ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
20, 1 (1988), pp. 1–13

M. F. NEUTS (Tucson, Arizona)

AN ALGORITHM FOR THE DISTRIBUTION OF THE TIME BETWEEN COINCIDENCES OF TWO INDEPENDENT *PH*-RENEWAL PROCESSES*

Abstract. It is shown that the distribution of the time between consecutive coincident renewals in two independent discrete *PH*-renewal processes has itself a discrete distribution of phase type. The efficient computation of that distribution and of its geometric asymptote are discussed. As side results, some arithmetic properties of discrete renewal processes are obtained.

1. Introduction. Consider two independent, discrete renewal processes with underlying lifetime densities $\{p_v(1)\}$ and $\{p_v(2)\}$. How often do renewals in both processes coincide? That question was recently examined by Kopocińska and Kopociński [2]. An exponential asymptotic result is presented in [3].

The point process of the successive times of coincidences is itself a renewal process. It suffices therefore to study the probability density $\{r_n\}$ of the random variable Y, the common value of the first partial sums $S_{N_1}(1)$ and $S_{N_1}(2)$ which agree. It is understood that $S_0(1) = S_0(2) = 0$, so that both renewal processes start from a renewal at time $n = 0$.

The density $\{r_n\}$ is a complicated functional of the densities $\{p_v(1)\}$ and $\{p_v(2)\}$. Explicit expressions for the probabilities r_n are available only for a few special cases and the construction of an efficient algorithm to compute these quantities is therefore a nice problem in algorithmic probability.

The following are brief restatements of properties of Y established in [2]:

a. The random variable Y is finite a.s.

b. If the densities $\{p_v(1)\}$ and $\{p_v(2)\}$ have finite means $\mu_1(1)$ and $\mu_1(2)$, respectively, then

* This research was supported in part by Grant Nr. ECS-8601203 from the National Science Foundation.

$$\mathrm{E}(Y) = \mu_1(1)\,\mu_1(2), \tag{1}$$

and in the contrary case, $\mathrm{E}(Y)$ is infinite.

c. Without loss of generality, we may assume that $p_0(1) = p_0(2) = 0$.

d. If one of the densities, say $\{p_\nu(2)\}$, is *geometric*, i.e., $p_\nu(2) = p_2^{\nu-1} q_2$ for $\nu \geqslant 1$, with $q_2 = 1 - p_2$, then $\{r_n\}$ is the geometric mixture

$$\sum_{k=1}^{\infty} p_2^{k-1} q_2 [p_n^{(k)}(1)]$$

of the successive convolutions of $\{p_\nu(1)\}$. The computation and other applications of such geometric mixtures are discussed in [5].

For general densities, Y may be viewed as the time until absorption in a Markov chain with a bivariate state space. That state space is generally infinite, and thus requires truncation in numerical computations. When both densities $\{p_\nu(1)\}$ and $\{p_\nu(2)\}$ have *finite support*, they are of phase type and, as we shall show, so is $\{r_n\}$. In fact, for any two *PH*-densities, not necessarily of finite support, the density $\{r_n\}$ is of phase type. As, in particular cases, the representation of $\{r_n\}$ may involve large matrices, it is necessary to exploit its structure and its asymptotic properties in the computation of $\{r_n\}$. This paper is devoted primarily to these aspects of practical computation.

Let us briefly discuss the periodicity of the density $\{r_n\}$. If the density $\{p_\nu(i)\}$ concentrates on the integers $a(i)+kd(i)$, $k \geqslant 0$, for $i = 1, 2$, the support of the density $\{r_n\}$ is given by the *Chinese Remainder Theorem*, which characterizes the common elements of the congruences $\{a(1)+kd(1)\}$ and $\{a(2)+kd(2)\}$. In particular, if both densities $\{p_\nu(1)\}$ and $\{p_\nu(2)\}$ are periodic with periods $d(1)$ and $d(2)$, we may first reduce $d(1)$ and $d(2)$ by redefining the densities so that $d(1)$ and $d(2)$ are relatively prime. The period d of the density $\{r_n\}$ is then the least common multiple of $d(1)$ and $d(2)$.

