COLLOQUIUM MATHEMATICUM

VOL. XXXVII 1977 FASC. 2

ITERATION OF RANDOM-VALUED FUNCTIONS ON THE UNIT INTERVAL

BY

KAROL BARON AND MAREK KUCZMA (KATOWICE)

Let $I = [0, 1]$ be the unit interval. If we are given a function $f\colon I \to I$, then we may form the iterates f^n of f and study the properties of the sequence $f^n(x)$ for $x \in I$. This is an important problem in the theory of equations and for many applications. However, it may happen that instead of the exact value of f at x we know only the probability distribution of this value. In the present paper we define iterates of such functions and give some simple results on the behaviour of these iterates.

In the sequel, $\boldsymbol{B}$ denotes the σ-algebra of the Borel subsets of I; and, for any σ-algebras $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, the σ-algebra generated by the sets $S_1 \times S_2$, $S_1 \in \boldsymbol{S}_1$, $S_2 \in \boldsymbol{S}_2$, is denoted by $\boldsymbol{S}_1 \mathsf{x} \boldsymbol{S}_2$.

1. We define a *random-valued function* (shortly, an rv-function) as a function $f\colon I \times \Omega \to I$, where $(\Omega, \boldsymbol{S}, \mathrm{P})$ is a probability space, such that for every set $B \in \boldsymbol{B}$ we have $f^{-1}(B) \in \boldsymbol{B}\mathsf{x}\boldsymbol{S}$.

Thus an rv-function is a measurable stochastic process for which the state space coincides with the time interval. For every fixed $x \in I$, $f(x) = f(x, \cdot)$ is a random variable on the space Ω of elementary events, with values in I.

Let $\mathscr{F}(x|\cdot)$ given by

$$\mathscr{F}(x|t) = \mathrm{P}[f(x) < t] \tag{1}$$

be the probability distribution of $f(x)$.

LEMMA 1. *For every fixed t, the function $\mathscr{F}(\cdot|t)$ is Borel measurable.*

Proof. It is enough to prove that for every set $A \in \boldsymbol{B}\mathsf{x}\boldsymbol{S}$ the function φ defined by

$$\varphi(x) = \mathrm{P}(A_x),$$

where $A_x = \{\omega \in \Omega\colon (x, \omega) \in A\}$, is Borel measurable. But this follows from the Fubini theorem (cf., e.g., [3], Theorem 35.A).

An rv-function is called *continuous at* $x_0 \in I$ if $x \to x_0$ implies that $f(x) \to f(x_0)$ in law, i.e., $\mathscr{F}(x|t) \to \mathscr{F}(x_0|t)$ at every point t at which $\mathscr{F}(x_0|\cdot)$ is continuous. f is called *continuous* if it is continuous at every point $x \in I$.

LEMMA 2. *If f is a continuous* rv-*function, with the probability distribution* (1), *and λ is a continuous real function on I, then the function a,*

$$a(x) = \int \lambda(y) d_y \mathscr{F}(x|y),$$

is continuous in I (¹).

This follows from the fact that the convergence in law is equivalent to the weak convergence.

Now let g and f be two rv-functions with the underlying probability spaces $(\Omega_1, \boldsymbol{S}_1, \mathrm{P}_1)$ and $(\Omega_2, \boldsymbol{S}_2, \mathrm{P}_2)$, and with the probability distributions

$$(2) \qquad \mathscr{G}(x|t) = \mathrm{P}_1[g(x) < t] \quad \text{and} \quad \mathscr{F}(x|t) = \mathrm{P}_2[f(x) < t],$$

respectively. We define the composition $f \circ g$ to be the function

$$f \circ g \colon I \times \Omega_1 \times \Omega_2 \to I$$

defined by

$$(3) \qquad (f \circ g)(x, \omega_1, \omega_2) = f\big(g(x, \omega_1), \omega_2\big).$$

LEMMA 3. *The composition $f \circ g$ is* rv-*function with the underlying probability space* $(\Omega_1 \times \Omega_2, \boldsymbol{S}_1 \times \boldsymbol{S}_2, \mathrm{P}_1 \times \mathrm{P}_2)$.

