MARKOV RENEWAL PROCESSES: DEFINITIONS AND PRELIMINARY PROPERTIES[1]

By Ronald Pyke[2]

University of Washington

1. Summary. This paper contains the definition of and some preliminary results on Markov Renewal processes and Semi-Markov processes. The close relationship between these two types of processes is described. The concept of regularity is introduced and characterized. A classification of the states of a Markov Renewal process is described and studied.

2. Introduction. At the International Congress of Mathematicians held at Amsterdam in 1954, Lévy [1] and Smith [2] independently presented papers in which a new class of stochastic processes, called Semi-Markov processes (S.-M.P.) by both authors, was defined. These processes are generalizations of both continuous and discrete parameter Markov processes with countable state spaces. In the case of Lévy, the suggestion of this possible generalization is credited to K. L. Chung. Also in 1954, Takács [3] introduced essentially the same type of stochastic process, and applied them to some problems in Counter theory.

A rough, yet descriptive, definition of an S.-M.P. would be that it is a stochastic process which moves from one to another of a countable number of states with the successive states visited forming a Markov chain, and that the process stays in a given state a random length of time, the distribution function (d.f.) of which may depend on this state as well as on the one to be visited next. It is thus a Markow Chain for which the time scale has been randomly transformed.

The family of stochastic processes to be defined and studied in this paper, called Markov Renewal processes (M.R.P.), may be shown to be equivalent to the family of S.-M.P.'s. An M.R.P. is one which records at each time t the number of times a particle has visited each of the possible states up to time t, if the particle moves from state to state according to a Markov Chain and if the time required for each successive move is a random variable (r.v.) whose d.f. may depend on the two states between which the move is being made.

It will be seen, after the definition of an M.R.P. has been formalized in Section 3 below, that a Renewal process (*i.e.*, a sequence of independent, identically distributed nonnegative r.v.'s) is equivalent to the special case of an M.R.P. with one state. However, as will become evident in the discussions below and in [4], the relationship between Renewal theory and that of M.R.P.'s is very much

Received January 18, 1960; revised May 9, 1961.

[1] The results of this paper were presented in an invited paper to the Annual Meeting of the Institute of Mathematical Statistics, December 1959.

[2] This research was supported by the Office of Naval Research under Contract Number Nonr-266(59), Project Number 042-205, Columbia University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

stronger than this fact alone indicates. Indeed, it would not be overexaggerating to describe the present theory as a marriage of the theories of Markov Chains and of Renewal processes. It is this close relationship which suggested the nomenclature, Markov Renewal process.

In Section 3, M.R.P.'s and S.-M.P.'s are defined, as well as some related processes. Although the processes studied here have been given "constructive" definitions, and hence are automatically separable, and have no instantaneous states (as may the S.-M.P.'s defined by Lévy [1]), there still exists the problem of whether or not an infinite number of transitions may be made in a finite interval of time. This problem is studied in Section 4, where a complete characterization is given of those M.R.P.'s for which only a finite number of transitions may be made in a finite interval of time. Such an M.R.P., with only a slight qualification, is said to be regular. It is proved that every M.R.P. with only finitely many states is regular. Several sufficient conditions for regularity are also given. In Section 5, an extension is made of the terminology used in classifying the states of a Markov Chain, to cover the case of M.R.P.'s. It is shown that the classification of any particular state in an M.R.P. is very greatly dependent upon the classification of this state in an embedded Markov Chain.

Many papers have been written on S.-M.P.'s and M.R.P.'s since 1954, mostly in the past two years. All papers known to this author which concern these processes and which are not referred to in the body of this paper are included in the supplementary references at the end of this paper, thus providing the reader with a complete list of references on this subject.

3. Definitions and notations. Bold face letters such as **F**, **Q**, **H**, **f**, **q**, **h**, are consistently used in this paper to denote (real) matrix-valued functions, with the capital letters having domain $(-\infty, \infty)$ and the lower-case letters having domain $(0, \infty)$. Mass functions (i.e., distribution functions whose total variations need not be equal to one) will be denoted by capital italic letters, whereas the corresponding lower-case letters will denote their respective Laplace-Stieltjes (L.-S.) transforms. For example, for $s \geqq 0$, $f(s) = \int_{-\infty}^{\infty} e^{-sx}\, dF(x)$, which may for the present be infinite. It will be convenient to introduce the degenerate d.f.'s, $U_c(x) = 1$ or 0, according as $x \geqq$ or $< c$. Unless otherwise stated, the subscripts i, j in a matrix (b_{ij}) or elsewhere will run through the integers greater than or equal to 1 and not greater than m, where m, fixed, is either a finite positive integer or plus infinity. The following convolution notation is used in this and subsequent papers. $K(t)*L(t) = \int_{0-}^{t} K(t-y)\, dL(y)$ if $t \geqq 0$ and $= 0$ if $t < 0$ for functions K and L for which the Lebesgue-Stieltjes integral is defined. Write $K*L$ for the function $K(\cdot)*L(\cdot)$, $K^{(0)} = U_0(\cdot)$, $K^{(n)} = K^{(n-1)}*K$, $(n = 1, 2, \cdots)$ and $K^{(-1)} = \sum_{n=0}^{\infty} K^{(n)}$ whenever the series converges.

