The Annals of Probability
1974, Vol. 2, No. 1, 103–111

THE LIMITING DISTRIBUTION OF MAXIMA OF RANDOM VARIABLES DEFINED ON A DENUMERABLE MARKOV CHAIN[1]

BY GEORGE O'BRIEN

York University

Gnedenko's work (1943) showing the existence of a limiting distribution for the maximum of the first n terms in a sequence of suitably normalized independent identically distributed random variables is extended to sequences of random variables defined on a positive-recurrent denumerable Markov chain.

1. Introduction. Let $J_0, J_1, \cdots$ be an aperiodic, irreducible positive-recurrent Markov chain with transition matrix P on the positive integers. Let $X_1, X_2, \cdots$ be real-valued random variables, defined on the same space as $\{J_n\}$, such that, for each i, j and n,

$$\begin{aligned} Q_{ij}(x) &= P(J_1 = j, X_1 \leqq x \mid J_0 = i) \\ &= P(J_n = j, X_n \leqq x \mid J_0, X_1, J_1, \cdots, X_{n-1}, J_{n-1} = i)\,. \end{aligned}$$

Then $P_{ij} = Q_{ij}(+\infty)$ for each i, j. If $P_{ij} \neq 0$, set $F_{ij}(x) = P_{ij}^{-1}Q_{ij}(x)$ for x real. If $P_{ij} = 0$, let F_{ij} be any distribution function.

Let β be a probability measure on the states of $\{J_n\}$. We consider β as a row-vector. Let R_{kl} be a real constant for each pair of positive integers (k, l). The conditional probability that $J_n = j$ and $X_\nu \leqq R_{J_{\nu-1}J_\nu}$, $\nu = 1, 2, \cdots, n$, given that J_0 has distribution β, is easily seen by induction to be the jth entry of the product β multiplied by the nth power of the matrix whose (k, l)th entry is $Q_{kl}(R_{kl})$.

If our Markov chain had only one state, $\{X_n\}$ would be independent and identically distributed, say with distribution function F. In this case, the quantity just described would be $(F(R))^n$, where R is a real constant. Gnedenko (1943) studied limits of quantities of this type, where R has a normalization depending on n. Our purpose is to obtain results like his in the present situation. The X_n's are dependent but they are conditionally independent, given $J_0, J_1, \cdots$.

The theorems to follow are analogues of those which Fabens and Neuts (1970) obtained for the case of aperiodic, irreducible, finite Markov chains.

2. Statement of the results. Let $\pi = (\pi_1, \pi_2, \cdots)$ be the stationary probability measure (row-vector) for P. Let β be any other probability measure (row-vector). For any row-vector α, let $||\alpha|| = \sum_{j=1}^{\infty} |\alpha_j|$ when the sum converges. Our main result is the following:

Received October 3, 1972; revised March 13, 1973.

[1] This research was supported in part by the National Research Council of Canada. Much of the work was done at the Summer Research Institute sponsored by the Canadian Mathematical Congress.

AMS 1970 *subject classifications.* Primary 60F05; Secondary 60J10, 60K15.

Key words and phrases. Limit distributions of maxima, denumerable Markov chain.

THEOREM 1. *Suppose that for every $i, j, n = 1, 2, \cdots$, there exist real constants a_{ijn} and b_{ijn} such that*

$$\lim_{n\to\infty} [F_{ij}(a_{ijn}x + b_{ijn})]^n = \Phi_{ij}(x) .$$

Let T_n be the (denumerable) matrix whose (i, j)th entry is $Q_{ij}(a_{ijn}R_{ij} + b_{ijn})$. Define

$$\chi = \prod_{i,j=1}^{\infty} [\Phi_{ij}(R_{ij})]^{\pi_i P_{ij}} ,$$

where we define $0^0 = 1$. (If this product diverges, it must go to 0; then we define $\chi = 0$.) If $||\pi T_n||^n \to \chi$, then $||\beta T_n^n - \chi\pi|| \to 0$ as $n \to \infty$.

The distribution functions Φ_{ij} may be some of Gnedenko's extreme value distributions or may be degenerate.

Let $A_n(i, j)$ be the event that $X_\nu \leqq a_{ijn}R_{ij} + b_{ijn}$ for those $\nu \leqq n$ such that $J_{\nu-1} = i$ and $J_\nu = j$. For "most" paths of length n, $J_{\nu-1} = i$ and $J_\nu = j$ about $n\pi_i P_{ij}$ times. Thus $P(A_n(i, j)) \approx F_{ij}^{n\pi_i P_{ij}}(a_{ijn}R_{ij} + b_{ijn}) = [F_{ij}^n(a_{ijn}R_{ij} + b_{ijn})]^{\pi_i P_{ij}} \to [\Phi_{ij}(R_{ij})]^{\pi_i P_{ij}}$. In view of the conditional independence of the X_ν's given $J_0, J_1, \cdots, J_n$, one might expect $P(\bigcap_{i,j} A_n(i, j)) \to \chi = \prod_{i,j} [\Phi_{ij}(R_{ij})]^{\pi_i P_{ij}}$. But $P(\bigcap_{i,j} A_n(i, j)) = \sum_j (\beta T_n^n)_j$; thus one might guess $\sum_j (\beta T_n^n)_j \to \chi$, which is implied by and related to the conclusion of the theorem.

