ISRAEL JOURNAL OF MATHEMATICS 84 (1993), 153–160

NONUNIQUENESS IN g-FUNCTIONS

BY

MAURY BRAMSON*

Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA

AND

STEVEN KALIKOW**

Mathematical Sciences Institute, Cornell University, Ithaca, NY 14853, USA

ABSTRACT

We give an example of two distinct stationary processes $\{X_n\}$ and $\{X'_n\}$ on $\{0,1\}$ for which $P[X_0 = 1 \mid X_{-1} = a_{-1}, X_{-2} = a_{-2}, \ldots] = P[X'_0 = 1 \mid X'_{-1} = a_{-1}, X'_{-2} = a_{-2}, \ldots]$ for all $\{a_i\}$, $i = -1, -2, \ldots$, even though these probabilities are bounded away from 0 and 1, and are continuous in $\{a_i\}$.

1. Introduction

One way of constructing a stationary process is to proceed inductively. Assuming that the entire past of the process has already been chosen, one can, according to some probability law, choose the present. Repeating this procedure indefinitely, one defines the process for all time. Markov chains provide motivation for this construction. Corresponding to the transition matrix of a Markov chain, one can employ g-functions for stationary processes. Here, we consider whether a g-function uniquely determines the stationary process under a natural "mixing" assumption.

* Supported in part by NSF Grant DMS 89-01545.

** Supported in part by the US Army Research Office.

Received February 27, 1992 and in revised form February 22, 1993

Let $\{X_n\}$, $n \in \mathbb{Z}$, be a stationary process; for simplicity, assume that $X_n \in \{0,1\}$. Using basic facts from martingales, one can define the conditional expectation f given the complete past

$$(1)\qquad \begin{aligned} f(a_{-1},a_{-2},\ldots) &= P[X_0 = 1 \mid X_{-1} = a_{-1}, X_{-2} = a_{-2},\ldots] \\ &= P[X_n = 1 \mid X_{n-1} = a_{-1}, X_{n-2} = a_{-2},\ldots], \end{aligned}$$

where $a_i \in \{0,1\}$, $i = -1,-2,\ldots$, and of course $0 \le f \le 1$. On the other hand, one can also start from functions f defined on sequences of 0's and 1's, and from them attempt to recover such stationary processes. Suppose that

(A) $0 \le f \le 1$ for all sequences $a_{-1}, a_{-2}, \ldots$.

(B) f is continuous, i.e., $\forall \epsilon > 0$, $\exists m \in \mathbb{Z}^+$ so that

$$|f(\alpha^1) - f(\alpha^2)| < \epsilon$$

for $\alpha^\ell = (a^\ell_{-1}, a^\ell_{-2}, \ldots)$, $\ell = 1,2$, with $a^1_i = a^2_i$ for $i \ge -m$.

(C) $\exists \epsilon \in \left(0, \frac{1}{2}\right)$ so that

$$\epsilon \le f(a_{-1}, a_{-2}, \ldots) \le 1 - \epsilon$$

for all $a_{-1}, a_{-2}, \ldots$.

Under Assumption (A), f is said to be a g-**function**; under (A) and (B), a **continuous** g-**function**. If f is a continuous g-function, then it is standard (and easy to show) that there exists a stationary process $\{X_n\}$ so that (1) holds. If (C) holds as well, we say f is a **regular** g-**function**. Does such a g-function necessarily correspond to a unique stationary measure? The purpose of this paper is to exhibit a regular g-function for which uniqueness does not hold.

g-functions were introduced by Doeblin and Fortet [3]. Harris [4] used g-functions to study the behavior of lumped state Markov chains. Keane [7] introduced the notion of continuous g-functions; he also gave conditions under which g-functions have unique measures which are mixing. Berbee [2] developed more general uniqueness criteria. Petit [8] extended Keane's work, with "continuous" being replaced by "differentiable"; he showed that all differentiable g-functions satisfying (C) have unique measures which are weak Bernoulli.