DEFINITION 3.1. Let $\mathbf{Q} = (Q_{ij})$ be a matrix-valued function on $(-\infty, \infty)$. $\mathbf{Q}$ is called a *matrix of transition distributions* if the Q_{ij} are mass functions satisfying (i) $Q_{ij}(t) = 0$ for $t \leqq 0$ and (ii) $\sum_{j=1}^{m} Q_{ij}(+\infty) = 1$, $(1 \leqq i < m + 1)$.

For each i and every real t, set $H_i(t) = \sum_{j=1}^{m} Q_{ij}(t)$. With this notation, (ii) of Definition 3.1 is equivalent to stating that every H_i is a d.f.

DEFINITION 3.2. The $m \times 1$ vector $\mathbf{A} = (a_1, a_2, \cdots, a_j, \cdots)$, is called a *vector of initial probabilities* if it satisfies (i) $a_j \geqq 0$ and (ii) $\sum_{i=1}^m a_i = 1$.

DEFINITION 3.3. The (J, X)-process[3] is defined as any two-dimensional stochastic process $\{(J_n, X_n); n \geqq 0\}$ defined on a complete probability space $(\Omega, \mathcal{B}, P)$, that satisfies $X_0 = 0$ a.s.,

$$P[J_0 = k] = a_k \tag{3.1}$$

and

$$P[J_n = k, X_n \leqq x \mid J_0, J_1, X_1, J_2, X_2, \cdots, J_{n-1}, X_{n-1}] \overset{\text{a.s.}}{=} Q_{J_{n-1},k}(x) \tag{3.2}$$

for all $x \varepsilon (-\infty, \infty)$ and $1 \leqq k < m + 1$.

Set $S_n = \sum_{i=0}^n X_i$ for $n \geqq 0$. The (J, X)-process defined above is closely related to a Markov process as shown in

LEMMA 3.1. *The two-dimensional (J, S)-process is a Markov process, and the J-process is a Markov Chain. In particular for $1 \leqq k < m + 1$, $n > 0$*

$$P[J_n = k, S_n \leqq y \mid J_0, J_1, S_1, \cdots, J_{n-1}, S_{n-1}] \overset{\text{a.s.}}{=} Q_{J_{n-1},k}(y - S_{n-1}) \tag{3.3}$$

and

$$P[J_n = k \mid J_0, J_1, \cdots, J_{n-1}] \overset{\text{a.s.}}{=} Q_{J_{n-1},k}(+\infty). \tag{3.4}$$

PROOF. That the J-process is a Markov Chain satisfying (3.4) is an immediate consequence of (3.1), (ii) of Definition 3.1 and the Lebesgue monotone convergence theorem applied to (3.2) when $x \to +\infty$. That the (J, S)-process is a Markov process is implied by (3.3) and (3.1). To verify (3.3), write the left-hand side of this expression as

$$P[J_n = k, X_n \leqq y - S_{n-1} \mid J_0, J_1, S_1, \cdots, J_{n-1}, S_{n-1}].$$

Since the conditioning σ-field of this conditional expectation is equal to that of the left-hand side of (3.2), and since this σ-field is generated by a finite number of r.v.'s it is known that a conditional probability distribution (as defined by Doob [5], p. 26) exists, by means of which it is easily seen that (3.2) implies (3.3).

Because of (3.4), it is natural to define $p_{ij} = Q_{ij}(+\infty)$ and $\mathbf{P} = (p_{ij})$. By Definition 3.1, $\mathbf{P}$ is a stochastic matrix. Furthermore, if $p_{ij} > 0$, define $F_{ij} = p_{ij}^{-1}Q_{ij}$, while if $p_{ij} = 0$ set $F_{ij} = U_1$. (Actually, when $p_{ij} = 0$, F_{ij} may be chosen arbitrarily. There is some notational advantage, however, in choosing a d.f. which has all moments finite, but the particular choice of a degenerate d.f. has no special merit.) Set $\mathbf{F} = (F_{ij})$. For convenience define $J_\infty = \infty$. Furthermore, introduce the following notation for moments.

$$\begin{aligned} b_{ij} &= \int_0^\infty t\, dF_{ij}(t), \qquad \eta_i = \int_0^\infty t\, dH_i(t) \\ \sigma_{ij}^2 &= \int_0^t (t - b_{ij})^2\, dF_{ij}(t), \qquad \sigma_i^2 = \int_0^\infty (t - \eta_i)^2\, dH_i(t). \end{aligned} \tag{3.5}$$

[3] As a convenient abbreviated notation, stochastic processes will be denoted by the letter(s) used to designate the corresponding r.v.'s.

The following easily verified consequences of the above definitions will be useful in later discussions.

$$
\begin{aligned}
&P[X_n \leqq x \mid J_0, \cdots, J_{n-1}] = H_{J_{n-1}}(x),\\
&P[J_n = j \mid J_0, \cdots, J_{n-1}] = p_{J_{n-1},j},\\
&P[X_n \leqq x \mid J_0, \cdots, J_n] = F_{J_{n-1},J_n}(x),\\
(3.6)\quad &P[X_{n_1} \leqq x_1, X_{n_2} \leqq x_2, \cdots, X_{n_k} \leqq x_k \mid J_n; n \geqq 0]\\
&\qquad = P[X_{n_1} \leqq x_1, \cdots, X_{n_k} \leqq x_k \mid J_0, J_1, \cdots, J_{n_k}]\\
&\qquad = \prod_{i=1}^{k} F_{J_{n_i-1},J_{n_i}}(x_i)
\end{aligned}
$$

for $0 < n_1 < \cdots < n_k$, all equalities holding with probability one. It follows from the last two relationships that $X_{n_1}, X_{n_2}, \cdots, X_{n_k}$ are mutually conditionally independent given $J_{n_1-1}, J_{n_1}, \cdots, J_{n_k-1}, J_{n_k}$ (*e.g.*, *cf.* [6], Definition 3).