Suppose R_{ij}, a_{ijn}, and b_{ijn} are independent of i and j; call them R, a_n, and b_n. In that case, what limiting distributions can $P(X_\nu \leqq a_n x + b_n,\ \nu = 1, 2, \cdots, n)$ have? Even without the assumption of the theorem that $(F_{ij}(a_n x + b_n))^n \to \Phi_{ij}(x)$, the only possible limits are the usual extreme value distributions and degenerate distributions. This is shown by Resnick and Neuts (1972) for finite Markov chains and by O'Brien and Denzel (1972) for denumerable chains.

The condition $||\pi T_n||^n \to \chi$ is discussed below in terms of Theorem 2. Following the approach of Fabens and Neuts (1970) we restate Theorem 1 in a more general context. Replace $F_{ij}(a_{ijn}R_{ij} + b_{ijn})$ by c_{ijn} and $\Phi_{ij}(R_{ij})$ by φ_{ij}.

THEOREM 2. *For each $i, j, n = 1, 2, \cdots$, let $c_{ijn} \in [0, 1]$ be such that $c_{ijn}^n \to \varphi_{ij}$ as $n \to \infty$. Let T_n be the matrix defined by $(T_n)_{ij} = c_{ijn}P_{ij}$. Let*

$$\chi = \prod_{i,j=1}^{\infty} \varphi_{ij}^{\pi_i P_{ij}}$$

(where $0^0 = 1$.) If $||\pi T_n||^n \to \chi$, then $||\beta T_n^n - \chi\pi|| \to 0$ as $n \to \infty$.

It is convenient in the proof of Theorem 2 to replace the condition $||\pi T_n||^n \to \chi$ by more workable conditions. We use the fact that $||\pi T_n|| = \sum_{ij} \pi_i P_{ij} c_{ijn}$.

LEMMA 1. *Suppose (as in Theorem 2) that $c_{ijn}^n \to \varphi_{ij}$ for all (i, j). Then $||\pi T_n||^n \to \chi$ and $\chi > 0$ are together equivalent to*

$$\lim_{n\to\infty} \sum_{i,j} \pi_i P_{ij}(nc_{ijn} - n - \log \varphi_{ij}) = 0 \tag{1}$$

(where $0 \log 0 = 0$). Also $\chi = 0$ implies $||\pi T_n||^n \to \chi$.

PROOF. First suppose $\chi > 0$. Then

$$\begin{aligned} ||\pi T_n||^n \to \chi &\Leftrightarrow \textstyle\sum_{i,j} \pi_i P_{ij} c_{ijn} = 1 + n^{-1} \log \chi + o(n^{-1}) \\ &\Leftrightarrow \textstyle\sum_{i,j} \pi_i P_{ij}(nc_{ijn} - n - \log \varphi_{ij}) \to 0 . \end{aligned}$$

Now assume (1) holds. The sum in (1) is finite for some n. Consequently, $\log \chi = \sum_{ij} \pi_i P_{ij} \log \varphi_{ij} > -\infty$, so that $\chi > 0$.

Now suppose $\chi = 0$. Let $\varepsilon > 0$. Take I large enough that $\chi' = \prod_{i,j=1}^{I} \varphi_{ij}^{\pi_i P_{ij}} < \varepsilon$. Let $\sum'$ denote sums with i and j going from 1 to I; let $\sum''$ denote sums with i or $j > I$. Then $\|\pi T_n\|^n = (\sum_{ij} \pi_i P_{ij} c_{ijn})^n \leqq [\sum' \pi_i P_{ij} c_{ijn} + \sum'' \pi_i P_{ij}]^n = [\sum' \pi_i P_{ij}(1 - n^{-1} \log \varphi_{ij} + o(n^{-1})) + \sum'' \pi_i P_{ij}]^n = [1 - n^{-1} \sum' \pi_i P_{ij} \log \varphi_{ij} + o(n^{-1})]^n \to \chi' < \varepsilon$. Therefore $\|\pi T_n\|^n \to \chi = 0$. This completes the proof of the lemma.

