ORC 64-16 (RR)
22 JULY 1964

LIMITING COVARIANCE IN MARKOV-RENEWAL PROCESSES

by

William S. Jewell

OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH

UNIVERSITY OF CALIFORNIA - BERKELEY

LIMITING COVARIANCE IN MARKOV-RENEWAL PROCESSES

by

William S. Jewell
Operations Research Center
University of California, Berkeley

22 July 1964 ORC 64-16(RR)

This research has been partially supported by the Office of Naval Research under Contact Nonr-222(83) with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ABSTRACT

General additive functions called rewards are defined on a "regular" finite-state Markov-renewal process. The asymptotic form of the mean total reward in $[0,t]$ has previously been obtained, and it is known that the total rewards are joint-normally distributed as $t \to \infty$. This paper finds the dominant asymptotic term in the covariance of the total rewards as a simple function of the moments of the per-transition rewards, and the "bias" term of the mean total rewards. Special formulas for the dominant covariance term of "number of visits", and "occupation time" in given states are also derived.

LIMITING COVARIANCE IN MARKOV - RENEWAL PROCESSES

Consider a finite-state Markov-renewal process[5] which moves through states $i_0, i_1, \ldots i_k, i_{k+1}, \ldots$ at times $S_0 = 0 < S_1 < \ldots S_k < S_{k+1} < \ldots$ If a reward is earned during each transition from state to state, and if successive rewards are additive, it is of interest to study the total reward earned during the interval $[0,t]$. Typically, the reward earned during a transition from i_k to i_{k+1} might be a random variable which depends upon the values of i_k , i_{k+1} , $S_{k+1} - S_k$, as well as on the "excess time" $t-S_k$ $(S_k \leq t < S_{k+1})$ of any uncompleted transition.

Thus, the total reward earned in $[0,t]$ is a random sum of additive random variables, and has a well-defined, though complicated distribution. The purpose of this paper is to summarize some known results on the asymptotic form of the <u>mean total reward</u>, and to present some new results on the dominant asymptotic term of the <u>(co-)variance of the total reward.</u> These results are useful primarily because a central limit theorem often holds for the distribution of total reward, as $t \to \infty$.

DEFINITIONS, NOTATION, AND SUMMARY OF RESULTS

The definition of a Markov-renewal process is that:

$$\Pr\{i_{k+1} = j \; ; \; S_{k+1} \leq x + S_k \mid i_k = i \; , \; i_{k-1}, \ldots i_1 \; , \; i_0 \; ; \; S_k, \; S_{k-1}, \ldots S_1, \; S_0 = 0\}$$

$$= Q_{ij}(x) = p_{ij}F_{ij}(x) \qquad \left\{ \begin{matrix} k=0,1,2,\ldots \\ i,j=1,2,\ldots M \\ x \geq 0 \end{matrix} \right\}$$

In other words, the process may be considered as an imbedded Markov chain in which the movement between the M states is governed by the transition probabilities p_{ij}, and in which the transition intervals, $\tau(i_k, i_{k+1}) = S_{k+1} - S_k$ are independent samples from the d.f. $F_{i_k, i_{k+1}}(\cdot)$. Thus, once the initial state i_0 is given, the bivariate distribution $Q_{ij}(\cdot)$ determines the entire process, since $p_{ij} = Q_{ij}(\infty)$. For simplicity, we assume $F_{ij}(0) = 0$.

Let $u_k = t - S_k$ be the elapsed time since the last transition $(0 \leq u_k \leq \tau(i_k, i_{k+1}))$ and define the random variables

$$a(i_k, i_{k+1}; u_k \mid \tau(i_k, i_{k+1})) \quad \text{and} \quad b(i_k, i_{k+1}; u_k \mid \tau(i_k, i_{k+1}))$$

for all $k=0,1,2,\ldots$ over the appropriate ranges of the arguments. a and b (for short) may be thought of as partial rewards which are accumulated after a time u_k has elapsed between a transition from state i_k to i_{k+1}; it is assumed that the joint distribution function of a and b is known, and that all its moments are finite for finite τ.