2. The case of *PH*-densities. Saying that $\{p_\nu(1)\}$ and $\{p_\nu(2)\}$ are *PH-densities* means that they may be considered as the absorption time densities in (m_1+1)- and (m_2+1)-state Markov chains with transition probability matrices

$$\begin{bmatrix} S(i) & S^0(i) \\ 0 & 1 \end{bmatrix}$$

with initial probability vectors $[\beta(i), 0]$, $i = 1, 2$. For a detailed discussion of *PH*-densities, see Chapter 2 of [4].

Any probability density $\{p_1, \ldots, p_K\}$ on the integers $\{1, \ldots, K\}$ is a *PH*-density. It is readily verified that $\{p_1, \ldots, p_K\}$ is the absorption time density from the state 0 in the $(K+1)$-state Markov chain with transition probability

matrix (illustrated here for $K = 7$)

$$\begin{bmatrix} 0 & p'_1 & 0 & 0 & 0 & 0 & 0 & q'_1 \\ 0 & 0 & p'_2 & 0 & 0 & 0 & 0 & q'_2 \\ 0 & 0 & 0 & p'_3 & 0 & 0 & 0 & q'_3 \\ 0 & 0 & 0 & 0 & p'_4 & 0 & 0 & q'_4 \\ 0 & 0 & 0 & 0 & 0 & p'_5 & 0 & q'_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & p'_6 & q'_6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where the quantities p'_i and q'_i $(1 \leqslant i \leqslant K-1)$ are defined by

$$p'_i = \Big[\sum_{j=i+1}^{K} p_j\Big]\Big[\sum_{j=i}^{K} p_j\Big]^{-1} \quad \text{and} \quad q'_j = 1 - p'_j.$$

Note that by replacing the last two elements 0 and 1 in the K-th row of that matrix by θ and $\theta' = 1-\theta$ with $0 \leqslant \theta \leqslant 1$, we may represent the modified geometric density

$$p_i^0 = p_i \ (1 \leqslant i < K), \qquad p_i^0 = p_K(1-\theta)\,\theta^{i-K} \ (i \geqslant K).$$

This is useful in reducing the truncation error in computations for general densities which are not of phase type and have unbounded support. We may truncate at a sufficiently high index K and put the remaining mass in a geometric tail of which the parameter θ is chosen so that the correct mean of the given density is matched.

The stochastic matrix $S^*(i) = S(i) + S^0(i)\,\beta(i)$, $i = 1, 2$, describes the Markov chain obtained by restarting the absorbing Markov chain immediately upon absorption, by choosing the new initial state according to independent multinomial trials with probabilities given by the components of the vector $\beta(i)$. The *restarting matrices* $S^0(i)\,\beta(i)$, $i = 1, 2$, give the probabilities $S_j^0(i)\,\beta_k(i)$ that the absorbing Markov chain reaches the absorbing state from the state j and is instantaneously restarted in the new initial state k $(1 \leqslant j, k \leqslant m_i)$. As discussed in [4], we may assume without loss of generality that the matrices $S^*(i)$ are irreducible. The invariant probability vector $\pi(i)$ of $S^*(i)$ is given by

$$\pi(i) = \frac{1}{\mu(i)}\,\beta(i)\,[I - S(i)]^{-1}, \qquad i = 1, 2. \tag{2}$$

We now consider an absorbing Markov chain with $m_1 m_2$ transient states (i_1, i_2), $1 \leqslant i_1 \leqslant m_1$, $1 \leqslant i_2 \leqslant m_2$, listed in the lexicographic order, and one absorbing state. Its initial transition probability vector is given by

$[\beta(1)\otimes\beta(2), 0]$ and its transition probability matrix by

$$\begin{bmatrix} L & L^0 \\ 0 & 1 \end{bmatrix}, \tag{3}$$

where

$$L = S(1)\otimes S(2)+S^0(1)\beta(1)\otimes S(2)+S(1)\otimes S^0(2)\beta(2) \tag{4}$$

and

$$L^0 = S^0(1)\otimes S^0(2). \tag{5}$$

The symbol $\otimes$ denotes the Kronecker product of matrices.

Setting $\gamma = \beta(1)\otimes\beta(2)$, we see that

$$L^* = L+L^0\gamma = S^*(1)\otimes S^*(2), \tag{6}$$

and since both $S^*(1)$ and $S^*(2)$ are irreducible, so is L^*. The invariant probability vector of the stochastic matrix L is given by $\pi(1)\otimes\pi(2)$.

