stichting
mathematisch
centrum

AFDELING MATHEMATISCHE STATISTIEK SW 23/73 DECEMBER

A. HORDIJK & K.M. Van HEE
A BAYES PROCESS

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS (MOS) subject classification scheme (1970): 60J07, 60J45, 62L10, 62L15

A Bayes process [†]

K.M. van Hee [*] and A. Hordijk

Summary

It is well-known that the posterior probabilities in sequential sampling constitute a Markov-process. This process which may be called a Bayes process is an important tool in applying the theory of optimal stopping to sequential decision problems. For the case of a simple nullhypothesis against a simple alternative we investigate the corresponding Bayes process. Several properties of this process are proved. It is shown that the class of continuous harmonic functions for this Bayes process is exactly the class of linear functions.

[†] This paper is not for review; it is meant for publication in a journal.

[*] Centraal Rekeninsituut, Rijksuniversiteit te Leiden. From December 1973: Technische Hogeschool Eindhoven.

1. *Introduction*

In the sequential testing of a simple nullhypothesis against a simple alternative the statistician has the freedom to look at a sequence of observations one at a time and to decide after each observation whether to stop sampling and to reject or accept the nullhypothesis on the base of the observations, or to continue sampling.

Suppose the hypotheses are $\theta = 0$ and $\theta = 1$ and the statistician may observe the sequence $\underline{x}_1, \underline{x}_2, \ldots$ of independent, identically distributed random variables with known probability densities $f_0(x)$ and $f_1(x)$, with respect to the measure μ, under $\theta = 0$ and $\theta = 1$ respectively. In the Bayesian approach of this problem it is assumed that the parameter θ is a random variable $\underline{t}$ which takes on the value 0 with probability y_0 for some $0 \le y_0 \le 1$ and the value 1 with probability $1 - y_0$, and that $\underline{x}_1, \underline{x}_2, \ldots$ are, conditionally given $\underline{t} = 0$, independent and identically distributed with density $f_\theta(x)$. In this paper we study the posterior probabilities of $\underline{t}$ given $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$

$$\underline{y}_n := P[\underline{t}=0 \mid \underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n], \qquad n=1,2,\ldots \; .$$

We shall prove that the sequence $\{\underline{y}_n,\ n=1,2,\ldots\}$ forms a Markov process. We call this process the Bayes process.

Using a Bayes process the well-known theory of optimal stopping of Markov processes can be applied to sequential decision problems. In another publication we proceed in this way. The authors could not find a place in the literature where the Bayes process was treated explicitly. This was the main motivation for this paper. In section 3 several properties of the Bayes process are derived. Although several of these results seem to be known we could not find an appropriate reference. As far as we know the characterization of continuous harmonic functions is a new result. It is well-known that this Bayes process is also a martingale which also may be derived from the fact that all lineair functions for this Bayes process are harmonic.

2. *Definition of the Bayes process*

The assumptions underlying this process are

Let (Ω, F, P) *be a probability space and let* $\underline{x}_n$, $n=1,2,\ldots$ *be a sequence of random variables defined on* (Ω, F, P) *with sample space* (X, B) *where* X *is a Borel subset of* $\mathbb{R}^1$ *and* B *is the corresponding Borel* σ*-algebra.*
Moreover, $\underline{t}$ *is a random variable on* (Ω, F, P) *with*

$$P[\underline{t}=0] = 1 - P[\underline{t}=1] = y_0 \quad \textit{for some } 0 \le y_0 \le 1.$$

The r.v.'s $\underline{x}_n$ *are, conditionally given* $\underline{t}=p$ $(p=0,1)$, *independent and identically distributed,* i.e.

$$f_{\underline{x}_1,\ldots,\underline{x}_n|\underline{t}=p}(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_p(x_i), \qquad n=1,2,\ldots,$$

where $f_{\underline{x}_1,\ldots,\underline{x}_n|\underline{t}=p}(\,\cdot\,)$ *denotes the conditional density of* $(\underline{x}_1,\ldots,\underline{x}_n)$ *with respect to the product measure* $\prod_{i=1}^{n} \mu(x_i)$, *with* μ *a fixed measure on* (X,B).