Kalikow [5] showed that a stationary, stochastic process can be represented as a random Markov chain precisely when it is a uniform martingale. The latter term means that the corresponding g-function is continuous; the former, that one can look a finite random distance into the past, and, using that portion of the past, apply a probability law to choose the present. The example of nonuniqueness which we present here will be a random Markov chain. (The question about nonuniqueness was first posed in [5].) We will not use this characterization, but note that following the proof of Theorem 7 in [5], it is easy to show that if the random distance one looks back has finite expectation for a random Markov chain also satisfying (C), then the corresponding g-function in fact uniquely determines the measure. (B. Weiss has since raised the question as to how much the assumption of finite expectation can be relaxed while preserving uniqueness.)

In [5], Kalikow also exhibited a uniform martingale which is K but not Bernoulli. Kalikow, Katznelson, and Weiss [6] have shown that every zero entropy transformation can be extended to a uniform martingale. We also note that the equilibrium measures of appropriately chosen one dimensional long-range Ising models are known not to be unique (see Aizenman, Chayes, Chayes, and Newman [1]). For these models, the spin at zero is influenced by sites arbitrarily far to the left and to the right.

2. Construction of the example

Let $\{p_j\}$, $j \in \mathbb{Z}^+$, be any decreasing sequence of numbers which satisfy $p_j \geq 0$, $\sum_{j=1}^{\infty} p_j = 1$, and

$$p_k \leq \frac{1}{2} \sum_{j>k} p_j \quad \text{for all } k. \tag{2}$$

For example, one can let $p_j = cr^j$, with $r \in \left(\frac{2}{3}, 1\right)$ and $c = (1-r)/r$. Let $\{m_j\}$, $j \in \mathbb{Z}^+$, be an increasing sequence of odd positive integers. We introduce the random variable N having distribution given by

$$P[N = m_j] = p_j \quad \text{for all } j. \tag{3}$$

From N, we define the random variable W on sequences $\alpha = (a_{-1}, a_{-2}, \ldots)$ of 0's and 1's by

$$W(\alpha) \begin{array}{ll} = 1-\epsilon & \text{if the majority of } \{a_{-1}, \ldots, a_{-N}\} \text{ are 1's,} \\ = \epsilon & \text{otherwise,} \end{array} \tag{4}$$

where $\epsilon \in \left(0, \frac{1}{2}\right)$.

Using W, we can construct the g-function

$$f(\alpha) = E[W(\alpha)]. \tag{5}$$

f prescribes looking back the random distance N into the past, and with probability $1-\epsilon$ adopting the state of the majority. It is easy to check that f satisfies properties (B) and (C), and so is regular. We will show that f is nevertheless the g-function of two distinct stationary processes, provided that $\{m_j\}$ increases rapidly enough.

We start by fixing the past to be $X_n^1 = 1$ and $X_n^0 = 0$, for $n = -1, -2, \ldots$, for the processes $\mathbf{X}^1 = \{X_n^1\}$ and $\mathbf{X}^0 = \{X_n^0\}$, and extending the construction for all n inductively according to (1). Of course, neither $\mathbf{X}^1$ nor $\mathbf{X}^0$ is stationary. We can nonetheless use the monotonicity of f to construct stationary limits. Specifically, let $X_n^{\ell,i} = X_{n+i}^{\ell}$ for $i \in \mathbb{Z}^+$ and $\ell = 0, 1$. Since $\{\mathbf{X}^{1,i}\}$ is stochastically decreasing as $i \to \infty$, it converges weakly to some stationary limit $\mathbf{X}^+$. Similarly, $\{\mathbf{X}^{0,i}\}$ converges to some stationary limit $\mathbf{X}^-$. One can check that both $\mathbf{X}^+$ and $\mathbf{X}^-$ have g-function f.

We will show that

$$\lim_{n\to\infty} P[X_n^1 = 1] > \tfrac{1}{2}, \tag{6}$$

provided $\{m_j\}$ increases rapidly. Similarly, $\lim_{n\to\infty} P[X_n^0 = 1] < \frac{1}{2}$. Since $P[X_0^1 = 1]$ and $P[X_0^0 = 1]$ are given by these limits, it will then follow that $\mathbf{X}^+$ and $\mathbf{X}^-$ are distinct. The regular g-function f defined in (5) therefore corresponds to distinct stationary processes.

3. Useful auxiliary processes

The remainder of the paper is devoted to demonstrating (6). Here, we construct processes $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$, which we later compare with $\mathbf{X}$.

