Transforming renewal processes for simulation of nonstationary arrival processes

Ira Gerhardt
Barry L. Nelson

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208-3119
i-gerhardt@northwestern.edu nelsonb@northwestern.edu

October 31, 2008

Abstract

Simulation models of real-life systems often assume stationary (homogeneous) Poisson arrivals. Therefore, when nonstationary arrival processes are required it is natural to assume Poisson arrivals with a time-varying arrival rate. For many systems, however, this provides an inaccurate representation of the arrival process which is either more or less variable than Poisson. In this paper we extend techniques that transform a stationary Poisson arrival process into a nonstationary Poisson arrival process (NSPP) by transforming a stationary renewal process into a nonstationary, non-Poisson (NSNP) arrival process. We show that the desired arrival rate is achieved, and that when the renewal base process is either more or less variable than Poisson, then the NSNP process is also more or less variable, respectively, than a NSPP. We also propose techniques for specifying the renewal base process when presented properties of, or data from, an arrival process and illustrate them by modeling real arrival data.

Keywords: Arrival counting process, Phase-type distribution, Nonstationary Poisson process

1. Introduction

Queueing models are frequently utilized as tools for performance analysis of real-life systems. Although the characteristics of such systems may vary in time, analytical models typically utilize stationary Poisson arrival processes. However, many situations exist where the assumptions of stationary arrivals, and of Poisson arrivals, will be inaccurate. By “stationary” we mean constant arrival rate.

For instance, there is clearly a need for nonstationary arrival processes in the quantitative management of call centers; see Gans et al. (2003) for background. Arrival rates to telephone
call centers vary widely with time of day or day of week, and fluctuations in call rates occur in response to advertising, seasonal trends, etc. (Testik et al., 2004). To simulate such systems without accounting for nonstationarity (e.g., using only the average arrival rate over a day) may lead to underestimation of key performance measures due to unidentified system congestion (Harrod and Kelton, 2006; Whitt, 1981).

In addition, real-world studies of telecommunication networks, manufacturing systems and consumer behavior have revealed that arrival processes may be either more variable or more regular than Poisson. For example, various Internet traffic studies have shown that network traffic is often too bursty to be accurately modeled by Poisson processes (Paxson and Floyd, 1995), while studies in consumer behavior notice the importance of models for buying incidence of frequently purchased goods that may be less variable than Poisson (Wu and Chen, 2000).

Thus, there exists the need for simulation input models for nonstationary, non-Poisson (NSNP) arrival processes. In this paper, we exploit two well-known methods that generate a nonstationary Poisson process (NSPP) by transforming a stationary Poisson process, to generate NSNP arrivals by transforming a stationary renewal process that is either more variable or more regular than Poisson. The two methods are inversion (Çınlar, 1975) and thinning (Lewis and Shedler, 1979). Our methods are easy to use and intuitive, but each only models one form of departure from “Poissonness.”

The remainder of this paper is organized as follows. In Section 2 we provide algorithms for applying the inversion and thinning methods to general stationary renewal processes; we analyze the resulting process, and discuss the NSPP as a special case. We describe advantages of using phase-type (Ph) distributions as specific choices for the stationary renewal process in Section 3. Section 4 contains techniques for specifying the renewal base process for both methods when provided properties of, or data on, the resulting nonstationary process. We also provide analysis of the variability of the fitted NSNP process in that section, examining the NSNP process empirically as well as performance statistics when the NSNP process acts as the arrival process to a queue. We conclude with suggestions for future research in Section 5.
2. Methods

2.1. Renewal Processes

We begin with a set of nonnegative interevent times \( \{X_n, n \geq 1\} \), where the subset \( \{X_n, n \geq 2\} \) are i.i.d. with cumulative distribution function \( G \), while \( X_1 \), the time until the first event, may not have distribution \( G \). We let \( S_n \) denote the time of the \( n^{th} \) event; that is, \( S_0 = 0 \) and \( S_n = \sum_{i=1}^{n} X_i \), for \( n = 1, 2, \ldots \). Let \( N(t) \) denote the number of events that have occurred on or before time \( t \); that is, \( N(t) = \max\{n \geq 0 : S_n \leq t\} \), for \( t \geq 0 \). Therefore, we have renewal sequence \( \{S_n, n \geq 1\} \) and delayed renewal process \( \{N(t), t \geq 0\} \) generated by interrenewal times \( \{X_n, n \geq 2\} \sim G \) (Cox and Lewis, 1966). We use the following shorthand throughout: \( G(t) = \Pr\{X_2 \leq t\} \), \( \tau = \mathbb{E}\{X_2\} \), \( \sigma^2 = \text{Var}\{X_2\} \), and \( \text{cv}^2 = \sigma^2/\tau^2 \). We assume that \( \lim_{h \downarrow 0} G(h) = 0 \), and that \( X_2 \) has density \( g(t) = \frac{dG}{dt} \), for \( t \geq 0 \).

If \( X_1 \) has the equilibrium distribution associated with \( G \), specifically

\[
G_e(t) \equiv \Pr\{X_1 \leq t\} = \frac{1}{\tau} \int_0^t [1 - G(u)]du,
\]

then \( N \) is an equilibrium renewal process (Kulkarni, 1995). Notice that \( \mathbb{E}\{X_1\} = (\tau^2 + \sigma^2)/(2\tau) \).

For an equilibrium renewal process,

\[
\mathbb{E}\{N(t)\} = \frac{t}{\tau}, \tag{1}
\]

and

\[
\text{Var}\{N(t)\} = \frac{t}{\tau} - \left(\frac{t}{\tau}\right)^2 + \frac{2}{\tau} \int_0^t \mathbb{E}\{N_s(u)\} du, \tag{2}
\]

for all \( t \geq 0 \), where \( N_s \) is the ordinary renewal process with i.i.d. interrenewal times \( \{X_n', n \geq 1\} \) from distribution \( G \) (i.e., \( X_1' \) is from the same distribution as the latter \( X_n'\); see Cox and Smith (1954)). Combining (2) with a result from Smith (1959, p. 1) regarding the mean count of an ordinary renewal process, we find that

\[
\text{Var}\{N(t)\} = \frac{\sigma^2}{\tau^3} t + o(t).
\]

Thus, for equilibrium renewal process \( N \),

\[
\text{Var}\{N(t)\} \approx \text{cv}^2 \mathbb{E}\{N(t)\}, \tag{3}
\]

3
for large $t$; by (3) we mean \( \lim_{t \to \infty} \frac{\text{Var}\{N(t)\}}{\mathbb{E}\{N(t)\}} = cv^2 \). Since $N$ is an equilibrium renewal process, $N$ is therefore a stationary point process. Based on our assumptions on $G$, the stationarity of $N$, and $\tau > 0$ (implying $\tau^{-1} < \infty$), we conclude that $N$ is regular (or orderly), implying only one renewal may occur at a time (Ross 1983).