As an immediate consequence of these assumptions we have that $\underline{x}_1,\ldots,\underline{x}_n$ are identically distributed with density $f_{\underline{x}_i}(x)$ with respect to $\mu(x)$. Indeed with $\nu(p)$ the counting measure on $\{0,1\}$ and $f_{\underline{t}}(p)$ the density of $\underline{t}$ with respect to $\nu(p)$ we have that

$$(1) \qquad f_{\underline{x}_i}(x) = \int f_{\underline{x}_i|\underline{t}=p}(x)\, f_{\underline{t}}(p)\, d\nu(p) = y_0 f_0(x) + (1-y_0) f_1(x),$$

and the joint density of $\underline{x}_1,\ldots,\underline{x}_n$ is

$$(2) \quad f_{\underline{x}_1,\ldots,\underline{x}_n}(x_1,\ldots,x_n) = \int f_{\underline{x}_1,\ldots,\underline{x}_n|\underline{t}=p}(x_1,\ldots,x_n)\, f_{\underline{t}}(p)\, d\nu(p) =$$

$$= y_0 \prod_{i=1}^{n} f_0(x_i) + (1-y_0) \prod_{i=1}^{n} f_1(x_i).$$

Hence, under the condition that $\mu(x|f_0(x) \neq f_1(x)) > 0$, $\underline{x}_1,\ldots,\underline{x}_n$ are only independent if $y_0 = 0$ or 1.

We define the sequence of posterior probabilities $\{\underline{y}_n,\ n=0,1,\ldots\}$ *by*

1) $\underline{y}_0 = y_0$ *a.s.*

2) $\underline{y}_n = P[\underline{t}=0|\underline{x}_1,\ldots,\underline{x}_n]$.

We denote $y_n(x_1,\ldots,x_n) = P[\underline{t}=0|\underline{x}_1=x_1,\ldots,\underline{x}_n=x_n]$.

To avoid unessential complications we suppose from now on that (2) is positive.

<u>Lemma 1</u>. *The functions* y_n *satisfy*

$$y_n(x_1,\ldots,x_n) = \frac{y_0 \prod_{i=1}^n f_0(x_i)}{y_0 \prod_{i=1}^n f_0(x_i) + (1-y_0) \prod_{i=1}^n f_1(x_i)},\quad n=0,1,\ldots \tag{3}$$

<u>Proof</u>. By Bayes' rule we have

$$y_n(x_1,\ldots,x_n) = P[\underline{t}=0|\underline{x}_1=x_1,\ldots,\underline{x}_n=x_n] =$$

$$= \frac{f_{\underline{x}_1,\ldots,\underline{x}_n|\underline{t}=0}(x_1,\ldots,x_n)\ f_{\underline{t}}(0)}{\int f_{\underline{x}_1,\ldots,\underline{x}_n|\underline{t}=p}(x_1,\ldots,x_n)\ f_{\underline{t}}(p)\ d\nu(p)} =$$

$$= \frac{y_0 \prod_{i=1}^n f_0(x_i)}{y_0 \prod_{i=1}^n f_0(x_i) + (1-y_0) \prod_{i=1}^n f_1(x_i)},$$

from which the statement follows. □

It is straightforward to verify that (3) implies

$$y_{n+1}(x_1,\dots,x_{n+1}) = \frac{y_n\ f_0(x_{n+1})}{y_n\ f_0(x_{n+1}) + (1-y_n)\ f_1(x_{n+1})},$$

where we write y_n for $y_n(x_1,\dots,x_n)$.

Lemma 2.

a) $$P[\underline{x}_{n+1} \in A \mid \underline{y}_n = y_n] = P[\underline{x}_{n+1} \in A \mid \underline{x}_1 = x_1,\dots,\underline{x}_n = x_n] =$$
$$= P[\underline{x}_{n+1} \in A \mid \underline{y}_1 = y_1,\dots,\underline{y}_n = y_n]$$

for all $A \in \mathcal{B}$.

b) $$P[\underline{x}_{n+1} \in A_1, \underline{x}_{n+2} \in A_2,\dots,\underline{x}_{n+k} \in A_k \mid \underline{x}_1 = x_1,\dots,\underline{x}_n = x_n] =$$
$$= y_n \prod_{i=1}^{k} \int_{A_i} f_0(x_{n+i})\ d\mu(x_{n+i}) + (1-y_n) \prod_{i=1}^{k} \int_{A_i} f_1(x_{n+i})\ d\mu(x_{n+i})$$

for $A_i \in \mathcal{B}$, $1 \le i \le k$.