Suppose i and j take only finitely many values and $c_{ijn}^n \to \varphi_{ij}$ for all (i, j), in the case $\chi > 0$. For each (i, j) we have $\pi_i P_{ij}(nc_{ijn} - n - \log \varphi_{ij}) \to 0$, which implies (1). In our present case, (1) is essentially a weak condition of "uniformity" on the convergence of c_{ijn}^n to φ_{ij}.

We also note that in the case $\chi > 0$, (1) cannot be completely dropped. In Section 5, we show that in one special case condition (1) is necessary for Theorem 2 to hold.

Condition (1) can be replaced by an apparently weaker condition (2) below. This can be seen from the proof of Theorem 2 or by the method indicated below:

COROLLARY 1. *In the case* $\chi > 0$, $\|\pi T_n\|^n \to \chi$ *may be replaced in Theorem* 2 *by*

$$\liminf_{n\to\infty} \sum_{i,j} \pi_i P_{ij}(nc_{ijn} - n - \log \varphi_{ij}) \geqq 0 . \tag{2}$$

PROOF. Assume (2). Then we have

$$\sum \pi_i P_{ij} \log \varphi_{ij} \leqq \liminf \sum \pi_i P_{ij}(nc_{ijn} - n) .$$

Consequently,

$$\begin{aligned} -\sum \pi_i P_{ij} \log \varphi_{ij} &\geqq -\liminf \sum \pi_i P_{ij}(nc_{ijn} - n) \\ &\geqq -\limsup \sum \pi_i P_{ij}(nc_{ijn} - n) \\ &= \liminf \sum \pi_i P_{ij}(n - nc_{ijn}) \\ &\geqq \sum \pi_i P_{ij}(-\log \varphi_{ij}) , \end{aligned}$$

by Fatou's lemma. This implies (1), as required.

The proof of Theorem 2 can be modified to yield:

COROLLARY 2. *In the case* $\chi = 0$, *it is sufficient that* $c_{ijn}^n \to \varphi_{ij}$ *for all pairs* (i, j) *in a set* A *such that*

$$\prod_{(i,j)\in A} \varphi_{ij}^{\pi_i P_{ij}} = 0 .$$

Before proceeding, we establish some conventions. Limits will mean as n approaches infinity, unless otherwise specified. Also sums and products will be from one to infinity. If A is a (denumerable) matrix, define the norm $\|A\| = \sup_i \sum_j |A_{ij}|$, when this is finite. If α is a row-vector and B is another matrix, then $\|AB\| \leqq \|A\|\,\|B\|$ and $\|\alpha A\| \leqq \|\alpha\|\,\|A\|$ when the right sides are defined.

We will freely use the language and properties of denumerable matrices as discussed in Kemeny, Snell and Knapp (1966). In addition, we use their version (due to Orey (1962)) of the convergence theorem for positive-recurrent Markov chains.

In Section 6, we give a theorem for the case of periodic Markov chains.

3. Some lemmas. We separate some parts of the proof of Theorem 2 as lemmas. The first of these is obvious.

LEMMA 2. *Suppose $P_{kl} = 0$ for some (k, l). Then we may assume $c_{kln} = 1$ for all n and $\varphi_{kl} = 1$.*

As a result of this lemma, we assume $P_{ij} > 0$ whenever $\varphi_{ij} = 0$. The following lemma is a special case of Theorem 2. We give it in a stronger form than required later, since it indicates the extent to which the method of Fabens and Neuts (1970) applies to denumerable chains.

LEMMA 3. *Suppose we replace* (1) *by the stronger condition*

$$\sup_i \textstyle\sum_j P_{ij}|nc_{ijn} - n - \log \varphi_{ij}| \to 0 \,. \tag{3}$$

Then $||\beta T_n{}^n - \chi\pi|| \to 0$.

PROOF. By (3), $|\sum_{i,j} \pi_i P_{ij}(nc_{ijn} - n - \log \varphi_{ij})| \leqq \sum_i \pi_i \sum_j P_{ij}|nc_{ijn} - n - \log \varphi_{ij}| \to 0$. Thus (1) holds. By Lemma 1, $\chi > 0$. Using Lemma 2, we assume w.l.o.g. that each $\varphi_{ij} > 0$. Define a matrix Γ by $\Gamma_{ij} = P_{ij} \log \varphi_{ij}$. Since the left side of (3) is finite for some n, we have

$$\begin{aligned} \gamma \equiv ||\Gamma|| &= \sup_i \textstyle\sum_j P_{ij}|\log \varphi_{ij}| \\ &\leqq \sup_i \textstyle\sum_j P_{ij}|\log \varphi_{ij} + n - nc_{ijn}| + \sup_i \sum_j P_{ij}|nc_{ijn} - n| \\ &< \infty \,. \end{aligned}$$