For the nonstationary process we desire, let $r(t)$ denote its rate function, and $R(t) = \int_0^t r(u)du$ the integrated rate function of the process. We assume that $r(t)$ is integrable.

### 2.2. The Inversion Method

In this section we first present an inversion algorithm for generating interarrival times for a NSNP process, and then analyze the mean and variance of that process. The inversion method is well-known when the base process is Poisson (with rate 1), but does not appear to have been studied for stationary renewal base processes. Since the stationary renewal process must have rate 1, we specify the distribution $G$ (and associated $G_e$) such that $\tau = 1$ and $cv^2 = \sigma^2$. Notice that, for $s \in \mathbb{R}$, $R^{-1}(s) = \inf\{t : R(t) \geq s\}$.

**Algorithm 2.1.** The Inversion Method for a Stationary Renewal Process.

1. Set $V_0 = 0$, index counter $n = 1$. Generate $S_1 \sim G_e$. Set $V_1 = R^{-1}(S_1)$.

2. Return interarrival time $W_n = V_n - V_{n-1}$.

3. Set $n = n + 1$. Generate $X_n \sim G$. Set $S_n = S_{n-1} + X_n$ and $V_n = R^{-1}(S_n)$.

4. Go to Step [2](#).

Thus, the sequence \{\(W_n, n \geq 1\)\} is a potential realization of interarrival times for the nonstationary process from the inversion method.

Let $\{I(t), t \geq 0\}$ denote the counting process generated by the inversion method; that is, $I(t) = \max\{n \geq 0 : V_n \leq t\}$. Then we have the following properties of $I(t)$.

**Result 2.1.** $\mathbb{E}\{I(t)\} = R(t)$, for all $t \geq 0$, and $\text{Var}\{I(t)\} \approx cv^2R(t)$, for large $t$. 

4
Proof. Since \( N \) is an equilibrium renewal process and \( \tau = \mathbb{E}\{X_2\} = 1 \), then \( \mathbb{E}\{N(t)\} = t \), for all \( t \geq 0 \), from (1), while \( \text{Var}\{N(t)\} \approx \text{cv}^2 t \), for large \( t \), from (3). Thus,

\[
\mathbb{E}\{I(t)\} = \mathbb{E}\{\mathbb{E}[I(t) | N(R(t))]\}
= \mathbb{E}\{N(R(t))\}
= R(t),
\]

for all \( t \geq 0 \), while

\[
\text{Var}\{I(t)\} = \mathbb{E}\{\text{Var}[I(t) | N(R(t))]\} + \mathbb{E}\{\mathbb{E}[I(t) | N(R(t))]\}
= 0 + \text{Var}\{N(R(t))\}
\approx \text{cv}^2 R(t),
\]

for large \( t \).

When \( N \) is Poisson with rate 1, the resulting process from the inversion method is a NSPP. In this case, \( \text{cv}^2 = 1 \) and the equilibrium distribution \( G_e(t) = G(t) = 1 - e^{-t} \), for \( t \geq 0 \). Result 2.1 holds exactly, since \( \text{Var}\{N(t)\} = t \) at all times, by (2). Thus, \( \text{Var}\{I(t)\}/\mathbb{E}\{I(t)\} = 1 \), for all \( t \geq 0 \).

The inversion method is useful when \( R(t) \) is easily invertible. Forms for the rate function \( r(t) \) that have utilized the inversion method for NSPP generation include piecewise-linear (Klein and Roberts 1984), trigonometric (Chen and Schmeiser 1992), and piecewise-constant (Harrod and Kelton 2006). The inversion method may not be useful if \( R(t) \) is difficult to invert or otherwise intractable. We discuss a second method that is unaffected by analytical complications in \( R(t) \) in Section 2.3.

Notice that \( \text{cv}^2 \) provides a measure of the deviation of the stationary renewal process from Poisson; Cox and Isham (1980) use \( \text{cv}^2 \) to classify renewal processes as either overdispersed (for \( \text{cv}^2 > 1 \)) or underdispersed (for \( \text{cv}^2 < 1 \)). From Result 2.1 we see that this deviation measure is effectively preserved by the inversion method; that is,

\[
\frac{\text{Var}\{I(t)\}}{\mathbb{E}\{I(t)\}} \approx \text{cv}^2 \approx \frac{\text{Var}\{N(t)\}}{\mathbb{E}\{N(t)\}},
\]
for large $t$. Thus, $cv^2$ provides a single parameter with which to construct arrival processes that are more or less variable than Poisson in a particularly intuitive way.

Of course, any method represents a choice as to how to depart from NSPP arrivals. For instance, the doubly stochastic Poisson process $D(t)$ (see, for instance, Avramidis et al. (2004)) is a NSPP with rate function $Zr(t)$, conditional on random variable $Z \geq 0$ with $E\{Z\} = 1$ and $\text{Var}\{Z\} = cv^2$. For this process it is easy to show that

$$\frac{\text{Var}\{D(t)\}}{\text{E}\{D(t)\}} = 1 + cv^2R(t),$$

which is a distinctly different choice since it becomes increasingly variable as time passes. The thinning method in the following section yields yet another choice.

2.3. The Thinning Method

As with the inversion method in Section 2.2, we present an algorithm for generating interarrival times by thinning a stationary renewal process to obtain a nonstationary process, and then offer analysis of the mean and variance of the generated NSNP. We begin by selecting a value $r^* \geq \max_{t \geq 0} r(t)$ which we assume to be finite. We assign the stationary renewal process the arrival rate $r^*$, and specify the distribution $G$ (and associated $G_e$) with $\tau = (r^*)^{-1}$ and $\sigma^2 = cv^2/(r^*)^2$.

**Algorithm 2.2. The Thinning Method for a Stationary Renewal Process.**

1. Set index counters $n = 1$, $k = 1$, and $T_0 = 0$. Generate $S_1 \sim G_e$.

2. Generate $U_1 \sim \text{Uniform}[0,1]$. If $U_1 \leq r(S_1)/r^*$, then

   (a) Set $T_1 = S_1$.

   (b) Return interarrival time $Y_1 = T_1 - T_0$.

   (c) Set $k = 2$.

3. Set $n = n + 1$. Generate $X_n \sim G$. Set $S_n = S_{n-1} + X_n$.

4. Generate $U_n \sim \text{Uniform}[0,1]$. If $U_n \leq r(S_n)/r^*$, then
(a) Set $T_k = S_n$.

(b) Return interarrival time $Y_k = T_k - T_{k-1}$.

(c) Set $k = k + 1$.

5. Go to Step 3.

Thus, the sequence $\{Y_k, k \geq 1\}$ is a potential realization of interarrival times for the nonstationary process generated from the thinning method.