In Renewal theory, the basic process studied is that which gives the number of partial sums or renewals in the intervals $(0, t]$ for all $t \geqq 0$. The natural analogues to this for the present theory are the counting processes defined now.

DEFINITION 3.4. The integer-valued stochastic processes $\{N(t); t \geqq 0\}$ and $\{N_j(t); t \geqq 0\}$ are defined by $N(t) = \sup\{n \geqq 0: S_n \leqq t\}$ and $N_j(t) =$ no. of times $J_k = j$ for $0 < k < N(t) + 1$.

Notice that without added restrictions on m and/or $\mathbf{Q}$, $N(t)$ may be infinite with positive probability. Notice also that the counting functions N_j are defined so as not to record the value of J_0. Setting $\mathbf{N}(t) = (N_1(t), N_2(t), \cdots, N_j(t), \cdots)$, the stochastic process $\{\mathbf{N}(t); t \geqq 0\}$ is called a *Markov Renewal Process* (M.R.P.) *determined by* $(m, \mathbf{A}, \mathbf{Q})$. Clearly $N(t) = \sum_{j=1}^{m} N_j(t)$ a.s.

Related to an M.R.P. is the stochastic process defined now which simply records the state of the process at each time point.

DEFINITION 3.5. The Z-process, $\{Z_t; t \geqq 0\}$ defined by $Z_t = J_{N(t)}$ is called a *Semi-Markov Process* (S.-M.P.) determined by $(m, \mathbf{A}, \mathbf{Q})$.

Let us introduce some additional vocabulary to facilitate later discussions. We shall say that a "transition" of an M.R.P. has occurred at each of the time points $S_0, S_1, S_2, \cdots$. The process (either an M.R.P. or an S.-M.P.) is said to be "in state i" at time t, if, and only if, $Z_t = i$.

As defined in Definition 3.4, an M.R.P. is a vector-valued process (infinite dimensional if $m = \infty$). It is clear that one could construct one-dimensional processes that are probabilistically equivalent to the $\mathbf{N}$-process. For example, the Y-process defined by $Y_t = j + 1 - 2^{-k+1}$ on the set $[J_{N(t)-n} = j, 0 \leqq n < k, J_{N(t)-k} \neq j]$ and $= \infty$ on the set $[N(t) = \infty]$ may be shown to be equivalent to the $\mathbf{N}$-process, since it records both the state of the M.R.P. and the number of preceding consecutive transitions to state i for each $t > 0$. For most discussions, the r.v.'s $N_j(t)$, and especially their expectations, play the central role, as does $N(t)$ for the special case of a Renewal process, namely the case $m = 1$. However, the Y-process representation serves to emphasize the relationships between an M.R.P. and an S.-M.P. The Y-process is always an S.-M.P., and is called the *associated S.-M.P.* of the given M.R.P. It is equal almost surely to the Z-process if and only if, $p_{ii} = 0$ for every state i which can be reached with positive probability. Otherwise the Y-process always has an infinity of states regardless of the value of m. It follows from this Y-process representation of an M.R.P. that, theoretically at least, any results about an M.R.P. may be derived from theorems concerning S.-M.P.'s. It is, however, both convenient and practical to keep these two kinds of processes distinct, and to use the process most natural for a given problem. For computation of moments of recurrence times (cf. [7]), it is natural to work this out for M.R.P.'s since most applications involve processes in which a transition from a state to itself is possible. It should be observed that although an M.R.P. has a finite number of states, the associated S.M.P. will in most applications have an infinite number of states, as is the case, for example, for a Renewal process ($m = 1$). On the other hand, for problems concerning the limiting stationarity of a given M.R.P., transitions from a state to itself play no essential role. One may then, without loss of generality, work with the related matrix of transition distributions $\mathbf{Q}^* = (Q_{ij}^*)$ defined by $Q_{ii}^* = Q_{ii}$ if $p_{ii} = 1$, $Q_{ii}^* = 0$ if $p_{ii} < 1$ and $Q_{ij}^* = Q_{ij}^*[1 - Q_{ii}]^{(-1)}$ if $i \neq j$ and $p_{ii} < 1$. One may verify

LEMMA 2. *Every* S.-M.P. *determined by* $(m, \mathbf{A}, \mathbf{Q})$ *has the same family of joint* d.f.'s *as every* S.-M.P. *determined by* $(m, \mathbf{A}, \mathbf{Q}^*)$.

Any S.-M.P. determined by $(m, \mathbf{A}, \mathbf{Q}^*)$ is called a *corresponding* S.-M.P. of the given M.R.P.