Proof. We shall prove that for every set $B \in \boldsymbol{B}$ we have

$$(f \circ g)^{-1}(B) \in \boldsymbol{B} \times \boldsymbol{S}_1 \times \boldsymbol{S}_2.$$

Define the function $G\colon I \times \Omega_1 \times \Omega_2 \to I \times \Omega_2$ by

$$G(x, \omega_1, \omega_2) = \big(g(x, \omega_1), \omega_2\big) = (g \times i)(x, \omega_1, \omega_2),$$

where $i\colon \Omega_2 \to \Omega_2$ is the identity. Thus $f \circ g = f \circ G$, where the second circle denotes the usual superposition of mappings. Hence

$$(f \circ g)^{-1}(B) = G^{-1}\big(f^{-1}(B)\big),$$

and the lemma will be proved when we show that

$$(4) \qquad G^{-1}(A) \in \boldsymbol{B} \times \boldsymbol{S}_1 \times \boldsymbol{S}_2$$

for every set $A \in \boldsymbol{B} \times \boldsymbol{S}_2$. Let $\boldsymbol{A}$ be the class of all those sets $A \subset I \times \Omega_2$ for which (4) holds. Then, clearly, $\boldsymbol{A}$ is a σ-algebra. Moreover, if $A = B \times S$ with $B \in \boldsymbol{B}$ and $S \in \boldsymbol{S}_2$, then

$$G^{-1}(A) = (g \times i)^{-1}(A) = g^{-1}(B) \times S \in \boldsymbol{B} \times \boldsymbol{S}_1 \times \boldsymbol{S}_2.$$

(¹) Unless otherwise indicated, the integration always extends over I.

Hence A contains all sets of the form $B \times S$ with $B \in \boldsymbol{B}$ and $S \in \boldsymbol{S}_2$, whence $\boldsymbol{B} \times \boldsymbol{S}_2 \subset \boldsymbol{A}$, which was to be proved.

Observe that the composition of rv-functions defined by (3) is associative.

LEMMA 4. *If g and f are rv-functions with the underlying probability spaces $(\Omega_1, \boldsymbol{S}_1, \mathrm{P}_1)$ and $(\Omega_2, \boldsymbol{S}_2, \mathrm{P}_2)$, and with the probability distributions* (2), *then the composition $f \circ g$ has the probability distribution given by*

$$(\mathrm{P}_1 \times \mathrm{P}_2)[(f \circ g)(x) < t] = \int \mathscr{F}(y\,|\,t)\, d_y \mathscr{G}(x\,|\,y). \tag{5}$$

Proof. With a fixed $x \in I$ and real t write

$$A = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 \colon f(g(x, \omega_1), \omega_2) < t\}$$

and

$$A_{\omega_1} = \{\omega_2 \in \Omega_2 \colon (\omega_1, \omega_2) \in A\}.$$

By the Fubini theorem we have

$$(\mathrm{P}_1 \times \mathrm{P}_2)(A) = \int_{\Omega_1} \mathrm{P}_2(A_{\omega_1})\, d\mathrm{P}_1 = \int_{\Omega_1} \mathscr{F}(g(x, \omega_1)\,|\,t)\, d\mathrm{P}_1.$$

Now Theorem 39.C in [3] yields

$$(\mathrm{P}_1 \times \mathrm{P}_2)(A) = \int \mathscr{F}(y\,|\,t)\, d_y(\mathrm{P}_1 \circ g^{-1}(x, y)) = \int \mathscr{F}(y\,|\,t)\, d_y \mathscr{G}(x\,|\,y),$$

which was to be proved.

The iterates f^n of an rv-function f are defined by the recurrence

$$f^1 = f, \quad f^{n+1} = f \circ f^n = f^n \circ f \ (n = 1, 2, \ldots).$$

Lemma 4 implies that the distribution functions $\mathscr{F}^n$ of the iterates f^n of an rv-function with distribution (1) are given by

$$\mathscr{F}^1(x\,|\,t) = \mathscr{F}(x\,|\,t),$$

(6)

$$\mathscr{F}^{n+1}(x\,|\,t) = \int \mathscr{F}(y\,|\,t)\, d_y \mathscr{F}^n(x\,|\,y) = \int \mathscr{F}^n(y\,|\,t)\, d_y \mathscr{F}(x\,|\,y) \quad (n = 1, 2, \ldots).$$

2. If $f \colon I \to I$ is a continuous real function fulfilling the inequality $f(x) < x$ for $x \in (0, 1]$, then, for every $x \in I$, the sequence $f^n(x)$ tends to zero (cf. [4], Theorem 0.4). Now we are going to extend this result to the case of random-valued functions.

