int_{-1}^1 \chi(\theta)K(\theta, \phi)\chi(\phi)\,d\theta\,d\theta > 0$$

or $\|\chi\| > 0 \Rightarrow \|K\chi\| > 0$ which is true if and only if K has an inverse.

For the final example let (Ω, F, λ) be a measure space and suppose that $P(x, \theta) \ll \lambda$ for all $\theta \in \Omega$. Put $p(x, \theta) = dP(x, \theta)/d\lambda$ and assume that $\int_X \int_\Omega p^2(x, \theta)\,d\mu(\theta)\,d\mu(x) < \infty$.

Consider the class $\mathcal{C}$ of all probability measures μ on $(\Omega, \mathcal{S})$ such that $\mu \ll \lambda$ and $\int_\Omega (d\mu/d\lambda)^2\,d\lambda < \infty$. We wish to know when mixtures of $P(x, \theta)$ with respect to elements of $\mathcal{C}$ are identifiable.

https://doi.org/10.1017/S1446788700024897 Published online by Cambridge University Press

If $\mu \in \mathcal{C}$ let $m = d\mu/d\lambda$ then the mixture is

$$Q(A) = \int_\Omega P(A,\theta)\,d\mu(\theta) = \int_A \left[\int_\Omega p(x,\theta)m(\theta)\,d\lambda(\theta)\right] d\lambda(x)$$

and

$$q(x) = \frac{dQ}{d\lambda} = \int_\Omega p(x,\theta)m(\theta)\,d\lambda(\theta).$$

Now Q is identifiable if and only if

$$\int_\Omega (m_1(\theta) - m_2(\theta))^2\,d\lambda(\theta) > 0 = \int_X [q_1(x) - q_2(x)]^2\,d\lambda(x) > 0.$$

Expand $p(x,\theta)$ as

$$p(x,\theta) = \sum_0^\infty \rho_n\phi_n(x)\psi_n(\theta)$$

where $\{\phi_n\}$ and $\{\psi_n\}$ are bi-orthonormal series and $\rho_n > 0$. We now have

$$\int_X [q_1(x) - q_2(x)]^2\,d\lambda(x) > 0$$

$\Leftrightarrow \sum_0^\infty \rho_n^2\alpha_n^2 > 0$, $\alpha_n = \int_\Omega \psi_n(\theta)(p_1(\theta) - p_2(\theta))\,d\lambda(\theta)$ and the last inequality holds if and only if $\{\psi_n\}$ is complete.

References

O. Barndoff-Nielsen (1965), 'Identifiability of mixtures of exponential families', *J. Math. Anal. Appl.* **12**, 115–121.

L. Breiman (1968), *Probability* (Addison-Wesley).

J. S. Maritz (1970), *Empirical Bayes methods* (Methuen, London).

G. P. Patil and S. Bildekar (1966), 'Identifiability of countable mixtures of discrete probability distributions using methods of infinite matrices', *Proc. Cambridge Philos. Soc.* **62**, 485–494.

K. R. Parthasarathy (1967), *Probability measures on metric spaces* (Academic Press, New York).

G. M. Tallis, 'The identifiability of mixtures of distributions', *J. Appl. Probability* **6**, 389–398.

H. Teicher (1963), 'Identifiability of finite mixtures', *Ann. Math. Statist.* **34**, 1265–1269.

Department of Statistics
University of Adelaide
Adelaide, S. A. 5001
Australia