A basic result on the integral for birth-death Markov processes

Carlos M. Hernández-Suárez

CGIC - Universidad de Colima, Gonzalo de Sandoval 444

Colima, Colima; 28045 México

Ph (331) 4-11-33, Fax (331) 2-75-81

e-mail : cmh2@cgic.ucol.mx

Carlos Castillo-Chavez

Biometrics Unit, Cornell University, Warren Hall 435, Ithaca, NY 14850

Phone: (607)-255-8103, e-mail : cc32@cornell.edu

Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas

Universidad Nacional Autónoma de México, México City, APDO Postal 20-726, México

e-mail: castillo@leibniz.iimas.unam.mx

Institute for Mathematics and its Applications, University of Minnesota

400 Lind Hall, 207 Church St. SE, Minneapolis, MN 55455-0463, 612-624-6066, FAX 612-626-7370

Department of Mathematics, Howard University, Washington, DC 20059

A basic result on the integral for birth-death Markov processes

1 Summary

In this paper a regenerative argument is used to derive an expression for the expectation of the integral for a birth-death Markov process up to extinction time as well as for the expected time to extinction. Some applications to classical birth-death processes are given.

ABSORBING MARKOV PROCESSES; BIRTH-DEATH PROCESSES; QUEUING THEORY; INTEGRALS OF MARKOV PROCESSES

2 Introduction

Integrals of nonnegative stochastic processes arise naturally in engineering, biology and inventories. Applications include queuing, storage, traffic and inventory systems (see 1-3). In biology it has been associated with total food consumption and production of toxins of a bacteria [4] as well as with total cost of epidemics [5-7]. Limiting properties for the integral of nonnegative stochastic processes have been studied in [8-13] among others, whereas McNeil [14] and Picard [15] have studied integral functionals. Here we introduce an alternative method to evaluate both the expectation of the integral for a birth-death Markov process and the expected time to absorption. The paper is organized as follows: section 3 discusses the theoretical results and section 4 presents some applications to specific birth-death Markov processes.

Let $X(t)$ be a birth-death process on a subset of $\mathcal{N} = 1, \infty, \in, \dots$, with birth and death rates λ_i , and μ_i respectively for state i . Assume that $X(0) = k$ for some $k > 0$, and define

$$Z = \inf \{t > 0 : X(t) = 0\}$$

and

$$Y = \int_0^Z X(t)dt,$$

thus Z is the time to absorption starting from state k and Y is the area under $X(t)$ up to the time when the process vanishes. Let S_r be the random

vector corresponding to the total amount of time spent in state r before absorption, $r = 1, 2, 3, \dots$. Observe that

$$Y = \int_0^Z X(t)dt = \sum_i i S_i \quad (1)$$

a result already pointed out in [16].

The absorbing state cause that the stationary distribution is degenerated with all its mass at zero, which makes very difficult to analyze the behavior of the process before absorption. This problem has been extensively studied in several contexts. Darroch and Seneta [17] defined the quasi-stationary distribution (QSD) as “the probability that the process is in state $j \in T$ at time t , given that has not yet been absorbed, and started from state $i \in T$ with probability π_i ”. Cavender [18] defined the QSD $q_n(t)$ as “the conditional probability of the system being in state n at time t , given that absorption has not occurred up until then”. More explicitly, Kryscio and Lefèvre [19] in the context of the SIS epidemic model defined the QSD as

$$Q_n = \lim_{t \rightarrow \infty} q_{i,n}(t)$$

where

$$q_{i,n}(t) = p_{i,n}(t)/(1 - p_{i,0}(t)) \quad \text{for } n = 1, 2, 3, \dots, N$$

where $p_{i,n}(t)$ is the probability that the process is in state n at time t given that it started from state i at time zero. They suggested two approximations to the QSD by modifying the transitions in the Markov process. One approximation consists in changing the death rate of state n , μ_n , from $n\mu$ to $(n-1)\mu$. The other approximation consists in equating μ_1 to zero. In both approximations the birth rates are left unchanged. Since these modifications are applied in the context of an SIS epidemic model, they are referred as “one permanently infected” and “reflecting state zero” approximations. Clearly, both of them lack of absorbing states and the modified process is ergodic. Kryscio and Lefèvre [19] showed that the “reflecting state zero” is stochastically larger than the QSD which in turn is stochastically larger than the “one permanently infected” approximation. Näsell [20] used Kryscio and Lefèvre

[19] approximations to the QSD of the Ross-malaria model and suggested other based on perturbation theory.

