Pavla Kunderová

On limit properties of the reward from a Markov replacement process

Sborník prací Přírodovědecké fakulty University Palackého v Olomouci. Matematika, Vol. 20 (1981), No. 1, 133--146

Persistent URL: http://dml.cz/dmlcz/120102

1981 — ACTA UNIVERSITATIS PALACKIANAE OLOMUCENSIS
FACULTAS RERUM NATURALIUM — TOM 69

*Katedra matematické analýzy a numerické matematiky přírodovědecké fakulty
University Palackého v Olomouci
Vedoucí katedry: prof. RNDr. Miroslav Laitoch, CSc.*

ON LIMIT PROPERTIES OF THE REWARD FROM A MARKOV REPLACEMENT PROCESS

PAVLA KUNDEROVÁ

(*Received March 31st, 1980*)

This paper is a close continuation of [7] and extends the validity of assertions proved there on replacement processes.

1. BASIC DEFINITIONS AND NOTATIONS

Let a homogeneous Markov process with rewards $\{X_t, t \geqq 0\}$ describing the evolution of a system in a state space $I = \{1, \dots, r\}$ be defined by exit intensities $(\mu(1), \dots, \mu(r))$, $0 < \mu(j) \leqq \infty$, $j = 1, \dots, r$ and by a matrix $\mathbf{P} = \| p(i,j) \|_{i,j=1}^{r}$ of transition probabilities in the moment of exit. Let us denote by $\mathbf{M} = \| \mu(i,j) \|_{i,j=1}^{r}$ the matrix of transition intensities of the process, where

$$\mu(i,j) = \mu(i)\, p(i,j) \text{ for } i \neq j,\ \mu(i,i) = -\mu(i) = -\sum_{j \neq i} \mu(i,j).$$

Consider a situation, where the development of the process may be influenced by an action called *replacement*. According to [5] we mean under a replacement of type $(i, +j)$ the instantaneous shift of the system from state i into state j. The complete history of this process is given by the following sequence

$$\omega = \{i_0, t_0, \delta_0; i_1, t_1, \delta_1; \dots; i_n, t_n, \delta_n; \dots\},$$

where $i_0, i_1, \dots, i_n, \dots$ are the states visited, $t_0, t_1, \dots, t_n, \dots$ the corresponding sojourn times and $\delta_0, \delta_1, \dots, \delta_n, \dots$ is the sequence of zeros and units, where $\delta_n = 0$ in case of $i_n \to i_{n+1}$ without interference and $\delta_n = 1$ in case of $i_n \to i_{n+1}$ being the replacement. We use in accordance with [5] the notation

$$\omega_n = \{i_0, \dots; i_{n-1}, t_{n-1}, \delta_{n-1}; i_n\}$$

for the history up to the n-th state change.

A replacement policy (see [5]) is a decision for all possible sequences ω_n for how long time the system will be left in i_n without shifting (maximal sojourn time) and in what state it is to be shifted. Since we do not to exclude the random choice of these quantities, we identify a replacement policy with a sequence of functions

$$F = \{{}^nF_k(t/\omega_n)\}, \qquad k = 1, \dots, r; n = 1, 2, \dots$$

where ${}^nF_k(t/\omega_n)$ is a probability that the maximal sojourn time in i_n will be less than t and the eventual shift will be into $k \neq i_n$.

Assumption 1.

Consider such replacement policies F only, where

a) *there exists only a finite number of replacements in every finite interval,*

b) *there are neither two or more replacements in the same moment, with probability 1.*

According to Assumption 1 there is assigned a trajectory $\{Y_t, t \geqq 0\}$ not left continuous at the time of transition and not right continuous at the time of replacement to almost every ω.

In what follows we denote by

$\sigma_0 = 0, \sigma_1, \sigma_2, \dots$ such moments in which the trajectory is discontinuous,

$$Y_t^- = Y_{t-}, \ t > 0; \ Y_0^- = Y_0; \ Y_t^+ = Y_{t+}, \ t \geqq 0;$$

$\mathscr{B}_t = \sigma a \ \{(Y_s = j), j \in I, s \in \langle 0, t\rangle; \text{ events of zero probability}\}$,

$$\mathscr{B}_t^+ = \bigcap_{s>t} \mathscr{B}_s,$$

E^F a mathematical expectation in a replacement process under the replacement policy F,

D a set of couples $(i, +j)$ meaning admissible replacements,

$$D_i = \{j: (i, +j) \in D\}.$$

The reward from the process is defined by the following sets of numbers:

$\varrho(i)$, $i \in I$, the reward per a time unit in state i,

$r(i,j)$, $i, j \in I$, the reward from the transition (i,j); we set $r(i, i) = 0$,

$v(i,j)$, $i, j \in I$, the reward from the replacement $(i, +j)$; we set $v(i, i) = 0$.