Let $\{p_j\}$, $\{m_j\}$, N, W, and $\alpha = (a_{-1}, a_{-2}, \ldots)$ be as before. For $k \in \mathbb{Z}^+$, introduce

$$\begin{aligned} W^k(\alpha) &= W(\alpha) \quad \text{if } N \in \{m_1, \ldots, m_{k-1}\}, \\ &= 1/2 \quad\quad \text{otherwise}, \\ f^k(\alpha) &= E[W^k(\alpha)]. \end{aligned} \tag{7}$$

Define $\mathbf{Y}^{k,\alpha}$ by fixing the past to be $Y_n^{k,\alpha} = a_n$ for $n = -1, -2, \ldots$, and using the analog of (1) for f^k to extend $Y_n^{k,\alpha}$ for nonnegative n. Similarly, set

$$(8)\qquad \begin{aligned} \hat{W}^k(\alpha) &= W(\alpha) \quad \text{if } N \in \{m_1, \ldots, m_{k-1}\},\\ &= \epsilon \qquad\quad \text{if } N = m_k,\\ &= 1-\epsilon \quad\ \text{if } N \in \{m_{k+1}, m_{k+2}, \ldots\},\\ \hat{f}^k(\alpha) &= E[\hat{W}^k(\alpha)]. \end{aligned}$$

Define $\mathbf{Z}^{k,\alpha}$ analogously, but by instead using $\hat{f}^k$. Note that f^k and $\hat{f}^k$ only look back a finite distance into the past, and do not depend on the choice of $m_k, m_{k+1}, \ldots$ (although $\hat{f}^k$ depends on p_k).

We first wish to compare $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$, for given k and α. On account of (2),

$$(9)\qquad \begin{aligned} \hat{f}^k(\alpha) - f^k(\alpha) &= \left(\tfrac{1}{2} - \epsilon\right)\Big(\sum_{j>k} p_j - p_k\Big)\\ &\geq \left(\tfrac{1}{2} - \epsilon\right) p_k \stackrel{\text{def.}}{=} \eta_k > 0. \end{aligned}$$

Owing to the monotonicity of f^k, it is not hard to couple $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$ so that $Y_n^{k,\alpha} \leq Z_n^{k,\alpha}$ for all n. Using (9) again, we obtain that

$$(10)\qquad E[Z_n^{k,\alpha}] - E[Y_n^{k,\alpha}] \geq \eta_k \quad \text{for all } n \geq 0.$$

Both f^k and $\hat{f}^k$ depend only on the first m_{k-1} coordinates $a_{-1}, \ldots, a_{-m_{k-1}}$. So both $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$ are m_{k-1}-step Markov chains on $\{0,1\}$. Equivalently, $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$ can be interpreted in terms of Markov chains on a $2^{m_{k-1}}$-point state space. Let $Y^{k,\alpha}[0,m)$ and $Z^{k,\alpha}[0,m)$ denote the number of 1's for either process in the interval $[0,m)$. Then by the ergodic theorem for Markov chains and the symmetry of f^k,

$$(11)\qquad \lim_{m\to\infty} \tfrac{1}{m} Y^{k,\alpha}[0,m) = \tfrac{1}{2} \quad \text{a.s.}$$

for fixed α. Again invoking the ergodic theorem, this time in conjunction with (10), (11), and bounded convergence, one obtains

$$(12)\qquad \lim_{m\to\infty} \tfrac{1}{m} Z^{k,\alpha}[0,m) \geq \tfrac{1}{2} + \eta_k \quad \text{a.s.}$$

Consequently, for given $\delta_k > 0$ and large enough m,

$$(13)\qquad P\left[\tfrac{1}{m} Z^{k,\alpha}[0,m) \leq (1+\eta_k)/2\right] \leq \delta_k.$$

Since $\mathbf{Z}^{k,\alpha}$ has finite memory, (13) in fact holds uniformly in α.