Let $A_n = P + n^{-1}\Gamma$. By (3),

$$||T_n - A_n|| = n^{-1} \sup_i \textstyle\sum_j |P_{ij} nc_{ijn} - nP_{ij} - P_{ij} \log \varphi_{ij}| = o(n^{-1}) \,.$$

Consequently, $||A_n|| \leqq ||T_n|| + ||A_n - T_n|| \leqq 1 + o(n^{-1})$. Thus

$$\begin{aligned} &||T_n{}^n - A_n{}^n|| \\ &\quad \leqq ||T_n{}^n - A_n T_n{}^{n-1}|| + ||A_n T_n{}^{n-1} - A_n{}^2 T_n{}^{n-2}|| + \cdots + ||A_n{}^{n-1} T_n - A_n{}^n|| \\ &\quad \leqq ||T_n - A_n||(||T_n||^{n-1} + ||T_n||^{n-2}||A_n|| + \cdots + ||A_n||^{n-1}) \\ &\quad \leqq o(n^{-1}) n (1 + o(n^{-1}))^{n-1} \,, \end{aligned}$$

where goes to 0.

Let B be the denumerable matrix each of whose rows is π. For each n and for $k = 0, 1, \cdots, n$, let $V(k, n) = n^{-k}\binom{n}{k} B\Gamma B\Gamma \cdots \Gamma B$, where the matrix Γ appears k times. Then $V_{ij}(k, n) = n^{-k}\binom{n}{k}(\log \chi)^k \pi_j$. As $n \to \infty$,

$$\begin{aligned} ||\textstyle\sum_{k=0}^n V(k, n) - \chi B|| &= ||\textstyle\sum_{k=0}^n n^{-k}\binom{n}{k}(\log \chi)^k \pi - \chi\pi|| \\ &= |(1 + n^{-1} \log \chi)^n - \chi| \, ||\pi|| \end{aligned}$$

which converges to 0. Now

$$\begin{aligned} ||\beta T_n{}^n - \chi\pi|| \leqq ||T_n{}^n - A_n{}^n|| &+ ||\beta A_n{}^n - \textstyle\sum_{k=0}^n \beta V(k, n)|| \\ &+ ||\textstyle\sum_{k=0}^n V(k, n) - \chi B|| \,. \end{aligned}$$

Since the first and last of these go to 0, it is sufficient to show the middle term goes to 0.

Let $U(k, n)$ be the sum of all the matrices expressible in the form $n^{-k}P^{m_0}\Gamma P^{m_1}\Gamma \cdots \Gamma P^{m_k}$ for which $m_0 + m_1 + \cdots + m_k = n - k$. There are $\binom{n}{k}$ such matrix products. We note that $A_n{}^n = (P + n^{-1}\Gamma)^n = U(0, n) + \cdots + U(k, n)$. We will compare $\beta U(k, n)$ and $\beta V(k, n)$.

Let $\varepsilon > 0$. Let $\eta = \varepsilon(1 + \gamma)^{-1}e^{-\gamma}$. If $\gamma = ||\Gamma|| > 0$, define a row-vector α by $\alpha_j = -||\pi\Gamma||^{-1}(\pi\Gamma)_j$. Note that $||\pi\Gamma|| > 0$ and $||\alpha|| = 1$. Let M be sufficiently large that for all $m \geqq M$. $||\alpha P^m - \pi|| < \eta$ and $||\beta P^m - \pi|| < \eta$ (see page 153 of Kemeny, Snell and Knapp (1966)). For $m \geqq M$, $||\pi\Gamma(P^m - B)|| \leqq ||\pi\Gamma||\,||\alpha P^m - \pi|| \leqq \gamma\eta$. If $\gamma = 0$, $||\pi\Gamma(P^m - B)|| = 0 \leqq \gamma\eta$.

A term $n^{-k}P^{m_0}\Gamma P^{m_1} \cdots \Gamma P^{m_k}$ of $U(k, n)$ is said to be "good" if each $m_i \geqq M$. There are $\binom{n-(k+1)M}{k}$ good terms. For a good term,

$$\begin{aligned}
&||n^{-k}\beta P^{m_0}\Gamma P^{m_1} \cdots \Gamma P^{m_k} - n^{-k}\beta B\Gamma\beta \cdots \Gamma B|| \\
&\quad\leqq n^{-k}||\beta P^{m_0}\Gamma \cdots \Gamma P^{m_k} - \beta B\Gamma P^{m_1} \cdots \Gamma P^{m_k}|| \\
&\qquad + n^{-k}||\beta B\Gamma P^{m_1} \cdots \Gamma P^{m_k} - \beta B\Gamma B\Gamma P^{m_2} \cdots \Gamma P^{m_k}|| + \cdots \\
&\qquad + n^{-k}||\beta B\Gamma \cdots B\Gamma P^{m_k} - \beta B\Gamma \cdots B\Gamma B|| \\
&\quad\leqq n^{-k}||\beta P^{m_0} - \pi||\,||\Gamma P^{m_1}\Gamma \cdots \Gamma P^{m_k}|| \\
&\qquad + n^{-k}||\beta||\,||\pi\Gamma(P^{m_1} - B)||\,||\Gamma P^{m_2}\Gamma \cdots \Gamma P^{m_k}|| + \cdots \\
&\qquad + n^{-k}||\beta||\,||B\Gamma B \cdots \Gamma||\,||\pi\Gamma(P^{m_k} - B)|| \\
&\quad\leqq (k + 1)n^{-k}\gamma^k\eta\,.
\end{aligned}$$