When $m = 1$, a Markov Renewal process becomes a Renewal process, the theory of which is extensive (*cf.* the survey paper on Renewal theory by Smith [8]). When the transition distributions are of the form $Q_{ij} = p_{ij}U_1$ for all i and j, the Markov Renewal process becomes a Markov Chain, and in this case is equivalent to its corresponding S.-M.P. by virtue of the constant transition times. Moreover, a continuous parameter Markov process with m states, *all of which are stable*, is a special case of an M.R.P. (in fact, of an S.-M.P.) for which the Q_{ij} are of the form

$$Q_{ij}(t) = p_{ij} \max (0, 1 - e^{-\lambda_i t}) \qquad (-\infty < t < \infty) \tag{3.7}$$

for constants $\lambda_i > 0$, and $p_{ii} = 0$ for every i.

4. Finiteness of $N(t)$ and regularity. It may easily be deduced from the constructive definitions of an M.R.P. and an S.-M.P. given in Section 3 that they are separable and that almost all sample functions of the Y-process, and of the Z-process, are step-functions over an interval of the form $[0, L)$ and identically equal to infinity over $[L, \infty)$, where $L > 0$ is a possibly infinite r.v. which is a Borel function of the Y-process. Clearly, the sets $[L < \infty]$ and $[N(t) = \infty, t \geqq L]$ differ only by a set of measure zero. It is important to be able to characterize those M.R.P.'s for which $L \overset{\text{a.s.}}{=} \infty$, or equivalently, those for which

$N(t) < \infty$ for all t. We shall first verify the intuitive result that in the case of $m < \infty$, the (a.s.) finiteness of $N(t)$ is always true.

LEMMA 4.1. *If $m < \infty$, then for all states i,*

$$P[N(t) < \infty, \quad \text{for all} \quad t \geqq 0] = 1. \tag{4.1}$$

PROOF. By Definition 3.4, $N(t)$ is nondecreasing. It suffices, therefore, to prove that $P[N(t) < \infty \mid Z_0 = i] = 1$ for each $t \geqq 0$, and for every i for which $a_i > 0$. Suppose $a_i > 0$. By Definition 3.4, (3.3) and (3.4) one obtains for $t \geqq 0$

$$\begin{aligned} P[N(t) \geqq n | Z_0 = i] = P[S_n \leqq t | J_0 = i] & \\ &= \sum_{\mathcal{S}_{n,i}} \mathop{*}_{j=0}^{n-1} Q_{\alpha_j \alpha_{j+1}}(t) \\ &= \sum_{\mathcal{S}_{n,i}} \left\{ \prod_{J=0}^{n=1} p_{\alpha_j \alpha_{j+1}} \right\} \mathop{*}_{j=0}^{n-1} F_{\alpha_j \alpha_{j+1}}(t) \end{aligned}$$

where $*$ denotes convolution of the indicated d.f.'s and where

$$\mathcal{S}_{n,i} = \{(\alpha_0, \alpha_1, \cdots, \alpha_n) : \alpha_0 = i, \alpha_j \ \text{an integer}, \ 1 \leqq \alpha_j \leqq m (1 \leqq j \leqq n)\} \tag{4.2}$$

is the set of all paths of length $n + 1$ of the J-process for which $J_0 = i$. Define $F = \max_{i,j} F_{ij}$. It is well known that for mass functions F_1, F_2, G_1 and G_2 for which $F_1 \leqq G_1$ and $F_2 \leqq G_2$, one has $F_1 * F_2 \leqq G_1 * G_2$. Consequently, it follows from (3.6) that

$$P[N(t) \geqq n \mid Z_0 = i] \leqq F^{(n)}(t) \sum_{\mathcal{S}_{n,i}} \prod_{j=0}^{n-1} p_{\alpha_j \alpha_{j+1}} = F^{(n)}(t).$$

Since $m < \infty$, one has by Definition 3.1 that $F(0) = 0$ and so for $t > 0$, $F^{(n)}(t) \to 1$ as $n \to +\infty$.

Any d.f. F satisfying $F(0) = 0$ which is an upper bound for every F_{ij}, would have sufficed in the proof of Lemma 4.1. An alternative choice of F which has a more intuitive interpretation than that used in the above proof is

$$F = 1 - \prod_{i,j} [1 - F_{ij}],$$

the d.f. of the minimum of a family of independent r.v.'s, one corresponding to each d.f. F_{ij}.

A consequence of Lemma 4.1 is that almost all path functions of a Y-process with $m < \infty$ are step-functions over $[0, \infty)$, as is also true for the corresponding S.-M.P.

Consider now the case of unrestricted m. For this case, it is necessary to impose restrictions in order to insure the (a.s.) finiteness of $N(t)$. To see this, the simplest example is the degenerate one for which $Q_{j,j+1} = U_{2^{-j}}(\cdot)(j \geqq 1)$ and all other $Q_{ij} = 0$. For this example, $N(t) \overset{\text{a.s.}}{=} n$, whenever $1 - 2^{-n} \leqq t < 1 - 2^{-n-1}$ for $n \geqq 0$, while $N(t) \overset{\text{a.s.}}{=} \infty$, whenever $t \geqq 1$. That is, $L = 1$(a.s.). In what follows, a necessary and sufficient condition for the (a.s.) finiteness of $N(t)$ for every $t \geqq 0$ is given, as well as several sufficient conditions which are applicable in the more common situations.

