Jñānabha Sect. A, Vol. 4, July 1974.

ON THE STATIONARY PROBABILITIES FOR A CERTAIN CLASS OF DENUMERABLE MARKOV CHAINS

By

Eric Brodheim*

New York Blood Centre, U. S. A.

Cyrus Derman†

Columbia University, New York, New York 10027, *U. S. A.*

AND

Barbara Keyfitz

Columbia University, New York, New York 10027, *U. S. A.*

(Received : July 19, 1974)

1. Introduction and Summary. Consider the Markov chain on the non-negative integers with matrix P_n of transition probabilities given by

$$p_{io}=1-\sum_{v=0}^{n+i-1} p_v = r_{n+i} \text{ (say)}, \; i=0,\, 1,\, \ldots,$$

$$p_{ij}=p_{n-j+i}, \qquad j=1,\, \ldots,\, i+n.$$

$$\quad =0 \qquad , \qquad j=i+n+1,\, \ldots,$$

where $p_i \geqslant 0$, $i=0,\, 1,\, \ldots,\ \sum_{i=0}^{\infty} p_i=1$, and n is a positive integer. This type of Markov chain arises in numerous contexts. For example, for

* The research reported in this paper was supported in part by Grant No. HL 09011 from the National Heart and Lung Institute of the National Institutes of Health, U. S. Dept. of Health, Education and Welfare and by Grant No. HL 09011 RR 05688 from the National Institutes of Health.

† Work supported in part by the Army, Navy under Office of Naval Research Contract N 00014—67—A—0108—0008 with Columbia University.

$n=1$, the determination of the stationary probabilities appears in Karlin ([2 p. 82) where P_1 arises in the context of an embedded Markov chain in a queuing model. For $n \geqslant 1$ the matrix P_n is relevant to the fluctuations of a non-backlogged inventory that is replenished periodically with n units and subject to demand for a period having distribution $\{p_v\}$. In [1] P_n is used as an approximating stochastic matrix to one that arises in the management of blood inventory where discarding of units in inventory takes place due to aging. P_n is, also, a transition matrix of a random walk that takes, at most, n steps to the right and has a reflecting barrier at the origin. Of interest to us are the stationary probabilities associated with P_n.

Throughout we shall assume

(i) $\sum_{v=0}^{\infty} v p_v > n$ (the positive recurrent case)

(ii) $p_0 > 0$

(iii) $\{p_v\}$ is aperiodic.

In section 2 the stationary probabilities are shown to be functions of the roots of equation (1) that lie within the unit circle. One root which is real dominates all the others. In section 3 the behavior of this root as a function of $\sum_{v=0}^{\infty} v p_v$ is explored for the Poisson case and more generally. In section 4 the special case of $n=2$ is summarized.

Related techniques can be found in Takács [6]. An alternative approach can be found in Keilson [3] together with a more inclusive discussion of the various contexts in which this type of process arises. See also Smith [5] and Lindley [4] for related discussions.

2. **The Main Results. Lemma 1.** Under (i), (ii), (iii) the equation

$$\sum_{v=0}^{\infty} p_v x^v = x^n \tag{1}$$

has precisely n (perhaps not distinct) complex roots of modulus less than 1, none of which is zero ; among these roots the one with the largest modulus is positive and real.

Proof. Let $g(x)=x^n$, $h(x)=\sum_{v=0}^{\infty} p_v x^v$

and $f(x)=g(x)-h(x)$. Since $g(1)=h(1)=1$, $g(0)=0$, $h(0)=p_0>0$, $g'(1)=n$, $h'(1)=\sum_{v=0}^{\infty} vp_v>n$ it follows that the graph of g lies above that of h when x $(x<1)$ is close to 1, but g lies below h at $x=0$. Hence, g and h must cross at some point in the open interval $(0, 1)$. However, they cross only a finite number of times, since, otherwise, $f(x)$ is identically zero, contradicting (i). Therefore there is a largest real root, x_0, $0<x_0<1$ of (1) and $g(x)>h(x)$ for all real x in $(x_0, 1)$. Let C be a circle about O of radius r, for any $r\varepsilon(x_0, 1)$. On C,

$$|h(Z)| = |\sum_{v=0}^{\infty} p_v Z^v|$$

$$\leqslant \sum_{v=0}^{\infty} p_v |Z|^v$$

$$< |Z|^n$$

$$=|g(Z)|.$$

From Rouche's Theorem $f(Z)$ and $g(Z)$, being analytic functions, have the same number of zeros, namely n, inside C. Since C can be arbitrarily close to the circle $|Z|=x_0$ the zeros lie in or on the circle $|Z|=x_0$.