In particular, denote the mean partial rewards and the second (cross-) moments of a and b by:

$$\rho^{A}_{ij}(u \mid \tau) = E\{a(i,j;u \mid \tau)\} \; ; \; \rho^{AA}_{ij}(u \mid \tau) = E\{[a(i,j;u \mid \tau)]^2\} \; ;$$
$$\rho^{AB}_{ij}(u \mid \tau) = E\{a(i,j;u \mid \tau)\, b(i,j;u \mid \tau)\} \qquad (1)$$

etc., and let

$$\rho_{ij}(\tau) = \rho_{ij}(\tau \mid \tau) \; ; \quad \rho_{ij} = \int_0^\infty \rho_{ij}(\tau)\, dF_{ij}(\tau) \; ; \quad \rho_i = \sum_{j=1}^{M} p_{ij}\rho_{ij} \; ; \qquad (2)$$

$$\text{and} \qquad \rho = \sum_{i=1}^{M} \pi_i \, \rho_i$$

be the final transition reward, and its various average values, for all moments of (1). The π_i are the stationary probabilities of the associated Markov chain probabilities, p_{ij} .

The <u>total reward of type</u> A earned in $[0,t]$ is defined as:

$$A_{i_0}(t) = \sum_{k=0}^{N(t)-1} a(i_k, i_{k+1}; \tau(i_k, i_{k+1}) \mid \tau(i_k, i_{k+1})) + \\ + a(i_{N(t)}, i_{N(t)+1}; u_{N(t)} \mid \tau(i_{N(t)}, i_{N(t)+1})) \tag{3}$$

where $N(t) = \sup\{k \geq 0: S_k \leq t\}$ is the number of transitions in $[0,t]$. (If $N(t) = 0$, the first term in (3) is zero). Thus, the total reward is the sum of all the total transition rewards accumulated during $N(t)$ transitions plus the partial reward earned during the excess time. A similar expression holds for $B_{i_0}(t)$.

We shall denote the mean and second (cross-)moments of the total rewards A and B by:

$$R_h^A(t) = E\{A_h(t)\} \quad ; \quad R_h^{AA}(t) = E\{[A_h(t)]^2\} \; ; \quad R_h^{AB}(t) = E\{A_h(t)B_h(t)\} \tag{4}$$

etc.

It is straightforward to calculate the joint distribution of $A(t)$ and $B(t)$ from the joint d.f. of a and b . However, in this paper, we shall concentrate on the limiting forms of the means and (co-)variances of these total rewards, as $t \to \infty$. These are important, since it is well-known that in most "regular" cases of an M.R.P. the limiting joint distribution of several additive functions such as $A(t)$ is the multivariate normal[4][6]. Thus, knowledge of the dominant terms in the means and covariances gives a very good approximation to the joint d.f. for long observation times; and, it is these terms which are useful in dynamic programming of an M.R.P. with infinite horizons[2].

We shall restrict our attention to the case in which all the states in the imbedded Markov chain are positive recurrent, and in which the first two moments of the first-passage time d.f.s, $G_{ij}(\cdot)$, are finite; it is also convenient to assume that all the $G_{ij}(\cdot)$ are non-lattice, although this is not restrictive if the limits are defined correctly[5][6]. For references to cases in which the limiting distribution may not be normal, see Reference 6.

The dominant term and the next term (here called the "gain (rate)" and the "bias", respectively) have previously been found for the mean of $A(t)$, as $t \to \infty$ [2][6]. In Reference 4, MILLER finds also expressions for the dominant covariance term for semi-Markov processes ($p_{ii} = 0$, all i) with $M=2$ and 3 states. In Reference 5, PYKE expresses the dominant covariance term for general M.R.P.s by finding a closed form for the second moment of that portion of $A(t)$ between successive returns to a given state; the results are expressed in terms of a renewal function for an associated M.R.P. with an absorbing state.

The main contribution of this paper is expression of the dominant covariance term as an explicit function of the bias term of the mean rewards, - that is, in terms of the first two moments of the first-passage d.f.s. This expression also includes the excess partial reward in (3), which is not considered in the other references. As special cases, explicit formulas are found for the variances and covariances of the number of times a given state is entered, and for the occupation times of a given state.

Special results which are needed for this development are given in the Appendix.

MEAN TOTAL REWARD

It is easy to show that the mean of (3) is given by the renewal equations:

$$R_h^A(t) = \rho_h^A - \sigma_h^A(t) + \sum_{j=1}^{M} \int_0^t R_j^A(t-x)\, dQ_{hj}(x) \tag{5}$$

for all $t \geq 0$, and for all starting states h. A similar expression holds for the mean total reward of type B.

The transient term in (5), $\sigma_h(t)$, is:

$$\sigma_h(t) = \sum_{j=1}^{M} \int_t^{\infty} [\rho_{hj}(x) - \rho_{hj}(t \mid x)]\, dQ_{hj}(x) \tag{6}$$

which we define for all moments of (1). Under the finiteness assumptions of those moments, we have $\lim_{t \to \infty} \sigma_h(t) = 0$; we set:

$$S_h = \int_0^{\infty} \sigma_h(t)\, dt \quad \text{and} \quad S = \sum_{j=1}^{M} \pi_j S_j \quad , \tag{7}$$

which are finite.