Let f be an rv-function with distribution (1). Let $m(x)$,

$$m(x) = \mathrm{E}[f(x)] = \int t\, d_t \mathscr{F}(x\,|\,t), \tag{7}$$

be the mean of $f(x)$, and $\Phi(x\,|\,\cdot)$,

$$(8)\qquad \Phi(x\,|\,s) = \mathrm{E}\big[\exp\big(-sf(x)\big)\big] = \int \exp(-st)\,d_t\mathscr{F}(x\,|\,t) \qquad (s \geqslant 0),$$

be the Laplace transform of $f(x)$.

THEOREM 1. *Let f be a continuous* rv-*function with distribution* (1), *and let any of the following three conditions be fulfilled*:

(i) $\mathscr{F}(x\,|\,x) = 1$ *for* $x \in (0,1]$;

(ii) $\Phi(x\,|\,s) > e^{-sx}$ *for* $x \in (0,1]$ *and* $s > 0$;

(iii) $m(x) < x$ *for* $x \in (0,1]$.

Then, for every $x \in I$, the sequence $f^n(x)$ converges in law to zero.

Before proceeding with the proof, we shall show the following

LEMMA 5. *For any* rv-*function f, condition* (i) *implies both* (ii) *and* (iii). *Condition* (ii) *implies*

(iv) $m(x) \leqslant x$ *for* $x \in (0,1]$.

Proof. Fix an $x \in (0,1]$. Condition (i) and the left continuity of $\mathscr{F}(x\,|\,\cdot)$ yield the existence of a point $t_x \in (0,x)$ such that

$$(9)\qquad \mathscr{F}(x\,|\,t_x) > 0.$$

Now for any $s > 0$ we have

$$\begin{aligned}\Phi(x\,|\,s) &= \int_{[0,x]} \exp(-st)\,d_t\mathscr{F}(x\,|\,t)\\ &= \int_{[0,t_x)} \exp(-st)\,d_t\mathscr{F}(x\,|\,t) + \int_{[t_x,x)} \exp(-st)\,d_t\mathscr{F}(x\,|\,t)\\ &\geqslant \exp(-st_x)\mathscr{F}(x\,|\,t_x) + \exp(-sx)\big(1-\mathscr{F}(x\,|\,t_x)\big),\end{aligned}$$

whence (ii) follows in view of (9). Inequality (iii) may be obtained in an analogous way.

Now, inequality (ii) implies that, with a fixed $x \in (0,1]$,

$$(10)\qquad \frac{1}{s}\big(\Phi(x\,|\,s)-1\big) > \frac{1}{s}(e^{-sx}-1) \qquad (s > 0).$$

Letting $s \to 0$ in (10) we get

$$-m(x) = \frac{d}{ds}\Phi(x\,|\,s)\Big|_{s=0} \geqslant -x,$$

which is equivalent to (iv).

Proof of Theorem 1. It follows from the continuity of f and from any of conditions (i)-(iii) that $f(0)$ is concentrated at zero, i.e.,

$$\mathrm{P}[f(0) = 0] = 1,$$

which implies that, for every n, $f^n(0) = 0$ with probability 1. Thus it is enough to prove the convergence $f^n(x) \to 0$ for $x \in (0, 1]$. Further, by Lemma 5, we need to consider only cases (ii) and (iii).