THEOREM 1. *The probability density r_n of the random variable Y is the PH-density with representation (γ, L).*

Proof. The Markov chain with the transition probability matrix L in (3) is constructed from the defining chains of the two *PH*-densities. If transitions between non-absorbing states occur in both chains, we have the transitions corresponding to the Kronecker product $S(1)\otimes S(2)$. Transitions in which absorption and restart occur in exactly one of the Markov chains correspond to the terms $S(1)\otimes S^0(2)\beta(2)$ and $S^0(1)\beta(1)\otimes S(2)$. Finally, coincident absorptions correspond to the column vector of probabilities $S^0(1)\otimes S^0(2)$.

Formally, the computation of $\{r_n\}$ is now straightforward. For easier interpretation of numerical results, we actually compute the quantities

$$P[Y > n] = \gamma L^n e = 1 - \sum_{\nu=0}^{n} r_\nu \tag{7}$$

for $n \geqslant 0$. These quantities are evaluated by recursively computing the vectors γL^n, followed by summation of their components. There is an easily implemented stopping criterion, since

$$\mathrm{E}(Y) = \sum_{n=0}^{\infty} P[Y > n] = \mu_1(1)\,\mu_1(2).$$

We stop as soon as the accumulated sum of the quantities $P[Y > n]$ is sufficiently close to the known value of $\mathrm{E}(Y)$. The numerical value of the sum is reported as the *computed mean.*

In practice, the simplicity of the recursive scheme is somewhat deceptive, and this for several reasons:

a. The matrix L is commonly of high order. In the interest of computational efficiency, its special structure needs to be exploited.

b. The quantities $P[Y > n]$ often decrease to zero very slowly. A crude implementation of the recursive scheme may result, after significant computational effort, in an uninformative table with several thousands of entries.

Fortunately, the sequences $\{P[Y > n]\}$ and $\{r_n\}$ have geometric asymptotes, that is

$$P[Y > n] = C\eta^n + o(\eta^n) \tag{8}$$

and

$$r_n = C_1 \eta^{n-1} + o(\eta^{n-1}) \tag{9}$$

as $n \to \infty$. The constants C and C_1 are non-negative and $0 < \eta < 1$. The computation of these asymptotic constants itself requires some effort and presents challenges to increase its efficiency.

3. Exploiting structure in computation. Because of the high dimensions of the resulting matrices, it is usually not advisible to work directly with Kronecker products. The following are well-known mathematical properties of Kronecker products which are very useful in algorithmic implementations.

Let A and B be matrices of dimensions $n_1 \times n_2$ and $n_1' \times n_2'$, respectively. Their Kronecker product is then of dimensions $n_1 n_1' \times n_2 n_2'$. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be row vectors of dimensions $n_1 n_1'$ and $n_2 n_2'$, respectively. The matrix U is obtained from the vector $\boldsymbol{u}$ by partitioning that vector into n_1 vectors of dimension n_1' and by writing these as the rows of the $(n_1 \times n_1')$-matrix U. The $(n_2 \times n_2')$-matrix V is similarly obtained from the row vector $\boldsymbol{v}$.

The product $\boldsymbol{u}(A \otimes B) = \boldsymbol{v}$ is now conveniently evaluated as

$$A^T U B = V. \tag{10}$$

Mutatis mutandis, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are column vectors of the same dimensions as before, then $(A \otimes B)\boldsymbol{v} = \boldsymbol{u}$ is evaluated by forming

$$A V^T B^T = U,$$

and $\boldsymbol{u}$ and $\boldsymbol{v}$ are now the direct sums of the columns of U and V.