The total difference between the good terms of $\beta A_n{}^n$ and the corresponding terms of $\sum_{k=0}^{n} \beta V(k, n)$ is bounded by

$$\sum_{k=0}^{n} \binom{n}{k}(k + 1)n^{-k}\gamma^k\eta \leqq \eta \sum_{k=0}^{\infty} (k + 1)\frac{\gamma^k}{k!} = \eta(1 + \gamma)e^{\gamma} = \varepsilon\,,$$

independently of n.

The total difference for the terms which are not included above is bounded by

$$\begin{aligned}
&2 \sum_{k=0}^{n} n^{-k}\gamma^k\left[\binom{n}{k} - \binom{n-(k+1)M}{k}\right] \\
&\quad\leqq 2 \sum_{k=0}^{J} n^{-k}\gamma^k \frac{1}{k!}\left[n^{(k)} - (n - (k + 1)M)^{(k)}\right] + 2 \sum_{k=J+1}^{\infty} \frac{\gamma^k}{k!}
\end{aligned}$$

since $n^k > n^{(k)} = n!/(n - k)! > (n - (k + 1)M)^{(k)}$. Choose J sufficiently large such that the second sum is less than ε. Then, for n sufficiently large, $n - (k + 1)M > 0$, so that $n^{(k)} - (n - (k + 1)M)^{(k)}$ is a polynomial in n of degree $k - 1$. Thus, the first sum goes to 0, so can be made less than ε. We have shown that for n sufficiently large,

$$||\beta A_n{}^n - \sum_{k=0}^{n} \beta V(k, n)|| < 3\varepsilon\,,$$

as required.

LEMMA 4. *If* $c_{ijn}^n \to \varphi_{ij}$ *for each* i *and* j, *then* $\limsup\,(\beta T_n{}^n)_j \leqq \chi\pi_j$ *for all* j.

PROOF. Let $\varepsilon > 0$. We divide the proof into two cases which we later combine. First suppose $\varphi_{kl} = 0$ for some pair (k, l). By Lemma 2, we may assume

$P_{kl} \neq 0$. For $(i, j) \neq (k, l)$, let $\Psi_{ij} = d_{ijn} = 1$ for all n. Let $\Psi_{kl} = \exp\{(\pi_k P_{kl})^{-1} \log \varepsilon\}$ and let $d_{kln} = \max(c_{kln}, \Psi_{kl}^{1/n})$. Let $\chi' = \varepsilon$.

Now suppose $\varphi_{ij} > 0$ for all (i, j). Let I be sufficiently large that $\chi' < \chi + \varepsilon$, where in this case $\chi' = \prod_{i,j=1}^{I} \varphi_{ij}^{\pi_i P_{ij}}$. For i and $j \leqq I$, let $\Psi_{ij} = \varphi_{ij}$ and $d_{ijn} = c_{ijn}$. For i or $j > I$, let $\Psi_{ij} = d_{ijn} = 1$, for all n.

In both cases, let $(S_n)_{ij} = P_{ij} d_{ijn}$. By Lemma 3, $\|\beta S_n{}^n - \chi'\pi\| \to 0$. For n sufficiently large, $(\beta T_n{}^n)_j \leqq (\beta S_n{}^n)_j \leqq \chi'\pi_j + \varepsilon \leqq \chi\pi_j + 2\varepsilon$. As ε is arbitrary, the result follows.

The next lemma shows that it is enough to consider the case for which β is the stationary measure π. The proof of Lemma 3 could have been slightly simplified by giving the next lemma first.

LEMMA 5. *If* $\chi > 0$, $\|\beta T_n{}^n - \pi T_n{}^n\| \to 0$ *as* $n \to \infty$.