c) *The conditional density of* $\underline{x}_{n+1}$ *given* $\underline{y}_n = y_n$ *is*

$$f_{\underline{x}_{n+1} \mid \underline{y}_n = y_n}(x) = y_n\ f_0(x) + (1-y_n)\ f_1(x).$$

Proof.

$$P[\underline{x}_{n+1} \in A_1,\dots,\underline{x}_{n+k} \in A_k \mid \underline{x}_1 = x_1,\dots,\underline{x}_n = x_n] =$$

$$= \int_{A_1 \times \dots \times A_k} \frac{\int f_{\underline{x}_1,\dots,\underline{x}_{n+k} \mid \underline{t}=p}(x_1,\dots,x_{n+k})\ f_{\underline{t}}(p)\ d\nu(p)}{\int f_{\underline{x}_1,\dots,\underline{x}_n \mid \underline{t}=p}(x_1,\dots,x_n)\ f_{\underline{t}}(p)\ d\nu(p)}\ d\mu(x_{n+1},\dots,x_{n+k}) =$$

$$= \int_{A_1\times\ldots\times A_k} \frac{y_0 \prod_{i=1}^{n+k} f_0(x_i) + (1-y_0)\prod_{i=1}^{n+k} f_1(x_i)}{y_0 \prod_{i=1}^{n} f_0(x_i) + (1-y_0) \prod_{i=1}^{n} f_1(x_i)}\, d\mu(x_{n+1},\ldots,x_{n+k}) \overset{\text{using (3)}}{=}$$

$$= \int_{A_1\times\ldots\times A_k} y_n \prod_{i=n+1}^{n+k} f_0(x_i)\, d\mu(x_{n+1},\ldots,x_{n+k}) +$$

$$+ \int_{A_1\times\ldots\times A_k} (1-y_n) \prod_{i=1}^{n+k} f_1(x_i)\, d\mu(x_{n+1},\ldots,x_{n+k}) =$$

$$= y_n \prod_{i=1}^{k} \int_{A_i} f_0(x_i)\, d\mu(x_i) + (1-y_n) \prod_{i=1}^{k} \int_{A_i} f_1(x_i)\, d\mu(x_i).$$

Since $\underline{y}_n$ is a function of $\underline{x}_1,\ldots,\underline{x}_n$ this proves the assertions a,b and c.□

Theorem 1. *The sequence* $\{\underline{y}_n,\ n=0,1,2,\ldots\}$ *is a stationary Markov process.*

Proof.

(4) $$P[\underline{y}_{n+1} \in B|\underline{y}_1=y_1,\ldots,\underline{y}_n=y_n] =$$

$$= \int P[\underline{y}_{n+1} \in B|\underline{y}_1=y_1,\ldots,\underline{y}_n=y_n,\underline{x}_{n+1}=x]dP_{\underline{x}_{n+1}|\underline{y}_1=y_1,\ldots,\underline{y}_n=y_n}(x) =$$

by lemma 2

$$= \int I_B\left(\frac{y_n f_0(x)}{y_n f_0(x) + (1-y_n) f_1(x)}\right) \{y_n f_0(x) + (1-y_n) f_1(x)\}\, d\mu(x).$$

Hence $P[\underline{y}_{n+1} \in B|\underline{y}_1=y_1,\ldots,\underline{y}_n=y_n]$ is a function of B and y_n only, which proves the Markov property. Moreover, since these conditional probabilities do not depend on n, the Markov process is stationary. □

Remark 1. From lemma 2a we see that $\underline{y}_n$ is sufficient, in the sense of Bayesian statistics, for the family $\{P_{\underline{x}_1,\dots,\underline{x}_n|\underline{t}=t}(\cdot)|t \text{ real}\}$ because $f_{\underline{t}|\underline{x}_1=x_1,\dots,\underline{x}_n=x_n}(t)$ depends on $x_1,\dots,x_n$ only through y_n.

Remark 2. Although in general the state space of the Bayes process is the interval [0,1], there are important situations for which a countable state space suffices. For example when the observed sequence $\underline{x}_1,\underline{x}_2,\dots$ is a sequence of Bernoulli trials. In this case the number of successes in combination with the number of failures can be used as state parameters of the process. Also when the likelihood ratio or the logarithm of the likelihood ratio takes on only rational values with positive probability then the state space is countable. In general, when from every state only a countable number of states can be reached and the initial distribution has a countable support, then the state space is countable.