4. Comparison between $\mathbf{X}^1$ and $\mathbf{Z}^{k,\alpha}$; derivation of (6)

So far, we have not stated anything about the sequence $\{m_j\}$ used to construct $\mathbf{X}^1$, $\mathbf{Y}^{k,\alpha}$ and $\mathbf{Z}^{k,\alpha}$, other than indicate that we want it to increase rapidly. We now define $\{m_j\}$ recursively as follows. Suppose that m_j has been given for $j < k$. Setting $\delta_k = 3^{-k-1}\eta_k$, we choose m_k large enough so that (13) holds for $m \geq m_k$. (Recall that $\mathbf{Z}^{k,\alpha}$ does not depend on m_j, $j \geq k$.) One obtains

$$P\left[\tfrac{1}{m_k} Z^{k,\alpha}[0,m_k) \leq (1+\eta_k)/2\right] \leq 3^{-k-1}\eta_k \tag{14}$$

for all k and α. Typically, $\{m_j\}$ will be increasing rapidly. We nevertheless explicitly assume as well that

$$m_j/m_{j+1} < \eta_j/2 \quad \text{for all } j. \tag{15}$$

We will demonstrate (6) by means of the following induction hypothesis. As before, $X^1[n_1,n_2)$ will denote the number of 1's in the interval $[n_1,n_2)$. We set

$$E_{j,n} = \{\omega\colon \tfrac{1}{m_j} X^1[n-m_j,n) \leq (1+\eta_j)/2\},\ \ j \geq 1, n \in \mathbb{Z}. \tag{16}$$

INDUCTION HYPOTHESIS. For all j and all $n \leq n_0$,

$$P[E_{j,n}] \leq 3^{-j}\eta_j. \tag{17}$$

Assume now that (17) holds for all n. One can easily check that

$$\frac{1}{m_j} \sum_{n'=n-m_j}^{n-1} P[X^1_{n'} = 1] \geq \frac{1}{2}(1+\eta_j/3) > 1/2. \tag{18}$$

As mentioned earlier, $P[X^1_{n'} = 1]$ is decreasing in n'. So selecting any j in (18) and letting $n \to \infty$, one obtains (6).

We first note that (17) holds trivially for $n_0 = 0$, since $X^1_{n'} = 1$ for $n' < 0$. Assume now that (17) holds for $n \leq n_0$ and all j. It suffices to show (17) for $n = n_0 + 1$ and all j. Setting $j = k$, k arbitrary, we will demonstrate (17) for $n = n_1 + m_k$, whenever $n_1 \leq n_0$. This will complete the proof.

For convenience, set $F_{k,n} = E_{k,n+m_k}$ and $G_{k,n} = \bigcup_{j>k} E_{j,n}$. One of course has

$$F_{k,n} \subset G_{k,n} \cup \left(F_{k,n} \cap G^c_{k,n}\right).$$

By the induction hypothesis,

$$P[G_{k,n_1}] \le \sum_{j>k} 3^{-j}\eta_j \le 2\cdot 3^{-k-1}\eta_k, \tag{19}$$

since η_k is decreasing. We will demonstrate

$$P\big[F_{k,n_1} \cap G^c_{k,n_1}\big] \le 3^{-k-1}\eta_k. \tag{20}$$

Together with (19), (20) implies that $P[F_{k,n_1}] \le 3^{-k}\eta_k$, which is the bound we want. To obtain (20), we will compare $\mathbf{X}^1$ with the processes $\mathbf{Z}^{k,\alpha}$. First note that under $G^c_{k,n_1} \cap \{N_{n_1} > m_k\}$, $X^1_n = 1$ for most of the sites $n \in [n_1 - N_{n_1}, n_1)$, where N_{n_1} is the random distance one looks back at the site n_1. The key point is that under $G^c_{k,n_1} \cap \{N_{n_1+\ell} > m_k\}$, $\ell = 0, 1, \ldots, m_k - 1$, $X^1_n = 1$ also holds for most of the sites $n \in [n_1 + \ell - N_{n_1+\ell}, n_1 + \ell)$. One is therefore able to compare f with $\hat{f}^k$ as in (23), and obtain (26) from (14).

Let $\bar{\alpha} = \{a_i\}$, $i \in \mathbb{Z}$, denote a doubly infinite sequence of 0's and 1's, and $r(\bar{\alpha}) = \{a_i\}$, $i < 0$, the corresponding restriction to the negative half-line. Denote by H_k the set of $\alpha = \{a_i\}$, $i < 0$, satisfying

$$\frac{1}{m_j} \sum_{i=-m_j}^{-1} a_i > (1 + \eta_j)/2, \quad j > k. \tag{21}$$

On account of (15), it follows that for $r(\bar{\alpha}) \in H_k$,

$$\frac{1}{m_j} \sum_{i=\ell-m_j}^{\ell-1} a_i > 1/2 \quad \text{for} \quad \ell = 0, 1, \ldots, m_k - 1, \;\; j > k. \tag{22}$$

That is, if a long interval (of length m_j) is shifted only a little (by less than m_k), then the proportion of 1's does not change by much. Setting $a_i = X^1_{n_1+i}$ in (21), we note that $\alpha \in H_k$ corresponds to $\omega \in G^c_{k,n_1}$.

