ON THE ASYMPTOTIC BEHAVIOR OF BAYES ESTIMATES IN THE DISCRETE CASE II[1]

By David A. Freedman

University of California, Berkeley

1. Introduction. This article proves: when sampling from a countably infinite population with unknown distribution, for all but a set of priors of the first category, Bayes estimates are consistent only at a set of distributions of the first category. That is, for essentially all priors the Bayes estimates are consistent essentially nowhere. This article is independent of (Freedman, 1963) although it uses the same definitions, essentially the same notation, and improves Remark 6, p. 1400.

2. Results. I, the population, is a countably infinite set with the discrete topology. Λ, the set of parameters, is the set of probabilities on I, with the weak* topology: $\lambda_n \to \lambda$ if and only if $\lambda_n(i) \to \lambda(i)$ for all $i \varepsilon I$. Finally, $\pi(\Lambda)$, the set of priors, is the set of probabilities on (the Borel subsets of) Λ, with the weak* topology: $\mu_n \to \mu$ if and only if $\int_\Lambda f\, d\mu_n \to \int_\Lambda f\, d\mu$ for all bounded, continuous, real functions f on Λ. It is well known that f can be restricted to polynomials in $\lambda \to \lambda(i)$: $i \varepsilon I$ without decreasing the topology. If $\lambda \varepsilon \Lambda$, then $\delta_\lambda \varepsilon \pi(\Lambda)$ puts mass 1 at λ. The spaces Λ and $\pi(\Lambda)$ are complete, separable, metric.

Consider the coordinate process $\{\xi_n : 1 \leqq n < \infty\}$ on the sample space I^∞ of I-sequences with the product topology. If $\lambda \varepsilon \Lambda$, then λ^∞ is the probability on (the Borel subsets of) I^∞ which makes the ξ_n independent with common distribution λ. If $\mu \varepsilon \pi(\Lambda)$, the *posterior* μ_n is this map from part of I^∞ to $\pi(\Lambda)$:

$$\mu_n(d\lambda) = [\prod_{i=1}^n \lambda(\xi_i)\mu(d\lambda)]/[\int_\Lambda \prod_{i=1}^n \lambda(\xi_i)\mu(d\lambda)],$$

defined when the denominator is positive. The pair (λ, μ) with $\lambda \varepsilon \Lambda$ and $\mu \varepsilon \pi(\Lambda)$ is *consistent* if $\lim_{n\to\infty} \mu_n = \delta_\lambda$ with λ^∞-probability 1.

A subset of a complete metric space is *first category* if it is a countable union of nowhere dense sets; a subset is *residual* if its complement is first category. For a discussion, see (Kuratowski, 1958, especially Sections 8–11). Topologically, first category sets are small, residual sets large.

The main result of this note is:

Theorem. *The set C of consistent pairs (λ, μ) is first category in the cartesian product of Λ and $\pi(\Lambda)$.*

Corollary. *There is a residual set of priors $\mu \varepsilon \pi(\Lambda)$ with the property*: $\{\lambda: \lambda \varepsilon \Lambda, (\lambda, \mu)$ *is consistent*$\}$ *is first category in* Λ.

It is convenient to prove a result stronger than the theorem:

Proposition. *The set R of pairs (λ, μ) with* $\limsup_{n\to\infty} \int_{I^\infty} \mu_n(U)\, d\lambda^\infty = 1$ *simultaneously for all nonempty open subsets U of Λ is residual.*

Received 15 October 1964.

[1] Prepared with the partial support of the National Science Foundation, Grant GP-2593.

Using the technique of Section 3, it is easy to prove that for essentially any pair of Bayesians, each thinks the other is crazy. Formally, for a residual set of pairs (μ, ν) in $\pi(\Lambda) \times \pi(\Lambda)$,

$$\limsup_{n\to\infty} \int_\Lambda \int_{I^\infty} \mu_n(U) d\lambda^\infty \, d\nu = 1$$

and

$$\limsup_{n\to\infty} \int_\Lambda \int_{I^\infty} \nu_n(U) d\lambda^\infty \, d\mu = 1,$$

simultaneously for all nonempty open subsets U of Λ.

