NON-EXISTENCE OF EVERYWHERE PROPER CONDITIONAL DISTRIBUTIONS[1]

By D. Blackwell and C. Ryll-Nardzewski

University of California, Berkeley

1. Introduction and summary. Let Ω be a Borel subset of a complete separable metric space, and denote by $\mathcal{B}$ the class of Borel subsets of Ω. For any probability measure P on $\mathcal{B}$ and any real-valued random variable f on Ω a *conditional distribution given* f is a real-valued function Q on $\Omega \times \mathcal{B}$ such that

(1) for each $\omega \varepsilon \Omega$, $Q(\omega, \cdot)$ is a probability measure on $\mathcal{B}$,

(2) for each $B \varepsilon \mathcal{B}$, $Q(\cdot, B)$ is an $\mathcal{A}$-measurable function on Ω, where $\mathcal{A}$ is the Borel field of f-sets, i.e., sets of the form $\{\omega : f(\omega) \varepsilon F\}$, where F is a linear Borel set, and

(3) for every $A \epsilon \mathcal{A}$, $B \epsilon \mathcal{B}$,

$$\int_A Q(\omega, B)\, dP(\omega) = P(A \cap B).$$

A conditional distribution Q will be called *proper* at ω_0 if

$$Q(\omega_0, A) = 1 \quad \text{for} \quad \omega_0 \varepsilon A \varepsilon \mathcal{A},$$

i.e., if, given that f has the value $f(\omega_0)$, we assign conditional probability 1 to the set of ω's at which f has the specified value. It is known [2] that conditional distributions always exist that are proper at almost all points of Ω, i.e., except at a set of points N with $P(N) = 0$. We shall show that, in general, the exceptional set N cannot be removed.

More precisely, we shall prove

Theorem 1. *Let Ω, $\mathcal{B}$, f, $\mathcal{A}$ be as above. A function Q with properties* (1), (2) *and*

$$(4) \qquad Q(\omega, A) = 1 \quad \text{for} \quad \omega \varepsilon A \varepsilon \mathcal{A}$$

exists if and only if there is an $\mathcal{A}$-measurable function g from Ω into Ω such that

$$(5) \qquad f(g(\omega)) = f(\omega) \quad \text{for all } \omega.$$

The existence of such a g implies that the range of f is a Borel set.

It follows from Theorem 1 that, whenever the range of f is not a Borel set, everywhere proper conditional distributions given f cannot exist.

The only difficult part of Theorem 1 will be a consequence of

Theorem 2. *Let X, Y be Borel subsets of complete separable metric spaces, let $\mathcal{A}$ be a countably-generated subfield of the field of Borel subsets of X and let $\mathcal{B}$ be the class of Borel subsets of Y. For any function μ on $X \times \mathcal{B}$ such that* (a) *$\mu(x, \cdot)$ is*

Received July 31, 1962.

[1] Prepared with the partial support of the National Science Foundation, Grants G-18792 and G-19673.

for each x a probability measure on $\mathcal{B}$ and (b) *for each $B \varepsilon \mathcal{B}$, $\mu(\cdot, B)$ is an $\mathcal{A}$-measurable function on X, and any set $S \varepsilon \mathcal{A} \times \mathcal{B}$ such that*

$$\mu(x, S_x) > 0 \quad \text{for all } x \varepsilon X,$$

where S_x denotes the x-section of S, i.e., $S_x = \{y: (x, y) \varepsilon S\}$, there is an $\mathcal{A}$-measurable function g from X into Y whose graph is a subset of S, i.e., $(x, g(x)) \varepsilon S$ for all $x \varepsilon X$.

The weaker hypothesis that S is $\mathcal{A} \times \mathcal{B}$ measurable with every x-section S_x non-empty is not sufficient to guarantee that S contains the graph of an $\mathcal{A}$-measurable function, as an example of Novikoff [3] shows.

2. Proofs. We first note how Theorem 1 follows from Theorem 2. If Ω, $\mathcal{B}$, f, $\mathcal{A}$ are as in Theorem 1, and Q has properties (1), (2), (4), we apply Theorem 2 with $X = Y = \Omega$, $\mu = Q$, and S the set of all pairs (x, y) for which $f(x) = f(y)$. Then $\mu(x, S_x) = Q(x, A(x)) = 1$, where $A(x) = \{\omega: f(\omega) = f(x)\}$, so that, from Theorem 2 there is an $\mathcal{A}$-measurable g from Ω into Ω such that $(\omega, g(\omega)) \varepsilon S$ for all ω, i.e.,

$$f(g(\omega)) = f(\omega) \quad \text{for all } \omega. \tag{5}$$

Conversely, for any $\mathcal{A}$-measurable g satisfying (5), we define

$$\begin{aligned} Q(\omega, B) &= 1 \quad \text{if} \quad g(\omega) \varepsilon B \\ &= 0 \quad \text{if} \quad g(\omega) \not\varepsilon B, \end{aligned}$$

and verify easily that Q satisfies (1), (2), and (4).

Moreover, since g is $\mathcal{A}$-measurable, there is [1] a Borel measurable function h from the real line into Ω such that

$$g(\omega) = h(f(\omega)) \quad \text{for all } \omega. \tag{6}$$

(5) and (6) together imply that the range of f is $\{y: f(h(y)) = y\}$, which is clearly a Borel set.