A lot of research on the long-term behavior of Markov chains with absorbing states has been done concerning the quasi-stationary distribution of the process. It may be interesting to develop expressions for π_i , the expected proportion of time the process spends on state i before going to absorption. This is certainly less stringent than the QSD since it does not condition in the process not being in the absorbing state, and as it will be show latter, $\underline{\Pi}' = \{\pi_1, \pi_2, \pi_3 \dots\}$ can be used to calculate the expected time to absorption and the expected integral under the stochastic path (1).

3 Methodology

If we substitute transitions to the absorbing state 0 by transitions to the initial state k , whenever the resulting process is ergodic a stationary distribution exists. Call this process the Modified Process and its stationary distribution the Modified Stationary Distribution, and denote this by $\underline{\Pi}' = \{\pi_1, \pi_2, \pi_3 \dots\}$. Clearly, if $k = 1$ then the modified process is the same as the “reflecting state zero” approximation of [19].

Let r be an arbitrary but fixed state $\in \mathcal{N}$, $r \neq \{0\}$. If the modified process is ergodic, then when $t \rightarrow \infty$ state r will be visited infinitely often. We will denote that a cycle has been completed every time a death occurs with the process being in state 1. also define $E\{S_r\}$ as the expected time spent in state r in a cycle. Hence, it is clear that $E\{S_r\}$ equals the expected time spent in state r before the process goes to absorption in the original process $\{X(t) : t > 0\}$.

Now define S_{rj} to be the time spent in state r in the j -th cycle in the modified process, $j = 1, 2, \dots$. By the strong law of large numbers, we have that

$$\pi_r = \lim_{n \rightarrow \infty} \frac{\sum_{j=1}^n S_{rj}}{\sum_{j=1}^n \sum_i S_{ij}} = \lim_{n \rightarrow \infty} \frac{n^{-1} \sum_{j=1}^n S_{rj}}{n^{-1} \sum_{j=1}^n \sum_i S_{ij}} = \frac{E\{S_r\}}{E\{\sum_i S_i\}}.$$

Since $E\{\sum_i S_i\} = E\{Z\}$ it follows that

$$E\{S_r\} = \pi_r E\{Z\}.$$

Furthermore, since $E \{S_1\} = 1/\mu_1$, (see appendix A) then

$$E \{Z\} = (\pi_1 \mu_1)^{-1}, \quad (2)$$

and from (1) and (2) one easily concludes that

$$E \{Y\} = E \left\{ \int_0^Z X(t) dt \right\} = \sum_i i E \{S_i\} = (\pi_1 \mu_1)^{-1} \sum_i i \pi_i. \quad (3)$$

Expression (2) for initial state $k = 1$ was already derived in [20] by simplyfing an expression of Norden [21], nevertheless, (2) is more general since here π_1 is unique for every initial state k .

Observe that we can apply this method to calculate the expected time for the process to reach any element in an arbitrary set of states $\mathcal{A}$, by changing transitions to elements in $\mathcal{A}$ to transitions to the initial state k . Clearly for $\mathcal{A} = \{0\}$ we are computing the expected time to absorption.

We now proceed to apply expressions (2) and (3) to some classical birth-death processes.

4 Examples

The following are applications of (3) to some birth-death Markov processes. In the first series of examples the initial state is state 1. As previously stated, the stationary distribution corresponds to the so-called “reflecting state 0 approximation” to the quasi-stationary distribution. This approximation is defined by the following system of equations

$$\pi_n \mu_n = \pi_{n-1} \lambda_{n-1} \quad n = 2, 3, \dots$$

which can be solved recursively to give the solution

$$\pi_n = \frac{\lambda_1 \lambda_2 \lambda_3 \dots \lambda_{n-1}}{\mu_2 \mu_3 \mu_4 \dots \mu_n} \pi_1.$$

Define $H(n)$ as

$$H(n) = \frac{\lambda_1 \lambda_2 \lambda_3 \dots \lambda_{n-1}}{\mu_2 \mu_3 \mu_4 \dots \mu_n}, \quad (4)$$

it is easy to see that $E\{Y\}$ in birth-death Markov processes with initial state 1 and 0 as an absorbing state, can be explicitly computed as

$$E\{Y\} = \frac{1}{\pi_1 \mu_1} \sum_{i=1}^{\infty} i H(i) \pi_i = \mu_1^{-1} \sum_{i=1}^{\infty} i H(i). \quad (5)$$

A simple result since it does not require the calculation of Π . The last equality will be used in the following examples.