REMARK 1. The set Λ_0 of $\lambda \varepsilon \Lambda$ with at least one coordinate zero is a dense, frontier F_σ, so first category.

REMARK 2. The set of $\mu \varepsilon \pi(\Lambda)$ with these properties is a dense, frontier G_δ, so residual: (i) $\mu\{\lambda\} = 0$ for all $\lambda \varepsilon \Lambda$; (ii) $\mu(\Lambda_0) = 0$; (iii) μ assigns positive mass to each nonempty open set.

For a discussion, see (Dubins and Freedman, 1964, especially 3.12–3.18) and (Kuratowski, 1958).

3. Proofs. To prove the proposition, let D be a countable, dense subset of $\Lambda - \Lambda_0$. If $v \varepsilon D$, then V is the set of $\mu \varepsilon \pi(\Lambda)$ assigning positive mass to v and concentrating on a finite number of points, all of which but v lie in Λ_0. Then

$$(1) \qquad V \text{ is dense in } \pi(\Lambda)$$

and

$$(2) \qquad \lambda^\infty[\lim_{n\to\infty} \mu_n = \delta_v] = 1, \quad \text{for} \quad \lambda \varepsilon \Lambda - \Lambda_0 \quad \text{and} \quad \mu \varepsilon V.$$

Let $\pi_+(\Lambda)$ be the set of $\mu \varepsilon \pi(\Lambda)$ with $\mu(\Lambda_0) < 1$. Then $\pi_+(\Lambda)$ is residual by Remark 2 (in fact, it is a dense G_δ), and μ_n is everywhere defined for $\mu \varepsilon \pi_+(\Lambda)$. Let $V_k : 1 \leqq k < \infty$ be a sequence of open subsets of Λ, with $V_k \supset$ closure V_{k+1}, and $\bigcap_k V_k = v$. Let v_k be a continuous function from Λ to [0, 1], equal to 1 on V_{k+1}, equal to 0 off V_k. From (2)

$$(3) \qquad \lim_{n\to\infty} \int_{I^\infty} \int_\Lambda v_k d\mu_n \, d\lambda^\infty = 1 \quad \text{for} \quad \lambda \varepsilon \Lambda - \Lambda_0 \quad \text{and} \quad \mu \varepsilon V,$$

and plainly

$$(4) \qquad (\lambda, \mu) \to \int_{I^\infty} \int_\Lambda v_k d\mu_n \, d\lambda^\infty \quad \text{is continuous on} \quad \Lambda \times \pi_+(\Lambda).$$

Then

$$[\Lambda \times \pi_+(\Lambda)] - R \subset \bigcup_{v\varepsilon D} \bigcup_{k=1}^\infty \bigcup_{j=1}^\infty \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty R_{vkjn}$$

where

$$R_{vkjn} = \{(\lambda, \mu) : \lambda \varepsilon \Lambda, \mu \varepsilon \pi_+(\Lambda), \int_{I^\infty} \int_\Lambda v_k d\mu_n \, d\lambda^\infty \leqq 1 - j^{-1}\}$$

is closed in $\Lambda \times \pi_+(\Lambda)$ by (4). Therefore, $\bigcap_{n=m}^\infty R_{vkjn}$ is closed in $\Lambda \times \pi_+(\Lambda)$, and has no interior in $\Lambda \times \pi_+(\Lambda)$ by Remark 1, Relations (1) and (3). Hence

$[\Lambda \times \pi_+(\Lambda)] - R$ is first category in $\Lambda \times \pi_+(\Lambda)$, from which the proposition follows easily.

The corollary is an immediate consequence of the "Fubini theorem" for category (Kurakowski, 1958, Section 24-vi), or (Oxtoby, 1960).

REFERENCES

DUBINS, LESTER E. and FREEDMAN, DAVID A. (1964). Measurable sets of measures. To appear in *Pac. J. Math.*

FREEDMAN, DAVID A. (1963). On the asymptotic behavior of Bayes estimates in the discrete case. *Ann. Math. Statist.* **34** 1386–1403.

KURATOWSKI, CASIMIR (1958). *Topologie* I, Hafner, New York.

OXTOBY, JOHN C. (1960). Cartesian products of Baire spaces. *Fund. Math.* **49** 157–166.