The Annals of Probability
1991, Vol. 19, No. 1, 396–400

SOME ORTHOGONALITY PRESERVING KERNELS WHICH ARE NOT COMPLETELY ORTHOGONAL

BY R. DANIEL MAULDIN[1] AND H. V. WEIZSÄCKER

University of North Texas and Universität Kaiserslautern

It is shown that a perturbed Bernoulli probability transition kernel yields an explicit example of an orthogonality preserving kernel which is not completely orthogonal. In statistical language, such a kernel defines models P_θ, $\theta \in [0,1]$, in which there is no estimate that estimates θ perfectly for all θ, but there is, for any given prior distribution on θ and hypothesis $H_0 \subset [0,1]$, a perfect test for H_0 against its complement $[0,1] \setminus H_0$. It is also demonstrated with an analysis and an application of sets and maps with the Baire property that there are continuum many nonisomorphic atomless orthogonality preserving transition kernels which are not completely orthogonal. Our methods may be regarded as refinements of those used by Blackwell.

The purposes of this note are to construct an explicit example of an orthogonality preserving kernel which is not completely orthogonal and to show that there are continuum many nonisomorphic atomless orthogonality preserving transition kernels which are not completely orthogonal. In statistical language such a kernel defines models P_θ, $\theta \in [0,1]$, in which there is no estimate that estimates θ perfectly for all θ, but there is, for any given prior distribution on θ and hypothesis $H_0 \subset [0,1]$, a perfect test for H_0 against its complement $[0,1] \setminus H_0$. Since the main technical problem is of a measure theoretic nature, we use measure theoretic rather than statistical terminology. Our methods may be regarded as refinements of those used by Blackwell (1980).

Let X and Y be Polish spaces. Recall that a transition kernel $x \to \mu_x \in \mathrm{Prob}(Y)$ is said to be completely orthogonal provided there is a Borel measurable map (or perfect statistic) $f\colon Y \to X$ such that for each x, $\mu_x(f^{-1}(x)) = 1$. A transition kernel $x \to \mu_x$ is orthogonality preserving means the mixture transformation, T, takes orthogonal measures on X to orthogonal measures on Y, where if ν is a measure on X and E is a Borel subset of Y, then

$$T(\nu)(E) = \int_X \mu_x(E)\, d\nu(x).$$

If Z and W are Polish spaces, two transition kernels $x \to \mu_x \in \mathrm{Prob}(Y)$ and $z \to \nu_z \in \mathrm{Prob}(W)$ are said to be isomorphic provided there are Borel isomorphisms $\varphi\colon X \to Z$ and $\psi\colon Y \to W$ such that $\mu_x(E) = \nu_{\varphi(x)}(\psi(E))$. Mauldin, Preiss and Weizsäcker (1983) showed that there is exactly one isomorphism

Received May 1989; revised September 1989.
[1]Research supported by NSF Grant DMS-88-03361.
AMS 1980 *subject classifications*. Primary 60A15, 62A15.
Key words and phrases. Orthogonal transition kernel, Bernoulli kernel, perfect test, Baire property.

class of atomless completely orthogonal kernels over uncountable Polish spaces. They also demonstrated that there exists an orthogonality preserving kernel which is not completely orthogonal. However, an explicit example of such a kernel was not given and it has also been unresolved until now whether any two such kernels are isomorphic, as recently discussed by Mauldin (1989).

1. Perturbed Bernoulli kernels. Let β_x for $0 \le x \le 1$ denote the Bernoulli measure on $Y = \{0, 1\}^N$. The kernel $(\beta_x)_{x \in [0,1]}$ is completely orthogonal, i.e., there is a Borel map $f: Y \to [0, 1]$ such that for each x, $\beta_x(f^{-1}(x)) = 1$. In this section, we explicitly modify this family so that the new kernel is still orthogonality preserving, but no longer completely orthogonal.