In the recursive computation to implement formula (7), we first store $S^*(1)$, so that

$$L = S^*(1) \otimes S(2) + S(1) \otimes S^0(2)\beta(2).$$

We first compute $\boldsymbol{u}[S^*(1) \otimes S(2)]$ by use of (10). The result is stored in V. Next we note that

$$S^T(1)\, U S^0(2)\, \beta(2) = \boldsymbol{Z}\beta(2),$$

where $\boldsymbol{Z}$ is a column vector of dimension m_1. We evaluate the vector $\boldsymbol{Z}$ and

add the easily computed matrix $\boldsymbol{Z}\beta(2)$ to V to obtain the vector $v = uL$. The vector v is the direct sum of the rows of the matrix V, but it is never directly used as such. We only operate with the matrices U and V and perform multiplications only on matrices of dimensions at most equal to $\max(m_1, m_2)$. Without further use of additional special structure, the evaluation of each term $P[Y > n]$ requires $2m_1^2 m_2^2 + 2m_1^2 m_2 + m_1 m_2$ multiplications.

It is of interest to look in detail at the case of modified geometric densities, of which densities with finite support are a special case. In order to show the structure of large matrices, we shall display these for particular but representative cases.

If $S(1)$ is of the form

$$\begin{bmatrix} 0 & p_1' & 0 & 0 & 0 \\ 0 & 0 & p_2' & 0 & 0 \\ 0 & 0 & 0 & p_3' & 0 \\ 0 & 0 & 0 & 0 & p_4' \\ 0 & 0 & 0 & 0 & \theta_1 \end{bmatrix},$$

then L has the block structure

$$\begin{bmatrix} q_1'(1)S(2) & p_1' S^*(2) & 0 & 0 & 0 \\ q_2'(1)S(2) & 0 & p_2' S^*(2) & 0 & 0 \\ q_3'(1)S(2) & 0 & 0 & p_3' S^*(2) & 0 \\ q_4'(1)S(2) & 0 & 0 & 0 & p_4' S^*(2) \\ \theta_1' S(2) & 0 & 0 & 0 & \theta_1 S^*(2) \end{bmatrix}$$

and the vector γ is given by $\gamma = [\beta(2), 0, 0, 0, 0]$.

Let v be a column vector of dimension $m_1 m_2$, partitioned into m_1 vectors $v(1), \ldots, v(m_2)$ of dimension m_2. We may compute Lv efficiently and store it in the same memory locations as used to store v. To do so, we first compute $x = S(2)v(1)$, and then successively for $i = 1, \ldots, m_1 - 1$

$$v(i) = p_i'(1) S^*(2) v(i+1) + q_i'(1) x$$

and, finally,

$$v(m_1) = \theta_1 S^*(2) v(m_1) + \theta_1' x.$$

By reserving the vector $S^*(2)v(m_1)$, most multiplications for this last equation may be avoided.

In order to compute the terms of the sequence $P[Y > n]$, it is efficient in this case to initialize v by e, and then premultiplying v successively by L. The term of interest is given by $\beta(2)v(1)$. We see that, merely by exploiting the special structure of the first density, we have reduced the number of multiplications per term to $(m_1 - 1)m_2(m_2+1) + m_2(m_2+3)$.

When $\beta(2) = (1, 0, \ldots, 0)$ and $S(2)$ has the same structure as $S(1)$, we see that the modules $S(2)\,w$ and $S^*(2)\,w$, which arise in the computation of Lv, are given by

$$S(2)\,w = \begin{bmatrix} p_1'(2)\,w_2 \\ p_2'(2)\,w_3 \\ \cdots\cdots\cdots \\ p_{m_2-1}'(2)\,w_{m_2} \\ \theta_2\,w_{m_2} \end{bmatrix}.$$

and

$$S^*(2)\,w = S(2)\,w + w_1\,S^0(2).$$

For each n, the probability $P[Y > n]$ is now simply the first component of the current vector v. Each term requires only $4m_1\,m_2 - m_2$ multiplications.

4. The geometric asymptote. The asymptotically geometric behavior stated in formulas (8) and (9) is easily proved when the matrix L is irreducible, but in some rather special cases L may be reducible. Let us first present the argument for the *irreducible* case.