Let $\{M(t), t \geq 0\}$ denote the counting process generated by the thinning method; that is, $M(t) = \max\{k \geq 0 : T_k \leq t\}$. Then we have the following property of $M(t)$.

**Result 2.2.** $E\{M(t)\} = R(t)$, for all $t \geq 0$.

The proof is presented in Appendix A.

When the renewal base process $N$ to be thinned is Poisson (with rate $r^*$), the resulting process $M$ is a NSPP (Lewis and Shedler, 1979). In this case, $\text{Var}\{M(t)\}/E\{M(t)\} = 1$, for all $t \geq 0$. If $N$ is not Poisson, an expression for $\text{Var}\{M(t)\}$ will depend on the interrenewal distribution $G$. The following result provides some insight on the effect of thinning on the arrival process variance.

**Result 2.3.** Suppose $r(t) = \bar{r} > 0$ for all $t \geq 0$. If $M$ is the resulting counting process when we thin equilibrium renewal process $N$ (with rate $r^* \geq \bar{r}$, and interrenewal variance $\sigma^2 = cv^2/(r^*)^2$), then

$$\frac{\text{Var}\{M(t)\}}{E\{M(t)\}} \approx \left(1 - \frac{\bar{r}}{r^*}\right) + \left(\frac{\bar{r}}{r^*}\right) cv^2,$$

for large $t$.

**Proof.**

$$\text{Var}\{M(t)\} = E\{\text{Var}[M(t)|N(t)]\} + \text{Var}\{E[M(t)|N(t)]\} \quad (5)$$

$$= E \left\{ \left(\frac{\bar{r}}{r^*}\right) \left(1 - \frac{\bar{r}}{r^*}\right) N(t) \right\} + \text{Var} \left\{ \left(\frac{\bar{r}}{r^*}\right) N(t) \right\} \quad (6)$$

$$= \left(\frac{\bar{r}}{r^*}\right) \left(1 - \frac{\bar{r}}{r^*}\right) E\{N(t)\} + \left(\frac{\bar{r}}{r^*}\right)^2 \text{Var}\{N(t)\}$$

$$\approx \left(\frac{\bar{r}}{r^*}\right) \left(1 - \frac{\bar{r}}{r^*}\right) r^* t + \left(\frac{\bar{r}}{r^*}\right)^2 cv^2 r^* t,$$
for large $t$, from (1) and (3). From Result 2.2, $\mathbb{E}\{M(t)\} = \bar{r}t$, for $t \geq 0$; Result 2.3 then follows by rearranging terms. To see how (5) yields (6), notice that we generate $M$ by thinning $N$ with constant probability $(1 - \bar{r}/r^*)$; thus, conditional on $N(t)$, $M(t) \sim \text{Bin}(N(t), \bar{r}/r^*)$, for $t \geq 0$. Therefore, $\mathbb{E}[M(t) | N(t)] = (\bar{r}/r^*)N(t)$ and $\text{Var}[M(t) | N(t)] = (\bar{r}/r^*)(1 - \bar{r}/r^*)N(t)$.

Much of the thinning literature extends Rényi’s result (1956) on Bernoulli thinning of point processes that converge to Poisson; see Gnedenko and Kovalenko (1989), Miller (1979), Rolski and Szekli (1991), and references therein. Shanthikumar (1986) utilizes thinning of simulated Poisson arrivals in estimating renewal functions, while Ogata (1981) extends Lewis and Shedler (1979) to simulate multivariate point processes. Manor (1998) provides complementary analysis to Result 2.3 here for the second moment of the resulting nonstationary interevent times when constant Bernoulli thinning is performed.

2.4. Are inversion and thinning equivalent?

Although the two methods proposed in this section can achieve the same arrival rate $r(t)$, they are not equivalent in general, as the following example shows.

**Example 2.1.** Let $R(t) = t/2$ and $N$ be stationary with rate 1 and interrenewal variance $cv^2$.

- **Inversion:** $V_n = R^{-1}(S_n) = 2S_n$.
  
  $\mathbb{E}\{V_n\} = \mathbb{E}\{2S_n\}$
  
  $= 2\mathbb{E}\left\{X_1 + \sum_{k=2}^{n} X_k\right\}$
  
  $= 2[\mathbb{E}\{X_1\} + (n - 1)\mathbb{E}\{X_2\}]$
  
  $= 2n + cv^2 - 1$.

- **Thinning:** Since $r(t) = 1/2$, we can generate $\{T_n\}$ by thinning $N$ with probability $1/2$. Therefore, $T_n = S_{B_n}$, where $B_n \sim \text{NegBin}(n, 1/2)$. Thus,
  
  $\mathbb{E}\{T_n\} = \mathbb{E}\{\mathbb{E}[T_n | B_n]\}$
  
  $= \mathbb{E}\left\{X_1 + \sum_{k=2}^{B_n} X_k\right\}$
  
  $= \mathbb{E}\{X_1\} + \mathbb{E}\{X_2\}\mathbb{E}\{B_n - 1\}$,
by Wald’s Lemma, and therefore,
\[ \mathbb{E}\{T_n\} = 2n + \frac{cv^2 - 1}{2}. \]

Even in this simple example, we see that \( \mathbb{E}\{V_n\} = \mathbb{E}\{T_n\} \) only in the case that \( cv^2 = 1 \). Of course they are equivalent when \( N \) is Poisson.

3. When the stationary renewal process is phase-type

Although we can use any renewal process as the input for either method described in the previous section, we prefer to use a process that is easy to initialize, easy to simulate, and has an interrenewal distribution easily fit to \( \tau \) and \( cv^2 \). In this section we discuss one such class of renewal processes known as the phase-type, or Ph, processes. We describe a representation for the Ph process, provide some analysis of the resulting nonstationary process from the thinning method, and detail benefits of using specific Ph processes as the input to both inversion and thinning.

3.1. The Ph Process

The Ph process (Neuts, 1981) has interrenewal times that describe the time it takes an underlying continuous-time Markov chain, or CTMC, to reach a single absorbing phase from a finite number \( m_T < \infty \) of transient phases. The most common representation for the Ph process is attributed to Lucantoni (1991). We utilize a related representation here that characterizes the Ph interrenewal distribution by transitions within the embedded discrete-time Markov chain, or DTMC, along with a vector of time-dependent rate functions (one for each transient phase) and a vector of the initial transient phase probabilities. Notice that along with the rate functions, the phase transition and initial phase probabilities may vary with time. This representation is used by Nelson and Taaffe (2004) and recounted here.

