

UNCLASSIFIED
AD 436547

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

64-12

IDENTIFIABILITY OF MIXTURES OF EXPONENTIAL FAMILIES

by

O. Barndorff-Nielsen

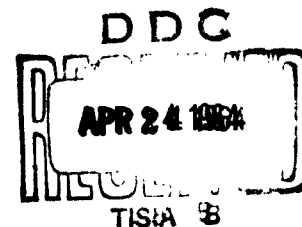
TECHNICAL REPORT NO. 38

February 17, 1964

PREPARED UNDER CONTRACT Nonr-225(28)

(NR-047-019)

OFFICE OF NAVAL RESEARCH



Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

MATHEMATICS DEPARTMENT

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

CATALOGED BY
AS AD NO. — 436547

436547

IDENTIFIABILITY OF MIXTURES OF EXPONENTIAL FAMILIES

by

O. Barndorff-Nielsen

Let $\mathcal{F}_0 = \{F(\cdot|\tau) : \tau \in T\}$ be a family of n -dimensional distribution functions (d.f.s.) depending on a m -dimensional parameter τ which ranges over a Borel set T in R^m , the m -dimensional Euclidian space. We assume that for each fixed $x = (x_1, \dots, x_n) \in R^n$ the function $F(x|\cdot)$ is Borel measurable. Let $\mathcal{F}(\mathcal{G})$ denote the set of all probability measures (p.m.s.) on the Borel field $\mathcal{G}^n$ of R^n ($\mathcal{G}^m$ of R^m) and let $\mathcal{A}_T$ denote the set of those $\gamma \in \mathcal{G}$ for which $\gamma(T) = 1$. The family $\mathcal{F}_0$ determines a mapping $\psi : \mathcal{A}_T \rightarrow \mathcal{F}$ by the relation

$$(1) \quad \psi(\gamma) = \int_T F(\cdot|\tau) d\gamma(\tau)$$

We speak of the d.f. $\psi(\gamma)$ as a mixture of $\mathcal{F}_0$ (w.r.t. γ). The mapping ψ is said to be identifiable if it is one to one. In certain connections (e.g. statistical estimation of γ) it is important to know whether ψ is identifiable. Various conditions for identifiability and nonidentifiability are known, see Teicher [4] and the references therein. Here we want to prove that, under mild restrictions, mixtures of exponential families $\mathcal{F}_0$ are identifiable. $\mathcal{F}_0$ is exponential (or of the Darrois-Koopman type) if for some σ -finite measure μ

$$(2) \quad dF(x|\tau) = a(\tau) b(x) e^{\sum_{j=1}^m \tau_j h_j(x)} d\mu(x)$$

for $x \in R^n$, $\tau = (\tau_1, \tau_2, \dots, \tau_m) \in T$, where $a(\tau) > 0$, $b(x) \geq 0$ and $a, b, h_j, j = 1, \dots, m$ are all measurable.

Let $\gamma_1, \gamma_2 \in \mathcal{A}_T$ and let

$$(3) \quad f_\nu(x) = \frac{d\gamma_\nu(\tau)}{d\mu} = b(x) \int_T a(\tau) e^{\sum_{j=1}^m \tau_j h_j(x)} d\gamma_\nu(\tau), \quad \nu = 1, 2.$$

Furthermore, let $\xi = \{x : f_1(x) = f_2(x) \neq 0\}$, let

$\eta = \{y = (h_1(x), \dots, h_m(x)) : x \in \xi\}$ and let

$$(4) \quad f_\nu^*(y) = \int_T a(\tau) e^{(\tau, y)} d\gamma_\nu(\tau), \quad \nu = 1, 2.$$

where (τ, y) denotes the inner product of $\tau \in T$ and $y \in R^m$. Then $f_1^*(y) = f_2^*(y)$ if $y \in \eta$; our aim is to show that under certain further restrictions this implies $\gamma_1 = \gamma_2$. Let $c(\eta)$ denote the convex hull of η . We shall distinguish between four cases.

(i) η is finite.

(ii) η is infinite, $c(\eta)$ is bounded and η does not have an accumulation point in the interior of $c(\eta)$.

(iii) As (ii) except that $c(\eta)$ is assumed unbounded.

(iv) η is infinite and η has an accumulation point in the interior of $c(\eta)$.

Case (i). The important example of this case is the binomial distribution. An analysis of the identifiability problem for that distribution can be found in [4].