For any $c > 0$, define the truncated moments $b_{ij}^{(c)} = \int_0^c t\, dF_{ij}(t)$. Clearly $b_{ij} = \lim_{c \to \infty} b_{ij}^{(c)}$. Define the family of integer sequences

$$\mathcal{S}_i = \{(\alpha_0, \alpha_1, \cdots) : \alpha_0 = i, 1 \leqq \alpha_j < m + 1(j \geqq 1)\}.$$

DEFINITION 4.1. A state i of an M.R.P. determined by $(m, \mathbf{A}, \mathbf{Q})$ is said to be *regular*-**A** if either $P[(J_0, J_1, \cdots) \varepsilon \mathcal{S}_i] \equiv a_i = 0$, or $a_i > 0$ and there exists a measurable subset $\mathcal{A} \subset \mathcal{S}_i$ such that $P[(J_0, J_1, \cdots) \varepsilon \mathcal{A} \mid J_0 = i] = 1$ and such that for every $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathcal{A}$ and for every $c > 0$ at least one of the series

$$\sum_{j=0}^{\infty} [1 - F_{\alpha_j \alpha_{j+1}}(c)], \qquad \sum_{j=0}^{\infty} b_{\alpha_j \alpha_{j+1}}^{(c)} \tag{4.3}$$

diverges. An M.R.P. determined by $(m, \mathbf{A}, \mathbf{Q})$ is said to be *regular*-**A** if each of its states is regular-**A**. If these properties hold for all initial distributions **A**, the state or the M.R.P. will be called *regular*. Since whether an M.R.P. is regular or not depends only upon the nature of **Q**, we shall alternatively speak of **Q** as being regular.

It would have sufficed in the above definition to have required the divergence of one of the series in (4.3) for only those sequences in $\mathcal{A}$ for which $p_{\alpha_j \alpha_{j+1}} > 0$ for every j. This is so because of the convention made earlier, that whenever $p_{ij} = 0$, $F_{ij} = U_1$ and hence $b_{ij}^{(c)} = 1 > 0$ for all $c \geqq 1$. It is shown in the following theorems that the concepts of the above definition may be used to characterize the (a.s.) finiteness of $N(t)$.

THEOREM 4.1. *For any given state i of an M.R.P. determined by* $(m, \mathbf{A}, \mathbf{Q})$,

$$P[J_0 = i, N(t) = \infty \quad \textit{for some} \quad t \geqq 0] = 0 \tag{4.4}$$

if and only if i is regular-**A**.

PROOF. The theorem is obvious whenever $a_i = 0$. Assume, therefore, that $a_i > 0$. For any $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathcal{S}_i$, one can show that

$$\begin{aligned} P[N(t) = \infty \mid J_k = \alpha_k, k \geqq 0] &= \lim_{n \to \infty} P[S_n \leqq t \mid J_k = \alpha_k, k \geqq 0] \\ &= \mathop{*}_{j=0}^{\infty} F_{\alpha_j \alpha_{j+1}}(t) \end{aligned} \tag{4.5}$$

by the last relationship of (3.6). It is known, and easily verified, that for nonnegative r.v.'s Kolmogorov's Three-Series criterion (*cf.* [9], p. 236) for a.s. convergence of a series of independent r.v.'s, becomes a "two-series" criterion, namely, "If $\{V_n : n \geqq 1\}$ is a sequence of nonnegative independent r.v.'s, then the series $\sum_{n=0}^{\infty} V_n < \infty$ (a.s.) if, and only if, for some finite $c > 0$,

$$\sum_{n=1}^{\infty} P[V_n > c] < \infty, \quad \text{and} \quad \sum_{n=1}^{\infty} E[\min(V_n, c)] < \infty.$$

Furthermore, if the series does not converge a.s. then it diverges a.s." Now (4.5) implies that with respect to the indicated conditional probability measure, the S_n's are representable as partial sums of independent nonnegative r.v.'s. Therefore, by the above version of Kolmogorov's theorem, one has that

$$\underset{j=0}{\overset{\infty}{*}} F_{\alpha_j \alpha_{j+1}}(t) = 0 \qquad (t \geqq 0) \tag{4.6}$$

if, and only if, for all $c > 0$, at least one of the series

$$\sum_{n=1}^{\infty} P[X_n > c \mid J_k = \alpha_k, k \geqq 0], \qquad \sum_{n=1}^{\infty} E[\min (X_n, c) \mid J_k = \alpha_k, k \geqq 1]$$

diverges, which is easily checked to be equivalent to specifying that at least one of the series given in (4.3) diverges. Therefore, if state i ($a_i > 0$) is regular-**A**, there exists a set $\mathfrak{A} \subset \mathcal{S}_i$ of conditional probability equal to one, such that for every $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathfrak{A}$, (4.6) is satisfied, and hence by (4.5)

$$P[N(t) = \infty \mid J_0 = i] = 0,$$

thus verifying (4.4). Conversely, if (4.4) is satisfied, then

$$0 = P[N(t) = \infty \mid J_0 = i] = E\left[\underset{n=0}{\overset{\infty}{*}} F_{J_n J_{n+1}}(t) \mid J_0 = i\right].$$

Because of the nonnegativeness of the integrand, this implies

$$P\left[\underset{n=0}{\overset{\infty}{*}} F_{J_n J_{n+1}}(t) = 0 \mid J_0 = i\right] = 1.$$

Consequently, in Definition 4.1, one may choose $\mathfrak{A} \subset \mathcal{S}_i$ to be the set of all α-sequences satisfying (4.6). Hence, state i is regular-**A**.

COROLLARY 4.1. *For an* M.R.P. *determined by* $(m, \mathbf{A}, \mathbf{Q})$,

$$P[N(t) < \infty \quad \text{for all} \quad t] = 1$$

*if, and only if, it is regular-***A**.