We let \( A(t) \) denote the time-dependent, one-step transition probability matrix of the embedded DTMC:
\[
A(t) = \begin{pmatrix}
A_1(t) & \tilde{A}_2(t) \\
\tilde{\alpha}(t) & 0
\end{pmatrix}.
\]
The $m_T \times m_T$ matrix $A_1(t)$ represents the time-dependent one-step transition probabilities between the $m_T$ transient phases, while the $m_T \times 1$ vector $\vec{A}_2(t)$ represents the time-dependent one-step transition probabilities from the transient phases to phase $m_T + 1$, the instantaneous absorbing phase representing an arrival. The $m_T \times 1$ vector $\vec{\alpha}(t)$ is the time-dependent initial probability vector for the next interarrival time.

We define the $m_T \times 1$ vector $\vec{\lambda}(t)$, whose $j^{th}$ argument is $\lambda_j(t)$, the time-dependent, integrable non-negative transition rate function corresponding to phase $j$, for $j = 1, 2, \ldots, m_T$. We use the convention $\lambda_{m_T+1}(t) = \infty$, for all $t$, corresponding to an instantaneous sojourn time in that phase.

The Ph arrival process is $(N(t), J(t))$, where $N(t)$ is the number of arrivals (i.e., renewals) by time $t$, and $J(t)$ is the current phase of the next arrival. Notice that $N(t)$ increases by 1 when the chain hits phase $m_T + 1$. The Nelson and Taaffe characterization for the Ph process $(N(t), J(t))$ is the pair $(A(t), \vec{\lambda}(t))$.

### 3.2. Adjustments to the algorithms for a stationary Ph process

We describe adjustments made to the algorithms presented in Sections 2.2 and 2.3 when the stationary renewal process is Ph with representation $(A, \vec{\lambda})$. Since the input Ph process is not time-dependent, we have dropped the ‘$(t)$’ from its Ph representation.

The distribution $G$ of interrenewal times is given by

$$G(t) = \Pr\{X_n \leq t\} = 1 - \vec{\alpha}^\top \exp\{L(A_1 - I)t\} \vec{e},$$

for $n \geq 2$, where $L$ is a diagonal matrix with nonzero elements $\lambda_j$, for $j = 1, 2, \ldots, m_T$, $I$ is the identity matrix, and $\vec{e}$ is a column vector with all coordinates equal to 1 (Kulkarni [1995]). From this we derive the equilibrium distribution

$$G_e(t) = \Pr\{X_1 \leq t\} = 1 - \vec{\pi}^\top \exp\{L(A_1 - I)t\} \vec{e},$$

where $\vec{\pi}$ is the stationary $m_T \times 1$ vector for the phase process $J(t)$; that is, $\vec{\pi}$ solves $\vec{\pi}^\top L(A_1 + \vec{A}_2 \vec{\alpha}) = \vec{\pi}^\top L$, such that $\vec{\pi}^\top \vec{e} = 1$. Notice that the parameters of the Ph arrival process $(N(t), J(t))$ must satisfy

$$\tau^{-1} = \vec{\pi}^\top L\vec{A}_2,$$

(7)
where $\tau = 1$ in the inversion algorithm or $\tau = (r^*)^{-1}$ in the thinning algorithm.

Many techniques exist for fitting a Ph process to a tuple of its first two interrenewal moments. These techniques involve establishing the parameters in a particular Ph subclass that yield the given moments. Two-moment methods typically consist of specifying an order-two hyperexponential or Coxian when $cv^2 > 1$ and specifying an Erlang (or mixture of Erlangs) when $cv^2 < 1$; e.g., see Marie (1980), Sauer and Chandy (1975), and Whitt (1981). Notice that the Poisson process is a special case of Ph where interrenewal times are exponential and $cv^2 = 1$.

The particular choices we recommend for the two-moment fit are a hyperexponential of order two with balanced means, or $h_2b$, when $cv^2 > 1$, and a mixture of two Erlangs of consecutive order and common rate, or MECon, when $cv^2 < 1$. Formulas for setting the parameters of these processes, given $\tau$ and $cv^2$, are presented in Appendix B. Benefits to using these particular Ph choices are two-fold. First, they provide coverage over the range of all $cv^2 > 0$. Second, simulating these choices is efficient in that at each renewal, the generation of the next interrenewal time can be done in a single step (i.e., the generated interrenewal time is either an exponential or an Erlang) rather than having to simulate each phase transition individually.

3.3. Properties of the nonstationary Ph process

If the base process for inversion is Ph, then the resulting NSNP is not necessarily Ph. However, when the renewal base process is a stationary Ph with representation $(A, \bar{\lambda})$, the process $(M(t), J(t))$, generated by the thinning method, is a nonstationary Ph process with representation $(C(t), \bar{\lambda})$, such that

$$C(t) = \begin{pmatrix} A_1 + \left(1 - \frac{r(t)}{r^*}\right) \bar{A}_2 \bar{\alpha}^\top \left(\frac{r(t)}{r^*}\right) \bar{A}_2 \\ \bar{\alpha}^\top \end{pmatrix}.$$ 

Further, we can show that for the counting process $M$,

$$\text{Var}\{M(t)\} = R(t) + \frac{2}{r^*} \int_0^t r(u) \int_0^u r(z) \left(\bar{\alpha}^\top \exp\left\{L \left(A_1 - I + \bar{A}_2 \bar{\alpha}^\top\right)\right\} (u - z)\right) L \bar{A}_2) \, dz \, du - R^2(t).$$

Let $\nu(t) \equiv \text{Var}\{M(t)\}/R(t)$, for $t \geq 0$. A useful consequence of (8) is that we can identify bounds for $\nu(t)$ when the renewal base process to thin is a specific stationary Ph process.
**Example 3.1.** Stationary Ph is balanced hyperexponential of order two, for unspecified \( r(t) \geq 0 \).

If \( X \sim h_2b(\lambda, \alpha) \), then \( \pi = (1/2, 1/2) \top \), and
\[
\text{cv}^2(X) = \frac{1 - 2\alpha + 2\alpha^2}{2\alpha(1 - \alpha)} > 1.
\]

To satisfy (7), \( \lambda = 2\alpha r^* \). The Ph representation for the fitted \( h_2b \) is then
\[
\begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\alpha & 1 - \alpha & 0
\end{pmatrix}, \quad \bar{\lambda} = (2\alpha r^*, 2(1 - \alpha)r^*) \top.
\]

Plugging these into (8), we find
\[
\text{Var}\{M(t)\} = R(t) + 4\alpha(1 - 2\alpha)^2 \int_0^t r(u) \int_0^u r(z) e^{-4\alpha(1-\alpha)r^*(u-z)} dz \ du.
\]

For unspecified \( r(t) \), we cannot find a closed-form expression for \( \text{Var}\{M(t)\} \), but by noticing that \( r(t) \leq r^* \), for all \( t \geq 0 \), we can provide an upper-bound for \( \nu(t) \):
\[
\text{Var}\{M(t)\} \leq R(t) + 4\alpha(1 - 2\alpha)^2 \int_0^t r(u) \int_0^u r(z) e^{-4\alpha(1-\alpha)r^*(u-z)} dz \ du
\leq R(t) + 4\alpha(1 - 2\alpha)^2 \int_0^t r(u) e^{-4\alpha(1-\alpha)r^*u} du
= \frac{1 - 2\alpha + 2\alpha^2}{2\alpha(1 - \alpha)} R(t) - \frac{(1 - 2\alpha)^2}{1 - \alpha} \int_0^t r(u) e^{-4\alpha(1-\alpha)r^*u} du
\leq \frac{1 - 2\alpha + 2\alpha^2}{2\alpha(1 - \alpha)} R(t).
\]

Therefore, \( R(t) \leq \text{Var}\{M(t)\} \leq [(1 - 2\alpha + 2\alpha^2)/(2\alpha(1 - \alpha))] R(t) \), and \( \nu(t) \in [1, \text{cv}^2(X)] \).