Construction. Let M be the set of all sequences of positive integers such that the ratio of two consecutive members converges to ∞. Let M and Y carry the usual product Borel structure. Let $(k_r)_{r \ge 1}$ be a sequence of nonnegative random integers such that

$$(1) \qquad k_r \to_{r \to \infty} \infty \quad \text{and} \quad P(\exists\, r \text{ such that } k_r = l) \to_{l \to \infty} 0,$$

for example, the k_r independent of each other according to the uniform distribution on $\{2^r, \dots, 2^{r+1} - 1\}$. For each $x \in [0, 1]$, let $(B_i^x)_{i=1}^\infty$ be a sequence of i.i.d. $B(1, x)$-random variables (so that the joint distribution of this sequence is β_x), independent of the (k_r).

For each $m \in M$, define the random subset N_m of N by

$$(2) \qquad N_m = \bigcup_{r=1}^{\infty} [m(k_r), m(k_r + 1)).$$

For each triple $(x, z, m) \in [0, 1] \times Y \times M$, let

$$(3) \qquad Y_i^{x,z,m} = \begin{cases} B_i^x, & \text{if } i \notin N_m, \\ z_i, & \text{if } i \in N_m, \end{cases}$$

and let $\rho_{x,z,m}$ be the joint distribution of $(Y_i^{x,z,m})_{i=1}^\infty$.

THEOREM 1. *Let g be a Borel map of $[0, 1]$ into $Y \times M$. Then the kernel $x \to \mu_x = \rho_{x, g(x)}$ is orthogonality preserving. If g is two-to-one and onto, then the kernel is not completely orthogonal.*

PROOF. In order to show that (μ_x) is orthogonality preserving it is sufficient to verify that the weak law of large numbers holds, that is,

$$\lim_{n \to \infty} \mu_x \left\{ y \in Y: \left| (1/n) \sum_{i=1}^{n} y_i - x \right| > \varepsilon \right\} = 0,$$

or equivalently,

$$(4) \qquad \lim_{n \to \infty} P\left\{ \left| (1/n) \sum_{i=1}^{n} Y_i^{x, g(x)} - x \right| > \varepsilon \right\} = 0.$$

That this condition is sufficient can be seen by passing to a suitable subsequence. A more general version of this implication is (vii) → (i) of Theorem 4.1 in Mauldin, Preiss and Weizsäcker (1983). For the proof of (4) observe that each term in the arithmetic mean in (4) is bounded by 1 and therefore it suffices to show that for each $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{(1/n)\#\{i \le n: Y_i^{x, g(x)} \ne B_i^x\} > \varepsilon\} = 0, \tag{5}$$

since (B_i^x) satisfies even the strong law of large numbers. Now, because $m_x(k+1)/m_x(k) \to \infty$ and in view of (2) and (3), the condition

$$(1/n)\#\{i \le n: Y_i^{x, g(x)} \ne B_i^x\} > \varepsilon$$

holds for large n only in case

$$m_x(k_r - 1) \le n < m_x(k_r + 1),$$

for some $r \in N$. Let $l(n)$ be the unique integer such that

$$m_x(l(n)) \le n < m_x(l(n) + 1).$$

Then $l(n) \to \infty$ as $n \to \infty$ and hence

$$P\{(1/n)\#\{i \le n: Y_i^{x, g(x)} \ne B_i^x\} > \varepsilon\}$$
$$\le P(\exists\ r \text{ such that } k_r = l(n) - 1 \text{ or } k_r = l(n)\} \to 0,$$

as $n \to \infty$, by assumption (1). In other words, (5) holds. □

2. Using the Baire property. In order to prove that the kernel (μ_x) constructed in Section 1 cannot be completely orthogonal, we need the following result which is a generalized version of a theorem of Blackwell (1980). For this particular form, see Lemma 5.1a of Mauldin, Preiss and Weizsäcker (1983).

LEMMA 2. *Let Y be the topological product of a sequence of finite spaces, $Y = \prod_{i=1}^{\infty} Y_i$. If a subset A of Y has the Baire property and is not meager and if L is an infinite set of positive integers, then there exist a point $z \in Y$ and an increasing sequence m of positive integers such that $m(1) = 1$, $\forall\ i > 1$, $m(i) \in L$, and the set*

$$A(z, m) = \{y \in Y: y_j = z_j, \textit{ for all } j \in [m(k), m(k+1)) \textit{ for } k = 1$$
$$\textit{and for infinitely many } k\}$$

is a subset of A.