The assumption (iii), however implies $|\sum_{v=0}^{\infty} p_v Z^v| < \sum_{v=0}^{\infty} p_v |Z|^v$ if Z is not real and positive ; hence, only one zero, the real one, is on the circle $|Z|=x_0$. Since $p_0>0$, it is obvious that no roots of (1) are zero. This completes the proof of lemma 1.

Lemma 2. Suppose $[p_{ij}]$ $i, j=0, 1, \ldots$ is any transition matrix.

Suppose the sequence of complex numbers $\{\pi_i\}$, $i=0, 1, \ldots\ldots$ satisfies

$$\sum_{i=0}^{\infty} |\pi_i| < \infty, \qquad \ldots(2)$$

$$\sum_{i=0}^{\infty} \pi_i \, p_{ij} = \pi_j, \; j=1, 2, \ldots\ldots \qquad \ldots(3)$$

then (3) holds for $j=0$.

Proof. We can, using (2), (3), and (4) write

$$\pi_0 = \sum_{j=0}^{\infty} \pi_j - \sum_{j=1}^{\infty} \pi_j = \sum_{j=0}^{\infty} \pi_j - \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} \pi_i p_{ij} = \sum_{j=0}^{\infty} \pi_j$$

$$- \sum_{i=0}^{\infty} \pi_i \sum_{j=1}^{\infty} p_{ij} = \sum_{j=0}^{\infty} \pi_j - \sum_{i=0}^{\infty} \pi_i (1-p_{i0}) = \sum_{i=0}^{\infty} \pi_i p_{i0}.$$

Label the roots of (1) $\theta_1, \ldots, \theta_n$, with $\theta_1 = x_0$, the largest root. Let

$$\pi_i = \sum_{r=1}^{n} K_r \theta_r^i, \quad i=0, 1, \ldots\ldots \tag{4}$$

where K_j ($j=1, \ldots, n$) are arbitrary constants, complex or real.

Lemma 3. The sequence $\{\pi_i, i=0, \ldots\}$ satisfies the equations (3) for $j \geqslant n$.

Proof. The proof is immediately verified using the fact that $\theta_1, \ldots, \theta_n$ are roots of (1).

Lemma 4. If $\theta_1, \ldots, \theta_n$ are distinct then there exist non-zero complex constants $K_1, \ldots, K_n$ such that $\sum_{i=0}^{n} K_i \neq 0$ and $\{\pi_i\}$ defined by (4) satisfy (3) for $j=1, \ldots, n-1$.

Proof. The equations (3) for $j=1, \ldots, n-1$ in terms of $K_1, \ldots, K_n$ are

$$\sum_{r=1}^{n} K_r \theta_r^j = \sum_{i=0}^{\infty} \sum_{r=1}^{n} K_r \theta_r^i p_{n-j+i} = \sum_{r=1}^{n} K_r \Big(\sum_{i=0}^{\infty} p_{n-j+i}$$

$$\theta_r^{i+n-j} \theta_r^{j-n} \Big) = \sum_{r=1}^{n} K_r \theta_r^{j-n} \sum_{v=n-j}^{\infty} p_v \theta_r^v$$

$$= \sum_{r=1}^{n} K_r \theta^{j-n} \Big(\theta_r^n - \sum_{v=0}^{n-j-1} p_v \theta_r^v \Big). \tag{5}$$

Cancelling the left side of (5) with the first term on the right we get

$$\sum_{v=0}^{n-j-1} p_v \sum_{r=1}^{n} K_r \theta_r^{v-n+j} = 0, \quad j=1, \ldots, n-1. \tag{6}$$