3. Some properties of the Bayes process

For the Bayes process $\{\underline{y}_n, n=0,1,2,\dots\}$ with state space $(E,\mathcal{B})$ where $E = [0,1]$ and $\mathcal{B}$ the Borel-σ-algebra on E, we shall define the transition probabilities $p^n(x,A)$ and on the Banach space $B(E,\mathcal{B})$ of bounded measurable functions with norm $\|f\| = \sup_{x\in E} |f(x)|$ we shall define a sequence of linear operators $\{P^n, n=0,1,2,\dots\}$.

Definitions.

1. $$p(y,A) = \int_X I_A(g_y(x))\{y\ f_0(x) + (1-y)\ f_1(x)\}\ d\mu(x),$$

where $$g_y(x) = \frac{y\ f_0(x)}{y\ f_0(x) + (1-y)f_1(x)}.$$

2. $$p^0(y,A) = I_A(y)\ ,\ \textit{with}\ I_A(y) = \begin{cases} 1 & \textit{if } y \in A \\ 0 & \textit{otherwise}. \end{cases}$$

3. $$p^n(y,A) = \int_E p(y,dz)\; p^{n-1}(z,A), \qquad n=2,3,4,\dots .$$

4. $$P^n f(y) = \int p^n(y,dz)\; f(z), \qquad f \in B(E,\mathcal{B}) \text{ and } n=0,1,2,\dots .$$

We note that according to (4) the above defined transition probabilities are the transition probabilities of the Bayes process introduced in section 2. Note also that $p^n(x,A)$ is a probability measure on $\mathcal{B}$ for fixed x and n, and $p^n(x,A)$ is a $\mathcal{B}$-measurable function for fixed A and n.

Lemma 3.

1. $$P^n f(x) = P(P^{n-1} f(x))$$

$$< \text{or: } \int_E p(x,dy) \int_E p^{n-1}(y,dz)\; f(z) = \int_E p^n(x,dz)\; f(z) > .$$

2. $$p^{n+m}(x,A) = \int_E p^n(x,dy)\; p^m(y,A).$$

3. $$p^n(x,A) = P[\underline{y}_{m+n} \in A \mid \underline{y}_m = x], \qquad m=1,2,3,\dots .$$

4. $$Pf(y) = \int_X f(g_y(x))\; \{y\, f_0(x) + (1-y)\, f_1(x)\}\; d\mu(x).$$

Proof. The assertion 2 follows from assertion 1 by induction on n. To prove assertion 1 take $f(z) = I_A(z)$, $A \in \mathcal{B}$, then

$$\int_E p(x,dy) \int_E p^{n-1}(y,dz)\; I_A(z) = \int_E p(x,dy)\; p^{n-1}(y,A) =$$

$$= p^n(x,A) = \int_E p^n(x,dz)\; I_A(z).$$

Hence assertion 1 is true for elementary functions. For nonnegative

$f \in B(E,\mathcal{B})$ there is an increasing sequence of elementary functions $f_k(z) \uparrow f(z)$; hence by the monotone convergence theorem

$$\int_E p(x,dy) \int_E p^{n-1}(y,dz)\ f_k(z) \to \int_E p(x,dy) \int_E p^{n-1}(y,dz)\ f(z).$$

On the other hand, also by the monotone convergence theorem

$$\int_E p(x,dy) \int_E p^{n-1}(y,dz)\ f_k(z) = \int_E p^n(y,dz)\ f_k(z) \to \int_E p^n(y,dz)\ f(z).$$

Hence assertion 1 is true for nonnegative functions and therefore for all elements of $B(E,\mathcal{B})$.

According to theorem 1 we have for $n = 2$

$$p^2(x,A) = \int_E p(x,dy)\ p(y,A) =$$

$$= \int_E P[\underline{y}_{m+1} \in A|\underline{y}_m = y]\ dP_{\underline{y}_m|\underline{y}_{m-1}=x}(y) =$$

$$= P[\underline{y}_{m+1} \in A|\underline{y}_{m-1}=x].$$

The assertion 3 now follows by induction on n.