Let $N_j(t)$ be the number of times a transition is made into state j during $[0,t]$, and define the <u>renewal function</u>, $M_{hj}(t) = E[N_j(t) \mid i_0 = h]$, and its (LaPlace-Stieltjes) transform, $\tilde{m}_{hj}(s) = \int_0^{\infty} \exp(-st)\, dM_{hj}(t)$.

From equation (A.7) of the Appendix, the (LaPlace) transform of (5) can then be written as:

$$\tilde{R}_h^A(s) = \sum_{j=1}^{M} \tilde{m}_{hj}(s)\ (\rho_j^A/s) - \sum_{j=1}^{M} \tilde{m}_{hj}(s)\ \tilde{\sigma}_j^A(s) \tag{8}$$

where the tilde is used for the transform of the appropriate functions.

It is well known[1] that:

$$\begin{aligned} &\lim_{t\to\infty}[M_{hj}(t) - (t/\mu_{jj})] = \omega_{hj} + o(1) \\ &\omega_{hj} = \frac{\mu_{jj}^{(2)}}{2(\mu_{jj})^2} - \frac{\mu_{hj}}{\mu_{jj}} + \delta_{hj} \end{aligned} \tag{9}$$

with the limit possibly being taken in the Cesàro sense. (In (9), $\mu_{ij}^{(k)}$ is the k^{th} moment of the first-passage-time d.f.; see the Appendix).

In terms of the renewal function bias, ω_{hj}, it is then possible to use a theorem due to Smith[7], or the usual Tauberian limit theorems of transform calculus to show that[2]:

$$\lim_{t\to\infty} [R_h^A(t) - g^A t] = w_h^A + o(1) \tag{10}$$

with the reward gain rate being given by:

$$g^A = \sum_{j=1}^{M} (\rho_j^A / \mu_{jj}) \tag{11}$$

and the reward bias as:

$$w_h^A = \sum_{j=1}^{M} \omega_{hj}\rho_j^A - \sum_{j=1}^{M} (s_j^A / \mu_{jj}) \quad . \tag{12}$$

A similar result holds for the asymptotic form of the mean total reward of type B. Notice that neither of the dominant terms in (9) or (10) depends upon the starting state; this is a consequence of the assumption that the imbedded Markov chain is ergodic, and may be generalized, if desired [5][6].

THE DOMINANT COVARIANCE TERM

By the use of arguments similar to those which led to (5), one may show that the second (cross-)moment of total reward is given by the renewal equations:

$$R_h^{AB}(t) = \rho_h^{AB} - \sigma_h^{AB}(t) + \sum_{j=1}^{M} \int_0^t \rho_{hj}^{A}(x)\, R_j^{B}(t-x)\, dQ_{hj}(x) + \sum_{j=1}^{M} \int_0^t \rho_{hj}^{B}(x)\, R_j^{A}(t-x)\, dQ_{hj}(x) + \sum_{j=1}^{M} \int_0^t R_j^{AB}(t-x)\, dQ_{hj}(x) \tag{13}$$

for all $t \geq 0$, and all starting states h.

By taking the transform of (13) and using (A.7) and (8), one may then find an explicit expression for $\tilde{R}_h^{AB}(s)$, in terms of $\tilde{m}_{ij}(s)$ and the various rewards similar to (8); we shall not reproduce it here since it includes at least six rather complex terms.

The procedure is then straightforward, although tedious: the limiting forms of each of the terms are then examined, and $R_h^{AB}(t)$ is found to be asymptotically of the form $k_2 t^2 + k_1 t + k_0 + o(1)$. The limiting form of $R_h^{A}(t)\, R_h^{B}(t)$, obtained from (10), is then subtracted, and the quadratic terms cancel. The dominant covariance term (14)

$$C_h^{AB} = \lim_{t\to\infty} \frac{\mathrm{Cov}\{A_h(t);\ B_h(t)\}}{t} = \lim_{t\to\infty} E\left\{ \left[\frac{A_h(t)-R_h^{A}(t)}{\sqrt{t}}\right]\left[\frac{B_h(t)-R_h^{B}(t)}{\sqrt{t}}\right] \right\}$$

is then found to be, after a great deal of algebraic manipulation, and use of the formulas of the Appendix:

$$C_h^{AB} = \frac{1}{\nu}\left\{\begin{bmatrix} \rho^{AB} \\ + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M}\pi_i p_{ij}\,[\rho_{ij}^{A}\omega_{jk}\rho_k^{B} + \rho_{ij}^{B}\omega_{jk}\rho_k^{A}] \\ - \left(\frac{\lambda^A\rho^B + \lambda^B\rho^A}{\nu}\right)\end{bmatrix}\right. \tag{15}$$

for all h , with

$$\lambda = \sum_{i=1}^{M}\sum_{j=1}^{M}\pi_i\int_0^{\infty} x\rho_{ij}(x)\, dQ_{ij}(x) \tag{16}$$

and

$$\nu = \sum_{i=1}^{M}\sum_{j=1}^{M}\pi_i p_{ij}\nu_{ij} \qquad ; \quad \nu_{ij} = E\{\tau(i,j)\} \tag{17}$$

As might be expected on intuitive grounds, $C_h^{AB} = C^{AB}$ is independent of the initial state h .

A somewhat simpler expression results if we substitute the appropriate gains and biases from (11) and (12):

$$C^{AB} = \frac{1}{\nu}\left\{\begin{bmatrix} \rho^{AB} + \sum_{i=1}^{M}\sum_{j=1}^{M}\pi_i p_{ij}[\rho_{ij}^{A}w_j^{B} + \rho_{ij}^{B}w_j^{A}] \\ + g^A(S^B - \lambda^B) + g^B(S^A - \lambda^A)\end{bmatrix}\right. \tag{18}$$

In either case, the correct dominant term for the variance, C^{AA} , is obtained by setting B equal to A .

NUMBER OF VISITS AND OCCUPATION TIMES

As important special cases, let us consider the asymptotic means and (co-)variances of: the number of times a state j is visited in $[0,t]$, $N_j(t)$; and the total occupation time of a state j in $[0,t]$, $T_j(t)$, defined as the total time for which state j is the last state visited. For definiteness, we shall consider only states 1 and 2, although the results can clearly be extended to any (sets of) states. The starting state is always h .

We have already indicated that:

$$\lim_{t\to\infty}\left\{(E[N_1(t)\mid i_0 = h] = M_{h1}(t)) - (\pi_1 t/\nu)\right\} = \omega_{h1} + o(1) \qquad (19)$$

where we have used the fact that $\mu_{jj} = \nu/\pi_j$, (A.2). A similar formula applies for state 2.

It is easy to show that: (20)

$$\lim_{t\to\infty}\{E[T_1(t)\mid i_0 = h] - (\pi_1\nu_1\, t/\nu)\} = \omega_{h1}\nu_1 - (\pi_1\nu_1^{(2)}/2\nu) + o(1)$$

where

$$\nu_i^{(k)} = \sum_{j=1}^{M} p_{ij}E\{[\tau(i,j)]^k\} \qquad (k=0,1,2,3\ldots)$$

To find all the (co-)variance terms, we substitute in the appropriate terms for the mean rewards. (For example, $\rho_{ij}^{N_1}(t\mid\tau) = \delta_{i1}$, and $\rho_{ij}^{T_1}(t\mid\tau) = \delta_{i1}t$, for all j , and all $0 < t \le \tau$);[$S^{N_1} = 0$; $S^{T_1} = \pi_1\nu_1^{(2)}/2$; $\lambda^{N_1} = \pi_1\nu_1$; $\lambda^{T_1} = \pi_1\nu_1^{(2)}$; and so on.] We find:

$$C^{N_1N_2} = \frac{1}{\nu}\{\pi_1\omega_{12} + \pi_2\omega_{21}\}$$

([illegible])

$$C^{N_1 N_1} = \frac{1}{\nu}\,(2\pi_1\omega_{11} - \pi_1) = \frac{\pi_1}{\nu}\left\{\frac{\mu_{11}^{(2)} - (\mu_{11})^2}{(\mu_{11})^2}\right\} \tag{22}$$

$$C^{T_1 T_2} = \frac{1}{\nu}\left\{\pi_1\nu_2\sum_{j=1}^{M} p_{1j}\nu_{1j}\omega_{j2} + \pi_2\nu_1\sum_{j=1}^{M} p_{2j}\nu_{2j}\omega_{j1} - \frac{\pi_1\pi_2}{\nu}\,(\nu_1\nu_2^{(2)} + \nu_2\nu_1^{(2)})\right\} \tag{23}$$