Considering the equations in (6) sequentially in reverse order, using (ii), we obtain the system

$$\sum_{r=1}^{n} K_r \, \theta_r^{-j} = 0, \; j=1, \ldots, n-1. \qquad \ldots(7)$$

The matrix $A = \begin{Bmatrix} 1 & \theta_1^{-1} & \theta_1^{-n+1} \\ \vdots & \vdots & \vdots \\ 1 & \theta_n^{-1} & \theta_n^{-n+1} \end{Bmatrix}$ is of rank n

since $\theta_1, \ldots, \theta_n$ are distinct and non-zero. The lemma follows from this fact.

Lemma 5. Let $\{p_{ij}^{(n)}\}$, $n=1, \ldots,$ be the n-step transition probability matrix of any Markov chain over the non-negative integers such that

$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} p_{ij}^{(n)} = a_j \geqslant 0, j=0, 1, \ldots$ independent of i. If $\{\pi_i\}$ $i=0, 1, \ldots$ is a sequence of complex numbers such that

(a) $\sum_{i=0}^{\infty} \pi_i \, p_{ij} = \pi_j, \; j=0, 1, \ldots$

(b) $\sum_{i=0}^{\infty} |\pi_i| < \infty,$

then

$$\pi_j = \sum_{i=0}^{\infty} \pi_i \, a_j, \; j=0, 1, \ldots$$

Proof. Follows from standard Markov chain arguments.

Corollary. If $\{\pi_i\}$ $i=0, 1, \ldots$satisfy the conditions of lemma 5 then either $\sum_{i=0}^{\infty} \pi_i \neq 0$ or $\pi_j \equiv 0$. In the former case if, $\pi_i' = \frac{\pi_i}{\sum_{j=0}^{\infty} \pi_j}$,

$i=0, 1, \ldots,$ then $\pi_j' = \sum_{i=0}^{\infty} \pi_i' \, p_{ij} \geqslant 0$ and $\sum_{j=0}^{\infty} \pi_j' = 1$; that is $\{\pi_i'\}$, $i=0, 1, \ldots$ are the stationary probabilities of $\{p_{ij}\}$.

We can then state

Theorem 1. If $\{p_n\}$ satisfies (i), (ii), (iii) there are n roots

$\theta_1, \ldots, \theta_n$, to (1) within the unit circle. If $\theta_1, \ldots, \theta_n$, are distinct, then,

$$\theta_1 > |\theta_i|, \; i=2, \ldots, n$$

and

$$\pi_i' = \frac{\sum_{r=1}^{n} K_r \, \theta_r^{i}}{\sum_{r=1}^{n} K_r \, \frac{1}{1-\theta_r}}, \quad i=0, 1, \ldots,$$

where $K_1, \ldots, K_n$ is a non-trivial solution to (7), are stationary probabilities of P_n.

Proof. The assertions concerning the roots follow from

Lemma 1. By Lemmas 3 and 4, $\{\pi_i\}$ $i=0, \ldots\ldots$satisfies (3) for $j \geqslant 1$; since $|\theta_r| < 1$, $r=1, \ldots, m$ lemma 2 asserts that (3) is also satisfied for $j=0$. Therefore (a) of lemma 5 is satisfied. Since (b) is also satisfied lemma 5 holds. By lemma 4, $\pi_0 \neq 0$; hence, by the corollary to lemma 5 $\sum_{i=0}^{\infty} \pi_i \neq 0$ and $\{\pi_j'\}$ are stationary probabilities of P_n.

Remark. By assumptions (i), (ii), and (iii), P_n is, in fact, the transition matrix of a single class positive recurrent Markov chain. Thus $a_j > 0, j=0, 1, \ldots$ and $\pi_j' > 0, j=0, 1, \ldots$ and $\{\pi_j'\}$ $j=0, 1, \ldots$ is the unique set of stationary probabilities for P_n.