A stationary replacement policy f is given by a function $f(j)$ defined on a subset $I_f \subset I$ and taking values in I such that $f(j) \in D_j$ for $j \in I_f, f(j) \neq j$. The replacement policy f is the prescription to realize instantaneously the replacement $j \to f(j)$ whenever there occurs a transition in state j. No replacements occur in states $j \notin I_f$.

Assumption 2.

$$(i, +j) \in D, (j, +k) \in D \Rightarrow (i, +k) \in D \qquad \text{or} \qquad i = k,$$
$$v(i,j) + v(j, k) \leqq v(i, k).$$

Let R_T be a *reward from the process* up to the time T. In accordance with our previous definitions

$$R_T = \int_0^T \varrho(Y_t)\,\mathrm{d}t + \sum_{n=0}^N [r(Y_{\sigma_n}^-, Y_{\sigma_n}) + v(Y_{\sigma_n}, Y_{\sigma_n}^+)], \qquad \sigma_N \leqq T < \sigma_{N+1}.$$

2. LIMIT PROPERTIES OF A REWARD

We demonstrate first some auxiliary assertions.

Lemma 1.

Let $g(i, k)$ be a function defined on $I \times I$, $g(i, i) = 0$, $i \in I$. Let

$$G_T = \sum_{n=1}^{N} g(Y_{\sigma_n}^-, Y_{\sigma_n}), \qquad \sigma_N \leqq T < \sigma_{N+1},$$

introduce

$$\gamma(i) = \sum_{k \neq i} \mu(i, k)\, g(i, k), \qquad \gamma_2(i) = \sum_{k \neq i} \mu(i, k)\, (g(i, k))^2.$$

Then it holds under an arbitrary replacement policy F for $0 \leqq t \leqq T$

$$E^F\{G_T - G_t/\mathscr{B}_t^+\} = E^F\{\int_t^T \gamma(Y_s)\, ds/\mathscr{B}_t^+\}, \tag{1}$$

$$E^F\{(G_T - G_t - \int_t^T \gamma(Y_s)\, ds)^2/\mathscr{B}_t^+\} = E^F\{\int_t^T \gamma_2(Y_s)\, ds/\mathscr{B}_t^+\}. \tag{2}$$

Proof: a) Since the conditional distribution describes a Markov replacement process under common replacement policy, the proof of (1) reduces to the verification of

$$E^{F'}(G_T) = E^{F'}\{\int_0^T \gamma(Y_s)\, ds\}, \qquad T \geqq 0,$$

for an arbitrary initial probability distribution and an arbitrary policy F'.

The proof of the above assertion proceeds similarly to that of Lemma 1 in [6].

b) Taking instead of $g(i, k)$ the function $g^2(i, k)$ throughout the proof of (1) we show that

$$E^{F'}(\sum_{n=1}^{N} g^2(Y_{\sigma_n}^-, Y_{\sigma_n})) = E^{F'}(\int_0^T \gamma_2(Y_s)\, ds), \qquad \sigma_N \leqq T < \sigma_{N+1}.$$

Then (2) will be established by proving

$$E^{F'}(G_T - \int_0^T \gamma(Y_s)\, ds)^2 = E^{F'}(\sum_{n=1}^{N} g^2(Y_{\sigma_n}^-, Y_{\sigma_n})),$$

under an arbitrary policy F' and an arbitrary initial distribution. The proof proceeds analogous to that of Corollary 1 in [6].

Lemma 2.

There exist constants K_{mT} such that

$$E^F |G_T|^m \leqq K_{mT}[\max_{i, j \in I} (|g(i, j)|)]^m, \qquad m = 1, 2, \ldots, \tag{3}$$

for an arbitrary replacement policy F.

Proof: We denote by $\mu = \max(\mu(1), \ldots, \mu(r))$, σ'_n the moment of the n-th transition (the n-the left discontinuity of the trajectory). We prove by induction

$$P^F(\sigma'_n \leqq t) \leqq H^{(n)}(t), \tag{4}$$

where $H^{(n)}(t)$ is the n-multiple convolution $H^{(1)}(t) = 1 - e^{-\mu t}$. We denote by N'_T the number of transitions in $\langle 0, T \rangle$. According to (4) it holds

$$E^F(N'_T)^m = \sum_{n=1}^{\infty} n^m [P^F(\sigma'_n \leqq T) - P^F(\sigma'_{n+1} \leqq T)] \leqq$$
$$\leqq \sum_{n=1}^{\infty} (n^m - (n-1)^m) H^{(n)}(T) = \sum_{n=1}^{\infty} (n^m - (n-1)^m) \frac{1}{(n-1)!} \int_0^{\mu T} x^{n-1} e^{-x} \, \mathrm{d}x =$$
$$= \sum_{n=1}^{\infty} n^m \frac{(\mu T)^n}{n!} e^{-\mu T} = K_{mT}.$$

