# On the Invariant Measures of Some Discrete-Time Markov Processes

ALFRED M. BRUCKSTEIN

***Abstract*—Expressions for the moments of invariant measures corresponding to a class of discrete-time Markov processes are given. The processes under consideration assume values in $R^+$ and have stationary transition kernels of exponential type, generalizing the Rayleigh and gamma distributions. The moments of their stationary distributions, obtained by extending a method due to Wold, are given in the form of convergent infinite products of gamma functions.**

## I. Introduction

Consider a real discrete-time Markov processes $\{X_i | i \in N\}$ with $X_i \geq 0$. The evolution of the process is, by assumption, completely specified by the initial density $p_0(x)$ of $X_0$ and a given transition kernel $K(a|\alpha)$. The transition kernel is a density in $a$ since, by definition,

$$K(a|\alpha)\, da = \Pr[X_i \in (a, a+da) | X_{i-1} = \alpha] \qquad (1)$$

and the probability density evolution equations are, for $i \geqslant 0$,

$$p_{i+1}(x) = \int_0^\infty K(x|y) p_i(y)\, dy. \qquad (2)$$

An interesting question that arises concerns the asymptotic behavior of the unconditional state distribution given a certain transition kernel. It is well-known that for not too "pathological" kernels (see, e.g., [1]) a Markov process approaches a steady-state distribution, and the corresponding density has the following invariance property

$$p_\infty(x) = \int_0^\infty K(x|y) p_\infty(y)\, dy. \qquad (3)$$

The problem of establishing easily verified conditions on the transition probabilities that are both necessary and sufficient for the existence of such an invariant distribution is still unsolved. One only has the so-called Doeblin theory, exposed in Doob [2], which establishes sufficient conditions for a kernel to have an invariant distribution. It is known however, that Markov processes approach steady-state distributions under far milder conditions than those required by the Doeblin condition. In the sequel we therefore assume that an invariant distribution exists for the particular transition kernels considered here and establish some of its properties under this assumption.

Let us now define the $r$th moment of the transition kernel by

$$m_r(\alpha) = \int_0^\infty a^r K(a|\alpha)\, da \qquad (4)$$

which is obviously a function of $\alpha$, and similarly the $r$th moment of the invariant density by

$$M_r = \int_0^\infty a^r p_\infty(a)\, da. \qquad (5)$$

An immediate result is that

$$M_r = E[m_r(\alpha)], \qquad (6)$$

where the expectation is taken under the invariant density $p_\infty(\cdot)$. This relation proves to be useful in obtaining further results when a particular class of transition kernels is considered. We note that in the above definitions $r$ can be an arbitrary real number.

Manuscript received February 8, 1983; revised July 1, 1983. This work was supported in part by the National Science Foundation under Grant ECS78-10003. This paper was presented at the IEEE International Symposium on Information Theory, St. Jovite, PQ, Canada, September 26–30, 1983.

The author is with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

## II. A General Class of Transition Kernels

We consider the following type of kernels, which often arise in modeling learning or adaptation processes, gain control systems, and stochastic point processes with memory (see, e.g., [3]–[7])

$$K(a|\alpha) = \frac{1}{\Gamma(1+\beta)} G^\beta(a, \alpha) e^{-G(a,\alpha)} \frac{d}{da} G(a, \alpha). \qquad (7)$$

In the above we assume that $a, \alpha, \beta > 0$, $\Gamma$ is the gamma function and $G(a, \alpha)$ is considered to be of the form

$$G(a, \alpha) = [a/\Phi(\alpha)]^s, \qquad s > 0, \qquad (8)$$

where $\Phi(\alpha)$ is some arbitrary function. The above defined class of transition kernels generalizes the gamma distributions ($s = 1$) and the Rayleigh distribution ($s = 2$, $\beta = 1$) and thus yields a rather flexible collection of discrete-time Markov processes.