COROLLARY 4.2. *For a given* m *and* $\mathbf{Q}$,

$$P[N(t) < \infty \quad \text{for all} \quad t] = 1$$

for all choices of a vector of initial probabilities if, and only if, $\mathbf{Q}$ *is regular.*

The foregoing theorems give a complete characterization of those M.R.P.'s having almost all sample functions equal to step-functions over $(0, \infty)$. In many practical situations, due to additional assumptions being stated, it is not necessary to check completely the conditions for regularity as given in Definition 4.1. In many instances, weaker sufficient conditions are available, and possibly are more easily checked. Some of these are given in the following discussion.

The simplest sufficient condition is a consequence of Lemma 4.1 namely that if $m < \infty$, then the M.R.P. is regular.

It is also easily shown that if for each α-sequence in a subset $\mathfrak{A} \subset \mathcal{S}_i$ of (conditional) probability one, there exists a finite $M > 0$, such that $F_{\alpha_j\alpha_{j+1}}(M) = 0$, then state i is regular. A particular application of this is to Markov Chains over discrete time, of which all states must, therefore, be regular.

If $\sum_{j=0}^{\infty} b_{\alpha_j\alpha_{j+1}} < \infty$ on some set of α-sequences in $\mathcal{S}_i$ of positive (conditional) measure, then i is not a regular state. This is a simple consequence of the fact that the convergence of the series of expectations of a sequence of independent r.v.'s, implies the convergence (a.s.) of the series of r.v.'s.

If for every α-sequence in a subset $\mathcal{A} \subset \mathcal{S}_i$ of (conditional) probability one, either $\sum_{j=0}^{\infty} \sigma^2_{\alpha_j\alpha_{j+1}} < \infty$, or there exists a finite $M > 0$ (possibly depending on the sequence) such that $F_{\alpha_j\alpha_{j+1}}(M) = 1$ $(j \geqq 0)$, then state i is regular if and only if $\sum_{j=0}^{\infty} b_{\alpha_j\alpha_{j+1}} = \infty$ for every $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathcal{A}$. This again is a simple consequence of known results on sums of (positive) independent r.v.'s (*cf.* [9], p. 236).

As a special case of the above, consider the following sufficient condition. If for every α-sequence in a subset $\mathcal{A} \subset \mathcal{S}_i$ of (conditional) probability one, either $\sum_{j=0}^{\infty} \sigma^2_{\alpha_j\alpha_{j+1}} < \infty$, or there exists a finite $M > 0$ (possibly depending on the sequence) such that $F_{\alpha_j\alpha_{j+1}}(M) = 1$ $(j \geqq 0)$, and there exist two real sequences $\{\delta_j\}, \{\eta_j\}$ of positive numbers such that $\sum_{j=0}^{\infty} \delta_j\eta_j = \infty$ and $F_{\alpha_j\alpha_{j+1}}(\delta_j) \leqq 1 - \eta_j$ $(j \geqq 0)$ for every $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathcal{A}$, then state i is regular. This follows immediately from the preceding since under these conditions $b_{\alpha_j\alpha_{j+1}} \geqq \delta_j\eta_j$. This condition is a corrected version of one due to Smith [2] (see also [10]). (When reading [2], the reader should note the different meaning of the word regular as it is used there.)

Consider now a condition designed primarily for continuous parameter Markov processes with an at most countable number of states. If there exists a set $\{\lambda_{ij}: 1 \leqq i, j < m + 1\}$ of finite positive numbers such that for every α-sequence in a subset $\mathcal{A} \subset \mathcal{S}_i$ of probability one, $F_{\alpha_j\alpha_{j+1}}(t) = 1 - \exp(-\lambda_{\alpha_j\alpha_{j+1}}t)$ for all $t \geqq 0$, then state i is regular if, and only if,

$$\sum_{j=0}^{\infty} \lambda^{-1}_{\alpha_j\alpha_{j+1}} = \infty \tag{4.7}$$

for all $(\alpha_0, \alpha_1, \cdots) \varepsilon \mathcal{A}$. That regularity implies (4.7) is immediate. It suffices to show that (4.7) implies the divergence of one of the series in (4.3). This is best seen by simply evaluating the series (4.3) to be

$$\sum_{j=0}^{\infty} \exp(-c\lambda_{\alpha_j\alpha_{j+1}}),$$

$$\sum_{j=0}^{\infty} \{\lambda^{-1}_{\alpha_j\alpha_{j+1}}[1 - \exp(-c\lambda_{\alpha_j\alpha_{j+1}})] - c\exp(-c\lambda_{\alpha_j\alpha_{j+1}})\}.$$

Assuming (4.7), one has that if the first series converges, then for j sufficiently large $1 - \exp(-c\lambda_{\alpha_j\alpha_{j+1}}) > \frac{1}{2}$, and so the second series must diverge. For $\lambda_{ij} = \lambda_i$, one obtains the known result for Markov processes, and if, moreover, one has $p_{j,j+1} = 1$, the above is the well known result for pure Birth processes (*cf.* [11] p. 349, [5] p. 271 and [12] p. 406).