Thus

$$E^F \mid G_T \mid^m = E^F(\mid \sum_{n=1}^{N} g(Y^-_{\sigma_n}, Y_{\sigma_n}) \mid^m) = E^F(\mid \sum_{j=1}^{N'_T} g(Y^-_{\sigma'_j}, Y_{\sigma'_j}) \mid^m) \leqq$$
$$\leqq E^F[(N'_T)^m (\max_{i,k \in I} \{\mid g(i, k) \mid\})^m] \leqq (\max_{i,k \in I} \{\mid g(i, k) \mid\})^m \, . \, K_{mT}. \square$$

Let f be a fixed chosen stationary replacement policy such that under it exists one recurrent class and eventually a transient class only. Let the constant Θ, $w(1), \ldots, w(r)$ be defined by the following equations

$$v(i, f(i)) + w(f(i)) - w(i) = 0, \qquad i \in I_f, \tag{5}$$
$$\varrho(i) + \sum_{k \neq i} \mu(i, k) [r(i, k) + w(k) - w(i)] - \Theta = 0, \qquad i \notin I_f.$$

According to [2] the system (5) uniquely determines the number Θ (Θ is the mean reward per a time unit from the process in using the replacement policy f), $w(1), \ldots, w(r)$ except for adding an arbitrary constant.

Denote for $i \in I$

$$\varphi(i) = \varrho(i) + \sum_{k \neq i} \mu(i, k) [r(i, k) + w(k) - w(i)] - \Theta,$$
$$\psi_1(i) = \sum_{k \neq i} \mu(i, k) [r(i, k) + w(k) - w(i)],$$
$$\psi_2(i) = \sum_{k \neq i} \mu(i, k) [r(i, k) + w(k) - w(i)]^2.$$

Let us introduce an auxiliary random process (see [3])

$$M_T = R_T - \Theta T + w(Y^+_T) - w(Y_0) - \int_0^T \varphi(Y_t) \, \mathrm{d}t -$$
$$- \sum_{n=0}^{N} [v(Y_{\sigma_n}, Y^+_{\sigma_n}) + w(Y^+_{\sigma_n}) - w(Y_{\sigma_n})], \qquad T \geqq 0, \sigma_N \leqq T < \sigma_{N+1}.$$

Lemma 3.

$\{M_T, T \geqq 0\}$ is a martingale with respect to $\{\mathscr{B}_T^+, T \geqq 0\}$ under an arbitrary policy F. It holds for $0 \leqq t \leqq T$

$$E^F\{(M_T - M_t)^2 / \mathscr{B}_t^+\} = E^F\{\int_t^T \psi_2(Y_s)\,ds / \mathscr{B}_t^+\} \; F\text{-almost everywhere.}$$

Proof: By substituting instead of R_T and $w(Y_T^+) - w(Y_0) = \sum_{n=0}^{N} [w(Y_{\sigma_n}) - w(Y_{\sigma_n}^-) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})]$, $\sigma_N \leqq T < \sigma_{N+1}$, into the expression for M_T we obtain

$$M_T = -\int_0^T \psi_1(Y_t)\,dt + \sum_{n=0}^{N} [r(Y_{\sigma_n}^-, Y_{\sigma_n}) + w(Y_{\sigma_n}) - w(Y_{\sigma_n}^-)].$$

The substitution of $g(i, k) = r(i, k) + w(k) - w(i)$ in (1) of Lemma 1 gives

$$E^F\{M_T - M_t / \mathscr{B}_t^+\} = E^F\{G_T - G_t - \int_t^T \gamma(Y_s)\,ds / \mathscr{B}_t^+\} = 0, \qquad t \leqq T,$$

and thus

$$E^F\{M_T / \mathscr{B}_t^+\} = E^F\{M_t / \mathscr{B}_t^+\} = M_t \qquad \text{for all } t \leqq T.$$

The other assertion proved follows analogous from (2), Lemma 1. □

Corollary.