The assertion 4 follows from a well-known theorem on the change of integration variables. Indeed, for arbitrary but fixed integer n

$$Pf(y) = \int_E f(z)\ dP_{\underline{y}_{n+1}|\underline{y}_n=y}(z) = \int_E f(g_y(x))\ dP_{\underline{x}_{n+1}|\underline{y}_n=y}(x) =$$

$$= \int_E f(g_y(x))\ \{y\ f_0(x) + (1-y)\ f_1(x)\}\ d\mu(x). \quad \square$$

Lemma 4. *Let* A_n *be the* σ*-algebra induced by* $(\underline{x}_{n+1}, \underline{x}_{n+2}, \ldots)$ *and let* $\underline{z}$ *be* A_n*-measurable. Then for* $m \le n$

$$E[\underline{z} \mid \underline{y}_m = y] = y\, E_0[\underline{z}] + (1-y)\, E_1[\underline{z}]$$

(the subscripts 0 and 1 mean the distribution given $\underline{t}=0$ or $\underline{t}=1$).

Proof.

$$E[\underline{z} \mid \underline{y}_m] = E[E[\underline{z} \mid \underline{t}, \underline{x}_1, \ldots, \underline{x}_m] \mid \underline{y}_m] = E[E[\underline{z} \mid \underline{t}] \mid \underline{y}_m] =$$

$$= \underline{y}_m\, E_0[\underline{z}] + (1-\underline{y}_m)\, E_1[\underline{z}].$$

The first equality follows from the fact that $\underline{y}_m$ is a function of $\underline{x}_1, \ldots, \underline{x}_m$. Because $\underline{z}$ and $\underline{x}_1, \ldots, \underline{x}_m$ are independent given $\underline{t}$, the second equality is true. $\square$

To avoid integrability questions, we assume in the sequel that all functions are elements of $B(E,\mathcal{B})$. It is straightforward to verify that $f \in B(E,\mathcal{B})$ implies $P^n f \in B(E,\mathcal{B})$ for all $n=1,2,3,\ldots$.

We call a function f *harmonic* if $Pf = f$. The function f is *excessive* if f is nonnegative and $Pf(x) \le f(x)$, $x \in E$. The function f is called a *potential* if for some nonnegative function ϕ (called the *charge*), $f = \sum_{n=0}^{\infty} P^n \phi$. We recall that a function f is *concave* if $f(\gamma a + (1-\gamma)b) \ge \gamma f(a) + (1-\gamma) f(b)$ for all $\gamma \in [0,1]$ and $a,b \in E$.

Note that if f is harmonic, then

$$(5) \qquad P^n f = f, \qquad \text{for } n=0,1,2,\ldots .$$

Theorem 2. *A linear function* f *on* E *is harmonic, in particular the function* I *for which* $I(x) = x$ *on* E.

Proof. Let $f(z) = az+b$, then $Pf(y) = \int_E p(y,dz)\ (az+b) = a \int_E p(y,dz)z + b =$ $= aPI(y)+b$. By lemma 3, point 4 we know

$$PI(y) = \int_X g_y(x)\ \{y\ f_0(x) + (1-y)\ f_1(x)\}\ d\mu(x) =$$

$$= \int_X y\ f_0(x)\ d\mu(x) = y\ ,$$

which proves the theorem. □

Theorem 3. *If* h *is concave then* Ph *is also concave.*

Proof. Let $y_0 = \lambda y_1+(1-\lambda)y_2$, $\lambda \in [0,1]$, $y_1,y_2 \in E$. Note that

$$g_{y_0}(x) = \frac{1}{y_0 f_0(x)+(1-y_0)f_1(x)} \cdot$$

$$\cdot \{\lambda(y_1 f_0(x) + (1-y_1)f_1(x))g_{y_1}(x) + (1-\lambda)(y_2 f_0(x)+(1-y_2)f_1(x))g_{y_2}(x)\}.$$