PROOF. Let $\varepsilon > 0$. Since $\varphi_{ij} > 0$ for all (i, j) (using Lemma 2), $c_{ijn} \to 1$ for each (i, j). Let m be sufficiently large that $\|\beta P^m - \pi\| < \varepsilon$. We define two probabilities on the set Ω of $(m + 1)$-tuples of positive integers. If $x = (i_0, i_1, \cdots, i_m)$ is in Ω, let $\mu(\{x\})$ (respectively $\nu(\{x\})$) be the probability that a Markov chain with transition matrix P and initial distribution π (respectively β) begins by going through the states $i_0, i_1, \cdots, i_m$ (in that order). Define $f_n : \Omega \to [0, 1]$ by $f_n(i_0, \cdots, i_m) = c_{i_0 i_1 n} c_{i_1 i_2 n} \cdots c_{i_{m-1} i_m n}$. Let $F = \{1, 2, \cdots, J\}^{m+1} \subseteqq \Omega$, where J is large enough that $\nu(F) > 1 - \varepsilon$ and $\mu(F) > 1 - \varepsilon$. Let $A_j = \{(i_0, \cdots, i_m) \in \Omega : i_m = j\}$. Finally, let $n \geqq m$ be sufficiently large that $f_n(x) \geqq 1 - \varepsilon$ for all $x \in F$. Note that $\int_{A_j} f_n d\nu = (\beta T_n{}^m)_j$, which implies

$$\begin{aligned}\|\beta T_n{}^n - \pi T_n{}^n\| &\leqq \|\beta T_n{}^m - \pi T_n{}^m\| \|T_n\|^{n-m} \leqq \|\beta T_n{}^m - \pi T_n{}^m\| \\ &= \textstyle\sum_j |\int_{A_j} f_n \, d(\nu - \mu)| \\ &\leqq \textstyle\sum_j |\int_{A_j \cap F} f_n \, d(\nu - \mu)| + \sum_j |\int_{A_j \cap F^c} f_n \, d(\nu - \mu)| \,.\end{aligned}$$

The second term is bounded by $(\nu + \mu)(F^c) < 2\varepsilon$. Also

$$\begin{aligned}\textstyle\int_{A_j \cap F} f_n \, d(\nu - \mu) &\leqq \nu(A_j \cap F) - (1 - \varepsilon)\mu(A_j \cap F) \\ &\leqq |\nu(A_j) - \mu(A_j)| + (\nu + \mu)(A_j \cap F^c) + \varepsilon(\mu + \nu)(A_j) \,.\end{aligned}$$

By symmetry, the same bound holds for $|\int_{A_j \cap F} f_n \, d(\nu - \mu)|$. Summing over j, we obtain

$$\begin{aligned}\|\beta T_n{}^n - \pi T_n{}^n\| &\leqq \textstyle\sum_j |\nu(A_j) - \mu(A_j)| + 6\varepsilon \\ &= \|\beta P^m - \pi\| + 6\varepsilon < 7\varepsilon \,,\end{aligned}$$

which gives the result.

4. Proof of Theorem 2. Assume first that $\chi = 0$. By the statement $\|\beta S_n{}^n - \chi'\pi\| \to 0$ in the proof of Lemma 4, we have for n sufficiently large that $\|\beta T_n{}^n - \chi\pi\| = \sum_j (\beta T_n{}^n)_j \leqq \sum_j (\beta S_n{}^n)_j < \chi' + \varepsilon \leqq 2\varepsilon$, so that $\lim \|\beta T_n{}^n - \chi\pi\| = 0$.

Now we consider the case $\chi > 0$. We may assume the initial distribution is π, by Lemma 5. Let $\varepsilon > 0$. Define d_{ijn}, Ψ_{ij}, S_n and χ' as in Lemma 4 (for the case when all $\varphi_{ij} > 0$), where we assume I is so large that $\log \chi' - \log \chi < \varepsilon$.

Define

$$\begin{aligned}\Delta(n) &= \|\pi S_n{}^n - \pi T_n{}^n\| \\ &= \textstyle\sum_{i,j} \pi_i[(S_n{}^n)_{ij} - (T_n{}^n)_{ij}] \\ &= \textstyle\sum_{i_0,i_1,\cdots,i_n} \pi_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} \\ &\quad \times [d_{i_0 i_1 n} \cdots d_{i_{n-1} i_n n} - c_{i_0 i_1 n} \cdots c_{i_{n-1} i_n n}] \,.\end{aligned} \tag{4}$$

The expression in the square brackets

$$\begin{aligned}&= (d_{i_0 i_1 n} \cdots d_{i_{n-1} i_n n} - c_{i_0 i_1 n} d_{i_1 i_2 n} \cdots d_{i_{n-1} i_n n}) \\ &\quad + (c_{i_0 i_1 n} d_{i_1 i_2 n} \cdots d_{i_{n-1} i_n n} - c_{i_0 i_1 n} c_{i_1 i_2 n} d_{i_2 i_3 n} \cdots d_{i_{n-1} i_n n}) + \cdots \\ &\quad + (c_{i_0 i_1 n} \cdots c_{i_{n-2} i_{n-1} n} d_{i_{n-1} i_n n} - c_{i_0 i_1 n} \cdots c_{i_{n-1} i_n n}) \\ &\leqq [(d_{i_0 i_1 n} - c_{i_0 i_1 n}) + (d_{i_1 i_2 n} - c_{i_1 i_2 n}) + \cdots + (d_{i_{n-1} i_n n} - c_{i_{n-1} i_n n})] \,.\end{aligned}$$