$$C^{T_1 T_1} = \frac{1}{\nu}\left\{2\pi_1\nu_1\sum_{j=1}^{M} p_{1j}\nu_{1j}\omega_{j1} - \frac{2(\pi_1)^2}{\nu}\;\nu_1\nu_1^{(2)} + \pi_1\nu_1^{(2)}\right\} \tag{24}$$

$$C^{N_1 T_2} = \frac{1}{\nu}\left\{\pi_1\nu_2\omega_{12} + \pi_2\sum_{j=1}^{M} p_{2j}\nu_{2j}\omega_{j1} - \frac{\pi_1\pi_2\nu_2^{(2)}}{\nu}\right\} \tag{25}$$

$$C^{N_1 T_1} = \frac{1}{\nu}\left\{\pi_1\nu_1\omega_{11} + \pi_1\sum_{j=1}^{M} p_{1j}\nu_{1j}\omega_{j1} - \frac{(\pi_1)^2\nu_1^{(2)}}{\nu}\right\} \tag{26}$$

To obtain the above forms, some formulas of the Appendix must be used, particularly (A.12):

$$\omega_{ij} = \delta_{ij} - \frac{\nu_i}{\nu}\pi_j + \sum_{k=1}^{M} p_{ik}\omega_{kj}$$

Formula (22) above is well-known, but it is believed the others are new.

For a 2-state semi-Markov process $(p_{11} = p_{22} = 0)$, formulas (22), (24), (25), and (26) agree with MILLER[4]. The others are not given by him.

MARKOV CHAINS WITH SINGLE-INDEX REWARDS

Consider a discrete-parameter Markov chain in which the rewards depend only on the state entered, i.e. $\rho_{ij}(t \mid \tau) = \rho_i$, for all $0 < t \leq \tau = 1$, and all j . Then (15) becomes:

$$C^{AB} = \rho^{AB} - 2 \sum_{i=1}^{M} \pi_i \rho_i^A \rho_i^B + \sum_{i=1}^{M} \sum_{j=1}^{M} \pi_i \omega_{ij} (\rho_i^A \rho_j^B + \rho_i^B \rho_j^A) \tag{27}$$

where, of course, the ω_{ij} now have a simpler solution, since all transition intervals have unity length. If we make the substitution:

$$c_{ij} = \pi_i \omega_{ij} + \pi_j \omega_{ji} - \pi_i \delta_{ij} \tag{28}$$

then

$$C^{AB} = \sum_{i=1}^{M} \pi_i (\rho_i^{AB} - \rho_i^A \rho_i^B) + \sum_{i=1}^{M} \sum_{j=1}^{M} \rho_i^A \, c_{ij} \, \rho_j^B \tag{29}$$

which is a slight generalization of a formula in KEMENY AND SNELL[3] (Theorem 4.6.3., p. 87).

We note that (9) does not agree with KEMENY and SNELL, since their limiting process is over the integers $k=0,1,2,\ldots$, and in our notation they obtain:

$$\lim_{k=0,1,2,\ldots\infty} [M_{hj}(k) - (k/\mu_{jj})] = \omega_{hj} + (1/2\,\mu_{jj}) + o(1) \tag{30}$$

and similarly for (10). This affects only the bias term .

NUMERICAL EXAMPLE

As a numerical example of the calculation of the variance, consider the example of Reference 2, in which a two-state alternating process ($p_{12} = p_{21} = 1$) represents a running machine (State 1) or one that has broken down. The two (maintenance, repair) policies which tied in gain rate, g, were:

I. (cheap maintenance, expensive repair) for which

$$a(1,2;\ u \mid \tau) = 100u \quad (0 \leq u \leq \tau)\ ; \quad \nu_{12} = 4 \text{ days}$$

$$a(2,1;\ u \mid \tau) = -100 - 200u \quad (0 < u \leq \tau)\ ; \quad \nu_{21} = 1 \text{ day}$$

and II. (expensive maintenance, expensive repair) which changes transition (1,2) rewards to:

$$a(1,2;\ u \mid \tau) = 84u \quad (0 \leq u \leq \tau)\ ; \quad \nu_{12} = 5 \text{ days.}$$

Note that these are deterministic, linear rewards, giving simple forms for (1) (2) (6) (7) and (17). Letting $\sigma_{ij}^2 = \nu_{ij}^{(2)} - [\nu_{ij}]^2$ be the variance of the transition time, we find the asymptotic forms of the mean reward (10) to be:

Policy I:
$$R_1(t) \approx 20t + 150 - 8\,\sigma_{12}^2 + 22\,\sigma_{21}^2 + o(1)$$
$$R_2(t) \approx 20t - 170 - 8\,\sigma_{12}^2 + 22\,\sigma_{21}^2 + o(1)$$

Policy II:
$$R_1(t) \approx 20t + 151\ 2/3 - 5\ 1/3\,\sigma_{12}^2 + 18\ 1/3\,\sigma_{21}^2 + o(1)$$
$$R_2(t) \approx 20t - 168\ 1/3 - 5\ 1/3\,\sigma_{12}^2 + 18\ 1/3\,\sigma_{21}^2 + o(1)$$

It is important to note that σ_{12}^2 would, in general, be <u>different</u> in Policies I and II.