**Example 3.2.** Stationary Ph is mixture of \( \text{Erlang}_2(\lambda) \) and \( \text{Erlang}_1(\lambda) \) (a MECon), for unspecified \( r(t) \geq 0 \).

If \( X \sim MECon_{2,1}(\lambda, \alpha) \), then
\[
\pi = \begin{pmatrix}
1 - \alpha & 1 - \alpha & \alpha \\
2 - \alpha & 2 - \alpha & 2 - \alpha
\end{pmatrix} \top,
\]
and
\[
\text{cv}^2(X) = \frac{2 - \alpha^2}{(2 - \alpha)^2} \in [1/2, 1].
\]
To satisfy (7), $\lambda = (2 - \alpha)r^*$. The Ph representation for the fitted $MECon_{2,1}$ is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 - \alpha & 0 & \alpha & 0 \end{pmatrix}, \quad \bar{\lambda} = ((2 - \alpha)r^*, (2 - \alpha)r^*, (2 - \alpha)r^*)^\top.$$

We plug these into (8) to derive a lower bound for $\nu(t)$:

$$\text{Var}\{M(t)\} = R(t) - 2(1 - \alpha)^2 \int_0^t r(u) \int_0^u r(z) e^{-r^*(2-\alpha)^2(u-z)} dz \, du$$

$$\geq R(t) - 2(1 - \alpha)^2 \int_0^t r(u) e^{-r^*(2-\alpha)^2u} \int_0^u r^* e^{r^*(2-\alpha)^2z} dz \, du$$

$$= R(t) \left( \frac{2 - \alpha^2}{(2 - \alpha)^2} \right) + 2 \left( \frac{1 - \alpha}{2 - \alpha} \right)^2 \int_0^t r(u) e^{-r^*(2-\alpha)^2u} du$$

$$\geq R(t) \left( \frac{2 - \alpha^2}{(2 - \alpha)^2} \right).$$

Therefore, $R(t) \geq \text{Var}\{M(t)\} \geq [(2 - \alpha^2)/(2 - \alpha)^2]R(t)$, and $\nu(t) \in [cv^2(X), 1]$.

4. Fitting a NSNP process

The minimum information required to define a NSNP process of the form described in this paper is a desired rate function $r(t)$ and a $cv^2$ for the renewal base process; the mean time between renewals $\tau$ for the base process is either 1 for inversion or $1/r^* = 1/\max_{t \geq 0} r(t)$ for thinning. See Appendix B for translating these properties into specific Ph renewal processes.

In this section we propose techniques for estimating $r(t)$ and $cv^2$ for both inversion and thinning when we have data on the actual arrival process, and illustrate them on a real data set. This is preliminary work, and we believe improvements should be possible. We leave open the important question of statistically verifying that the observed data are well represented by transforming a stationary renewal process via inversion or thinning.

4.1. Fitting to data using the inversion method

Several techniques, both parametric and non-parametric, have been explored for estimating the rate function (or integrated rate function) of a NSPP given a set of $n$ realizations of the observed arrival process. In parametric estimation, a form for the rate function is assumed, and maximum likelihood or other techniques are used to solve for the parameters of the specified rate function;
see Kuhl et al. (2006), Massey et al. (1996) and related papers. In non-parametric estimation, the
data are used to create a piece-wise constant rate function, either by defining a single interval for
each data point (Leemis, 1991) or by setting the respective interval lengths exogenously (Harrod
and Kelton, 2006). Henderson (2003) provides an analysis of the asymptotic behavior of non-
parametric estimators of the latter form, both when the interval length $\delta$ is constant in $n$ and when
$\delta$ decreases with $n$. The non-parametric techniques apply to NSNP data as well, so we adopt them
here. Therefore, our primary task is estimating $c v^2$.

We assume we have $n > 1$ i.i.d. realizations on time interval $[0, T_E]$, for known constant $T_E > 0$;
each realization represents the sequence of times at which arrivals occur. We let $V_{kj}$ denote the time
of the $j$th observed arrival in the $k$th realization (with $V_{k0} = 0$), and define $I_k(t)$ to be the number
of arrivals that have occurred in the $k$th realization by time $t$; that is, $I_k(t) = \max\{j \geq 0 : V_{kj} \leq t\}$,
for $k = 1, 2, \ldots, n, t \in [0, T_E]$. We further define

$$\hat{R}(t) = \frac{1}{n} \sum_{k=1}^{n} I_k(t),$$

and

$$\hat{V}(t) = \frac{1}{n-1} \left[ \left( \sum_{k=1}^{n} I_k(t)^2 \right) - n\hat{R}(t)^2 \right].$$

Thus, $\hat{R}(t)$ and $\hat{V}(t)$ are the sample mean and variance for the number of arrivals that have occurred
on or before time $t \in [0, T_E]$.

If we assume that the observed data were created by a process of the form $V_n = R^{-1}(S_n)$,
then $\text{Var}\{I(t)\}/R(t) \approx cv^2$, for large $t$, by Result 2.1. This suggests that we select a set of times
$0 = t_0 < t_1 < t_2 < \cdots < t_m = T_E$ and estimate $cv^2$ by fitting the line $\hat{V}(t) = cv^2\hat{R}(t) + \varepsilon_t$ to the
set of sample pairs $(\hat{R}(t_i), \hat{V}(t_i))$, for $i = 1, 2, \ldots, m$.