(a) $\lambda_n = \lambda, \mu_n = \mu, k = 1$.

This is the random walk with absorbing barrier at zero or $M/M/1$ queue. This process lacks of absorbing states but by reflecting state zero to the initial state k one can obtain information on important quantities, for instance, the total work performed by the server in a busy period and the expected time to the first visit to state zero. Let $\rho = \lambda/\mu$. From (4) we have $H(n) = \rho^{n-1}$. The expected time to absorption is $E\{Z\} = (\mu - \lambda)^{-1}$, and applying (5) we have that

$$\begin{aligned} E\{Y\} &= \mu^{-1} \sum_{i=1}^{\infty} i \rho^{i-1} \\ &= \frac{1}{\mu(1 - \rho)^2} \end{aligned}$$

clearly the last expression holds for $\rho = \lambda/\mu < 1$.

(b) $\lambda_n = \lambda, \mu_n = \mu n, k = 1$

This is the $M/M/\infty$ queue or immigration-death process. Again, by reflecting state zero to the initial state, Y corresponds to the total amount of work performed by the server in a busy period and $E\{Z\}$ is the expected time to the first visit to zero. In this case we have $H(n) = \rho^{n-1}/n!$. $E\{Z\}$ can be shown to be $(e^\rho - 1)/\rho\mu$ and

$$E\{Y\} = \mu^{-1} \sum_{i=1}^{\infty} i \rho^{i-1}/i! = \mu^{-1} e^\rho$$

(c) $\lambda_n = \lambda n, \mu_n = \mu n, k = 1$

This is the linear birth-death process. We have $H(n) = \rho^{n-1}/n$. Here $E\{Z\} = -\log(1 - \rho)\lambda^{-1}$ and

$$\begin{aligned} E\{Y\} &= \mu^{-1} \sum_{i=1}^{\infty} i \rho^{i-1} / i \\ &= (\mu - \lambda)^{-1}. \end{aligned}$$

for $\rho < 1$.

(d) $\lambda_n = \lambda n(N - n)/N, \mu_n = \mu n, k = 1$.

This is the stochastic logistic epidemic model, or the Susceptible -Infected -Susceptible (SIS) model. Here Y corresponds to the sum of the total duration times of all infections, which is often referred to as the cost of an epidemic [7]. In this case it can be shown that

$$\pi_j = \frac{\pi_1}{j} \frac{(N - 1)!}{(N - j)!} \left(\frac{\rho}{N} \right)^{j-1}, \quad \text{for } j = 2, 3, \dots, N$$

using (3) we get

$$E\{Y\} = \mu_1^{-1} \sum_{j=1}^N \frac{(N - 1)!}{(N - j)!} \left(\frac{\rho}{N} \right)^{j-1}.$$

which can not be expressed in a simple form for general ρ , but at least can be solved numerically. In the particular case $\rho < 1$ and $N \rightarrow \infty$ it is shown in [21] that

$$\pi_j \approx \frac{1}{j} \rho^{j-1} \pi_1.$$

Using this result, we can see that the limiting distribution can be approximated with that of the birth-death process (see 22), therefore $E\{Z\} = -\log(1 - \rho)\lambda^{-1}$. Thus an approximation to $E\{Y\}$ under the specified conditions is $(\mu - \lambda)^{-1}$.

Initial State $k > 1$.

Expressions (2) and (3) still apply for an initial state $k > 1$, as long as the limiting distribution exists. Since absorptions to state 0 in the original process are being substituted by transitions to the initial state k , the recurrence relationship that defines this process becomes (see Appendix B):

$$\begin{aligned}\mu_n \pi_n &= \pi_{n-1} \lambda_{n-1} + \pi_1 \mu_1, & n \leq k \\ \mu_n \pi_n &= \pi_{n-1} \lambda_{n-1}, & n > k\end{aligned}\tag{6}$$

This system of recursions for the π_j 's is a complex set which is not possible to write in a simple form. Here we analyze the random walk with absorbing barrier at zero or $M/M/1$ queue (case a) and the birth-death process (case c). We also present numerical results for the SIS epidemic model (case d).