Case (ii). From the viewpoint of statistics (ii) is the case of least interest. We have obtained no general results. The problem is essentially this: ($n = m = 1$). Let γ_1 and γ_2 be two p.m.'s on $(R, \mathcal{B})$ whose Laplace transforms $\varphi_1(z)$ and $\varphi_2(z)$ both exist in a strip $0 \leq \operatorname{Re} z \leq \rho$, $\rho > 0$. Let $\{x_n\}$ be a sequence of real numbers such that $0 < x_n \leq \rho$ for all n and $x_n \rightarrow 0$ as $n \rightarrow \infty$. Find conditions under which $\varphi_1(x_n) = \varphi_2(x_n)$ for all n implies $\varphi_1(it) = \varphi_2(it)$ for all real t (i.e., identity of the Fourier transforms of γ_1 and γ_2 and hence identity of γ_1 and γ_2).

Case (iii). We shall treat the subcase:

(iii)'. η contains the set I^+ of all lattice points in R^m with nonnegative components, i.e., $I^+ = \{k = (k_1, \dots, k_m) : k_j \text{ is a non-negative integer, } j = 1, \dots, m\}$.

We have, since $0 = (0, \dots, 0) \in I^+$

$$(5) \quad f_1^*(0) = \int_T a(\tau) d\gamma_1(\tau) = \int_T a(\tau) d\gamma_2(\tau) = f_2^*(0).$$

Let us denote the common (positive) value in (5) by c and let us introduce the p.m.'s γ_v^* , $v = 1, 2$, by $d\gamma_v^*(\tau) = c^{-1} a(\tau) d\gamma_v(\tau)$. Thus

$$(6) \quad f_1^*(k) = \int_T e^{(\tau, k)} d\gamma_1^*(\tau) = \int_T e^{(\tau, k)} d\gamma_2^*(\tau) = f_2^*(k) \quad \forall k \in I^+.$$

Let w be the transformation: $\tau \rightarrow \lambda = w(\tau)$ where $\lambda = (\lambda_1, \dots, \lambda_m) = (e^{\tau_1}, \dots, e^{\tau_m})$; let $\Lambda = w(T)$ and $\pi_v = \gamma_v^* w^{-1}$, $v = 1, 2$. We obtain from (6)

$$(7) \quad \mu_{k_1 \dots k_m} = \int_{\Lambda} \lambda_1^{k_1} \dots \lambda_m^{k_m} d\pi_1(\lambda) = \int_{\Lambda} \lambda_1^{k_1} \dots \lambda_m^{k_m} d\pi_2(\lambda)$$

$$\forall k = (k_1, \dots, k_m) \in I^+.$$

We can draw the following conclusion.

Proposition 1. Suppose that assumption (iii)' is satisfied and suppose that π_1 and π_2 are uniquely determined by their moments (7). Then $\pi_1 = \pi_2$ and consequently $\gamma_1 = \gamma_2$.

In order to derive a sufficient condition for $\gamma_1 = \gamma_2$ which is more useful than that of Proposition 1 we state the following lemma.

Lemma 1. Let π be an arbitrary p.m. on $(R^m, \mathcal{G}^m)$ with $\pi(R^{+m}) = 1$ where R^+ is the set of nonnegative reals and with all moments

$$(8) \quad \mu_{k_1 \dots k_m} = \int_{R^m} \lambda_1^{k_1} \dots \lambda_m^{k_m} d\pi(\lambda), \quad k \in I^+$$

finite. If there exists a positive number ρ such that the series

$$(9) \quad \sum_{k \in I^+} \mu_{k_1} \dots \mu_{k_m} \frac{\rho^{k_1 + \dots + k_m}}{k_1! \dots k_m!}$$

is convergent then π is the unique p.m. with these moments.

The lemma and its proof are straightforward generalizations of a result in the book of Cramer [2; 176].

Let us apply the lemma to (7). We find (dropping the subscript v)

$$\begin{aligned} 0 &\leq \sum_k \mu_{k_1} \dots \mu_{k_m} \frac{\rho^{k_1 + \dots + k_m}}{k_1! \dots k_m!} \\ &= \int_{\Lambda} \sum_k \prod_{j=1}^m \frac{(\lambda_j \rho)^{k_j}}{k_j!} d\pi \\ &= \int_{\Lambda} \prod_{j=1}^m \left(\sum_{l=1}^{\infty} \frac{(\lambda_j \rho)^l}{l!} \right) d\pi \\ &= \frac{1}{c} \int_T a(\tau) e^{\rho \sum e^{\tau_j}} d\gamma(\tau) \\ &\leq \frac{1}{c} \sup_{\tau \in T} a(\tau) e^{\rho \sum e^{\tau_j}}. \end{aligned}$$

Therefore

Proposition 2. Suppose that assumption (iii)' is satisfied and suppose that

$$(10) \quad \sup_{\tau \in T} a(\tau) e^{\rho \sum e^{\tau_j}} < \infty$$

for some $\rho > 0$. Then $\gamma_1 = \gamma_2$.