For the proof of Theorem 2, we need the

Lemma. *If X, Y, $\mathcal{A}$, $\mathcal{B}$ are as in Theorem 2 and μ is a function on $X \times \mathcal{B}$ such that* (a) *$\mu(x, \cdot)$ is for each x a nonnegative, finite measure on $\mathcal{B}$ and* (b) *$\mu(\cdot, B)$ is $\mathcal{A}$-measurable in x for each $B \varepsilon \mathcal{B}$, then for every $\mathcal{A} \times \mathcal{B}$-measurable subset S of $X \times Y$ and every θ, $0 \leqq \theta < 1$, there is an $\mathcal{A} \times \mathcal{B}$ set $\tilde{S} \subset S$ such that $\tilde{S}$ has closed x-sections, and $\mu(x, \tilde{S}_x) \geqq \theta\mu(x, S_x)$ for all x.*

Proof. If $S = A \times B$ where $A \varepsilon \mathcal{A}$ and B is closed, we may choose $\tilde{S} = S$. The class of sets S for which the lemma holds is clearly closed under finite union. We must show that the class of S for which the lemma holds is closed under monotone union and monotone intersection. If S_n is increasing and the lemma holds for each S_n, choose $\tilde{S}_n \subset S_n$ such that $\tilde{S}_n$ has closed x-sections and

$$\mu(x, \tilde{S}_{nx}) \geqq \theta^{\frac{1}{2}}\mu(x, S_{nx}) \quad \textit{for all } x.$$

We may suppose, replacing $\tilde{S}_n$ by $\tilde{S}_1 \cup \cdots \cup \tilde{S}_n$, that $\tilde{S}_n \subset \tilde{S}_{n+1}$. Define $T_n =$

$\{x: \mu(x, \tilde{S}_{nx}) \geqq \theta\mu(x, S_x)\}$, where $S = \bigcup S_n$. The T_n are monotone increasing and $\bigcup T_n = X$, so that the set $\tilde{S}$ whose x-section is $\tilde{S}_{nx}$ for $x \in T_n - T_{n-1}$ has the required properties.

If S_n is decreasing and the lemma holds for each S_n, let $\bigcap S_n = S$, and define λ on $X \times \mathcal{B}$ by

$$\lambda(x, B) = \mu(x, S_x \cap B)/\mu(x, S_x) \quad \text{if} \quad \mu(x, S_x) > 0,$$

$$\lambda(x, B) = 0 \quad \text{if} \quad \mu(x, S_x) = 0.$$

Choose $\tilde{S}_n \subset S_n$, with closed x-sections, such that $\lambda(x, \tilde{S}_{nx}) \geqq \theta_n\lambda(x, S_{nx})$ for all x, n, where $\theta_n = 1 - (1 - \theta)/2^n$. We assert that $\tilde{S} = \bigcap_n \tilde{S}_n$ has the required properties. Its x-sections are clearly closed, and $\tilde{S} \subset S$. If $\mu(x, S_x) > 0$, then $\lambda(x, S_{nx}) = 1$ and $\lambda(x, \tilde{S}_{nx}) \geqq \theta_n$. Then

$$\lambda(x, \tilde{S}_x) \geqq 1 - \sum_n (1 - \theta_n) = \theta$$

i.e., $\mu(x, \tilde{S}_x) \geqq \theta\mu(x, S_x)$. If $\mu(x, S_x) = 0$, this inequality is trivially true, and the lemma is proved.

We turn to the proof of Theorem 2. Applying the lemma with Y replaced by its completion, we obtain a set $S_1 \subset S$ with closed x-sections and $\mu(x, S_{1x}) > 0$ for all x. For any $\epsilon > 0$, we cover Y with a sequence $F_1, F_2, \cdots$ of closed sets, each of diameter $< \epsilon$, define $n(x)$ as the smallest integer k for which $\mu(x, S_{1x} \cap F_k) > 0$, and denote by S_2 the set whose x-section is $S_{1x} \cap F_k$ for $n(x) = k$ and $\epsilon = 1$. Applying the same construction to S_2 with $\epsilon = \frac{1}{2}$ yields $S_3 \subset S_2$, etc. We obtain a sequence of sets $S \supset S_1 \supset S_2, \cdots$, with $\mu(x, S_{nx}) > 0$, and S_{nx} closed of diameter $< 1/n - 1$. The set $S^* = \bigcap S_n$ is then $\mathcal{A} \times \mathcal{B}$-measurable, and each S_x^* contains exactly one point, so that S^* is the graph of a function g. According to a theorem of Sierpinski [4], any function whose graph is a Borel set is Borel measurable, so that g is Borel measurable. Finally, $\mathcal{A} \times \mathcal{B}$-measurability of S^* implies that, for any Borel measurable function h on Ω such that $\mathcal{A}$ is the field of h-sets, the value of g is determined by that of h, and this, with Borel measurability of g, implies $\mathcal{A}$-measurability of g. This completes the proof.

REFERENCES

[1] Blackwell D. (1955). On a class of probability spaces. *Proc. Third Berkeley Symp. Math. Statist. Prob.* II 1–6.
[2] Loève M. (1960). *Probability Theory.* Van Nostrand.
[3] Nivikoff P. (1931). Sur les fonctions implicites measurables B. *Fund. Math.* **17** 8–25.
[4] Sierpinski W. (1921). Sur les images des fonctions representables analytiquement. *Fund. Math.* **2** 74–80.