Let the initial state $X(0) = k$, $k > 0$. It can be shown by using (6) that for the $M/M/1$ queue,

$$\pi_j = \pi_1 \frac{1 - \rho^j}{(1 - \rho)}, \quad j = 1, 2, \dots, k$$

and that

$$\pi_j = \pi_1 \frac{1 - \rho^k}{(1 - \rho)} \rho^{j-k}, \quad j = k+1, k+2, \dots$$

Thus we have $E\{Z\} = k/(\mu - \lambda)$ and

$$E\{Y\} = \frac{k(1 + k + \rho(1 - k))}{2(1 - \rho)^2}.$$

For the birth-death process, the limiting distribution using (6) is given by

$$\pi_j = \pi_1 \frac{1 - \rho^j}{j(1 - \rho)}, \quad j = 1, 2, \dots, k$$

and

$$\pi_j = \pi_1 \frac{1 - \rho^k}{j(1 - \rho)} k \rho^{j-k}, \quad j = k, k+1, k+2, \dots$$

The expression for $E\{Z\}$ is rather complicated, but for $E\{Y\}$ becomes

$$E\{Y\} = \frac{k - \rho + \rho^{k+1}(1 - k)}{\mu(1 - \rho)^2}.$$

Now we analyze numerically the behavior of $E\{Y\}$ and $E\{Z\}$ for the SIS epidemic model for different values of ρ keeping N constant. For a specified value of ρ we obtain an estimate of $E\{Y\}$ and $E\{Z\}$ via simulations with MATLAB and compare it with the calculated with (2) and (3). The results are shown in Table 1.

5 Conclusions

The methodology proposed here, summarized in result (5), uses simple ergodic and regenerative arguments to calculate analytically the expected value of the integral under $X(t)$ up to extinction time for birth-death Markov processes. Although in general, the expressions for some birth-death processes may be complicated, they can be solved numerically since they are based in a simple system of recursions. Extensions of this methodology to a general class of Markov processes in continuous time with a set of absorbing states $\mathcal{A}$ are worth to explore, where the transitions to elements in $\mathcal{A}$ are substituted with transitions to the initial state.

Acknowledgements

The authors wish to thank Charles E. McCulloch, Jorge Velasco-Hernández, Rich Durrett and Sid Resnick for valuable comments. We also want to thank three anonymous referees for suggestions to improve this work.

References

- [1] P. A. P. Moran, The Theory of Storage. Methuen, London (1959).
- [2] D. P. Gaver, Highway delays resulting from flow-stopping incidents, J. Appl. Prob. 6 (1969) 137-153.

- [3] J. Gani, First emptiness problems in queuing, storage and traffic theory. Proc. 6th Berkeley Symp. Math. Stat. Prob. 3 (1970) 515-532.
- [4] P. S. Puri, A class of stochastic models of response after infection in the absence of defense mechanism, Proc. 5 th Berkeley Symp. Math. Stat. Prob. 4 (1967) 537-547.
- [5] D. Jerwood, A note on the cost of the simple epidemic, J. Appl. Prob. 3 (1970) 339-352.
- [6] F. Downtown, The area under the infectives trajectory of the general stochastic epidemic, J. Appl. Prob. 9 (1972).414-417.
- [7] J. Gani, D. Jerwood, The cost of a general stochastic epidemic, J. Appl. Prob. 9 (1972) 257-269.
- [8] D.A. Darling, M. Kac, On occupation times for Markoff processes, Trans. Amer. Math. Soc. 84 (1957) 444-458.
- [9] L. Takács, On a sojourn time problem, Theor. Prob. Appl. 3 (1958) 58-65.
- [10] L. Takács, On a sojourn time problem in the theory of stochastic processes, Trans. Amer. Math. Soc. 93 (1959) 531-540.
- [11] H. Kesten, Occupation times for Markov and semi-Markov processes, Trans. Amer. Math. Soc. 103 (1962) 82-112.
- [12] R. Pyke, R. A. Schaufele, Limit theorems for Markov renewal processes, Ann. Math. Stat. 35 (1964) 1746-1764.
- [13] N. Kaplan, Limit theorems for the integral of a population process with immigration, Stoc. Proc. Appl. 2 (1974) 281-294.
- [14] D. R. McNeil, Integral functionals of birth-death processes and related limiting distributions, Ann. Math. Stat. 41 (1970) 480-485.
- [15] P. Picard, On the integral functionals of linear birth-death immigration processes, Stat. Decis. 2 (1985) 105-109.
- [16] P. S. Puri, On the homogeneous birth-and-death process and its integral, Biometrika 53 (1966) 61-71.