Under an arbitrary replacement policy F

$$\lim_{n \to \infty} \frac{1}{T} M_T = 0 \qquad F\text{-almost everywhere.} \tag{6}$$

Proof: 1. We can write $M_n = \sum_{k=1}^{n} (M_k - M_{k-1})$. According to Lemma 3

$$\sum_{n=1}^{\infty} \frac{1}{n^2} E(M_n - M_{n-1})^2 =$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^2} E(\int_{n-1}^{n} \psi_2(Y_s)\,ds) \leqq \sum_{n=1}^{\infty} \frac{1}{n^2} (\max_{i \in I} \{\psi_2(i)\}),$$

and $\{M_n, n = 1, 2, \ldots\}$ being a martingale, it is by [4], page 407

$$\lim_{n \to \infty} \frac{1}{n} M_n = 0 \qquad F\text{-almost everywhere.} \tag{7}$$

2. Let $n \leqq T < n + 1$, then

$$\left| \frac{1}{T} M_T \right| \leqq \frac{1}{n} \sup_{n \leqq T < n+1} | M_T - M_n | + \frac{1}{n} | M_n |.$$

According to (7) it suffices to prove that

$$\lim_{n \to \infty} \frac{1}{n} \sup_{n \leqq T < n+1} | M_T - M_n | = 0 \qquad F\text{-almost everywhere.} \tag{8}$$

Denote by

$$c = \max_{i \in I} \{\psi_1(i)\}, \qquad k = \max_{i, j \in I} \{| r(i, j) + w(j) - w(i) |\},$$

X_n the number of transitions during the time $\langle n, n + 1)$. Then

$$\sup_{n \leqq T < n+1} | M_T - M_n | \leqq c + kX_n. \tag{9}$$

As the series $E(\sum_{n=1}^{\infty} \frac{1}{n^2} X_n^2)$ converges, it is $\lim_{n \to \infty} \frac{1}{n^2} X_n^2 = 0$ F-almost everywhere. Hence

$$\lim_{n \to \infty} \frac{1}{n} (c + kX_n) = 0 \qquad F\text{-almost everywhere.}$$

This due to (9) proves (8). □

Theorem 1.

Let the optimality equation (see [2]) for the replacement policy f hold, i.e.

$$\max \{v(j, k) + w(k) - w(j), k \in D_j; \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k) - w(j)] - \Theta\} = 0,$$

$$j \in I. \tag{10}$$

Then under an arbitrary policy F

$$\limsup_{n \to \infty} \frac{1}{T} R_T \leqq \Theta \qquad F\text{-almost everywhere.}$$

Proof:

It follows from assumption (10) that $\varphi(j) \leqq 0$ for all $j \in I$, i.e.

$$-\int_0^T \varphi(Y_t) \, dt \geqq 0.$$

Likewise, we have from (10)

$$-\sum_{n=0}^{N} [v(Y_{\sigma_n}, Y_{\sigma_n}^+) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})] \geqq 0.$$

Thus

$$M_T \geqq R_T - \Theta T + w(Y_T^+) - w(Y_0).$$

Since

$$\lim_{T \to \infty} \frac{1}{T} [w(Y_T^+) - w(Y_0)] = 0, \tag{11}$$

it holds

$$\limsup_{T \to \infty} \frac{1}{T} M_T \geqq \limsup_{T \to \infty} \frac{1}{T} R_T - \Theta \qquad F\text{-almost everywhere}$$

whence the statement follows from Corollary 3. □

Definitions.

We call the *state $i \in I$ consistent with the policy f*, if $\varphi(i) = 0$. We call *the replacement $i \to k$ consistent with f*, if $v(i, k) + w(k) - w(i) = 0$.

Denote by

Q_T the whole sojourn time in the inconsistent states in $\langle 0, T \rangle$,

$\bar{Q}_T$ the whole sojourn time in states I_f in the interval $\langle 0, T \rangle$,

O_T the whole number of inconsistent replacements in $\langle 0, T \rangle$,

$\bar{O}_T$ the whole number of replacements diferent from $i \to f(i)$ in $\langle 0, T \rangle$.

Obviously

$$\bar{Q}_T \geqq Q_T, \qquad \bar{O}_T \geqq O_T.$$

Theorem 2.

Let F be a replacement policy. If

$$\lim_{T \to \infty} \frac{1}{T} Q_T = \lim_{T \to \infty} \frac{1}{T} O_T = 0 \qquad F\text{-almost everywhere } (F\text{-in probability}) \tag{12}$$

then

$$\lim_{T \to \infty} \frac{1}{T} R_T = \Theta \qquad F\text{-almost everywhere } (F\text{-in probability}). \tag{13}$$

If the equation of optimality (10) *is valid, then* (12) *is necessary for the validity of* (13) *as well.*