Let $\eta > 0$ be the Perron–Frobenius eigenvalue of L, and u and v the corresponding left and right eigenvectors, normalized so that $ue = 1$ and $uv = 1$. It is known that this uniquely determines the vectors u and v, which are both strictly positive. Moreover, as $n \to \infty$,

$$L^n = \eta^n vu + o(\eta^n).$$

It now immediately follows that

$$P[Y > n] = \eta^n \gamma v + o(\eta^n) \tag{11}$$

and

$$r_n = \eta^{n-1}(\gamma v)(uL^0) + o(\eta^{n-1}). \tag{12}$$

The inner product uL^0 may be explicitly evaluated. It is given by

$$uL^0 = 1 - \eta. \tag{13}$$

To see this, we postmultiply by e in the equation $uL = \eta u$ to obtain

$$\eta = \eta ue = uLe = ue - uL^0 = 1 - uL^0.$$

Setting $\gamma v = C$, we rewrite (10) as

$$r_n = C(1-\eta)\,\eta^{n-1} + o(\eta^{n-1}).$$

The vector u, the eigenvalue η and the equality (13) have noteworthy probabilistic interpretations. Since

$$\frac{\gamma L^n}{\gamma L^n e} \to u \quad \text{as } n \to \infty, \tag{14}$$

u may be interpreted as the vector of conditional probabilities of the transient states in the Markov chain (3), given that absorption takes a "long" time. From (13) we see that η may be viewed as the conditional probability that a coincidence will *not* occur at the next time unit, given that the time since the last coincidence is long. The equality (13) also serves as an internal accuracy check in numerical computation.

Assuming for the time being that η, u and v have been computed, the asymptotic formula (9) is used as follows. Having computed $u(n) = \gamma L^n$, we evaluate $u(n+1) = u(n)L$ and check whether

$$\max_i |u_i(n+1) - \eta u_i(n)| < \frac{\varepsilon}{m_1 m_2}.$$

If that inequality holds for $n = N$, then it is readily verified that, for all $n > N$,

$$|P[Y > n] - C\eta^n| < \varepsilon.$$

As an alternative, we may stop when that last inequality holds for K (say, ten or more) successive values of n. While that criterion does not guarantee that the inequality holds for all higher n, it requires much less effort to check and is usually adequate.

As an internal accuracy check on the asymptotic formula, we report the computed mean of the density $\{r_n\}$ if the asymptotic formula is used from $n = N+1$ onward. That quantity is given by

$$\sum_{n=0}^{N} P[Y > n] + C\eta^{n+1}(1-\eta)^{-1}$$

and should be satisfactorily close to $E(Y)$.

Reducibility of L arises when *both PH*-densities $\{p_v(1)\}$ and $\{p_v(2)\}$ require lead-in phases which are visited *only once*. These arise, for example, with densities of finite support for which some initial terms $p_1(i), p_2(i), \ldots$ $(i = 1, 2)$ are zero. The corresponding matrix $S(i)$ has lead-in phases when its upper left-hand corner consists of a block of the form

$$\begin{bmatrix} 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where the asterisks indicate positive elements. The effect of lead-in phases in *both PH*-densities is to cause the matrix L to be (after relabeling of the

states) to be of the form

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

where the square matrix D is irreducible and A is a square non-negative matrix of spectral radius zero. If A is of order g, then A^g (or possibly already a lower power of A) vanishes. This implies that L^g is of the form

$$\begin{bmatrix} 0 & E \\ 0 & D^g \end{bmatrix}.$$

We handled the difficulty of reducibility of L by identifying the lead-in phases as part of our algorithm to compute η, which is now also the Perron–Frobenius eigenvalue of the matrix D. In applying the asymptotic formulas, we now write, for $n \geqslant g$,

$$\gamma L^n e = \gamma L^g (L^{n-g} e)$$

and note that γL^g is a non-negative vector with zero components corresponding to the lead-in phases. If we write γ^* for the vector γ with those components deleted, we see that, for $n \geqslant g$,

$$\gamma L^n e = \gamma^* D^{n-g} e = \eta^{n-g} \gamma^* v + o(\eta^{n-g})$$

as $n \to \infty$.

The eigenvalue η and the vectors u and v can be computed by several iterative procedures. The block structure of the matrix L appears to make the present case an ideal candidate for the *aggregation-disaggregation* method of Schweitzer [6]. In our computations, we implemented Elsner's algorithm [1], which is easier to program, as no code for either algorithm appears to be readily available.

5. Some arithmetic properties of discrete renewal processes. By choosing the probability density $\{p_v(1)\}$ to be degenerate at m_1, we obtain highly tractable results on certain arithmetic properties of discrete *PH*-renewal processes. In displaying the structure of matrices, we shall use $m_1 = 4$.

