Z. Wahrscheinlichkeitstheorie verw. Geb. 12, 56–58 (1969)

Lévy Homeomorphic Parametrization and Exponential Families

Ole Barndorff-Nielsen

Summary. Let $\mathscr{P}=\{P_\theta\colon \theta\in\Theta\}$ be an exponential family of probability distributions with θ the canonical parameter and consider the one to one mapping $\varphi\colon P_\theta\to\theta$. It is shown that, under mild regularity assumptions, φ and φ^{-1} are continuous with respect to the Lévy metric in $\mathscr{P}$ and Euclidean metric in Θ.

Let $(\mathscr{X},\mathscr{B},\mathscr{P})$ be a Euclidean statistical field, i.e. $\mathscr{X}$ is a measurable subset of some Euclidean space, $\mathscr{B}$ is the σ-algebra of its Borel subsets and $\mathscr{P}$ is a family of probability distributions on $\mathscr{B}$. $\mathscr{P}$ is said to be *parametrized* by φ if φ is a one to one mapping from $\mathscr{P}$ onto a subset Θ of a Euclidean space. Such a mapping φ will be called *Lévy homeomorphic* if it is a homeomorphism with respect to the Lévy metric in $\mathscr{P}$ and the usual, Euclidean metric in Θ.

$\mathscr{P}$ is *exponential* if it is dominated by a σ-finite measure μ and if there exist a positive integer k, real functions $a|\mathscr{P},\alpha_1|\mathscr{P},\dots,\alpha_k|\mathscr{P}$ and real, measurable functions $b|\mathscr{X}, T_1|\mathscr{X},\dots,T_k|\mathscr{X}$ such that μ-a.e.

$$\frac{dP}{d\mu}(x)=a(P)\,b(x)\,e^{\alpha(P)\cdot T(x)},\qquad P\in\mathscr{P}, \tag{1}$$

where $\alpha=(\alpha_1,\dots,\alpha_k)$, $T=(T_1,\dots,T_k)$ and where $\cdot$ denotes inner product. The smallest k for which there exists functions a,α,b and T so that $dP/d\mu$ can be written in the form (1), is the *order* of $\mathscr{P}$; the order does not depend on which dominating measure μ is considered. In an exponential family the probability measures are mutually absolutely continuous.

It is assumed in the sequel that $\mathscr{P}$ is exponential and that k is equal to the order of $\mathscr{P}$. Under this assumption the functions $1, T_1,\dots,T_k$ are linearly independent in the sense that if for some constants $\gamma_0,\gamma_1,\dots,\gamma_k$ and some (and hence all) $P\in\mathscr{P}$ there holds

$$P\{\gamma_0+\gamma_1 T_1+\cdots+\gamma_k T_k=0\}=1$$

then $\gamma_0=\gamma_1=\cdots=\gamma_k=0$.

Consider the mapping

$$\varphi\colon\ P\to\alpha(P),\qquad P\in\mathscr{P},$$

and let

$$\Theta=\{\alpha(P)\colon\ P\in\mathscr{P}\}\ (\subset R^k).$$

φ is one to one and hence $\mathscr{P}$ is parametrized by φ.

The purpose of this note is to prove the following

Proposition[1]. *If Θ is open and T is continuous, then φ is Lévy homeomorphic.*

Remark. As is apparent from the following proof, φ is continuous whether Θ is open or not.

Proof. Let $P_\theta = \varphi^{-1}(\theta)$ and write $a(\theta)$ for $a(P_\theta)$. Then, $\mu-$a.e.

$$\frac{dP_\theta}{d\mu}(x) = a(\theta)\, b(x)\, e^{\theta \cdot T(x)}, \qquad \theta \in \Theta,$$

where

$$a^{-1}(\theta) = \int e^{\theta \cdot T(x)}\, b(x)\, d\mu.$$

a is a continuous function of θ because the multidimensional Laplace transform is continuous at all inner points of its domain of convergence.

Thus $\theta_n \to \theta$ $(\theta, \theta_n \in \Theta)$ entails

$$\frac{dP_{\theta_n}}{d\mu} \to \frac{dP_\theta}{d\mu} \quad \mu-\text{a.e.}$$

and this implies, by a wellknown corollary to the dominated convergence theorem,

$$\frac{dP_{\theta_n}}{d\mu} \to \frac{dP_\theta}{d\mu} \qquad \text{in mean w.r.t. } \mu.$$

Consequently $P_{\theta_n} \to P_\theta$ (weak convergence) and the continuity of φ^{-1} is established.