Substituting this back into (4) and performing the summation where possible, we obtain

$$\Delta(n) \leqq n \textstyle\sum_{i,j} \pi_i P_{ij}(d_{ijn} - c_{ijn}) \,,$$

by the stationarity of π. Now by (1),

$$\textstyle\sum_{i,j} \pi_i P_{ij}(nc_{ijn} - n) \to \sum_{i,j} \pi_i P_{ij} \log \varphi_{ij} = \log \chi \,.$$

Subtracting this from the corresponding expression for d_{ijn}, Ψ_{ij} and χ', we obtain

$$n \textstyle\sum_{i,j} \pi_i P_{ij}(d_{ijn} - c_{ijn}) \to \log \chi' - \log \chi < \varepsilon \,.$$

Thus, for n sufficiently large, $\Delta(n) < 2\varepsilon$. By Lemma 3, $\lim \sum_j (\pi S_n{}^n)_j = \chi'$. Thus $\chi' - \liminf \sum_j (\pi T_n{}^n)_j = \limsup \Delta(n) \leqq 2\varepsilon \leqq 2\varepsilon + \chi' - \chi$. As ε is arbitrary,

$$\liminf \textstyle\sum_j (\pi T_n{}^n)_j \geqq \chi \,. \tag{5}$$

Let $\{n'\}$ be a subsequence of positive integers such that $\pi T_{n'}^{n'}$ converges componentwise, say to a vector l. By Lemma 4, we have

$$l_j \leqq \chi \pi_j \tag{6}$$

for all j. Also $0 \leqq (\pi T_n{}^n)_j \leqq \pi_j$ for all j and n. By dominated convergence (for the counting measure), $\|\pi T_{n'}^{n'} - l\| \to 0$. By (5) and (6), we must have $l = \chi\pi$ and $\|\pi T_n{}^n - \chi\pi\| \to 0$.

5. Example. In one very special case, $\|\pi T_n\|^n \to \chi$ is a necessary condition for Theorem 2 to hold. Assume $P_{ij} = \pi_j$, $c_{ijn} = c_{jn}$ and $\varphi_{ij} = \varphi_j$ (all independently of i). Assume $\|\pi T_n\|^n \not\to \chi$. Then $\chi \neq 0$ and (1) fails, by Lemma 1. Thus (2) also fails; there is a subsequence $\{n'\}$ of positive integers such that $\sum_j \pi_j(n'c_{jn'} - n' - \log \varphi_j) \to \delta < 0$.

$$\begin{aligned}\textstyle\sum_j(\pi T_n{}^n)_j &= \textstyle\sum_{i_1,i_2,\cdots,i_n} \pi_{i_1} \cdots \pi_{i_n} c_{i_1 n} \cdots c_{i_n n} \\ &= (\textstyle\sum_i \pi_i c_{in})^n \\ &= [1 + n^{-1}(\textstyle\sum_i \pi_i(nc_{in} - n))]^n \\ &< \exp(\log \chi + \tfrac{1}{2}\delta)\end{aligned}$$

for n sufficiently large in the subsequence. Thus,

$$\liminf ||\pi T_n^{\,n}|| < ||\chi\pi|| = \chi ,$$

which implies that $||\pi T_n^{\,n} - \chi\pi|| \nrightarrow 0$.

More specifically, let $\varphi_j = 1$ for all j. Let $\alpha \in (0, 1)$. Let $k(n)$ be a sequence such that $k(n) \to \infty$ and $\sum_{i=1}^{k(n)} \pi_i \leqq \alpha^{1/n}$. Define $c_{jn} = 1$ if $j \leqq k(n)$ and 0 otherwise. Since $k(n) \to \infty$, $c_{jn}^n \to 1 = \varphi_j$. But $(\sum_{ij} c_{ijn}\pi_i P_{ij})^n = (\sum_j c_{jn}\pi_j)^n = (\sum_{j=1}^{k(n)} \pi_j)^n \leqq \alpha < 1 = \chi$.

To obtain an example in the language of Theorem 1, let $P_{ij} = \pi_j$; let $F_{ij}(0) = 0$ and $F_{ij}(1) = 1$ for all (i, j); let $a_{ijn} = 0$ and let $b_{ijn} = c_{jn}$ as defined above.