As an application, let us consider the instance where $n = m$, $h_j(x) = x_j$ (j -th coordinate of x ; $j = 1, 2, \dots, m$) and where the measure μ in (2) is concentrated on I^+ ; then without loss of generality we can and will assume μ to be counting measure on I^+ . Hence the family $\mathcal{F}_0$ is given by

$$(11) \quad F(x|\tau) = \begin{cases} \sum_{k=0}^{[x]} a(\tau) b(k) e^{(\tau, k)} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

in an obvious notation. Assumption (iii)' becomes: $b(k) > 0 \forall k \in I^+$ and we have

Corollary 1. If the family $\mathcal{F}_0$ given by (11) satisfies $b(k) > 0 \forall k \in I^+$ and

$$(12) \quad \sup_{\tau \in T} a(\tau) e^{\rho \sum e^{\tau_j}} < \infty$$

for some $\rho > 0$ then ψ is identifiable.

Specializing still further we obtain (Feller [3])

Corollary 2. The mapping ψ determined by the Poisson family

$\mathcal{F}_0 = \{F(\cdot|\tau) : -\infty < \tau < \infty\}$, where

$$F(x|\tau) = \sum_{k=0}^{[x]} e^{-\lambda} \frac{\lambda^k}{k!}, \quad x \geq 0, \quad \lambda = e^\tau$$

is identifiable.

Case (iv). We shall prove that $\gamma_1 = \gamma_2$ provided

(iv)'. There exists an accumulation point $y^{(0)} = (y_1^{(0)}, \dots, y_m^{(0)})$ of η in the interior of $c(\eta)$ with the following property. If two arbitrary complex power series

$$\sum a_{j_1 j_2 \dots j_m}^{(\nu)} (z_1 - y_1^{(0)})^{j_1} (z_2 - y_2^{(0)})^{j_2} \dots (z_m - y_m^{(0)})^{j_m},$$

$\nu = 1, 2$

coincide for all $z = (z_1, \dots, z_m) \in \eta \cap V$ for some neighborhood V of $y^{(0)}$, then they have identical coefficients.

We note that assumption (iv)' is equal to (iv) if $m = 1$. A sufficient condition for (iv)' is that η be dense in some open subset of R^m .

Proposition 3. Suppose that assumption (iv)' is satisfied. Then $\gamma_1 = \gamma_2$.

Proof. Without loss of generality we can and will assume that the origin 0 is in η and that there is a neighborhood

$K = \{y : |y_j| < \rho, j = 1, \dots, m\}$ of 0 for which $K \subset c(\eta)$ and K contains $y^{(0)}$. Then

$$(13) \quad f_1^*(0) = \int_T a(\tau) d\gamma_1(\tau) = \int_T a(\tau) d\gamma_2(\tau) = f_2^*(0) .$$

Let us denote the common value in (13) by c and let us define the p.m.s. γ_v^* , $v = 1, 2$ by $d\gamma_v^*(\tau) = \frac{1}{c} a(\tau) d\gamma_v(\tau)$. Furthermore, let ϕ_v , $v = 1, 2$ denote the Laplace transform of γ_v

$$\phi_v(z) = \int_T e^{(\tau, z)} d\gamma_v^*(\tau)$$

where $z = (z_1, \dots, z_m)$, $z_j = u_j + iv_j$ ($j = 1, \dots, m$). ϕ_v exists for all $z \in K' = \{z | u = (u_1, \dots, u_m) \in K\}$. In fact, for any such z , $|\exp((\tau, z))| \leq \exp((\tau, u))$ and a moments reflection shows that there exists a $y \in \eta$ with $(\tau, u) \leq (\tau, y)$; thus

$$\int_T |e^{(\tau, z)}| d\gamma_v^*(\tau) \leq \frac{1}{c} \int_T a(\tau) e^{(\tau, y)} d\gamma_v(\tau) < \infty .$$