In Reference 2, we resolved the tie in gain rates for deterministic transition intervals (all $\sigma_{ij}^2 = 0$) in terms of the bias of Policy 2. ($151\frac{2}{3} > 150$) If we now compute the dominant term of the variance of reward, we obtain:

Policy I:
$$\operatorname{Var} A_h(t) \approx (1{,}280\,\sigma_{12}^2 + 9{,}680\,\sigma_{21}^2)t + o(t) \qquad (h=1,2)$$

Policy II:
$$\operatorname{Var} A_h(t) \approx (682\ 2/3\,\sigma_{12}^2 + 8{,}066\ 2/3\,\sigma_{21}^2)t + o(t)$$

which are of rather large magnitude for moderate t (100, say), and exponential transition time d.f.s. If we attempt to resolve the tie on the basis of minimum variance, we see that different values for σ_{12}^2 in the two policies could break the tie either way. Of course, deterministic transition times give zero for the dominant variance term.

CONCLUSION

As indicated in the beginning, it is possible to extend these results to more general M.R.P.; however, it is clear that soon certain terms retain their dependence on the initial state, or cease to exist. It is also possible to find the next ("bias") term of the covariance explicitly; however, this term depends upon the initial conditions, and appears to have little practical interest.

The primary application of these results would seem to be in the fact that variables $\{[A_h(t) - R_h^A(t)]/\sqrt{t}\}$; $\{[B_h(t) - R_h^B(t)]/\sqrt{t}\}$;.... have a limiting multivariate distribution with zero mean and (co-)-variances C^{AA} , C^{AB} , etc. For instance, this might make certain problems of estimation easier. Another possibility might be the selection of "independent" rewards as linear combinations of other conflicting, covariant goals, for a given Markov-renewal process.

APPENDIX

We summarize certain formulas on M.R.P.s which are needed in the text. Some of these formulas are from References 1 and 5. (In all that follows $i,j, = 1,2,\ldots M$).

Let $\nu_{ij}^{(k)}$ and $\mu_{ij}^{(k)}$ $(k=0,1,2,.)$ be the k^{th} moments of the transition interval d.f., $F_{ij}(.)$, and the first-passage time d.f., $G_{ij}(.)$, respectively; the superscript of the first moment is suppressed, and it is assumed that the first two moments are finite. We define the averaged moments successively as:

$$\nu_i^{(k)} = \sum_{j=1}^{M} p_{ij}\nu_{ij}^{(k)} \qquad \text{and} \qquad \nu^{(k)} = \sum_{i=1}^{M} \pi_i \nu_i^{(k)}$$

where the π_i are the stationary probabilities associated with the imbedded Markov chain. These two sets of moments are related through the p_{ij} by:

$$\mu_{ij} = \nu_i + \sum_{\substack{k=1\\k\neq j}}^{M} p_{ik}\mu_{kj} \tag{A.1}$$

$$\mu_{ij}^{(2)} = \nu_i^{(2)} + 2\sum_{\substack{k=1\\k\neq j}}^{M} p_{ik}\nu_{ik}\mu_{kj} + \sum_{\substack{k=1\\k\neq j}}^{M} p_{ik}\mu_{kj}^{(2)} \tag{A.2}$$

By summing (A.1) and (A.2) when multiplied by π_i, the simpler diagonal moments follow:

$$\mu_{jj} = (\nu/\pi_j) \tag{A.2}$$

$$\mu_{jj}^{(2)} = (1/\pi_j)\left\{ \nu^{(2)} + 2 \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \neq j}}^{M} \pi_i p_{ik} \nu_{ik} \mu_{kj} \right\} \tag{A.3}$$

Let $N_j(t)$ be the number of times state j is visited in $[0,t]$, and define the renewal function $M_{hj}(t) = E\{N_j(t) \mid i_o = h\}$ (This is δ_{hj} more than the renewal function used in References 1, 2, and 5).