For the kernels under consideration, it is easy to obtain their moments, and we have

$$m_r(\alpha) = \frac{\Gamma(1+\beta+r/s)}{\Gamma(1+\beta)} \Phi^r(\alpha) = C_\beta(r/s)\Phi^r(\alpha) \qquad (9)$$

with $C_\beta(\cdot)$ defined accordingly. The moments of the invariant distribution, provided it exists, are therefore given by

$$M_r = C_\beta(r/s) E[\Phi^r(\alpha)]. \qquad (10)$$

Not unexpectedly the function $\Phi(\cdot)$, which determines the conditional mean value of the process by (9), turns out to be very important in determining the asymptotic behavior. If this function is known to be either convex or concave one immediately has, from Jensen's inequality, the following inequalities involving the asymptotic expected value of the process $X$

$$\mu = E[X_\infty] \begin{cases} \geq C_\beta(1/s)\Phi(\mu), & \text{if } \Phi \text{ convex;} \\ \leq C_\beta(1/s)\Phi(\mu), & \text{if } \Phi \text{ concave.} \end{cases} \qquad (11)$$

This provides the intervals to which $\mu$ may belong. If, for example, $\Phi(\cdot)$ is concave and increasing, then $\mu$ will always be less than or equal to the largest positive solution of $x = C_\beta(1/s)\Phi(x)$. On the other hand, if $\phi(\cdot)$ is convex and decreasing then the expected value of the invariant density has to be greater than or equal to the solution of that same equation.

## III. Product Expansions for Moments of the Invariant Measure

Wold analyzed a point process with Markovian interevent intervals having a simple exponential kernel (in our notation $\beta = s = 1$) with $\Phi(\alpha) = \alpha^{1/2}$. Using an ingenious trick he succeeded in deriving expressions for the moments of the stationary interval distributions [3]. We are now in a position to extend his results for the class of kernels defined by (7) for which

$$\Phi(\alpha) = A\alpha^\xi, \qquad \text{for } \xi \in (0,1). \qquad (12)$$

This function is concave and increasing in $\alpha$ and consequently we will have

$$\mu \leq \mu^* = [C_\beta(1/s)A]^{1/(1-\xi)}. \qquad (13)$$

From (10) we further obtain

$$M_r = C_\beta(r/s) A^r E[\alpha^{r\xi}]. \qquad (14)$$

and for all $r$ for which $r\xi < 1$ we can write, invoking successively

0018-9448/84/0100-0125$01.00 

the concavity and monotonicity of $\alpha^{r\xi}$,

$$M_r \le C_\beta(r/s)A^r\mu^{r\xi} \le C_\beta(r/s)A^r(\mu^*)^{r\xi}. \tag{15}$$

This provides immediate upper bounds on the moments up to $r = [1/\xi]$.

The trick used by Wold for the analysis of his Markov point process is based on the observation that (14) may be regarded as a recursive relation between generalized moments of the stationary distribution. Indeed if one reads (14) as

$$M_r = C_\beta(r/s)A^r M_{r\xi}, \tag{16}$$

it becomes clear that iteration of (16) provides the following expression for the $r$th moment

$$M_r = A^{r/(1-\xi)} \prod_{i=0}^{\infty} C_\beta(r\xi^i/s). \tag{17}$$

Note that (17) holds in general, and not only for $r < [1/\xi]$, the highest moment for which we had the bound in (15). Thus, we can obtain an infinite product development for *all* the moments of the stationary distribution. We still have to prove, however, that infinite products of the type

$$Z = \prod_{i=0}^{\infty} \frac{\Gamma\left(1 + \beta + \frac{r}{s}\xi^i\right)}{\Gamma(1+\beta)} \tag{18}$$

are convergent. This is done in the Appendix.