Note that in this lemma it is no restriction to assume $m \in M$ since if $A(z, m) \subset A$, then this relation also holds for any subsequence of m, provided $m(1)$ is unchanged.

Now, suppose (μ_x) is completely orthogonal, separated by the map f: $Y \to [0, 1]$. Let G be a Borel set such that (z, m) is a bijection on G and on $[0, 1] \setminus G$. Recall every Borel subset D of a Polish space has the Baire property and either D or its complement is not meager. So, switching, if necessary between A and its complement, we may assume that $A = f^{-1}(G)$ is not meager. There is, according to the lemma, a point $(z, m) \in Y \times S$ such that $A(z, m) \subset A$. Now, choose $x \in [0,1] \setminus G$ such that $(z, m) = g(x)$. The construction of the sequence $Y_i^{x, g(x)}$ implies $(Y_i^{x, g(x)})_{i \geq 1} \in A(z, m)$ and thus, $\mu_x(A(z, m)) = 1$ and $\mu_x(A) = 1$, while $f^{-1}(\{x\}) \cap A = f^{-1}(\{x\} \cap G) = \varnothing$. This is a contradiction to $\mu_x(f^{-1}(x)) = 1$, concluding the proof of Theorem 1.

3. Orthogonality preserving kernels not completely orthogonal. First, we give a general condition insuring that a kernel is not completely orthogonal. It shows that in Theorem 1 the condition "two-to-one" is superfluous.

THEOREM 3. *Let Y be the topological product of a sequence of finite spaces, $Y = \prod_{i=1}^{\infty} Y_i$, with $\operatorname{card}(Y_i) \geq 2$. Let $x \to \mu_x \in \operatorname{Prob}(Y)$ be an atomless transition kernel such that for each (z, m), there is some x such that there is a Borel subset K_x of $A(z, m)$ of full μ_x-measure. Then the kernel μ_x is not completely orthogonal.*

PROOF. According to the lemma this assumption implies that for every Borel set A in Y there is some x such that $\mu_x(A) = 1$ or $\mu_x(Y \setminus A) = 1$. We show no completely orthogonal kernel can have this property. Let us assume μ_x is completely orthogonal. According to Theorem 2.2 of Mauldin, Preiss and Weizsäcker (1983), there is a Borel isomorphism h of Y onto $[0,1] \times [0,1]$ such that $\mu_x(E) = \nu_x(h(E))$, where $x \to \nu_x \in \operatorname{Prob}([0,1] \times [0,1])$ is the canonical atomless completely orthogonal kernel given by $\nu_x(E) = \varepsilon_x \times \lambda(E) = \lambda(E_x)$, where λ is Lebesgue measure on the unit interval. Let $A = h^{-1}([0,1] \times (\frac{1}{2}, 1])$. Then we have $\mu_x(A) = \nu_x([0,1] \times (\frac{1}{2} \times 1]) = \frac{1}{2}$, for every x which is a contradiction. □

4. Nonisomorphic kernels. Finally, we construct continuum many nonisomorphic atomless orthogonality preserving transition kernels. Throughout this section, let $Y = \prod_{i=1}^{\infty} Y_i$, where $Y_i = \{j/2^i\colon 1 \leq j \leq 2^i\}$ and let φ_n be the projection map of Y onto the nth coordinate. For each (z, m), let

$$\begin{aligned} K(z, m) = \{y \in Y\colon y_j = z_j, &\text{ if } 1 \leq j < m(2) \text{ and } \forall\, r \geq 1,\\ &\exists\, i, 0 \leq i \leq 2^r - 1, \text{ such that } y_j = z_j,\\ &\text{for } m(2^r + i) \leq j < m(2^r + i + 1)\}. \end{aligned}$$

Note that $K(z, m) = K(z', m')$ if and only if $(z, m) = (z', m')$.