The matrix L is then given by

$$L = \begin{bmatrix} 0 & S^*(2) & 0 & 0 \\ 0 & 0 & S^*(2) & 0 \\ 0 & 0 & 0 & S^*(2) \\ S(2) & 0 & 0 & 0 \end{bmatrix},$$

and $\gamma = [\beta(2), 0, 0, 0]$. In what follows, we may drop the index 2 on the matrices S and S^*. Also, r_n may differ from zero only when n is a multiple of m_1. It is therefore sufficient to study the probabilities $P[Y > km_1]$ for $k \geqslant 1$.

It readily follows from the structure of the matrix L that

$$P[Y > km_1] = \beta(S^{*m_1-1}S)^k e, \quad k \geqslant 0, \tag{15}$$

so that the random variable Y/m_1 has the *PH*-density with representation $[\beta, S^{*m_1-1}S]$. The random variable Y is the *first partial sum of the sequence of integer-valued lifetimes which is a multiple of* m_1.

In a discrete renewal process, the embedded sequence of renewal times, which are multiples of m_1, is itself a renewal process. The mean time between successive visits to multiples of m_1 is given by

$$\mu(m_1) = m_1 \beta[I - S^{*m_1-1}S]^{-1} e,$$

but by virtue of formula (1), this quantity should also be equal to

$$m_1 \mu(1) = m_1 \beta(I-S)^{-1} e.$$

This has the interesting consequence that, in the *stationary* renewal process, the fraction $\mu(1)/\mu(m_1)$ of renewals occurring at multiples of m_1 is given by $1/m_1$.

The equality

$$\beta[I - S^{*m_1-1}S]^{-1} e = \beta(I-S)^{-1} e \tag{16}$$

may also be established by a simple matrix calculation. We set

$$w(m_1) = [I - S^{*m_1-1}S]^{-1} e$$

so that

$$w(m_1) = e + S^{*m_1-1} S w(m_1).$$

Now premultiplying by the vector

$$\pi = \frac{1}{\mu(1)} \beta(I-S)^{-1},$$

which is the left invariant probability vector of the irreducible stochastic matrix S^*, we obtain

$$\pi w(m_1) = 1 + \pi S w(m_1),$$

which implies that $\beta w(m_1) = \mu(1)$ and that is tantamount to (16).

The *PH*-densities with representations $[\beta, S^{*m_1-1}S]$ all have mean $\mu(1)$, but plots of these densities or of selected quantiles can be used to illustrate how the renewal times of the *PH*-renewal process favor or disfavor multiples of various integers m_1.

The successive visits of a discrete *PH*-renewal process to a residue class $\{km_1 + r\}$, with $1 \leqslant r \leqslant m-1$, do not form a renewal process. We may,

however, obtain the probability distribution of the first passage time Y to the set $\{km_1+r,\ k \geqslant 0\}$. The analogue of formula (15) is

$$P[Y > km_1+r] = \beta S^{*r-1} S[S^{*m-1} S]^k e \quad \text{for } k \geqslant 0.$$

This formula, as well as (15), may be obtained by a direct probability argument. In order for the event $\{Y > km_1+r\}$ to occur, the Markov chain with transition probability matrix S^* should *not* have a restart (renewal) at the time points vm_1+r, $0 \leqslant v \leqslant k$. At all other times, phase transitions (described by S) or restarts (described by $S^0\beta$) may occur.

6. The associated counting random variables. In some applications, we may also be interested in the numbers N_1 and N_2 of renewals which have occurred in both *PH*-renewal processes at the time Y of the first coincidence. These random variables are also briefly considered in [2].

We consider the joint probability density

$$r(v_1, v_2; n) = P[N_1 = v_1,\ N_2 = v_2;\ Y = n] \quad \text{for } v_1 \geqslant 1,\ v_2 \geqslant 1,\ n \geqslant 1.$$

The computation of the trivariate sequence $\{r(v_1, v_2; n\}$ is a monumental task, even for *PH*-renewal processes. If necessary, it can be carried out efficiently by the following recursive scheme.