Let $\bar{\alpha}^\ell = \{a^\ell_i\}$, with $a^\ell_i = a_{\ell+i}$, $i \in \mathbb{Z}$. Then $r(\bar{\alpha}^\ell)$ is the sequence $\bar{\alpha}$ shifted ℓ units to the left and truncated at 0. It follows from (22) that for $r(\bar{\alpha}) \in H_k$, and $\ell = 0, \ldots, m_k - 1$,

$$W(r(\bar{\alpha}^\ell)) = 1 - \epsilon \quad \text{if} \quad N \in \{m_{k+1}, m_{k+2}, \ldots\}.$$

From the definitions of f and $\hat{f}^k$ in (5) and (8), one therefore sees that

$$\hat{f}^k(r(\bar{\alpha}^\ell)) \le f(r(\bar{\alpha}^\ell)) \quad \text{for} \quad r(\bar{\alpha}) \in H_k, \;\; \ell = 0, \ldots, m_k - 1. \tag{23}$$

Introduce the processes $\mathbf{X}^{n,\alpha}$ obtained by conditioning $\mathbf{X}^1$ on $\mathbf{X}^1_{n+i} = a_i$ for $i < 0$. Also, recall the process $\mathbf{Z}^{k,\alpha}$ introduced after (8). Clearly,

$$Z_i^{k,\alpha} = X_{n_1+i}^{n_1,\alpha} \qquad \text{for} \quad i < 0. \tag{24}$$

Owing to the monotonicity of f and (23), it is therefore not difficult to couple $\mathbf{X}^{n_1,\alpha}$ and $\mathbf{Z}^{k,\alpha}$, for $\alpha \in H_k$, so that

$$Z_i^{k,\alpha} \leq X_{n_1+i}^{n_1,\alpha} \quad \text{for} \quad i = 0, \ldots, m_k - 1. \tag{25}$$

Together, (14) and (25) imply that for all k and all $\alpha \in H_k$,

$$P\left[\tfrac{1}{m_k} X^{n_1,\alpha}[n_1, n_1 + m_k) \leq (1 + \eta_1)/2\right] \leq 3^{-k-1}\eta_k. \tag{26}$$

Denote the above events by F^α_{k,n_1}. It follows that

$$\begin{aligned} P\left[F_{k,n_1} \cap G^c_{k,n_1}\right] &= \sum_{\alpha \in H_k} P\left[F^\alpha_{k,n_1}\right] P\left[X^1_{n_1+i} = a_i,\ i < 0\right] \\ &\leq 3^{-k-1}\eta_k, \end{aligned}$$

which implies (20). Consequently (17), and hence (6), hold.

References

[1] M. Aizenman, J. Chayes, L. Chayes and C. Newman, *Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models,* J. Stat. Physics **50** (1988), 1–40.

[2] H. Berbee, *Chains with infinite connections: uniqueness and Markov representation,* Probability Theory and Related Fields **76** (1987), 243–253.

[3] W. Doeblin and R. Fortet, *Sur les chaines a liaisons completes,* Bull. Soc. Math. Fr. **65** (1937), 132–148.

[4] T. Harris, *On chains of infinite order,* Pacific Journal Math. **5** (1955), 707–724.

[5] S. Kalikow, *Random Markov processes and uniform martingales,* Israel Journal of Mathematics **71** (1990), 33–54.

[6] S. Kalikow, Y. Katznelson and B. Weiss, *Finitarily deterministic generators for zero entropy systems,* Israel J. Math. **79** (1992), 33–45.

[7] M. Keane, *Strongly mixing g-measures,* Invent. Math. **16** (1972), 309–324.

[8] B. Petit, *Schemas de Bernoulli et g-mesures,* C. R. Acad. Sci. Paris, Ser. A **280** (1975), 17–20.