Proof:

$$M_T = R_T - \Theta T + w(Y_T^+) - w(Y_0) - \int_0^T \varphi(Y_t) \, \mathrm{d}t -$$
$$- \sum_{n=0}^{N} [v(Y_{\sigma_n}, Y_{\sigma_n}^+) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})], \qquad \sigma_N \leqq T < \sigma_{N+1}.$$

a) The function $\varphi(.)$ is constant in any interval $\langle \sigma_{j-1}, \sigma_j)$. If i is a consistent state with f, then $\varphi(i) = 0$ and thus

$$\min_{i \in I} \{\varphi(i)\} Q_T \leqq \int_0^T \varphi(Y_t) \, \mathrm{d}t \leqq \max_{i \in I} \{\varphi(i)\} Q_T.$$

There are nonzero addends in the last sum of the expression M_T in those moments σ_n only, where an inconsistent replacement with f occurs, hence

$$\min_{i, j \in I} \{v(i, j) + w(j) - w(i)\} O_T \leqq \sum_{n=0}^{N} [v(Y_{\sigma_n}, Y_{\sigma_n}^+) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})] \leqq$$
$$\leqq \max_{i, j \in I} \{v(i, j) + w(j) - w(i)\} O_T.$$

The above relations prove together with (6) and (11) that (12) follows from (13).

b) Let (13) hold and let f fulfils (10). If i is the state consistent with f, then $\varphi(i) = 0$. In the opposite case then $i \in I_f$ and according to (10) $\varphi(i) < 0$.

Denote by I_0 the set of inconsistent states with f. According to (13)

$$0 \geqq \max_{i \in I_0} \{\varphi(i)\} \frac{Q_T}{T} \geqq \frac{1}{T} \int_0^T \varphi(Y_t)\,dt \to 0 \qquad \text{for } T \to \infty.$$

The nonzero expressions are in the sum

$$\sum_{n=0}^{N} [v(Y_{\sigma_n}, Y_{\sigma_n}^+) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})]$$

in those moments σ_n, if there is in F a transition or a replacement consistent with f. If (10), (13) hold, then

$$0 \geqq \max_{i \to k \text{ replacements inconsistent with } f} \{v(i, k) + w(k) - w(i)\} \frac{O_T}{T} \geqq$$

$$\geqq \frac{1}{T} \sum_{n=0}^{N} [v(Y_{\sigma_n}, Y_{\sigma_n}^+) + w(Y_{\sigma_n}^+) - w(Y_{\sigma_n})] \to 0 \qquad \text{for } T \to \infty.$$

Hence, if (10) holds, then (12) is necessary for (13) to be fulfilled. □

Theorem 3.

Let F be a replacement policy. Let

$$\lim_{T \to \infty} \frac{1}{\sqrt{T}} \bar{Q}_T = 0 = \lim_{T \to \infty} \frac{1}{\sqrt{T}} \bar{O}_T \qquad F\text{-in probability} \tag{14}$$

then

$$\frac{R_T - \Theta T}{\sqrt{T}}$$

has for $T \to \infty$ asymptotically normal distribution $N(0, \zeta)$, where ζ is determined by equations

$$w_2(f(i)) - w_2(i) = 0, \qquad i \in I_f,$$

$$\psi_2(i) + \sum_{k \neq i} \mu(i, k) [w_2(k) - w_2(i)] - \zeta = 0, \qquad i \notin I_f,$$

containing auxiliary constants $w_2(1), \ldots, w_2(r)$.

Proof: We prove this theorem in several steps.

I. *We prove first that it follows from* (14)

$$\lim_{n \to \infty} \left(\frac{R_n - \Theta n}{\sqrt{n}} - \frac{M_n}{\sqrt{n}} \right) = 0 \qquad F\text{-in probability.}$$

According to definition

$$M_n = R_n - \Theta n + w(Y_n^+) - w(Y_0) - \int_0^n \varphi(Y_t)\,dt -$$

$$- \sum_{j=0}^{N} [v(Y_{\sigma_j}, Y_{\sigma_j}^+) + w(Y_{\sigma_j}^+) - w(Y_{\sigma_j})], \qquad \sigma_N \leqq n < \sigma_{N+1}.$$

Obviously

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} [w(Y_n^+) - w(Y_0)] = 0.$$

Since

$$\min_{i\in I} \{\varphi(i)\} Q_n \leqq \int_0^n \varphi(Y_t)\,dt \leqq \max_{i\in I} \{\varphi(i)\} Q_n,$$

$$\min_{i,k\in I} \{v(i, k) + w(k) - w(i)\} O_n \leqq \sum_{j=0}^{N} [v(Y_{\sigma_j}, Y_{\sigma_j}^+) + w(Y_{\sigma_j}^+) - w(Y_{\sigma_j})] \leqq$$
$$\leqq \max_{i,k\in I} \{v(i, k) + w(k) - w(i)\} O_n,$$

(see the proof of Theorem 2)
assertion I follows from (14) by using $Q_n \leqq \bar{Q}_n$, $O_n \leqq \bar{O}_n$.