While we might be tempted to use ordinary least-squares regression, the residuals $\varepsilon_t$ are neither
independent nor have equal variance. To account for the latter concern, we perform weighted least-
squares regression, weighting the $i$th squared error in the sum by an amount inversely proportional
to the variance of the dependent variable $\hat{V}(t_i)$, for $i = 1, 2, \ldots, m$. Notice that $\text{Var}\{\hat{V}(t)\}$ depends
on moments of $I_k(t)$ higher than the second, but if we treat $\hat{V}(t)$ as approximately $\chi^2$ (that is,
\( \hat{V}(t) \sim \text{Var}\{I(t)\} \chi^2_{n-1}/(n-1) \), then

\[
\text{Var}\{\hat{V}(t)\} = \frac{\text{Var}\{I(t)\}^2}{(n-1)^2} - 2(n-1) = \frac{2(\text{Var}\{I(t)\})^2}{n-1} \approx \frac{2(c_v^2)^2}{n-1} - R(t)^2;
\]  

(11)

the ‘\( \approx \)’ in (11) indicates that this first relationship holds exactly if \( \hat{V}(t) \) is \( \chi^2 \) with these parameters. Thus, \( \text{Var}\{\hat{V}(t)\} \propto R(t)^2 \); therefore, we weight each residual by \( 1/R(t)^2 \), using \( 1/\hat{R}(t)^2 \) as a substitute. Our estimator is

\[
\hat{c}_v^2 = \arg \min_{c_v^2} \left\{ \sum_{i=1}^m \frac{1}{R(t_i)^2} \left[ \hat{V}(t_i) - c_v^2 \hat{R}(t_i) \right]^2 \right\}.
\]  

(12)

A closed-form solution exists for \( \hat{c}_v^2 \) in (12), namely

\[
\hat{c}_v^2 = m^{-1} \sum_{i=1}^m \left[ \hat{V}(t_i)/\hat{R}(t_i) \right].
\]  

This is intuitive, as the fitted \( c_v^2 \) is equal to the average value of \( \hat{V}(t)/\hat{R}(t) \) across the chosen sample of times.

We can extend the idea for estimating \( c_v^2 \) from a set of sample points to minimize the cumulative weighted residual along the entire \([0, T_E]\) interval. Here, the estimator is

\[
\hat{c}_v^2 = \arg \min_{c_v^2} \left\{ \int_0^{T_E} \frac{1}{R(u)^2} \left[ \hat{V}(u) - c_v^2 \hat{R}(u) \right]^2 \, du \right\},
\]  

(13)

with closed-form solution

\[
\hat{c}_v^2 = \frac{1}{T_E} \int_0^{T_E} \left( \frac{\hat{V}(u)}{R(u)} \right) \, du.
\]

As expected, \( \hat{c}_v^2 \) in (13) is the average value of \( \hat{V}(t)/\hat{R}(t) \) across \([0, T_E]\).

Using \( \hat{c}_v^2 \) in (13) rather than \( \hat{c}_v^2 \) in (12) eliminates the dependence of the fitting technique on the selection of \( \{t_i, i = 1, 2, \ldots, m\} \). However, this advantage disappears when only counts over time intervals are available (which frequently occurs in the literature; e.g., see Jongbloed and Koole (2001)), since we must treat \( \hat{R}(t) \) and \( \hat{V}(t) \) as constant during the intervals. With these ideas in mind, we use \( \hat{c}_v^2 \) in (13) as the estimator for \( c_v^2 \) in the presence of individual arrival time data; when only counts over time intervals are available, we use \( \hat{c}_v^2 \) in (12). We leave an investigation comparing the two estimators for future research.

To assess the quality of fit we can employ standard regression measures, such as \( R^2 \) or confidence intervals on the parameter estimate, with the caution that these will be approximate since the residuals are not independent and may well be non-normal. Future improvements to the fitting
technique may incorporate models for autocorrelation within the residuals, utilizing ideas such as those in Channouf et al. (2007) and references therein.

Notice that \( \hat{R}(t) \), as defined in (9), could be an estimate of the integrated rate function \( R(t) \), for \( t \in [0, T_E] \), by Result 2.1. However, \( \hat{R}(t) \) is a step function that increases in value only at the observed arrival times (i.e., when \( t \in \{V_{kj}\} \)). Thus, the inversion algorithm may return interarrival times of size zero for the NSNP using \( \hat{R}(t) \) as the estimator for \( R(t) \). Instead, we suggest an estimator that utilizes linear interpolation, similar to one proposed by Leemis (1991).

Let \( A_k \) denote the number of observed arrivals on \([0, T_E]\) in the \( k \)th realization, with \( A_T = \sum_{k=1}^{n} A_k \). Let \( T'_q \) denote the \( q \)th smallest observed arrival time \( V_{kj} \) across all \( n \) realizations (we assume no ties), \( q = 1, 2, \ldots, A_T \), with \( T'_0 = 0 \) and \( T'_{A_T+1} = T_E \). We suggest using the linear interpolation

\[
\tilde{R}(t) = \hat{R}(T'_{q-1}) + \frac{t - T'_{q-1}}{T'_q - T'_{q-1}} \left( \hat{R}(T'_q) - \hat{R}(T'_{q-1}) \right),
\]

for \( t \in (T'_{q-1}, T'_q), \ q = 1, 2, \ldots, A_T + 1 \).

Therefore, given a set of \( n \) realizations of the observed process, we have defined a technique that provides an estimate \( \tilde{R}(t) \) for the integrated rate function and an estimate \( \hat{cv}^2 \) in (12) (or \( \hat{cv}^2 \) in (13), when possible) for \( cv^2 \) of the renewal base process for the inversion method. Given interrenewal mean \( \tau = 1 \), we specify a Ph process, as described in Section 3.2 and Appendix B, to match \( \tau \) and \( cv^2 \).

**4.2. Example: Specifying the renewal process for inversion**

We apply the fitting technique for the inversion method to a set of internet traffic data describing arrivals to an e-mail server. We show here that the observed arrival process is more variable than Poisson, and utilize our fitting technique to specify a Ph renewal base process with \( cv^2 > 1 \) to simulate this arrival process via the inversion method.

The data consists of the timestamps for connections to the iems.northwestern.edu department server’s inbound mail port between 8:00AM and 8:00PM (local time) on Tuesday, Wednesday and
Thursday in three consecutive weeks. Thus, we have \( n = 9 \) realizations, with \( T_E = 720 \) minutes. We calculate \( I_k(t) \), for \( t \in [0, T_E] \), and define \( \tilde{R}(t) \) and \( \tilde{V}(t) \), as in (9) and (10), respectively, and \( \tilde{R}(t) \), as in (14).

Since we have individual arrival times, we use \( \hat{c}v^2 \), as in (13); we find \( \hat{c}v^2 = 124.3 \), with \( R^2 \) from the regression of 0.98 indicating a very good fit. Thus, we specify a renewal base process with \( \tau = 1 \) and \( cv^2 = 124.3 \), and use \( \tilde{R}(t) \) as an estimate for the integrated rate function \( R(t) \), for \( t \in [0, T_E] \). Since \( cv^2 > 1 \), the base process is an \( h_2b \) with rate \( \lambda = 1.99 \) connections per minute and mixing probability \( \alpha = 0.996 \) (see Appendix B).