Hence, since h is concave,

$$h(g_{y_0}(x)) \geq \frac{1}{y_0 f_0(x) + (1-y_0)f_1(x)} \cdot$$

$$\cdot \{\lambda(y_1 f_0(x)+(1-y_1)f_1(x))h(g_{y_1}(x)) +$$

$$+ (1-\lambda)(y_2 f_0(x)+(1-y_2)f_1(x))h(g_{y_2}(x))\}.$$

By lemma 3, point 4 we have

$$Ph(y_0) = \int h(g_{y_0}(x))\ \{y_0 f_0(x)+(1-y_0)f_1(x)\}\ d\mu(x) \geq$$

$$\geq \lambda \int h(g_{y_1}(x))\ \{y_1 f_0(x)+(1-y_1)f_1(x)\}\ d\mu(x) +$$

$$+ (1-\lambda)\int h(g_{y_2}(x))\ \{y_2 f_0(x)+(1-y_2)f_1(x)\}\ d\mu(x) =$$

$$= \lambda\ Ph(y_1) + (1-\lambda)\ Ph(y_2).\quad \square$$

Theorem 4. *If* h *is concave then* $Ph \le h$.

Proof. By Jensen's inequality: $Ph(y) = \int p(y,dz)\, h(z) \le h\, (\int p(y,dz) z) =$
$= h(y)$. □

Theorem 5. *Suppose*

$$\mu(\{x \mid f_0(x) \neq f_1(x)\}) > 0. \tag{6}$$

If h *is harmonic and* $\lim_{x \downarrow 0} h(x) = \lim_{x \uparrow 1} h(x) = h(0) = h(1) = 0$,
then $h(x) = 0$ *for all* $x \in E$.

Proof. Call $m = \sup_{x \in E} h(x)$. Because $h \in B(E,\mathcal{B})$ we have that $m < \infty$. There exists a sequence $\{y_i\}$, with $y_i \in E$, such that $h(y_i) \uparrow m$ as $i \to \infty$. Moreover, there is a subsequence $\{y_k\} \subset \{y_i\}$ and a point $y_0 \in E$ such that $y_k \to y_0$ as $k \to \infty$.

Let us assume for a moment that

(7) for each interval $[a,b] \subset E$ with $0 < a \le b < 1$, there are numbers n and $\varepsilon > 0$ such that $p^n(y,[a,b]^c) > \varepsilon$ for all $y \in [a,b]$.

Let $y_0 \neq 0$ or 1. Let interval $[a,b]$ be such that $0 < a \le b < 1$ and $y_0 \in (a,b)$. We shall prove that (7) implies $\sup_{x \in [a,b]^c} h(x) = m$. Note that $\sup_{x \in [a,b]} h(x) = m$. Suppose $\sup_{x \in [a,b]^c} h(x) = m-\alpha$, with $\alpha > 0$. For k large enough we have $y_k \in [a,b]$, hence in view of (7)

$$h(y_k) = \int p^n(y_k,dx)\, h(x) \le$$

$$\le m\, p^n(y_k,[a,b]) + (m-\alpha)\, p^n(y_k,[a,b]^c) = m-\alpha\varepsilon.$$

So, $\lim_{k\to\infty} h(y_k) \le m-\alpha\varepsilon$, which is in contradiction with $\lim_{k\to\infty} h(y_k) = m$.
Hence $\sup_{x \in [a,b]^c} h(x) = m$. By the continuity of $h(x)$ in the points 0 and 1 it follows that $m = 0$. If $y_0 = 0$ or 1 then also $m = 0$. Hence $h(x) \le 0$ for all $x \in E$.

In the same way, reasoning with $g(y) = -h(y)$, it is easy to see that $h(y) \geq 0$ for $y \in E$. Hence $h(y) = 0$ for all $y \in E$.

In order to prove (7) assume that $\mu(\{x | f_0(x) > f_1(x)\}) > 0$. Let $y \in [a,b]$ with $0 < a,b < 1$,

$$p^n(y,[a,b]^c) \geq p^n(y,(b,1]) =$$

$$= y\, P_0\left[\prod_{i=1}^{n} \frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)} < \frac{1/b-1}{1/y-1}\right] + (1-y)\, P_1\left[\prod_{i=1}^{n} \frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)} < \frac{1/b-1}{1/y-1}\right] \geq$$

$$\geq y\, P_0\left[\prod_{i=1}^{n} \frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)} < \frac{1/b-1}{1/y-1}\right] \geq$$

$$\geq y\, P_0\left[\prod_{i=1}^{n} \left\{\frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)} < \left(\frac{1/b-1}{1/y-1}\right)^{1/n}\right\}\right] =$$

$$= y \prod_{i=1}^{n} P_0\left[\frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)} < \left(\frac{1/b-1}{1/y-1}\right)^{1/n}\right] =$$

$$= y \prod_{i=1}^{n} \int_A f_0(x)\, d\mu(x),$$

where

$$A = \left\{ x \,\middle|\, \frac{f_1(x)}{f_0(x)} < \left(\frac{1/b-1}{1/y-1}\right)^{1/n} \right\}.$$

Then for n sufficiently large $\mu(A) > 0$. Hence

$$\prod_{i=1}^{n} \int_A f_0(x)\, d\mu(x) = \varepsilon > 0$$

and $p^n(y,[a,b]^c) > a\varepsilon$ for all $y \in [a,b]$.