(ii) Let $\Phi_n(x|\cdot)$,

$$\Phi_n(x|s) = \int e^{-st} d_t \mathscr{F}^n(x|t),$$

be the Laplace transform of $f^n(x)$. By (6) we have

$$\Phi_{n+1}(x|s) = \int e^{-st} d_t \int \mathscr{F}(y|t) d_y \mathscr{F}^n(x|y).$$

In virtue of Theorem 10.2 in [7], Chapter XI, and of Lemma 1 we get

$$\Phi_{n+1}(x|s) = \int \left(\int e^{-st} d_t \mathscr{F}(y|t)\right) d_y \mathscr{F}^n(x|y) = \int \Phi(y|s) d_y \mathscr{F}^n(x|y),$$

whence

$$\Phi_{n+1}(x|s) = \Phi_n(x|s) + \int \left(\Phi(y|s) - e^{-sy}\right) d_y \mathscr{F}^n(x|y). \tag{11}$$

Condition (ii) and Lemma 2 yield the inequality $\Phi(y|s) \geqslant e^{-sy}$ for $y \in I$, $s \geqslant 0$, thus we infer from (11) that the sequence $\Phi_n(x|s)$ is increasing, and hence convergent. By the continuity theorem ([2], Chapter XIII) the sequence of random variables $f^n(x)$ tends in law to a random variable $f_0(x)$ with a distribution $\mathscr{F}_0(x|\cdot)$. Letting $n \to \infty$ in (11), we get, in virtue of the continuity of $\Phi(\cdot|s)$ (Lemma 2),

$$\int \left(\Phi(y|s) - e^{-sy}\right) d_y \mathscr{F}_0(x|y) = 0. \tag{12}$$

Now, (12) and (ii) imply that

$$\mathscr{F}_0(x|t) = \begin{cases} 0 & \text{for } t \leqslant 0, \\ 1 & \text{for } t > 0, \end{cases} \tag{13}$$

i.e., $f_0(x) = 0$ with probability 1. This completes the proof in case (ii).

(iii) Let $m_n(x)$,

$$m_n(x) = \mathrm{E}[f^n(x)] = \int t d_t \mathscr{F}^n(x|t), \tag{14}$$

be the mean of $f^n(x)$. By an argument similar as in case (ii) we arrive at the relation

$$m_{n+1}(x) = m_n(x) - \int \left(y - m(y)\right) d_y \mathscr{F}^n(x|y), \tag{15}$$

which implies the monotonicity, and hence the convergence of the sequence $m_n(x)$. Now, by Helly's theorem [7], from every subsequence $\mathscr{F}^{n_k}(x|\cdot)$ of $\mathscr{F}^n(x|\cdot)$ we may choose a subsubsequence $\mathscr{F}^{n_{k_l}}(x|\cdot)$ which converges weakly to a probability distribution $\mathscr{F}_0(x|\cdot)$ (which *a priori* may depend on the subsequence n_k). Replacing in (15) n by n_{k_l} and letting $l \to \infty$ we obtain

$$\int \big(y - m(y)\big) d_y \mathscr{F}_0(x|y) = 0.$$

From (iii) it follows that $\mathscr{F}_0(x|\cdot)$ must be of form (13) and, consequently, is independent of the particular sequence n_k. Thus the sequence $\mathscr{F}^n(x|\cdot)$ must itself converge weakly to $\mathscr{F}_0(x|\cdot)$ given by (13), which completes the proof.

3. The n-th iterate $f^n(x)$ is a random variable on the probability space $(\Omega^n, \boldsymbol{S}^n, \mathrm{P}^n)$. However, all those product spaces may be embedded in the infinite product $(\Omega^\infty, \boldsymbol{S}^\infty, \mathrm{P}^\infty)$ (cf. [3], Theorem 38.B), and the functions f^n can be extended, in a natural way, onto $I \times \Omega^\infty$ by putting

$$\tilde{f}^n(x, \omega_1, \ldots, \omega_n, \ldots) = f^n(x, \omega_1, \ldots, \omega_n).$$

Thus $f^n(x)$ may be regarded as a Markov chain of random variables on the same probability space $(\Omega^\infty, \boldsymbol{S}^\infty, \mathrm{P}^\infty)$.

Denote by $\boldsymbol{F}_n$ the σ-algebra of sets $A \in \boldsymbol{S}^\infty$ which are of the form

$$A = A_n \times \mathop{\mathsf{X}}_{i=n+1}^{\infty} \Omega_i, \tag{16}$$

where $A_n \in \boldsymbol{S}^n$ and $\Omega_i = \Omega$ for $i = n+1, \ldots$ Clearly, $\boldsymbol{F}_n \subset \boldsymbol{F}_{n+1}$.