By direct arguments:

$$M_{hj}(t) = \delta_{hj} + \sum_{k=1}^{M} \int_0^t M_{kj}(t-x)\, dQ_{hk}(x) \tag{A.4}$$

Define $\tilde{f}(s)$ as the LaPlace transform of $f(t)$, or the LaPlace-Stieltjestransform of $F(t) = \int_0^t f(x)dx$. Then (A.4) has the transform:

$$\tilde{m}_{hj}(s) = \delta_{hj} + \sum_{k=1}^{M} \tilde{q}_{hk}(s)\, \tilde{m}_{kj}(s) \tag{A.5}$$

By direct arguments, the indices under the summation can be changed so that:

$$\sum_{k=1}^{M} \tilde{q}_{hk}(s)\, \tilde{m}_{kj}(s) = \sum_{k=1}^{M} \tilde{m}_{hk}(s)\, \tilde{q}_{kj}(s) \tag{A.6}$$

for all h, j, and $s > 0$. Denoting the corresponding matrices by dropping the subscripts and the transform argument, (A.6) reads: $\tilde{q}\,\tilde{m} = \tilde{m}\,\tilde{q}$, and from (A.5) we get the inverse matrix:

$$[I - \tilde{q}]^{-1} = \tilde{m} \qquad (s > 0) \tag{A.7}$$

with I the identity matrix. (A.7) is particularly useful in solving the renewal equations of the text.

It is well known that, under the assumption of an ergodic imbedded

Markov chain, non-lattice $G_{ij}(.)$ with finite first and second moments, the limiting renewal function has the form:

$$\lim_{t \to \infty} [M_{hj}(t) - (t/\mu_{jj})] = \omega_{hj} + o(1) \tag{A.8}$$

with

$$\omega_{hj} = \frac{\mu_{jj}^{(2)}}{2(\mu_{jj})^2} - \frac{\mu_{hj}}{\mu_{jj}} + \delta_{hj} \tag{A.9}$$

(With lattice $G_{ij}(.)$, (A.8) still holds as a Cesàro limit.)

It then follows from (A.5), (A.6), and some Tauberian arguments that:

$$\frac{1}{\mu_{jj}} = \sum_{k=1}^{M} p_{kj} \left(\frac{1}{\mu_{kk}} \right) \tag{A.10}$$

$$\sum_{k=1}^{M} p_{hk}\omega_{kj} = \sum_{k=1}^{M} \omega_{hk} p_{kj} + \frac{\nu_h}{\mu_{jj}} - \sum_{k=1}^{M} \frac{1}{\mu_{kk}} p_{kj}\nu_{kj} \tag{A.11}$$

$$\omega_{hj} = \delta_{hj} - \frac{\nu_h}{\nu} \pi_j + \sum_{k=1}^{M} p_{hk}\omega_{kj} \tag{A.12}$$

$$\sum_{i=1}^{M} \sum_{j=1}^{M} \pi_i p_{ij} \nu_{ij} \omega_{jk} = \frac{1}{2} (\nu^{(2)}/\mu_{kk}) \tag{A.13}$$

These last four formulas are believed to be original; they are particularly useful in reducing special forms of (15).

Similar forms obtain for the gain, g, and bias term, w_h of the mean reward (10), (11), and (12). We have:

$$g = \sum_{k=1}^{M} (\rho_k/\mu_{kk}) = \frac{\rho}{\nu} \tag{A.14}$$

$$w_h = \sum_{k=1}^{M} \omega_{hk}\rho_k - (S/\nu) \tag{A.15}$$

$$w_h + g\nu_h = \rho_h + \sum_{k=1}^{M} p_{hk}w_k \tag{A.16}$$

$$\sum_{k=1}^{M} p_{hk}w_k = \sum_{k=1}^{M}\sum_{\ell=1}^{M} \omega_{hk}p_{k\ell}\rho_\ell + g\nu_h - S/\nu$$
$$- \frac{1}{\nu}\sum_{k=1}^{M}\sum_{\ell=1}^{M} \pi_k p_{k\ell}\nu_{k\ell}\rho_\ell \tag{A.17}$$

$$\sum_{i=1}^{M}\sum_{j=1}^{M} \pi_i p_{ij}\nu_{ij}w_j = \frac{1}{2} g\nu^{(2)} - S \tag{A.18}$$

Equation (A.16) is used in the policy-improvement portion of an algorithm for dynamic programming in a Markov-renewal process [2] .

