ELSEVIER

Statistics & Probability Letters 54 (2001) 183–187

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

A note on a theorem of Karlin☆

Örjan Stenflo

Department of Mathematics, Umeå University, S-90187 Umeå, Sweden

Received August 2000

Abstract

We give an example of place-dependent random iterations with two affine contractions on the unit interval generating a Markov chain with more than one stationary probability measure. The probability function is continuous and strictly positive. This constitutes a counterexample to a conjecture raised by an incomplete proof by Karlin (Pacific J. Math. 3 (1953) 725–756). © 2001 Elsevier Science B.V. All rights reserved

MSC: 28D05; 28A80; 37H99; 60J05; 60G10

Keywords: Iterated Function Systems (IFS); Random systems with complete connections; Markov chains; Invariant measures; Stationary measures; g-measures

1. Introduction

Let f_0 and f_1 be two maps from $[0,1]$ into itself defined by

$$f_0(x) = \sigma x \quad \text{and} \quad f_1(x) = \alpha + (1-\alpha)x, \tag{1}$$

where both σ and α satisfy $0 < \sigma,\ \alpha < 1$, and let p be a real-valued continuous function on $[0,1]$ with $0 < p(x) < 1$.

Suppose a point $x \in [0,1]$ moves randomly with probability $p(x)$ to $f_0(x)$ and with probability $1 - p(x)$ to $f_1(x)$. This procedure generates a Markov chain on $[0,1]$.

It is a natural question to ask whether a Markov chain generated in this way necessarily has a unique stationary probability measure. In this paper we give an example based on a result by Bramson and Kalikow (1993) showing that this is not necessarily the case.

The history of this question dates back to the paper by Karlin (1953). At that time Markov chains of this form were mostly known as learning models. Nowadays, 'Random systems with complete connections' or 'Iterated Function Systems with place-dependent probabilities', are more widely used terms.

☆ Supported by Hellmuth Hertz' Foundation.

E-mail address: stenflo@math.umu.se (Ö. Stenflo).

PII: S0167-7152(01)00045-1

In Karlin (1953, Theorem 36), a theorem was stated under the above assumptions which (for any fixed choice of parameter values σ and α) would imply an attractive, and thus necessarily unique, stationary probability distribution. The proof of this theorem contained a gap in that an argument requiring that the function $p(x)$ has bounded derivative was used. We may in this context regard the case of a unique stationary probability distribution (for any choice of parameter values σ and α) under merely continuity assumptions on p as Karlin's conjecture. The result in the present paper may thus be regarded as a negative answer to Karlin's conjecture. See e.g. Keane (1972), Kaijser (1981), Barnsley et al. (1988) and Kaijser (1994) for further comments on Karlin's paper.

For p as above, define the modulus of uniform continuity

$$\Delta_p(t) = \sup\{p(x) - p(y) : |x - y| < t\}.$$

We say that $p(x)$ is Dini-continuous if

$$\int_0^1 \frac{\Delta_p(t)}{t}\,\mathrm{d}t < \infty,$$

or equivalently

$$\sum_{n=1}^{\infty} \Delta_p(c^n) < \infty,$$

for some (and thus all) $0 < c < 1$.

In 1937 Doeblin and Fortet (1937) proved a result implying uniqueness of the stationary probability measure (for any choice of parameter values σ and α) under the assumption that $p(x)$ is Lipschitz-continuous. Kaijser (1981), Barnsley et al. (1988) and Li (1995) all contain different proofs generalising this result (for any choice of parameter values σ and α) to the Dini-continuous $p(x)$ case. (In Li, 1995, Theorem 2.2.3, the theorem was stated in a form merely assuming a continuous $p(x)$, but also these further assumptions on $p(x)$ are needed.) A proof of Karlin's theorem in the case when $p(x)$ is Dini-continuous also follows from Walters (1975), see Fan and Lau (1999). All the above mentioned papers allow a finite set of maps. The papers Kaijser (1981) and Barnsley et al. (1988) are more general in that also average contractive systems are covered, and Fan and Lau (1999) is more general in that the "probabilistic" assumption on $p(x)$ is relaxed. Note that Dini-continuity is weaker than Hölder continuity.

Also the Dini-condition can be relaxed. Harris (1955) (see also Kaijser (1994)) proved that if

$$\sum_{m=0}^{\infty} \prod_{n=0}^{m} (1 - \Delta_p(c^n)) = \infty,$$

where $c = \max(\sigma, 1 - \alpha)$, then there exists a unique invariant probability measure. A related condition for uniqueness can also be found in Berbee (1987). It also follows from Theorem 7.1. in Kaijser (1979) and Theorem 1 in Burton and Keller (1993), that Karlin's theorem holds if $p(x)$ is non-increasing and continuous. (In Burton and Keller, 1993, Theorem 1, assumption (v) it appears to be a typographical error in that all inequalities except the first inequality should be reversed.)