II. $\frac{M_n}{\sqrt{n}}$ *has for* $n \to \infty$ *asymptotically normal distribution* $N(0, \zeta)$.

The proof of the above statement lies in the verification of assumptions of the central limit theorem for martingales below (see [1], [7]):

Let $\{M_n = \sum_{m=0}^{n-1} Y_m,\ n = 1, 2, \ldots\}$ be a martingale with respect to the class of σ-algebras $\{\mathscr{F}_n, n = 1, 2, \ldots\}$. Let

(i) $\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\{Y_m^2 \cdot \chi_{\{|Y_m| \geqq \varepsilon\sqrt{n}\}}/\mathscr{F}_m\} = 0$ in probability for all $\varepsilon > 0$,

(ii) $\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\{Y_m^2/\mathscr{F}_m\} = \zeta$ in probability, where ζ is a constant,

then $\frac{M_n}{\sqrt{n}}$ is asymptotically normal $N(0, \zeta)$ for $n \to \infty$.

In our case we have $M_n = \sum_{m=0}^{n-1} (M_{m+1} - M_m)$. By Lemma 3 $\{M_n, n = 1, 2, \ldots\}$ is a martingale with respect to the class of σ-algebras $\{\mathscr{B}_n^+, n = 1, 2, \ldots\}$.

1. Let $\varepsilon > 0$ be an arbitrary number. Then

$$E\{(M_{m+1} - M_m)^2 \chi_{\{|M_{m+1} - M_m| \geqq \varepsilon\sqrt{n}\}}/\mathscr{B}_m^+\} \leqq \frac{1}{\varepsilon\sqrt{n}} E\{| M_{m+1} - M_m |^3/\mathscr{B}_m^+\}.$$

To the proof of

$$E\{| M_{m+1} - M_m |^3/\mathscr{B}_m^+\} \leqq c, \qquad c \text{ constant}, \tag{15}$$

it is sufficient to show that under an arbitrary replacement policy F' and under an arbitrary distribution

$$E^{F'}(| M_1 - M_0 |^3) \leqq c.$$

As $M_0 = 0$, we have (using the notation of the proof in Lemma 3)

$$E^{F'}(| M_1 |^3) = E^{F'}(| G_1 - \int_0^1 \gamma(Y_s)\,ds |^3) \leqq K_{1,3} + 3kK_{1,2} + 3k^2K_{1,1} + k^3 = c,$$

where $k = \max_{i \in I} \{| \gamma(i) |\}$ and where according to Lemma 2 $E^{F'}(| G_1 |^m) \leqq K_{1,m}$, $m = 1, 2, \ldots$ The realization of (i) follows then from (15).

2. Let the numbers $w_2(1), \ldots, w_2(r)$, ζ be solutions of the system of equations from the statement of the theorem. Let us define to the verification of (ii)

$$\varphi_2(i) = \psi_2(i) + \sum_{k \neq i} \mu(i, k) [w_2(k) - w_2(i)] - \zeta, \qquad i \in I.$$

a) We prove that under an arbitrary policy F

$$U_T = \int_0^T \psi_2(Y_t) \, \mathrm{d}t - \zeta T + w_2(Y_T^+) - w_2(Y_0) - \int_0^T \varphi_2(Y_t) \, \mathrm{d}t - \sum_{n=0}^N [w_2(Y_{\sigma_n}^+) - w_2(Y_{\sigma_n})],$$
$$T \geqq 0, \qquad \sigma_N \leqq T < \sigma_{N+1},$$

is a martingale with respect to $\{\mathscr{B}_T^+, T \geqq 0\}$ satisfying the law of large numbers. Denote

$$\xi_1(i) = \sum_{k \neq i} \mu(i, k) [w_2(k) - w_2(i)], \qquad i \in I.$$

On substituting and modifying we get

$$U_T = \sum_{n=0}^N [w_2(Y_{\sigma_n}) - w_2(Y_{\sigma_n}^-)] - \int_0^T \xi_1(Y_t) \, \mathrm{d}t.$$