When multiple roots are present, the form of the stationary probabilities change, but in a classical pattern. Suppose θ is a multiple root of (1) of order μ. Then it is readily seen that θ is a common root of $\sum j^k p_j \theta^j - n^k \theta^n = 0$, $k=0, 1, \ldots, \mu-1$.

From this, it can be shown that $\pi_j^{(r)} = j^r \theta^j$ satisfies (3) for $j \geqslant n$ for $r=0, \ldots, \mu-1$. Hence, if $\theta_1, \theta_2, \ldots, \theta_m$ are the roots within the unit circle of (1) of multiplicity μ_i, $i=1, \ldots, m$, θ_1 the dominant, perhaps multiple, root, then

$$\pi_j = \sum_{i=1}^{\infty} \sum_{r=1}^{\mu_i - 1} K_r(i) \, \pi^{(r)}(i). \; j=0, 1, \ldots \tag{8}$$

where $\pi_j^{(r)}(i)=j^r\,\theta_i^j$, $i=1, \ldots, m$, and $K_r(i)$, $r=1, \ldots, \mu_i-1$;

$i=1, \ldots, m$ are arbitrary non-trivial constants that satisfy lemma 3. It can also be shown that lemma 4 holds; the matrix corresponding to A is

$$\tilde{A}=\begin{Bmatrix} 1 & \theta_1^{-1} & \theta_1^{-2} & \ldots & \theta_1^{-(n-1)} \\ 1 & -\theta_1^{-1} & -2\theta_1^{-2} & \ldots & -(n-1)\,\theta_1^{-(n-1)} \\ 1 & (-1)^2\,\theta_1^{-1} & (-2)^2\,\theta_1^{-2} & \ldots & -(1-n)^2\,\theta_1^{-(n-1)} \\ \vdots & & & & \\ 1 & (-1)^{\mu_1-1}\,\theta_1^{-1} & (-2)^{\mu_1-1}\,\theta_1^{-2} & \ldots & (1-n)^{\mu_1-1}\,\theta_1^{-(n-1)} \\ \vdots & & & & \\ 1 & (-1)^{\mu_m-1}\,\theta_m^{-1} & (-2)^{\mu_m-1}\,\theta_m^{-2} & \ldots & (1-n)^{\mu_m-1}\,\theta_m^{-(n-1)} \end{Bmatrix}$$

All other aspects of the proof of theorem 1 are unchanged by the presence of multiple roots of (1). Therefore, we can state the following, of which theorem 1 is a special case.

Theorem 2. If $\{P_n\}$ satisfies (i), (ii), and (iii) there are m ($m\leqslant n$) roots $\theta_1, \ldots, \theta_n$ to (1) within the unit circle. If μ_i is the multiplicity of θ_i, then $\sum_{i=1}^{m}\mu_i=n$. The quantities

$$\pi_i'=\pi_i\left(\sum_{j=0}^{\infty}\pi_j\right)^{-1},\ i=0,\ 1,\ \ldots,$$

where π_j, $j=0, 1, \ldots,$ are given by (8) are stationary probabilities of P_n.

The remark following theorem 1 holds here as well.

3. **The behaviour of the dominant root.** Special Case of Poisson Distribution. Suppose $\{p_v\}$ is a Poisson distribution with mean $\lambda=n+\varepsilon$, $\varepsilon>0$. Then if x_0 is the dominant root of (1) and $y=1-x_0$, taking logarithms in (1) we get

$$-(1+\varepsilon/n)\,y=\log\,(1-y) \tag{9}$$

Expanding $\log(1-y)$ and dividing through (9) by y yields

$$\varepsilon/n = y/2 + y^2/3 + y^4/4 + \ldots\ldots \qquad (10)$$

From (10) it is clear that y, $0<y\leqslant 1$, is the unique solution to (9). Let $\eta = 2\varepsilon/n$. From (10) we get the inequalities

$$\eta > y \text{ and } \eta > y + 2/3y^2 \qquad (11)$$

The second unequality in (11) provides an upper bound for y; namely

$$\begin{aligned} y &< 3/4\,(\sqrt{1+8/3\eta}-1) \\ &= \eta - 2/3n^2 + \ldots \end{aligned} \qquad (12)$$