REFERENCES

1. R. E. Barlow, "Applications of Semi-Markov Processes to Counter Problems", Studies in Applied Probability and Management Science, Chapter 3, (K. J. Arrow, S. Karlin, and H. Scarf, eds.), Stanford University Press, Stanford, California, 1962.

2. W. S. Jewell, "Markov-Renewal Programming. I: Formulation, Finite Return Models", and "Markov-Renewal Programming. II: Infinite Return Models, Example", Operations Research, Vol. 11, No. 6 (1963), pp. 938-948, and pp. 949-971.

3. J. G. Kemeny and J. L. Snell, Finite Markov Chains, D. Van Nostrand, Princeton, New Jersey, 1960.

4. R. G. Miller, Jr., "Asymptotic Multivariate Occupation Time Distribution for Semi-Markov Processes", Technical Report number 70, Stanford University, Contract Nonr 225(52) (NR-342-002), (1961).

5. R. Pyke, "Markov Renewal Processes: Definitions and Preliminary Properties", and "Markov Renewal Processes with Finitely Many States", Annals of Mathematical Statistics, Vol. 32, No. 4 (1961), pp. 1231-1242, and pp. 1243-1259.

6. R. Pyke, "Limit Theorems for Markov Renewal Processes", Technical Report CU-24-61, Columbia University, Contract Nonr-266(59), February 28, 1961; revised and reissued as Technical Report No. 6, University

of Washington, National Science Foundation Grant P-11008, June 7, 1963.

7. W. L. Smith, "Asymptotic Renewal Theorems," *Proceedings of the Royal Society of Edinburgh*, Series A, Vol. 64 (1954), pp 9-48.

BASIC DISTRIBUTION LIST

FOR UNCLASSIFIED TECHNICAL REPORTS

Head, Logistics and Mathematical
Statistics Branch
Office of Naval Research
Washington 25, D. C

C. O., ONR Office
Navy No. 100, Box 39, F. P. O.
New York, New York

ASTIA Document Service Center
Building No. 5
Cameron Station
Alexandra, Virginia

Institute for Defense Analyses
Communications Research Division
von Neumann Hall
Princeton, New Jersey

Technical Information Officer
Naval Research Laboratory
Washington 25, D. C.

C. O., ONR Branch Office
1030 East Green Street
Pasadena 1, California
ATTN: Dr. A. R. Laufer

Bureau of Supplies and Accounts
Code OW, Department of the Navy
Washington 25, D. C.

Professor Russell Ackoff
Operations Research Group
Case Institute of Technology
Cleveland 6, Ohio

Professor Kenneth J. Arrow
Serra House, Stanford University
Stanford, California

Professor G. L. Bach
Carnegie Institute of Technology
Planning and Control of Industrial
Operations, Schenley Park
Pittsburgh 13, Pennsylvania

Professor L. W. Cohen
Mathematics Department
University of Maryland
College Park, Maryland

Professor Donald Eckman
Director, Systems Research Center
Case Institute of Technology
Cleveland, Ohio

Professor Lawrence E. Fouraker
Graduate School of Business
Harvard University
Cambridge, Massachusetts

Professor David Gale
Department of Mathematics
Brown University
Providence 12, Rhode Island

Dr. Murray Geisler
The RAND Corporation
1700 Main Street
Santa Monica, California

Professor L. Hurwicz
School of Business Administration
University of Minnesota
Minneapolis 15, Minnesota

Dr. James R. Jackson, Director
Western Management Sciences Institute
University of California
Los Angeles 24, California

Professor Samuel Karlin
Mathematics Department
Stanford University
Stanford, California

Professor C. E. Lemke
Department of Mathematics
Rensselaer Polytechnic Institute
Troy, New York

Professor W. H. Marlow
Logistics Research Project
The George Washington University
707 - 22nd Street, N. W.
Washington 7, D. C.

BASIC DISTRIBUTION LIST

FOR UNCLASSIFIED TECHNICAL REPORTS

Professor A. Charnes
The Technological Institute
Northwestern University
Evanston, Illinois

Professor R. Radner
Department of Economics
University of California
Berkeley 4, California

Professor Murray Rosenblatt
Department of Mathematics
Brown University
Providence 12, Rhode Island

Mr. J. R. Simpson
Bureau of Supplies and Accounts
Navy Department (Code W31)
Washington 25, D. C.

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Lincoln Hall, Cornell University
Ithaca 1, New York

C. O. ONR Branch Office
346 Broadway
New York 13, New York
ATTN: J. Laderman

Professor Oskar Morgenstern
Economics Research Project
Princeton University
92 A Nassau Street
Princeton, New Jersey