The proof of (7) for $\mu(\{x | f_0(x) < f_1(x)\})$ proceeds in a similar way. □

Theorem 6. *Under condition* (6) *any continuous harmonic function is a linear function.*

Proof. Let h be an arbitrary continuous harmonic function. Call $b = h(0)$ and $a = h(1)-h(0)$; $g(y) = ay+b$. The function g is by theorem 2 harmonic. A linear combination of harmonic functions is also harmonic. Hence $f = h-g$ is harmonic, continuous and $f(0) = f(1) = 0$. According to theorem 5 we have $f \equiv 0$ and consequently $h(y) = ay+b$. □

Theorem 7. *If relation* (6) *is true and* h *is a nonnegative concave function with* $h(0) = h(1) = 0$ *then* $\lim_{n\to\infty} P^n h = 0$ *and* h *is a potential with charge* $h-Ph$.

Proof. From theorem 3 it follows by induction that $P^n h$ is concave. Since $P^n h \geq 0$ for all $n \geq 1$ we have from theorem 4 that $0 \leq P^{n+1}h \leq P^n h$ for $n \geq 1$. Hence $g(x) = \lim_{n\to\infty} P^n h(x)$ exists. Since $g(x)$ is bounded and concave $g(x)$ is continuous. Also $g(x)$ is harmonic, since

$$Pg(x) = \int p(x,dy) \lim_{n\to\infty} P^n h(y) = \lim_{n\to\infty} \int p(x,dy)\, P^n h(y) = g(x),$$

by the boundedness of h. Furthermore $0 \leq g(0) \leq h(0)$, so $g(0) = 0$ and similar $g(1) = 0$. Applying theorem 5 we obtain $g(x) \equiv 0$, which proves the first assertion of the theorem.

Let $\phi(x) = h(x)-Ph(x)$. By theorem 4, $\phi(x) \geq 0$.

$$\sum_{n=0}^{N} P^n\phi(x) = \sum_{n=0}^{N} P^n h(x) - \sum_{n=0}^{N} P^{n+1}h(x) =$$

$$= h(x) - P^{N+1}h(x) \to h(x) \quad \text{as } N\to\infty,$$

which proves that $h(x)$ is a potential with charge $\phi(x)$. □

Remark 3. It is well-known that the likelihood ratios, $\underline{z}_n = \prod_{i=1}^{n} \frac{f_1(\underline{x}_i)}{f_0(\underline{x}_i)}$ form a martingale (cf. [Feller], page 211). Hence the sequence $\{\underline{y}_n, n=0,1,2,\ldots\}$ forms also a martingale because $y_n(x_1,\ldots,x_n)$ is a one-one correspondence

$$y_n(x_1,\ldots,x_n) = \left[1 + \frac{1-y_0}{y_0} \prod_{i=1}^{n} \frac{f_1(x_i)}{f_0(x_i)}\right]^{-1}$$

Theorem 2 is an immediate consequence of this.

It is also well-known (cf. [Doob]) that the likelihood ratios converge a.s. if condition (6) is true, $\underline{z}_n \to 0$ a.s. on $\underline{t} = 0$ and $\underline{z}_n \to \infty$ a.s. on $\underline{t} = 1$. From the martingale convergence theorem it follows that there exists a $\underline{y}_\infty$ such that $\underline{y}_n \to \underline{y}_\infty$ a.s. Hence $P[\underline{y}_\infty=0] = 1 - P[\underline{y}_\infty=1] = y_0$. Theorem 5 is a direct consequence of this.

References

Doob, J.L., *Stochastic processes*, Wiley, New York, 1953.

Feller, W.F., *An introduction to probability theory and its applications*, Part II, second edition, Wiley, New York, 1966.