On substituting η for y in all terms but the linear term in (10) and using the first inequality of (11) we get

$$\begin{aligned} \eta &< y + 2/3\eta^3 + 2/4\eta^2 + \ldots \\ &= y + 2/\eta\,(1/3\eta^3 + 1/4\eta^4 + \ldots + \eta + 1/2\eta^2 - \eta - 1/2\eta^2) \\ &= y + 2/\eta\,(-\log(1-\eta) - \eta - 1/2\eta^2); \end{aligned}$$

that is

$$\begin{aligned} y &> 2(1 + \eta + 1/\eta \log(1-\eta)\,) \\ &= \eta - 2/3\eta^2 - 2/4\eta^3 - \ldots \end{aligned} \qquad (13)$$

(**Note.** The two approximations are the same in the first two terms.)

We have proved

Theorem 3. If $\{p_v\}$ is Poisson with mean $\lambda = n+\varepsilon$, $\varepsilon>0$, then x_0, $0\leqslant x_0<1$, is the unique real solution to (1), upper and lower bounds on $1-x_0$ are given by (12) and (13), and $\lim_{\varepsilon\to 0} x_0 = 1$.

More general case. Suppose we have a family of distributions $\{p_v(\varepsilon),\ 0<\varepsilon\leqslant L\}$ such that

(iv) $\lambda(\varepsilon) = \sum_{i=0}^{\infty} i\,p_i(\varepsilon) = n+\varepsilon.$

(v) $\sigma^2(\varepsilon) = \sum_{i=0} (i-\lambda(\varepsilon))^2\, p_i(\varepsilon) < \infty,\ \forall\varepsilon$

(vi) $\exists\ \alpha, \beta \ni 0<\beta<p_0(\varepsilon)<\alpha<1,\ \forall\varepsilon.$

Let $x(\varepsilon)$, $0<x(\varepsilon)<1$ satisfy (1) when $\{p_v(\varepsilon)\}$ is the distribution.

Then

$$p_0(\varepsilon)+p_n\, x^n(\varepsilon)-x^n(\varepsilon)\leqslant 0$$

yielding

$$x(\varepsilon)\geqslant\left(\frac{p_0(\varepsilon)}{1-p_n(\varepsilon)}\right)^{1/n}\geqslant\beta^{1/n}=\gamma \text{ (say)}, \forall\varepsilon.$$

Thus $1>\gamma>\theta$ provides a uniform lower bound for $x(\varepsilon)$. Let $x=e^{-\theta}$.

Then $\theta\leqslant-\log\gamma$. Let

$$M_\varepsilon(\theta)=\sum_{i=0}^{\infty} p_i(\varepsilon)\, e^{-\theta i}.$$

Then for x satisfying (1) and on taking logarithms we have that θ satisfies

$$-n\theta=\log M_\varepsilon(\theta). \tag{14}$$

Since, $\sigma^2(\varepsilon)=\log'' M_\varepsilon(0)<\infty$, $\log M_\varepsilon(\theta)$ can be expanded, using Taylor's theorem, to

$$\log M_\varepsilon(\theta)=\log M_\varepsilon(0)+\log' M_\varepsilon(0)\,\theta+1/2\log'' M_\varepsilon(\tau)\,\theta^2$$

for some $0\leqslant\tau\leqslant\theta$. Evaluating each term and substituting in (14) we get

$$-n\theta=-\lambda(\varepsilon)\,\theta+$$

$$\frac{1/2\sum_{i=0}^{\infty} p_i(\varepsilon)^{-i\tau}\,\Sigma i^2 p(\varepsilon)\,e^{-i\tau}-\left(\sum_{i=0}^{\infty} ip_i(\varepsilon)\,e^{-i\tau}\right)^2\theta^2}{M_\varepsilon^2(\tau)} \tag{15}$$