Let us define row vectors $\boldsymbol{K}(v_1, v_2; n)$ of dimension $m_1 m_2$. The component of $\boldsymbol{K}(v_1, v_2; n)$, indexed by (j_1, j_2) (with $1 \leqslant j_1 \leqslant m_1$, $1 \leqslant j_2 \leqslant m_2$, and these pairs listed in the lexicographic order), is the probability that at time n no coincidence has yet occurred, that respectively v_1 and v_2 renewals have occurred in the first and second processes, that the first process is in its phase j_1, and the second in its phase j_2. For $v_1 < 0$ or $v_2 < 0$, these vectors are clearly zero. For other values of $\boldsymbol{K}(v_1, v_2; n)$, they satisfy the recurrence relations

$$\begin{aligned} \boldsymbol{K}(0, 0, 0) &= \gamma = \beta(1) \otimes \beta(2), \\ \boldsymbol{K}(v_1, v_2; n) &= \boldsymbol{K}(v_1, v_2; n-1)[S(1) \otimes S(2)] \\ &\quad + \boldsymbol{K}(v_1-1, v_2; n-1)[\boldsymbol{S}^0(1)\beta \otimes S(2)] \\ &\quad + \boldsymbol{K}(v_1; v_2-1; n-1)[S(1) \otimes \boldsymbol{S}^0(2)\beta(2)]. \end{aligned} \tag{17}$$

The vector generating function

$$\boldsymbol{K}^*(z_1, z_2; w) = \sum_{n=0}^{\infty} \sum_{v_1=0}^{\infty} \sum_{v_2=0}^{\infty} \boldsymbol{K}(v_1, v_2; n) z_1^{v_1} z_2^{v_2} w^n$$

is, after routine calculations, found to be given by

$$\begin{aligned} \boldsymbol{K}^*(z_1, z_2; w) = \gamma[I - wS(1) \otimes S(2) - wz_1 \boldsymbol{S}^0(1)\beta(1) \otimes S(2) \\ - wz_2 S(1) \otimes \boldsymbol{S}^0(2)\beta(2)]^{-1}. \end{aligned} \tag{18}$$

It is now clear that for $v_1 \geqslant 1$, $v_2 \geqslant 1$, $n \geqslant 1$, the probability $r(v_1, v_2; n)$ is given by

$$r(v_1, v_2; n) = \boldsymbol{K}(v_1-1, v_2-1; n-1)[\boldsymbol{S}^0(1)\otimes\boldsymbol{S}^0(2)], \tag{19}$$

and the generating function $R(z_1, z_2; w) = \mathrm{E}(z_1^{N_1} z_2^{N_2} w^Y)$ by

$$R(z_1, z_2; w) = wz_1 z_2 \boldsymbol{K}^*(z_1, z_2; w)[\boldsymbol{S}^0(1)\otimes\boldsymbol{S}^0(2)]. \tag{20}$$

Formulas (18) and (20) are primarily useful in deriving expressions for the means and, in principle, the higher moments of N_1, N_2 and Y. For example, after routine differentiations and manipulations, we obtain

$$\mathrm{E}(N_1)$$
$$= \gamma(I-L)^{-1}\{I-(\boldsymbol{S}^0(1)\boldsymbol{\beta}(1)\otimes\boldsymbol{S}(2))[I-\boldsymbol{S}(1)\otimes\boldsymbol{S}(2)]^{-1}\}^{-1}(\boldsymbol{S}^0(1)\otimes\boldsymbol{S}^0(2)).$$

For $\mathrm{E}(N_2)$ we obtain a similar formula with the matrix $\boldsymbol{S}^0(1)\boldsymbol{\beta}(1)\otimes\boldsymbol{S}(2)$ replaced by $\boldsymbol{S}(1)\otimes\boldsymbol{S}^0(2)\boldsymbol{\beta}(2)$. It should be stressed that these expressions for $\mathrm{E}(N_1)$ and $\mathrm{E}(N_2)$ are in their simplest analytic forms. Appreciable simplifications arise only in very special cases. If the density $\{p_v(1)\}$ is geometric, then by setting $\boldsymbol{S}(1) = p_1$ and $\boldsymbol{S}^0(1) = q_1 = 1-p_1$ we obtain expressions for $\mathrm{E}(N_1)$ and $\mathrm{E}(N_2)$ which, after considerable matrix algebra, reduce to

$$\mathrm{E}(N_1) = p_1 + q_1\,\mu(2)$$

and

$$\mathrm{E}(N_2) = \frac{1}{1-p_1\,\boldsymbol{\beta}(2)[I-p_1\,\boldsymbol{S}(2)]^{-1}\,\boldsymbol{S}^0(2)}.$$