6. The periodic case. We now drop the assumption that the Markov chain is aperiodic. Let its period be d. Fix one state f and separate the states into classes $e_1, e_2, \cdots, e_d$: $j \in e_k$ iff $(P^{nd+k})_{fj} > 0$ for some n. For each k, there is a probability measure $\pi(k) = (\pi_1(k), \pi_2(k), \cdots)$ which is concentrated on e_k, such that, if β is a probability measure concentrated on e_l, then $||\beta P^{dn-l+k} - \pi(k)|| \to 0$. Let $\pi = d^{-1}(\pi(1) + \cdots + \pi(d))$. Then, for any probability β,

$$||n^{-1} \textstyle\sum_{j=0}^{n-1} (\beta P^j) - \pi|| \to 0 .$$

The theorem stated below is analogous to Lemma 3 in the aperiodic case. It is possible to weaken the hypothesis (7) somewhat, but only at the cost of considerable complication. The theorem is proved by watching the process at every dth time and thereby reducing the situation to that of Theorem 2.

THEOREM 3. *For each i, j, and n, let $c_{ijn} \in [0, 1]$ be such that $c_{ijn}^n \to \varphi_{ij}$ as $n \to \infty$. Let T_n be the matrix defined by $(T_n)_{ij} = c_{ijn}P_{ij}$. Let*

$$\chi = \pi_{i,j}\varphi_{ij}^{\pi_i P_{ij}} .$$

If $\chi = 0$ or if

$$\sup_i \textstyle\sum_j P_{ij}|nc_{ijn} - n - \log \varphi_{ij}| \to 0 \tag{7}$$

then

$$||n^{-1} \textstyle\sum_{j=1}^{n} \beta T_j^j - \chi\pi|| \to 0 \tag{8}$$

for any probability measure β.

OUTLINE OF THE PROOF. Assume that $\chi > 0$. The other case follows from this, much as in Lemma 4. Let $Q = P^d$ and $S_n = T_{nd}^d$. Define $g_{ijn} = 1$ if $Q_{ij} = 0$ and $(S_n)_{ij}(Q_{ij})^{-1}$ otherwise. In the latter case, we may apply (7) to obtain (letting $m = nd$)

$$\begin{aligned} g_{ijn} &= (Q_{ij})^{-1} \textstyle\sum_{i_1,\cdots,i_{d-1}} P_{ii_1} \cdots P_{i_{d-1}j}(c_{ii_1m} \cdots c_{i_{d-1}jm}) \\ &= 1 + m^{-1}(Q_{ij})^{-1} \textstyle\sum_{i_1,\cdots,i_{d-1}} P_{ii_1} \cdots P_{i_{d-1}j} \log(\varphi_{ii_1}\varphi_{i_1i_2} \cdots \varphi_{i_{d-1}j}) + o(m^{-1}) . \end{aligned}$$

Thus g_{ijn}^n converges. Condition (1) of Theorem 2 also follows from (7). Moreover, we have

$$\begin{aligned} \log \chi &= \textstyle\sum_{i_0,i_d} \pi_{i_0}(k)Q_{i_0i_d} \log(\lim g_{i_0i_dn}^n) \\ &= \textstyle\sum_{i,j} \pi_i P_{ij} \log \varphi_{ij} , \end{aligned}$$

as desired. This is independent of k. By Theorem 2,

$$||\beta T_{dn}^{dn} - \chi\pi(k)|| = ||\beta S_n^{\ n} - \chi\pi(k)|| \to 0$$

is β is a probability concentrated on e_k.

To obtain (8), we first show (again using (7))

$$||\beta T_{dn+i}^{dn+i} - \chi\pi(k+i)|| \to 0$$

(where $k + i$ is calculated modulo d) for β concentrated on e_k. This is enough to obtain the desired result.

Acknowledgment. I am indebted to Professor R. A. Schaufele for suggesting the problem and for offering many useful comments on the solution.

REFERENCES

[1] Fabens, A. J. and Neuts, M. F. (1970). The limiting distribution of the maximum term in a sequence of random variables defined on a Markov chain. *J. Appl. Probability* **7** 754–760.

[2] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44** 423–453.

[3] Kemeny, J. G., Snell, J. L. and Knapp, A. W. (1966). *Denumerable Markov Chains*. Van Nostrand, Princeton.

[4] O'Brien, G. L. and Denzel, G. E. (1972). Limit distributions for extreme values of chain-dependent processes. To appear.

[5] Orey, S., (1962). An ergodic theorem for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1** 174–176.

[6] Resnick, S. I. and Neuts, M. F. (1970). Limit laws for maxima of a sequence of random variables defined on a Markov chain. *Advances in Appl. Probability* **2** 323–343.

Department of Mathematics
York University
4700 Keele Street
Downsview 463, Ontario
Canada