A closely related question within the theory of dynamical systems concerns uniqueness/non-uniqueness of g-measures. A regular g-function (see Bramson and Kalikow, 1993) is a function satisfying certain conditions corresponding to the properties of our p. A g-measure is an invariant probability measure with conditional probabilities determined by the function g. (See e.g. Quas (1996) for details.) The concept of g-measures was introduced by Keane (1972). The reader is referred to Quas (1996) for an account on the history of sufficient conditions ensuring uniqueness of g-measures. In an ingenious paper by Bramson and Kalikow (1993), a regular g-function with more than one g-measure on the space $\Sigma := \{0,1\}^{\mathbb{N}}$ was constructed. Based on their methods, Quas (1996) gave an example constructing a function g such that the map $T(x) = 10x \bmod 1$ on the unit circle has more than one g-measure.

Since we may (for parameter values such that $f_0([0,1]) \cap f_1([0,1]) = \emptyset$) look at our maps as inverse branches of a 2-1 expanding local homeomorphism T (defined on a Cantor set), and g as a function equivalent to p, Quas' example strongly indicated that a counterexample to Karlin's conjecture (with $p(x)$ being merely continuous) would exist.

In this paper we will show that a counterexample to Karlin's conjecture directly follows from Bramson and Kalikow (1993).

2. Construction of the counterexample

For a compact metric space X, random iterations with two functions, h_0 and h_1 from X into itself, according to a probability function $P : X \to (0,1)$ generates a Markov chain by the rule that a point $x \in X$ moves randomly with probability $P(x)$ to $h_0(x)$ and with probability $1 - P(x)$ to $h_1(x)$.

Our main result is the following;

Theorem 1. *Let f_0 and f_1 be maps from $[0,1]$ into itself defined by*

$$f_0(x) = \frac{x}{3} \quad \textit{and} \quad f_1(x) = \frac{x}{3} + \frac{2}{3}.$$

Then there exist a real-valued continuous function p on $[0,1]$ with $0 < p(x) < 1$ such that the generated Markov chain has more than one stationary probability measure.

Proof. We are going to use a result by Bramson and Kalikow (1993) in a crucial way. In order to state their result in a form fitting our context, we need to introduce some notation. Let $\Sigma := \{0,1\}^{\mathbb{N}}$ and introduce a topology on Σ induced by the metric

$$d(\mathbf{i},\mathbf{j}) := \begin{cases} 2^{-n}, & \text{if } \mathbf{i} \text{ and } \mathbf{j} \text{ differ for the first time in the } n\text{th digit} \\ 0, & \text{if } \mathbf{i} = \mathbf{j}. \end{cases}$$

The space (Σ, d) is a compact metric space.

For $a \in \{0,1\}$ and $\mathbf{i} = i_0 i_1 i_2 \ldots \in \Sigma$, let $a\mathbf{i} = a i_0 i_1 i_2 \ldots$

The following result was proved in Bramson and Kalikow (1993):

Theorem 2 (Bramson and Kalikow (1993)). *Let $\hat{f}_0$ and $\hat{f}_1$ be maps from Σ into itself defined by*

$$\hat{f}_0(\mathbf{i}) = 0\mathbf{i} \quad \textit{and} \quad \hat{f}_1(\mathbf{i}) = 1\mathbf{i}.$$

Then there exist a real-valued continuous function $\hat{p}$ on Σ with $0 < \hat{p}(\mathbf{i}) < 1$ such that the generated Markov chain has more than one stationary probability measure.

The reader is referred to Bramson and Kalikow (1993) for a proof of this result and details concerning the construction of $\hat{p}$.

We are going to use this result to first construct our p on the middle-third Cantor set and then extend p to the remaining points in $[0,1]$ by linear interpolation.

For a sequence $\mathbf{i} = i_0 i_1 i_2 \ldots \in \Sigma$, define

$$\hat{Z}(\mathbf{i}) = \lim_{n\to\infty} f_{i_0} \circ f_{i_1} \circ \cdots \circ f_{i_n}(0). \tag{2}$$

The map $\hat{Z} : \Sigma \to [0,1]$ is $1-1$ and the image of Σ is the middle-third Cantor set, $\mathscr{C}$. It is also readily checked that $\hat{Z}$ regarded as a bijective map from Σ onto $\mathscr{C}$ is continuous if we consider $\mathscr{C}$ with its subspace topology. Define, for $x \in \mathscr{C}$, $p(x) := \hat{p}(\hat{Z}^{-1}(x))$.

For $x \in [0,1] \setminus \mathscr{C}$ define $a_1 = \sup\{y \in \mathscr{C}: y < x\}$ and $a_2 = \inf\{y \in \mathscr{C}: y > x\}$. Note that since $\mathscr{C}$ is closed, a_1 and a_2 belong to $\mathscr{C}$, and $a_1 < x < a_2$. Define $p(x) = p(a_1) + ((x - a_1)/(a_2 - a_1))(p(a_2) - p(a_1))$.