- [17] J.N. Darroch, E. Seneta, On quasi-stationary distributions in absorbing continuous-time finite Markov chains, J. Appl. Prob. 4 (1967) 192-196.
- [18] J. A. Cavender, Quasi-stationary distribution of birth-and-death processes, Adv. Appl. prob. 10 (1978) 570-586.
- [19] R.J. Kryscio, C. Lefèvre, On the extinction of the S-I-S stochastic logistic epidemic, J. Appl. Prob. 27 (1989) 685-694.
- [20] I. N. Nåsell, The quasi-stationary distribution of the closed endemic SIS model, Adv. Appl. prob. 28 (1996) 895-932.
- [21] R. H. Norden, On the distribution of the time to extinction in the stochastic logistic population model, Adv. Appl. Prob. 14 (1982) 687-708.
- [22] N.T. Bailey, The Elements of Stochastic Processes, Wiley, NY, (1964).

Appendix A

Given that process will visit state 1 with probability 1, we can calculate $E(S_1)$, the expected time spent in this state before extinction, assuming that the process is already in state 1. Using first step analysis, we have that the expected time in state 1 is $(\lambda + \mu)^{-1}$ if a death occurs in state 1, whereas by the memoryless property it is $(\lambda + \mu)^{-1} + E(S_1)$ if the process goes to state 2 (a birth occurs). Therefore, by the law of total probability

$$E(S_1) = E(S_1 | \text{A death occurs in state 1}) P(\text{A death occurs in state 1}) + E(S_1 | \text{A birth occurs in state 1}) P(\text{A birth occurs in state 1}).$$

Now, given that

$$P(\text{A death occurs in state 1}) = \frac{\mu}{\lambda + \mu} = 1 - P(\text{A birth occurs in state 1})$$

and that

$$E(S_1 | \text{A death occurs in state 1}) = (\lambda + \mu)^{-1},$$

and

$$E(S_1 | \text{A birth occurs in state 1}) = (\lambda + \mu)^{-1} + E(S_1),$$

then we have

$$E(S_1) = (\lambda + \mu)^{-1} \frac{\mu}{\lambda + \mu} + [(\lambda + \mu)^{-1} + E(S_1)] \frac{\lambda}{\lambda + \mu}.$$

Solving for $E(S_1)$ yields

$$E(S_1) = 1/\mu.$$

Appendix B

Let $N_{i,j}(t)$ be the number of transitions from state i to state j by time t . Then

$$\lim_{t \rightarrow \infty} N_{i,j}(t) = \pi_i R(i, j)$$

where $R(i, j)$ is the rate at which the process moves from state i to state j . Clearly in a birth death process

$$\lim_{t \rightarrow \infty} N_{n,n+1}(t) = \pi_n \lambda_n,$$

and

$$\lim_{t \rightarrow \infty} N_{n,n-1}(t) = \pi_n \mu_n.$$

For any state $n > k$ there is only one way to reach state n from state $n - 1$, thus when $t \rightarrow \infty$ the number of transitions from $n - 1$ to n equals that of n to $n - 1$:

$$\lim_{t \rightarrow \infty} N_{n,n-1}(t) = \lim_{t \rightarrow \infty} N_{n-1,n}(t)$$

or $\pi_n \mu_n = \pi_{n-1} \lambda_{n-1}$ which proves the second equality of (7). Since transitions to state zero are substituted with transitions to the initial state k , for $n \leq k$ every transition from n to $n - 1$ will eventually result in a transition from $n - 1$ to n or a transition from 1 to the initial state k , therefore

$$\lim_{t \rightarrow \infty} N_{n,n-1}(t) = \lim_{t \rightarrow \infty} N_{n-1,n}(t) + \lim_{t \rightarrow \infty} N_{1,k}(t)$$

or $\pi_n \mu_n = \pi_{n-1} \lambda_{n-1} + \pi_1 \mu_1$ which proves the first equality of (7).