Lastly, we mention a sufficient condition which is in terms of the underlying Markov Chain. If state i, considered as a state of the J-process, a Markov Chain by Lemma 3.1, is such that with (conditional) probability one, the J-process, starting in state i, will reach a recurrent state (*cf.* [12] p. 353), then state i is regular. To see this let j (possibly equal to i) be the first recurrent state that is reached, and suppose it was reached at the n_0th transition (n_0 may be zero). Let $n_1, n_2, \cdots$ be the successive integers n at which $J_n = j$. Set $T_0 = S_{n_0}$ and $T_k = S_{n_k} - S_{n_{k-1}}$. By assumption, the r.v. n_k are finite a.s., and hence, so are the T_k's. Furthermore, $\{T_k : k \geqq 1\}$ forms a Renewal process and so $\sum_{k=1}^{\infty} T_k = \infty$ (a.s.). Since $\sum_{n=1}^{\infty} X_n = \sum_{k=0}^{\infty} T_k$ (a.s.), the former series diverges and hence $N(t) < \infty$ (a.s.) for every t, as required.

5. Classification of states. In this section, the states of an M.R.P. will be classified in much the same manner as is done for Markov Chains, the terminology of the latter being retained. The reader is referred to Feller [12] and to Chung [13] for material on Markov Chains. To facilitate the definitions to follow, it is assumed that for all M.R.P.'s considered below, every initial probability is positive. Consider the notation defined, for all i, j and $t \geqq 0$, by

$$P_{ij}(t) = \begin{cases} P[Z_t = j \mid Z_0 = i] & \text{if } t \geqq 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{5.1}$$

$$G_{ij}(t) = \begin{cases} P[N_j(t) > 0 \mid Z_0 = i] & \text{if } t \geqq 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{5.2}$$

and let the moments (possibly infinite) of G_{ij} be denoted by μ_{ij}. According to these definitions, $P_{ij}(t)$ is the probability that an M.R.P., initially in state i, is in state j at time t, while G_{ij} is a mass function representing the probability distribution of the time (first passage time) until the next transition into state j of a process which is initially in state i. Notice that when $i = j$, this definition does not require the process to leave state i during the first passage time. μ_{ii} will be called the mean recurrence time of state i.

Definition 5.1. (a) States i and j are said to *communicate* if, and only if, either $G_{ij}(\infty)G_{ji}(\infty) > 0$ or $i = j$.

(b) Communication is an equivalence relation, and the disjoint equivalence classes are called *classes* and are denoted by C_i (whenever $i \varepsilon C_i$).

(c) An M.R.P. is said to be *irreducible* if, and only if, there is only one class.

(d) A class C_i is said to be *essential* (or closed) if, and only if, for all $j \varepsilon C_i$ and for all $t \geqq 0$, $\sum_{k \varepsilon C_i} P_{jk}(t) = 1$.

(e) State i is said to be *recurrent* (or persistent (Smith [8]) or ergodic (Lévy [11])) if and only if $G_{ii}(\infty) = 1$, and is said to be *transient* otherwise.

(f) State i is said to be *null recurrent* (weakly ergodic (Lévy [11]) if, and only if, it is recurrent and $\mu_{ii} = \infty$. State i is said to be *positive* (ergodic (Feller [12]), strongly ergodic (Lévy [11])) if, and only if, it is recurrent and $\mu_{ii} < \infty$.

As should be expected, the properties defined here for M.R.P.'s are very closely related to those of the corresponding Markov Chains (c.M.C.) determined by the same m, **A**, and **P**, namely, the J-processes. This is illustrated by the following theorem.

THEOREM 5.1. *For a given* M.R.P.: (a) *state i is recurrent (is transient) [communicates with state j], if and only if state i is recurrent (is transient) [communicates with state j] in the* c.M.C.; (b) *a class is essential if and only if in the* c.M.C. *it is essential; and* (c) *if $m < \infty$, then state i is positive if and only if state i, in the* c.M.C., *is positive and $\eta_j < \infty$ for all $j \varepsilon C_i$.*

PROOF: (a) From (5.2), it follows that

$$G_{ij}(\infty) = P[J_n = j \text{ for some } n > 0 \mid J_0 = i]. \tag{5.3}$$

This relation suffices to verify (a) since the properties of recurrence, transience and communication involve only the quantities $G_{ij}(\infty)$, and since (5.3) shows that these quantities are identical to the analogous quantities of the c.M.C. (b) is immediate. To prove (c), let $\mathcal{S}_{n,i,j}$ denote the subset of $(\alpha_0, \cdots, \alpha_n) \varepsilon \mathcal{S}_{n,i}$ for which $\alpha_n = j$, where $\mathcal{S}_{n,i}$ is as defined in (4.2). Then one may straightforwardly show (for arbitrary m) that

$$\mu_{ij} = \sum_{n=1}^{\infty} \sum_{\mathcal{S}_{n,i,j}}^{\infty} \prod_{k=0}^{n-1} p_{\alpha_k \alpha_{k+1}} (b_{\alpha_0 \alpha_1} + \cdots + b_{\alpha_{n-1} \alpha_n}) \tag{5.4}$$

Now assuming $m < \infty$ and $i = j$, one can write

$$(\min_{k, j \varepsilon C_i} b_{kj}) \mu_{ii}^* \leqq \mu_{ii} \leqq (\max_{k, j \varepsilon C_i} b_{kj}) \mu_{ii}^*$$

where

$$u_{ii}^* = \sum_{n=1}^{\infty} n \sum_{\mathcal{S}_{n,i,j}} \prod_{k=0}^{n-1} p_{\alpha_k \alpha_{k+1}} \tag{5.5}$$

is the mean recurrence time of state i in the c.M.C. Since $m < \infty$ the minimum shown above is positive and since $m < \infty$ and $\eta_j < \infty$ for all $j \varepsilon C_i$, the maximum shown above is finite.