Let $\hat{v}_1$ and $\hat{v}_2$ be two distinct stationary probability measures on Σ for the Markov chain generated by $\hat{f}_0, \hat{f}_1$ and $\hat{p}$, existing according to the theorem by Bramson and Kalikow stated above. That is, for any real valued continuous function $\hat{h}$ on Σ, we have that $\hat{v}_i$, $i = 1,2$, satisfy the invariance equation

$$\int_\Sigma \hat{h}\,\mathrm{d}\hat{v}_i = \int_\Sigma (\hat{p}(\mathbf{i})\hat{h}(\hat{f}_0(\mathbf{i})) + (1-\hat{p}(\mathbf{i}))\hat{h}(\hat{f}_1(\mathbf{i})))\,\mathrm{d}\hat{v}_i(\mathbf{i}). \tag{3}$$

Define $v_1 := \hat{v}_1 \circ \hat{Z}^{-1}$ and $v_2 := \hat{v}_2 \circ \hat{Z}^{-1}$. From (3) and by changing variables it follows that for any real valued continuous function h on $[0,1]$ the probability measures v_i, $i = 1,2$, satisfy the invariance equation

$$\begin{aligned}
\int_{[0,1]} h\,\mathrm{d}v_i &= \int_\Sigma (h \circ \hat{Z})\,\mathrm{d}\hat{v}_i \\
&= \int_\Sigma (\hat{p}(\mathbf{i}) h \circ \hat{Z}(\hat{f}_0(\mathbf{i})) + (1-\hat{p}(\mathbf{i})) h \circ \hat{Z}(\hat{f}_1(\mathbf{i})))\,\mathrm{d}\hat{v}_i(\mathbf{i}) \\
&= \int_\Sigma (p(\hat{Z}(\mathbf{i})) h \circ f_0(\hat{Z}(\mathbf{i})) + (1 - p(\hat{Z}(\mathbf{i}))) h \circ f_1(\hat{Z}(\mathbf{i})))\,\mathrm{d}\hat{v}_i(\mathbf{i}) \\
&= \int_{[0,1]} (p(x) h \circ f_0(x) + (1-p(x)) h \circ f_1(x))\,\mathrm{d}v_i(x).
\end{aligned}$$

Thus v_1 and v_2 are two distinct stationary probability measures for the Markov chain generated by f_0, f_1 and p.

This completes the proof of the theorem. □

Remark. Note that a crucial point in the proof above was that $\hat{Z}$ defined in (2) was $1-1$ and continuous. Thus all choices of parameter-values with $\sigma < \alpha$ in (1) will work and there is thus nothing special with our choice of parameter values corresponding to $\sigma = 1/3$ and $\alpha = 2/3$.

The stationary probability measures constructed in our example are singular w.r.t. Lebesgue measure. It is an interesting problem to find counterexamples in the 'overlapping' cases when $\sigma \geqslant \alpha$ and counterexamples giving stationary measures absolute continuous w.r.t. Lebesgue measure. A paper closely related to the latter question is the paper by Quas (1996).

Acknowledgements

I am grateful to John Elton, Jeff Geronimo and Jeff Steif for helpful conversations. Thanks also to Thomas Kaijser and Anthony Quas for clarifying remarks. This paper was written during a postdoctoral visit at the School of Mathematics of Georgia Institute of Technology. I am grateful to The Royal Physiographic Society in Lund (Hellmuth Hertz' Foundation) for financial support.

References

Barnsley, M.F., Demko, S.G., Elton, J.H., Geronimo, J.S., 1988. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. Ann. Inst. H. Poincaré Probab. Statist. 24, 367–394. (Erratum: 25:589–590, 1989).

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

Bramson, M., Kalikow, S., 1993. Nonuniqueness in g-functions. Israel J. Math. 84, 153–160.

Burton, R.M., Keller, G., 1993. Stationary measures for randomly chosen maps. J. Theoret. Probab. 6, 1–16.

Doeblin, W., Fortet, R., 1937. Sur des chaines a liaisons completès. Bull. Soc. Math. France 65, 132–148.

Fan, A.H., Lau, K-S., 1999. Iterated function system and Ruelle operator. J. Math. Anal. Appl. 231 (2), 319–344.

Harris, T.E., 1955. On chains of infinite order. Pacific J. Math. 5, 707–724.

Kaijser, T., 1979. Another central limit theorem for random systems with complete connections. Rev. Roumaine Math. Pures Appl. 24, 383–412.

Kaijser, T., 1981. On a new contraction condition for random systems with complete connections. Rev. Roumaine Math. Pures Appl. 26, 1075–1117.

Kaijser, T., 1994. On a theorem of Karlin. Acta Appl. Math. 34, 51–69.

Karlin, S., 1953. Some random walks arising in learning models. I. Pacific J. Math. 3, 725–756.

Keane, M., 1972. Strongly mixing g-measures. Invent. Math. 16, 309–324.

Li, Y., 1995. Limit theorems in reflected Brownian motions and in Markov chains associated with iterated function systems. Doctoral Thesis. The University of Connecticut.

Quas, A.N., 1996. Non-ergodicity for C^1 expanding maps and g-measures. Ergodic Theory Dynam. Systems 16, 531–543.

Walters, P., 1975. Ruelle's operator theorem and g-measures. Trans. Amer. Math. Soc. 214, 375–387.