Notice that the denominator in this last expression is $1-P_2(p_1)$, where $P_2(z)$ is the probability generating function of the density $\{p_v(2)\}$. For this special case, the expressions

$$\mathrm{E}(N_1) = p_1 + q_1\,\mu(2) \quad \text{and} \quad \mathrm{E}(N_2) = [1-P_2(p_1)]^{-1}$$

are general and are not limited to the case where $\{p_v(2)\}$ is of phase type.

In using the recurrence relation (17), we note that it is essentially similar to the elementary Pascal triangle except that much more storage is required. For any n, we need, in principle, store all the $m_1 m_2$-vectors $\boldsymbol{K}(v_1, v_2; n)$ with $0 \leqslant v_1+v_2 \leqslant n$, and from these, the vectors for which $v_1+v_2 = n+1$ can be computed. For each value of n, we compute and store the corresponding probabilities $r(v_1+1, v_2+1; n+1)$ and store the vectors $\boldsymbol{K}$ corresponding to $n+1$ in the same memory locations as previously used. Clearly, such a massive recursive computation should be planned with the greatest care and negligible vectors $\boldsymbol{K}$ should be adaptively trimmed.

7. A numerical example. We wrote a FORTRAN code to implement the recursive computation of the probabilities $P[Y > n]$ by means of formula (7). An option of the program allows implementation of the asymptotic formula

(9). When that option is called, the Perron–Frobenius eigenvalue η and the eigenvectors u and v are iteratively computed by an (improvable) application of Elsner's algorithm. We tested the code on a large number of examples. To report one such example, we let the densities $\{p_\nu(1)\}$ and $\{p_\nu(2)\}$ be negative binomial with generating function $(qz(1-pz)^{-1})^m$, with common parameter $p = 0.35$ and m equal to 5 for the first density and 6 for the second. The means and standard deviations of these densities are respectively

$$\mu(1) = 7.6923, \quad \sigma(1) = 2.0352$$

and

$$\mu(2) = 9.2308, \quad \sigma(2) = 2.2294.$$

The direct recursive scheme halts when 758 terms have been computed. The computed mean $\sum_{n=0}^{758} P[Y > n]$ equals 71.005, while the exact mean is 71.006.

Evaluated to five decimal places of accuracy the maximal eigenvalue η of the (30×30)-matrix L is equal to 0.98537 and the constant C is found to be 1.0387. If we stop the recursive computation as soon as ten successive terms are within 5×10^{-5} of the values given by the asymptotic formula, only **forty** terms are computed exactly. Replacing the complementary distribution by its geometric asymptote for $n > 40$ results in a computed mean of 71.003. These results are fairly typical, but it is of course easy to construct examples for which the approach to the geometric asymptote is much slower. In such cases, the exact recursion usually also evaluates many more terms before the computed mean is close to the known exact mean.

References

[1] L. Elsner, *Verfahren zur Berechnung des Spektralradius nicht-negativer irreduzibler Matrizen. I*, Computing 8 (1971), pp. 32–39.

[2] I. Kopocińska and B. Kopociński, *On coincidences in renewal streams*, Zastos. Mat. 19 (1987), pp. 169–180.

[3] B. Kopociński and K. Topolski, *Limit distribution of the time of coincidence in renewal streams*, this fascicle, pp. 15–21.

[4] M. F. Neuts, *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*, Johns Hopkins Univ. Press, Baltimore 1981.

[5] — *Computation and applications of geometric mixtures of convolutions*, Preprint, Dept. of Systems and Industrial Engineering, Univ. of Arizona, Tucson, Arizona, 1986.

[6] P. Schweitzer, *An iterative aggregation-disaggregation algorithm for computing the spectral radius of a non-negative matrix*, Working Paper No. QM 8433, Graduate School of Management, Univ. of Rochester, 1985.

SYSTEMS AND INDUSTRIAL ENGINEERING
THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721, U.S.A.

Received on 1986.12.15