To gain a sense of the difference between our fitted NSNP process and a NSPP with the same rate, Figure 1 shows sample paths generated from inversion with \( cv^2 = 124.3 \) (i.e., the fitted model) and \( cv^2 = 1 \), a NSPP. The increased variability is immediately apparent, and the consequences of ignoring it are severe. To illustrate this, we simulated 2000 replications of NSNP/M/1 and \( M(t)/M/1 \) queues with integrated rate function \( \tilde{R}(t) \) but \( cv^2 = 124.3 \) and 1, respectively, and service rate of 6 connections per minute in both cases. The top plot in Figure 2 is an estimate of the time-dependent mean number of entities at the node for both cases. In the model with \( cv^2 = 124.3 \), the mean number is an order of magnitude larger than the corresponding values in the model with \( cv^2 = 1 \). The effect of misspecifying \( cv^2 \) is even more dramatic when we examine the time-dependent variance of the number of entities at the node, seen in the bottom graph in Figure 2. Standard errors in both plots are roughly 2% of the estimated values.

4.3. Fitting to data using the thinning method

In this section we propose a technique for specifying the rate \( r^* \) and the \( cv^2 \) of a (potentially) non-Poisson stationary renewal base process for the thinning method, assuming the same sort of data as in Section 4.1 are available. We again calculate \( I_k(t) \) and define \( \tilde{R}(t) \) and \( \tilde{V}(t) \) as in (9) and (10), respectively.

We select a set of times \( 0 = t_0 < t_1 < t_2 < \cdots < t_m = T_E \) such that within each interval we believe the arrival rate is approximately constant. We define \( \hat{r}_i = [\tilde{R}(t_i) - \tilde{R}(t_{i-1})]/(t_i - t_{i-1}) \), for \( i = 1, 2, \ldots, m \). This provides an estimate \( \hat{r}(t) \) for the rate function \( r(t) \), where \( \hat{r}(t) = \hat{r}_i \), for
Figure 1: Ten sample paths of a NSNP process from inversion for a base process with $\text{cv}^2 = 124.3$ (top) and NSPP ($\text{cv}^2 = 1$, bottom).
Figure 2: Time-dependent mean (top) and variance (bottom) of number of customers at a NSNP/M/1 node with $c v^2 = 124.3$ (solid line) and $c v^2 = 1$ (dashed line) using inversion.
\( t \in (t_{i-1}, t_{i}], i = 1, 2, \ldots, m \). We set the arrival rate for the renewal process to be thinned at 
\( r^* = \max_{1 \leq i \leq m} \tilde{r}_i \).

We previously utilized Result 2.1 in our fitting technique for the inversion method; that result quantifies the asymptotic relationship between the variation of the resulting counting process \( I(t) \) and the \( \text{cv}^2 \) of its renewal base process. We have no analogous result for the thinning method for general rate function \( r(t) \). However, we can utilize Result 2.3 in specifying \( \text{cv}^2 \) if we assume that we are thinning the renewal process to achieve a target constant arrival rate. To exploit this, we find the largest time-window in \([0, T_E]\) during which the rate function \( \tilde{r}(t) \) is nearly constant. That is, we examine the set of interval arrival rates \( \{\tilde{r}_i, i = 1, 2, \ldots, m\} \), with the goal of selecting the largest subset of consecutive intervals across which the values of \( \tilde{r}_i \) are approximately equal.

Let \( P \subset \{1, 2, \ldots, m\} \) denote the selected subset of interval indices of maximum size across which the arrival rate is approximately constant. Define \( j' = |P| \), and \( i' = \min\{1 \leq i \leq m : i \in P\} \). Let \( M_k \) denote the number of observed arrival times in the \( k^{th} \) realization during these \( j' \) consecutive intervals; that is, \( M_k = I_k(t_{i'+j'-1}) - I_k(t_{i'-1}), \) for \( k = 1, 2, \ldots, n \). Let \( \bar{R} \) and \( \bar{V} \) denote the sample mean and variance, respectively, of \( M_k \); that is, \( \bar{R} = n^{-1} \sum_{k=1}^{n} M_k \), and \( \bar{V} = (n - 1)^{-1}(\sum_{k=1}^{n} M_k^2 - n\bar{R}^2) \). Finally, let \( \bar{r} \) denote the average arrival rate over these \( j' \) consecutive intervals; therefore, \( \bar{r} = \bar{R}/(t_{i'+j'-1} - t_{i'-1}) \).

We now assume that we want to thin the renewal process with arrival rate \( r^* \) to achieve that target arrival rate \( \bar{r} \). We replace \( \text{Var}\{M(t)\} \) with \( \bar{V} \) and \( \mathbb{E}\{M(t)\} \) with \( \bar{R} \) in (4), and solve for \( \text{cv}^2 \):

\[
\text{cv}^2 = \frac{r^*}{\bar{r}} \left[ \frac{\bar{V}}{\bar{R}} - \left(1 - \frac{\bar{r}}{r^*}\right) \right].
\] (15)

Given interrenewal mean \( \tau = (r^*)^{-1} \), we then specify a Ph process to match \( \tau \) and \( \text{cv}^2 \), as described in Section 3.2 and Appendix B.

Several potential limitations exist for this fitting technique. If the arrival rates \( \tilde{r}_i \) vary widely across intervals, it may prove impossible to find a subset of consecutive intervals with similar arrival rates of size greater than one. Another limitation is that the resulting NSNP process from thinning may not be a good fit for those arrivals occurring during intervals with indices not in \( P \); this would be particularly problematic if the time interval \([t_{i'-1}, t_{i'+j'-1}]\), representing those intervals with
indices in $\mathcal{P}$, is only a small fraction of the total interval $[0, T_E]$. Although this technique will do a good job approximating the arrival rate, by Result 2.2 the variation of the resulting NSNP process on the other intervals may not be representative of the actual arrival process.

4.4. Example: Specifying the renewal process for thinning

We apply the fitting technique for the thinning method described here on the example data set that we previously examined in Section 4.2 and generate analogous plots. Upon examining the data, we find that a 15-minute time window is reasonable for observing approximately constant arrival rates. Thus, we have $n = 9$ realizations, $m = 48$ sample points, $T_E = 720$, and interval length $\delta = 15$; we select sample points $t_i = i\delta$, for $i = 1, 2, \ldots, m$, and calculate $I_k(t_i)$ (for $k = 1, 2, \ldots, n$), $\hat{R}(t_i)$ and $\hat{V}(t_i)$. From these we derive interval arrival rates $\tilde{r}_i$, for $i = 1, 2, \ldots, m$, and define $\tilde{r}(t)$, for $t \in [0, T_E]$, as in Section 4.3. We set $r^* = 11.32$ connections per minute.