Let B be an uncountable Borel set of increasing elements of N^N such that the sets $\sigma(N)$, $\sigma \in B$, are almost disjoint. [There are many ways to construct such sets. One favorite method is the following: For each $0 < \theta < +\infty$, let $M(\theta)$ be the set of all points of the positive lattice $N \times N$ whose distance to the line $y = \theta x$ is less than 1. Clearly, if $\theta \neq \theta'$, then $M(\theta) \cap M(\theta')$ is finite and $\theta \to M(\theta)$ is a Borel measurable map into $\{0,1\}^{N \times N}$. The remainder of the construction is a matter of coding.] Let D be the set of all increasing sequences in N^N which begin with 1. Let $W = \{K(z, m): (z, m) \in Y \times D\}$. W is a Borel subset of $\mathscr{K}(Y)$, the space of compact subsets of Y. It follows from Lemma 2 that W can be partitioned into Borel sets $W_1, W_2, W_3, \dots$ such that, for each n and each nonmeager Baire set A, there is some $K(z, m) \in W_n$ with $K(z, m) \subset A$. For each $\sigma \in B$, let g_σ be a Borel measurable map of $[0, 1]$ onto $Y \times D$ such that for each p, if $K(z, m) \in W_p$, then $|g_\sigma^{-1}(z, m)| = \sigma(p)$. Let $x \to \mu_{\sigma, x} = \rho_{x, g_\sigma(x)}$ as in Theorem 1.

THEOREM 4. *For each* $\sigma \in B$, $\mu_{\sigma, x}$ *is an atomless, orthogonality preserving kernel, but not completely orthogonal. If* σ, $\tau \in B$ *and* $\sigma \neq \tau$, *then* $\mu_{\sigma, x}$ *and* $\mu_{\tau, x}$ *are not isomorphic.*

PROOF. If $(z, m) = g_\sigma(x)$, then by construction $\operatorname{supp} \mu_{\sigma, x} = K(z, m)$. Therefore, the kernels $\mu_{\sigma, x}$ are not completely orthogonal according to the preceding theorem. But, they are orthogonality preserving according to Theorem 1. Suppose $\sigma \neq \tau$ and there are Borel isomorphisms φ and ψ such that $\mu_{\sigma, x}(E) = \mu_{\tau, \varphi(x)}(\psi(E))$. Let A be a comeager set such that $\psi|_A$ is continuous [see Kuratowski (1966), page 400]. Choose k_1 and k_2 such that $\sigma([k_1, \infty))$ and $\tau([k_2, \infty))$ are disjoint with $\tau(k_2) > \sigma(k_1)$. Choose s such that $\sigma(s) > \tau(k_2)$. Find $K \in W_s$ such that $K \subset A$ and $\psi|_K$ is continuous. Let $x_1, \dots, x_{\sigma(s)}$ be the distinct elements of $[0, 1]$ such that $\operatorname{supp} \mu_{\sigma, x_i} = K$. Since $\psi|_K$ is continuous, $\psi(K) = \operatorname{supp} \mu_{\tau, \varphi(x_i)}$. Thus, $\psi(K) \in W_q$, where $\tau(q) \geq \sigma(s)$. Since $\tau(q) \neq \sigma(s)$, there is some $x \notin \{x_1, \dots, x_{\sigma(s)}\}$ such that $\operatorname{supp} \mu_{\tau, \varphi(x)} = \psi(K)$. This implies the contradiction $K = \operatorname{supp} \mu_{\sigma, x}$. □

REFERENCES

BLACKWELL, D. (1980). There are no Borel SPLIFs. *Ann. Probab.* **8** 1189–1190.

MAULDIN, R. D. (1989). One-to-one selections and orthogonal transition kernels. In *Measure and Measurable Dynamics* (R. D. Mauldin, R. M. Shortt and C. E. Silva, eds.). *Comp. Math.* **94** 185–190.

MAULDIN, R. D., PREISS, D. and WEIZSÄCKER, H. v. (1983). Orthogonal transition kernels. *Ann. Probab.* **11** 970–988.

KURATOWSKI, K. (1966). *Topology* **1**. Academic, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NORTH TEXAS
BOX 5116
DENTON, TEXAS 76203

FACHBEREICH MATHEMATIK
UNIVERSITÄT KAISERSLAUTERN
6750 KAISERSLAUTERN
GERMANY