Using Lemma 1 for $\gamma(i, k) = w_2(k) - w_2(i)$ gives

$$E(U_T - U_t / \mathscr{B}_t^+) = E(G_T - G_t - \int_t^T \gamma(Y_s) \, \mathrm{d}s / \mathscr{B}_t^+) = 0, \qquad t \leqq T, \tag{16}$$

$$E\{(U_T - U_t)^2 / \mathscr{B}_t^+\} = E\{\int_t^T \xi_2(Y_s) \, \mathrm{d}s / \mathscr{B}_t^+\}, \qquad t \leqq T, \tag{17}$$

where

$$\xi_2(i) = \sum_{k \neq i} \mu(i, k) [w_2(k) - w_2(i)]^2, \qquad i \in I.$$

It follows from (16) that $\{U_T, T \geqq 0\}$ is a martingale with respect to $\{\mathscr{B}_t^+, t \geqq 0\}$ and thus from (17) in the same manner as in the Corollary of Lemma 3

$$\lim_{T \to \infty} \frac{1}{T} U_T = 0, \qquad F\text{-almost everywhere.} \tag{18}$$

We can prove sililarly as in the proof of Theorem 2 that under the validity of (14)

$$\lim_{T \to \infty} \frac{1}{T} \{w_2(Y_T^+) - w_2(Y_0) - \int_0^T \varphi_2(Y_t) \, \mathrm{d}t - \sum_{n=0}^N [w_2(Y_{\sigma_n}^+) - w_2(Y_{\sigma_n})]\} = 0,$$
$$F\text{-in probability} \tag{19}$$

and thus from (18) and from the definition of U_T

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \psi_2(Y_t) \, \mathrm{d}t = \zeta, \qquad F\text{-in probability.} \tag{20}$$

b) We shall prove further that $\{D_n, n = 1, 2, \ldots\}$, where (see [7])

$$D_n = \int_0^n \psi_2(Y_t)\,\mathrm{d}t - \sum_{m=0}^{n-1} E\{(M_{m+1} - M_m)^2/\mathscr{B}_m^+\},$$

is a martingale with respect to $\{\mathscr{B}_n^+, n = 1, 2, \ldots\}$, for which the law of large numbers holds.

According to Lemma 3 we can write

$$D_n = \int_0^n \psi_2(Y_t)\,\mathrm{d}t - \sum_{m=0}^{n-1} E\{\int_m^{m+1} \psi_2(Y_t)\,\mathrm{d}t/\mathscr{B}_m^+\}.$$

For each $m \leqq n$ natural numbers

$$E\{D_n/\mathscr{B}_m^+\} = \int_0^m \psi_2(I_t)\,\mathrm{d}t - \sum_{j=0}^{m-1} E\{\int_j^{j+1} \psi_2(Y_t)\,\mathrm{d}t/\mathscr{B}_j^+\} = D_m.$$

If we denote

$$Y_m = \int_m^{m+1} \psi_2(Y_t)\,\mathrm{d}t - E\{\int_m^{m+1} \psi_2(Y_t)\,\mathrm{d}t/\mathscr{B}_m^+\},$$

then

$$D_n = \sum_{m=0}^{n-1} Y_m.$$

As for arbitrary $m = 0, 1, \ldots$

$$EY_m^2 \leqq E(\int_m^{m+1} \psi_2(Y_t)\,\mathrm{d}t)^2 \leqq c^2,$$

where

$$c = \max_{i \in I} \{\psi_2(i)\},$$

is the series

$$\sum_{m=0}^{\infty} \frac{EY_m^2}{(m+1)^2}$$

convergent and by [4], page 407

$$\lim_{n\to\infty} \frac{1}{n} D_n = 0 \qquad F\text{-almost everywhere.} \tag{21}$$

It is obvious from (20) and (21) that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\{(M_{m+1} - M_m)^2/\mathscr{B}_m^+\} = \zeta \qquad F\text{-in probability.}$$

it means the assumption (ii) for martingale $\{M_n, n = 1, 2, \ldots\}$ is valid.

In parts I and II of the proof we have proved the following assertion: Let (14) be valid, then $\frac{R_n - \Theta n}{\sqrt{n}}$ has for $n \to \infty$ asymptotically normal distribution $N(0, \zeta)$.

III. *Analogous to part I of this proof we can verify that*

$$\lim_{T\to\infty}\left(\frac{R_T-\Theta T}{\sqrt{T}}-\frac{M_T}{\sqrt{T}}\right)=0 \qquad F\text{-in probability.}$$

IV. *To conclude the proof we establish* $\frac{M_T}{\sqrt{T}}$ *having for* $T\to\infty$ *asymptotically normal distribution* $N(0, \zeta)$.