LEMMA 6. *If f is an* rv*-function, then mean* (7) *is Borel measurable.*

This follows from the Fubini theorem (cf., e.g., [6], Chapter II, Theorem 14, and [7], Chapter VIII, Theorem 7.1).

LEMMA 7. *The conditional expectation of $f^{n+1}(x)$ with respect to $\boldsymbol{F}_n$ is given by*

$$\mathrm{E}[f^{n+1}(x)|\boldsymbol{F}_n] = m\circ\big(f^n(x)\big).$$

Proof. It follows from Lemma 6 that $m\circ\big(f^n(x)\big)$ is measurable $\boldsymbol{F}_n$. Further, for any set $A \in \boldsymbol{F}_n$ (cf. (16)),

$$\int_A m\big(f^n(x, \omega_1, \omega_2, \ldots)\big) d\mathrm{P}^\infty = \int_{A_n} m\big(f^n(x, \omega_1, \ldots, \omega_n)\big) d\mathrm{P}^n$$

$$= \int_{A_n}\int_{\Omega} f\big(f^n(x, \omega_1, \ldots, \omega_n), \omega\big) d\mathrm{P}\, d\mathrm{P}^n = \int_{A_n \times \Omega} f^{n+1}(x, \omega_1, \ldots, \omega_n, \omega_{n+1}) d\mathrm{P}^{n+1}$$

$$= \int_A f^{n+1}(x, \omega_1, \omega_2, \ldots) d\mathrm{P}^\infty,$$

which completes the proof.

Lemma 7 allows us to improve the convergence occurring in Theorem 1. Namely, we have the following

THEOREM 2. *If f is an* rv*-function with mean* (7) *and if $m(x) \leqslant x$ for $x \in I$, then, for every $x \in I$, the sequence $f^n(x)$ converges with probability* 1.

Proof. In view of Lemma 7 we have

$$E[f^{n+1}(x) \mid \boldsymbol{F}_n] \leqslant f^n(x),$$

i.e., the sequence $f^n(x)$ is a supermartingale. Similarly as in the proof of Theorem 1 we arrive at formula (15) which shows that the sequence of means (14) is decreasing, and hence bounded. Thus Theorem 2 follows from the theorem on the convergence of supermartingales (cf. [5], p. 393).

Note that without assuming the continuity of f the sequence f^n need not converge to zero. However, Theorems 1 and 2 imply, in view of Lemma 5, the following

THEOREM 3. *Under assumptions of Theorem* 1, *for every* $x \in I$ *the sequence* $f^n(x)$ *converges to* 0 *with probability* 1.

Remarks. The weak convergence of the sequence $\mathscr{F}^n(x \mid \cdot)$ in the proof of Theorem 1 results immediately from Theorem 2. But the direct argument given there is almost so simple, and does not rely on the deep martingale theorem.

There are also possible other interpretations of the notions and results presented above. Distribution (1) may be interpreted as the transition probability from the state x to the interval $(-\infty, t)$; and f^n become iterates of the transition operator (cf. [1] and [2]).

The problem of iteration of rv-functions is different from that of stochastic approximation (cf. [8]). The latter consists in the investigation of the convergence of sequences of random variables X_n given by the recurrent relation $X_{n+1} = f(X_n)$, where, however, the shape of the function f is known. In our case we do not know the exact form of the function f either.

The authors are indebted to Mr. C. Ferens for his interesting discussions and valuable remarks.

REFERENCES

[1] Е. Б. Дынкин, *Марковские процессы*, Москва 1963.
[2] W. Feller, *An introduction to probability theory and its applications*, Vol. 2, New York 1966.
[3] P. Halmos, *Measure theory*, New York 1950.
[4] M. Kuczma, *Functional equations in a single variable*, Monografie Matematyczne 46, Warszawa 1968.
[5] M. Loève, *Probability theory*, second edition, New York 1960.
[6] P. A. Meyer, *Probability and potentials*, London 1966.
[7] R. Sikorski, *Funkcje rzeczywiste*, Vol. 1, Monografie Matematyczne 35, Warszawa 1958.
[8] M. T. Wasan, *Stochastic approximation*, Cambridge Tracts 58, Cambridge 1969.

Reçu par la Rédaction le 28. 6. 1974