We note that the above method of finding the moments will work for any kernel which leads to relations of the form

$$M_r = F(r)M_{f(r)} \tag{19}$$

and for which iteration results in converging infinite products of the form

$$F(r)F(f(r))F(f(f(r)))\cdots. \tag{20}$$

## IV. Concluding Remarks

This note generalizes a method due to Wold which yields the moments of the invariant distribution for a class of discrete-time Markov processes. Once its moments are available, the stationary density $p_\infty(x)$ can, in principle, be obtained by appealing to the well-known formula

$$p_\infty(x) = L^{-1}\left\{\sum_{i=1}^{\infty} \frac{(-s)^i}{i!} M_i\right\}, \tag{21}$$

where $L^{-1}$ denotes the inverse Laplace transform. Indeed, note that the moment sequence determined by (17) obeys the following condition

$$\lim_{r\to\infty} \frac{1}{r} M_r^{1/r} = 0 \tag{22}$$

which can easily be proved using the bounds given by (A4) in the Appendix. According to Breiman [8, p. 182], this is a sufficient condition for the existence of a unique distribution which has the given moment sequence.

Although infinite product expansions are involved in the expressions for the moments, their numerical evaluation is not a problem since they converge at rather fast rates as the factors in the product approach 1. Also, the rate of convergence is faster for lower order moments which are usually the ones we are more interested in.

## Appendix

We prove that products of the type

$$Z = \prod_{i=0}^{\infty} \frac{\Gamma\left(1 + \beta + \frac{r}{s}\xi^i\right)}{\Gamma(1+\beta)} \tag{A1}$$

are always convergent using a rather general result on positive convex functions.

*Lemma:* If $g(\cdot)$ is a differentiable, positive and convex function and $\delta_n$ is a sequence of positive numbers for which $\sum_0^\infty \delta_n < \infty$ then

$$\prod_{i=0}^{\infty} \frac{g(\Delta + \delta_n)}{g(\Delta)} < \infty. \tag{A2}$$

*Proof:* One has for all $n > N$, when $\delta_n < 1$ that

$$g(\Delta) + g'(\Delta)\delta_n \le g(\Delta + \delta_n) \le g(\Delta) + (g(\Delta+1) - g(\Delta))\delta_n. \tag{A3}$$

This implies that

$$1 + \frac{g'(\Delta)}{g(\Delta)}\delta_n \le \frac{g(\Delta+\delta_n)}{g(\Delta)} \le 1 + \frac{(g(\Delta+1) - g(\Delta))}{g(\Delta)}\delta_n \tag{A4}$$

and, since the ratio $g(\Delta + \delta_n)/g(\Delta)$ is upper and lower bounded by a series whose product converges (this is a consequence of the fact that if $\sum|\eta_n| < \infty$ then $\prod(1 + \eta_n) < \infty$), it follows that

$$\prod_{n=0}^{\infty} \frac{g(\Delta + \delta_n)}{g(\Delta)} < \infty. \tag{A5}$$

Q.E.D.

Now setting $g(\cdot) = \Gamma(\cdot)$, $\Delta = 1 + \beta$ and $\delta_n = (r/s)\xi^n$, the convergence of $Z$ above follows immediately.

## Acknowledgment

The author would like to thank Prof. E. Masry and the two referees for their valuable comments.

## References

[1] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I–II. New York: Wiley, 1966.
[2] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.
[3] D. J. Lawrance, "Some models for stationary series of univariate events", in *Stochastic Point Processes*, P. A. W. Lewis, Ed. New York: Wiley, 1972.
[4] C-P. Wang and S. S. Rappaport, "Adaptive gain control," *IEEE Trans. Inform Theory*, vol. IT-17, pp. 625–627, Sept. 1971.
[5] M. F. Norman, *Markov Processes and Learning Models*. New York: Academic, 1972.
[6] A. C. Sanderson, W. M. Kozak, and T. W. Calvert, "Distribution coding in the visual pathway," *Biophysics J.*, vol. 13, pp. 218–244, 1973.
[7] A. M. Bruckstein, "Models of coding in the nervous system," MSc. thesis, Technion I.I.T., Haifa, Israel, Oct. 1980.
[8] L. Breiman, *Probability*. Reading, MA: Addison Wesley, 1968.

0018-9448/84/0100-0126$01.00 