Let $n \leqq T < n + 1$. We know (see the proof of Lemma 3) that

$$E(M_T - M_n)^2 \leqq \max_{i\in I}\{\psi_2(i)\} = c,$$

and thus

$$EM_T^2 \leqq cT.$$

Hence

$$E\left(\frac{M_T}{\sqrt{T}}-\frac{M_n}{\sqrt{n}}\right)^2 = E\left[M_T\left(\frac{1}{\sqrt{T}}-\frac{1}{\sqrt{n}}\right)+\frac{1}{\sqrt{n}}(M_T-M_n)\right]^2 \leqq$$

$$\leqq 2\left[\frac{1}{T}\left(1-\sqrt{\frac{T}{n}}\right)^2 cT+\frac{c}{n}\right] \leqq 2c\left[\left(1-\sqrt{1-\frac{1}{T}}\right)^2+\frac{1}{T-1}\right],$$

and thus

$$\lim_{T\to\infty} E\left(\frac{M_T}{\sqrt{T}}-\frac{M_n}{\sqrt{n}}\right)^2 = 0.$$

Using Chebyshev inequality we get

$$\lim_{T\to\infty}\left(\frac{M_T}{\sqrt{T}}-\frac{M_n}{\sqrt{n}}\right)=0 \qquad F\text{-in probability}$$

and the assertion IV. follows from assertion II.

Theorem 3 is proved by III. and IV. □

In writing this article I have benefited from the advice and criticism given by dr. Petr Mandl, DrSc. whom I wish to acknowledge my gratitude.

REFERENCES

[1] Brown, P. M., Eagleson, G. K.: *Martingale convergence to infinitely divisible laws with finite variences.* Trans. Amer. Math. Soc. 162, p. 449.

[2] Kunderová, P.: *On a mean reward from Markov replacement process with only one isolated class of recurent states,* Acta UP Olomucensis, F.R.N. 1979, Tom 55.

[3] Kunderová, P.: *Kandidátská disertační práce,* UK Praha, 1977.

[4] Loève, M.: *Probability Theory.* Princeton. (Russian translation Moscow 1962).

[5] Mandl, P.: *An identity for Markovian replacement processes.* J. Appl. Prob., Vol. 6, No 2, pp. 348—354.

[6] Mandl, P.: *On the adaptive Control of Finite State Markov Processes,* Z. Wahrscheinlichkeitstheorie verw. Geb. 27, 263—276 (1973).

[7] Mandl, P.: *Some applications of martingales in controlled Markov processes.* Bulletin of the International Statistical Institute, Vol. 4, 1973, Vienna.

SOUHRN

LIMITNÍ VLASTNOSTI VÝNOSU Z MARKOVOVA PROCESU S OBNOVAMI

PAVLA KUNDEROVÁ

Článek úzce navazuje na [7] a rozšiřuje platnost tam uvedených tvrzení pro procesy s obnovami (viz [5]). Nechť R_T je výnos z procesu za dobu $\langle 0, T \rangle$, Θ průměrný výnos na jednotku času při užití stacionární strategie f při níž existuje pouze jedna třída rekurentních stavů. Je dokázáno (věta 1), že je-li f optimální (viz [2]), je při libovolné strategii obnovy F

$$\limsup_{T \to \infty} \frac{1}{T} R_T \leqq \Theta \qquad F \text{ — skoro všude.}$$

Zavádí se pojem souhlasné obnovy a souhlasného stavu se strategií f. Věta 2 uvádí podmínky postačující resp. nutné k tomu, aby

$$\lim_{T \to \infty} \frac{1}{T} R_T = \Theta \qquad F \text{ – skoro všude } (F \text{ – podle pravděpodobnosti}).$$

Jsou formulovány podmínky (věta 3), za nichž má $\frac{R_T - \Theta T}{\sqrt{T}}$ pro $T \to \infty$ asymptoticky normální rozdělení $N(0, \zeta)$, kde ζ je jistá konstanta.

РЕЗЮМЕ

ПРЕДЕЛЬНЫЕ КАЧЕСТВА ДОХОДА ИЗ ПРОЦЕССА МАРКОВА С ВОССТАНОВЛЕНИЯМИ

ПАВЛА КУНДЕРОВА

В работе обобщаются теоремы формулированные в [7] для управляемых процессов Маркова. Пусть R_T доход из процесса в течение интервала $\langle 0, T \rangle$, Θ средний доход за единицу времени, когда множество состояний процесса при использовании стационарной стратегии f имеет единственный класс возвратных состояний. Показано достаточное условие для того, чтобы для любой стратегии F

$$\limsup_{T \to \infty} \frac{1}{T} R_T \leqq \Theta$$

F — почти наверное (теорема 1). Определены согласное восстановление и согласное состояние со стратегией f. Решение проблемы об асимптотическом распределении $\frac{R_T - \Theta T}{\sqrt{T}}$ при $T \to \infty$ находится в теореме 3. Теорема 2 устанавливает условия для того чтобы

$$\lim_{T\to\infty} \frac{1}{T} R_T = \Theta \qquad F \text{ — почти наверное } (F \text{ — по вероятности}).$$