More is true: ϕ_v is an analytic function of $z = (z_1, \dots, z_m)$ in the domain K' . To prove this it suffices to show that ϕ_v is analytic in each of the variables z_j , $j = 1, \dots, m$ (see [1]). Hence let us consider

$$(14) \quad \frac{\varphi_\nu(z + he_j) - \varphi_\nu(z)}{h} = \int_T e^{(\tau, z)} \frac{e^{\tau_j h} - 1}{h} d\gamma_\nu^*(\tau)$$

where $z = u + iv \in K'$, e_j denotes the j -th unit vector in R^m and h is an arbitrary complex number. Let $\delta > 0$ be so small that $z + he_j \in K'$ for all h such that $|h| \leq \delta$. Using the (well known) inequality

$$\left| \frac{e^{\tau_j h} - 1}{h} \right| \leq \frac{e^{|\tau_j| \delta}}{\delta} \quad \text{for } |h| \leq \delta$$

we find that the integrand in (14) is dominated by

$$\frac{1}{\delta} \left(e^{(\tau, u + \delta e_j)} + e^{(\tau, u - \delta e_j)} \right)$$

and since the integral of this quantity is finite we may pass to the limit $h \rightarrow 0$ under the integration sign in (14) to obtain

$$\frac{\varphi(z + he_j) - \varphi(z)}{h} \rightarrow \int_T \tau_j e^{(\tau, z)} d\gamma_\nu^*(\tau) \quad \text{as } h \rightarrow 0.$$

We have thus shown that φ_ν is analytic in K' . Consequently φ_ν can be expanded in a power series around $z^{(0)} = y^{(0)}$

$$\varphi_\nu(z) = \sum a_{j_1 j_2 \dots j_m}^{(\nu)} (z_1 - y_1^{(0)})^{j_1} (z_2 - y_2^{(0)})^{j_2} \dots (z_m - y_m^{(0)})^{j_m}$$

the expansion being valid in some neighborhood V of $y^{(0)}$. We have $\varphi_1(z) = \varphi_2(z) \forall z \in \eta$ and hence, by assumption (iv)' and uniqueness of analytic continuation, $\varphi_1(z) = \varphi_2(z) \forall z \in K'$. In particular $\varphi_1(z) = \varphi_2(z)$ for all purely imaginary $z = iv = (iv_1, \dots, iv_m)$, i.e., the characteristic functions of γ_1^* and γ_2^* coincide, hence $\gamma_1^* = \gamma_2^*$ or, equivalently, $\gamma_1 = \gamma_2$. q.e.d.

By the remark preceding Proposition 3, we obtain

Corollary 3. Suppose that in the representation (2): (a) μ is n -dimensional Lebesgue measure, (b) the functions h_j , $j = 1, \dots, m$ are all continuous, (c) the set $\{y : y = (h_1(x), \dots, h_m(x)), b(x) > 0, x \in \mathbb{R}^n\}$ contains a (nonempty) open set. Then ψ is identifiable.

Specializing still further we get

Corollary 4. Suppose that $\mathcal{F}_0$ is the Gaussian family

$$\mathcal{F}_0 = \{F(\cdot | \tau) | \tau = (\tau_1, \tau_2), -\infty < \tau_1 < \infty, 0 < \tau_2 < \infty\},$$

$$(15) \quad \frac{dF(x_1, \dots, x_n | \tau_1, \tau_2)}{d\mu} = (2\pi \sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{\ell=1}^n (x_\ell - \frac{1}{2})^2\right) \\ = \left(\frac{\tau_2}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{n}{2} \frac{\tau_1^2}{\tau_2}\right) e^{h_1(x)\tau_1 + h_2(x)\tau_2}$$

where μ is n -dimensional Lebesgue measure, $\tau_1 = \frac{1}{2} \sigma^{-2}$, $\tau_2 = \sigma^{-2}$, $h_1(x) = \sum x_\ell$ and $h_2(x) = -\frac{1}{2} \sum x_\ell^2$. If $n > 1$, then ψ is identifiable (Teicher has shown, see [5], that ψ is not identifiable if $n = 1$).

REFERENCES

- [1] Bochner, S. and Martin, W.T., "Several complex variables," Princeton University Press (1948).
- [2] Cramér, H., "Mathematical methods of statistics," Princeton University Press (1957).
- [3] Feller, W., On a general class of contagious distributions, Ann. Math. Stat. 14 (1943) 389-399.
- [4] Teicher, H., Identifiability of finite mixtures, Ann. Math. Stat. 34 (1963) 1265-1269.
- [5] Teicher, H., On the mixture of distributions, Ann. Math. Stat. 31 (1960) 55-73.