There does not seem to be any simple necessary and sufficient condition for a positive state in the case of $m = \infty$. Examples may readily be constructed to show that a state of an M.R.P. may be positive (null recurrent), while the same state in the c.M.C. is null recurrent (positive). One sufficient condition for the positivity of state i is that the state be positive in the c.M.P. and $\sum_{j \varepsilon C_i} \eta_j < \infty$. The proof of this, as well as further discussion along these lines, is contained in [7] where explicit computations of the μ_{ij} in terms of the μ_{ij}^* is made.

REFERENCES

[1] PAUL LÉVY, "Systèmes Semi-Markoviens à au plus une infinité dénombrable d'états possibles," *Proc. Int. Congr. Math.*, Amsterdam, Vol. 2 (1954), p. 294. "Processus Semi-Markoviens," *ibid.*, Vol. 3 (1954), pp. 416–426.

[2] W. L. SMITH, "Regenerative stochastic processes," *Proc. Roy. Soc. (London)*, Ser. A, Vol. 232 (1955), pp. 6–31 [cf. the abstract of this paper in *Proc. Int. Congr. Math.*, Amsterdam, Vol. 2 (1954), pp. 304–305].

[3] LAJOS TAKÁCS, "Some investigations concerning recurrent stochastic processes of a certain type," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, Vol. 3 (1954), pp. 115–128 [Hungarian. English Summary. Cf. *Math. Reviews*, Vol. 17 (1956), p. 866.]

[4] RONALD PYKE, "Markov Renewal processes with finitely many states," *Ann. Math. Stat.*, Vol. 32 (1961), pp. 1243–1259.

[5] J. L. DOOB, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[6] M. V. JOHNS AND RONALD PYKE, "On conditional expectation and quasi-rings," *Pacific J. Math.*, Vol. 9 (1959), pp. 715–722.

[7] RONALD PYKE, "Limit theorems for Markov Renewal processes," Technical Report No. 24 (1961), Contract No. Nonr-266(59) Project No. 042-205. Columbia University.

[8] W. L. SMITH, "Renewal theory and its ramifications," *J. Roy. Stat. Soc.*, Ser. B, Vol. 20 (1958), pp. 243–302.

[9] M. LOÈVE, *Probability Theory*, 1st Edition, D. Van Nostrand Co., Princeton, N. J., 1955.

[10] W. L. SMITH, "Remarks on the paper 'Regenerative stochastic processes'," *Proc. Roy. Soc. (London)*, Ser. A, Vol. 256 (1960), pp. 496–501.

[11] PAUL LÉVY, "Systèmes Markoviens et stationaires. Cas dénombrables," *Ann. École Norm. Suppl.*, Vol. 68 (1951), pp. 327–381.

[12] W. FELLER, *An Introduction to Probability Theory*, 2nd Edition, John Wiley and Sons, New York, 1950.

[13] K. L. CHUNG, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin, 1960.

SUPPLEMENTARY REFERENCES

[14] PHILIP M. ANSELONE, "Ergodic theory for discrete Semi-Markov chains," *Duke Math. J.*, Vol. 27 (1960), pp. 33–40.

[15] RICHARD BARLOW, "Applications of Semi-Markov processes to counter and reliability problems," Technical Report No. 57 (1960), Stanford University, Contract N6onr-25140 (NR-342-022).

[16] VIOLET R. CANE, "Behaviour sequences as Semi-Markov chains," *J. Roy. Stat. Soc.*, Ser. B, Vol. 21 (1959), pp. 36–58.

[17] AUGUSTUS J. FABENS, "The solution of queueing and inventory models by Semi-Markov processes," Technical Report No. 20 (1959), Stanford University, Contract Nonr-225(28), (NR-047-019).

[18] HARRY KESTEN, "Occupation times for Markov and Semi-Markov chains," *cf.* Abstract No. 61T-77, *Notices Amer. Math. Soc.*, Vol. 8 (1961), p. 179.

[19] RUPERT G. MILLER, JR., "Asymptotic multivariate occupation time distributions for Semi-Markov processes," Technical Report No. 70 (1961), Stanford University, Contract Nonr-225(52), (NR-342-022).

[20] STEVEN OREY, "Change of time scale for Markov processes," Technical Report, University of Minnesota (1959), Contract No. AF 49(638)-617.

[21] RONALD PYKE, "Markov Renewal processes of zero order and their application to counter theory," Technical Report No. 64 (1960), Stanford University, Contract Nonr-255(52), (NR-342-022); to appear in a volume edited by Samuel Karlin, Stanford University Press, Calif.

[22] RONALD PYKE, "Doeblin Ratio limit theorems for Markov Renewal processes," to appear [*cf.*, Abstract, *Ann. Math. Stat.*, Vol. 31 (1960), pp. 245–246].

[23] GEORGE H. WEISS, "On a Semi-Markovian process with a particular application to reliability theory," NAVORD Report 4351, White Oak, Maryland.

[24] CYRUS DERMAN, "Remark concerning two-state semi-Markov processes," *Ann. Math. Stat.*, Vol. 32 (1961), pp. 615–616.