To prove continuity of φ it must be shown that if $P_{\theta_n} \to P_\theta$, then $\theta_n \to \theta$. Let $C_0(R^k)$ denote the space of continuous functions on R^k with compact support and let $T \circ P$ $(P \in \mathscr{P})$ be the measure on R^k induced from $(\mathscr{X}, \mathscr{B}, P)$ by the mapping T. From $P_{\theta_n} \to P_\theta$ and the continuity of T it follows that $T \circ P_{\theta_n} \to T \circ P_\theta$, i.e.

$$\int f(t)\, dT \circ P_{\theta_n} \to \int f(t)\, dT \circ P_\theta, \qquad f \in C_0(R^k).$$

For shortness, let $\pi = T \circ P_\theta$. Now,

$$\int f(t)\, dT \circ P_{\theta_n} = \int f(t)\, \frac{a(\theta_n)}{a(\theta)}\, e^{(\theta_n - \theta) \cdot t}\, d\pi$$

and thus it suffices to prove that if $0 \in \Theta$ and if for some sequence a_n

$$a_n \int f(t)\, e^{\theta_n \cdot t}\, d\pi \to \int f(t)\, d\pi, \qquad f \in C_0(R^k), \tag{2}$$

then $\theta_n \to 0$.

On account of the compactness of the surface of the unit sphere in R^k, one can, without loss of generality, assume

$$\frac{\theta_n}{\|\theta_n\|} \to e, \qquad \text{where } \|e\| = 1.$$

Let

$$F(c) = \pi\{t \in R^k \colon\ t \cdot e \leqq c\}, \qquad c \in R.$$

F is a distribution function with at least two points of increase, c_1 and $c_2 > c_1$. In fact, if F had only one point of increase c, then

$$1 = \pi\{t \cdot e = c\} = P_\theta\{T \cdot e = c\}$$

1. This proposition gives an answer to the question raised by Nölle [1; p. 77].

in contradiction to the linear independence of 1, $T_1, \ldots, T_k$. Choose a $\delta \in \left(0, \frac{c_2 - c_1}{5}\right)$ and let $J_1 = \{t\colon\ c_1 - \delta < t \cdot e < c_1 + \delta\}$. Furthermore, let f be a nonnegative function in $C_0(R^k)$ which satisfies

$$f(t) = 0, \qquad t \notin J_1$$

and

$$\int_{J_1} f(t)\, d\pi > 0.$$

Such f certainly exists. Let S_f denote the (compact) support of f. For n sufficiently large one has

$$\theta_n \cdot t = \|\theta_n\| \left[\left(\frac{\theta_n}{\|\theta_n\|} - e\right) \cdot t + e \cdot t\right] \leqq \|\theta_n\| (c_1 + 2\delta), \qquad t \in S_f$$

and hence

$$a_n \int_{R^k} f(t)\, e^{\theta_n \cdot t}\, d\pi \leqq a_n \int_{S_f} f(t)\, e^{\|\theta_n\|\,(c_1 + 2\delta)}\, d\pi = a_n\, e^{\|\theta_n\|\,(c_1 + 2\delta)} \int_{R^k} f(t)\, d\pi.$$

By letting $n \to \infty$ and employing (2), one obtains

$$1 \leqq \liminf a_n\, e^{\|\theta_n\|\,(c_1 + 2\delta)}. \tag{3}$$

In a similar way it may be shown that

$$1 \geqq \limsup a_n\, e^{\|\theta_n\|\,(c_2 - 2\delta)}. \tag{4}$$

(3), (4) and the inequality $c_1 + 2\delta < c_2 - 2\delta$ together imply $\theta_n \to 0$.

Acknowledgement. I am indebted to Karl Pedersen for a useful remark which led to a shortening of my original proof of the continuity of φ.

Reference

1. Nölle, G.: Zur Theorie der bedingten Tests. Inaugural-Dissertation an der Universität Münster (1966).

Ole Barndorff-Nielsen
Matematisk Institut
Aarhus Universitet
Ny Munkegade
8000 Aarhus C, Danmark

(Received February 15, 1968)