Let

$$q_{i,\varepsilon}(\theta)=\frac{p_i(\varepsilon)\,e^{-i\theta}}{M_\varepsilon(\theta)},$$

$$\lambda(\varepsilon,\theta)=\sum_{i=0}^{\infty} iq_{i,\varepsilon}(\theta). \quad \text{Then}$$

the coefficient of $\theta^2/2$ in (15) becomes

$$\sum_{i=0}^{\infty} i^2 q_{i,\varepsilon}(\tau)-\left[\sum_{i=0}^{\infty} iq_{i,\varepsilon}(\tau)\right]^2=\sum_{i=0}^{\infty}[i-\lambda(\varepsilon,\tau)]^2 q_{i,\varepsilon}(\tau)$$

$$\geqslant q_{0,\varepsilon}(\tau)\,\lambda^2(\varepsilon,\tau)$$

$$\geqslant\beta\lambda^2(\varepsilon,\tau),\ \forall\varepsilon. \tag{16}$$

We now show that $\lambda(\varepsilon, \theta)$ is bounded away from 0 for every $0<\varepsilon<L$ and $0\leqslant\theta\leqslant-\log\gamma$.

$$\lambda(\varepsilon, \theta)=\frac{\sum_{i=0}^{\infty} i\, p_i(\varepsilon)\, e^{-\theta i}}{M_\varepsilon(\theta)}$$

$$\geqslant \sum_{i=0}^{\infty} i\, p_i(\varepsilon)\, e^{i \log \gamma}$$

$$\geqslant \sum_{i=1}^{k(n+L)} p_i(\varepsilon)\, e^{k(n+L) \log \gamma}.$$

If $k\geqslant 1$, using the Chebychev-Markov inequality and (vi), we have

$$\sum_{i=1}^{k(n+L)} p_i(\varepsilon)\, e^{k(n+L)\log\gamma}\geqslant[1-1/k-p_0(\varepsilon)]\, e^{k(n+L)\log\gamma}.$$

$$\geqslant(1-1/k-\alpha)\, e^{k(n+1)\log\gamma}.$$

Thus, choosing k large enough we see that

$$\inf_{\substack{0<\varepsilon<L\\ 0\leqslant\theta\leqslant-\log\gamma}} \lambda(\varepsilon, \theta)=\underline{\lambda}>0.$$ From (15) and (iv),

$$\varepsilon\geqslant 1/2\, \beta\underline{\lambda}^2\, \theta;\ \forall\varepsilon;$$

i.e., $$x(\varepsilon)\geqslant e^{\frac{-2\varepsilon}{\alpha\underline{\lambda}^2}}.\quad \forall\quad \varepsilon.$$

Consequently $\lim_{\varepsilon\to 0} X(\varepsilon)=1$. Summarizing we have proved.

Theorem 4. If $\{p_v(\varepsilon)\}$ $0<\varepsilon\leqslant L$, is a family of distributions satisfying (i), (ii), ..., (vi), then there exists a constant $\Delta\geqslant\underline{\lambda}^2>0$ such that $x(\varepsilon)\geqslant e^{-2\varepsilon/\Delta}$ for all $0<\varepsilon\leqslant L$.

4. Special Case of n=2. For $n=2$, θ_1 and θ_2 are distinct. Also, since θ_1 is real, θ_2 must be real. It can be shown that

$$\pi_i=\frac{\left|\frac{\theta_1}{\theta_2}\right|\ \theta_1^i+\theta_2^i}{\left|\frac{\theta_1}{\theta_2}\right|\ \frac{1}{1-\theta_1}+\frac{1}{1-\theta_2}},\ i=0, 1, \dots$$

REFERENCES

1. Brodheim, E., Derman, C., Prostocos, G., *On the Evaluation of a Class of Inventory Policies for Perishable Products such as Blood, submitted for publication in* Management Science.

2. Karlin, S., *A First Course in Stochastic Processes*, Academic Press, 1966.

3. Keilson J., *Green's Function Methods in Probability Theory*, Griffin (1965).

4. Lindley, D. V., *The theory of queues with a single server*, Proc. Cambridge Phil. Soc., 48, 277-289, 1952.

5. Smith, W. L., *On the distribution of queueing times*, Proc. Camb. Phil. Soc., 49, 449-461, 1953.

6. Takács, L., *Introduction to the Theory of Queues*, Oxford University Press, 1962.