In finding $\mathcal{P} \subset \{1, 2, \ldots, m\}$, we notice that intervals 29 through 34 have similar arrival rates (approximately 4.6 connections per minute). Thus, $\mathcal{P} = \{29, 30, 31, 32, 33, 34\}$, $j' = 6$, and $i' = 29$. Setting these parameters yields $\bar{R} = 413.67$, $\bar{V} = 8,296.25$, and $\bar{r} = 4.596$; therefore, $c\bar{v}^2 = 47.96$, from (15), which is smaller than we estimated for the inversion method. Thus, we specify the renewal base process with arrival rate $r^* = 11.32$ and $c\bar{v}^2 = 47.96$, and use $\tilde{r}(t)$ as an estimate for the rate function $r(t)$, for $t \in [0, T_E]$. Since $c\bar{v}^2 > 1$, the renewal base process we specify is an $h_2b$ with rate $\lambda = 22.92$ and mixing probability $\alpha = 0.99$ (see Appendix B).

Figures 3 and 4 show that the sample paths and the impact on queue performance are dramatically different for the fitted NSNP process and a NSPP with the same arrival rate.

5. Conclusions

In this paper we have shown how to generate a NSNP arrival process by transforming a stationary renewal process. If the renewal base process is more or less variable than Poisson, then the resulting nonstationary process will be more or less variable than a NSPP. Further, we have shown that the $cv^2$ of the interrenewal distribution $G$ provides specific information on the variation of the NSNP process, either asymptotically (with inversion) or in the form of a bound (with thinning). Finally,
Figure 3: Ten sample paths of a NSNP process from thinning for base process with $cv^2 = 47.96$ (top) and NSPP ($cv^2 = 1$, bottom).
Figure 4: Time-dependent mean (top) and variance (bottom) of number of customers at NSNP/$M$/1 node with \( \text{cv}^2 = 47.96 \) (solid line) and \( \text{cv}^2 = 1 \) (dashed line) using thinning.
we have proposed a technique for specifying a renewal base process for each method when presented properties of, or data from, the nonstationary arrival process.

One direction for future research is to move beyond two-moment techniques and provide analogous methods for simulating when we have third moment (or higher) information that we desire in the resulting NSNP process. We may also pursue improvements to the fitting techniques presented here, particularly for the thinning method, as well as tools for validating the fit for both methods.

Acknowledgments

The authors thank Jeremy Staum and Mike Taaffe for helpful discussions, Johnathan Gaetz for providing the data set, and the anonymous referees for useful comments, including their suggestion of an alternative proof of Result 2.2. This work is supported by National Science Foundation Grant DMII-0521857.

References


Appendices

A. Proof of Result 2.2

Result 2.2: \( E\{M(t)\} = R(t), \text{ for all } t \geq 0.\)

Proof. Define the sequence \( \{B_n, n \geq 0\} \), such that \( B_0 = 0 \) and \( B_n = 1 \) if the \( n^{th} \) renewal is accepted as a nonstationary arrival, while \( B_n = 0 \) otherwise, for \( n \geq 1 \). Then, \( M(t) = \sum_{n=0}^{\infty} I_{[S_n \leq t]} B_n \), and

\[
E\{M(t)\} = E\left\{ \sum_{n=0}^{\infty} I_{[S_n \leq t]} B_n \right\} = \sum_{n=0}^{\infty} E\{I_{[S_n \leq t]} B_n\},
\]

where the interchange can be justified using the monotone convergence theorem. Therefore,

\[
E\{M(t)\} = \sum_{n=0}^{\infty} E\{I_{[S_n \leq t]} B_n\} = \sum_{n=0}^{\infty} E\left\{ E\left[ I_{[S_n \leq t]} B_n \mid S_n \right] \right\} = \sum_{n=0}^{\infty} E\left\{ I_{[S_n \leq t]} \frac{r(S_n)}{r^*} \right\},
\]

since, conditional on \( \{S_n, n \geq 1\} \), the \( B_n \)'s are independent, while \( E[B_n|S_n] = \Pr\{B_n = 1|S_n\} = \frac{r(S_n)}{r^*}, \) for all \( n \geq 1 \). Therefore,

\[
E\{M(t)\} = \sum_{n=0}^{\infty} E\left\{ I_{[S_n \leq t]} \frac{r(S_n)}{r^*} \right\} = \frac{1}{r^*} \sum_{n=0}^{\infty} E\{I_{[S_n \leq t]} r(S_n)\} = \frac{1}{r^*} E\left\{ \sum_{n=0}^{\infty} I_{[S_n \leq t]} r(S_n) \right\}
\]

27
(where this interchange can be similarly justified). For stationary renewal process $N$, we can show that

$$
\mathbb{E}\{M(t)\} = \frac{1}{r^*} \mathbb{E}\left\{ \sum_{n=0}^{\infty} I_{S_n \leq t} r(S_n) \right\}
= \frac{1}{r^*} \int_0^\infty I_{u \leq t} r(u) r^* du,
$$

by Proposition 9.1.14 in Činlar (1975), since the derivative of the renewal function is $r^*$ for all $t \geq 0$. Finally,

$$
\mathbb{E}\{M(t)\} = \frac{1}{r^*} \int_0^\infty I_{u \leq t} r(u) r^* du
= \int_0^\infty I_{u \leq t} r(u) du
= \int_0^t r(u) du
= R(t),
$$

for all $t \geq 0$.

\[\square\]

**B. Specifying a Ph distribution to match $\tau$ and $cv^2$**

We cite the follow techniques for choosing a Ph distribution to match $\tau$ and $cv^2$:

- If $cv^2 > 1$, then we specify an $h_2b$ (Sauer and Chandy, 1975), which implies that $X$ is exponentially distributed with mean $\lambda^{-1}$ with probability $\alpha$, or exponentially distributed with mean $\lambda_2^{-1}$ with probability $1-\alpha$. We say $h_2b$ has “balanced means” if $\alpha/\lambda = (1-\alpha)/\lambda_2$.

Thus, $h_2b$ has only two free parameters: $\alpha$ and $\lambda$. We back these out of the expressions for the mean and $cv^2$ of an $h_2b$ giving

$$
\alpha = \frac{1}{2} \left( 1 + \sqrt{\frac{cv^2 - 1}{cv^2 + 1}} \right), \quad \lambda = \frac{2\alpha}{\tau}.
$$

- If $cv^2 < 1$, then we use a MECon distribution (Tijms, 1994). First, we find integer $K$ such that $1/K \leq cv^2 < 1/(K-1)$, since $cv^2$ for an Erlang of order $K$ (denoted by $E_K(\lambda)$) is $1/K$.

Then $X$ is $E_{K-1}(\lambda)$ distributed with probability $\alpha$, or $E_K(\lambda)$ distributed with probability
Again, this leaves only two free parameters: the mixing probability $\alpha$ and the common rate $\lambda$. We back these out of the expressions for the mean and $cv^2$ of a MECon giving

$$\alpha = \frac{1}{1 + cv^2} \left( K \cdot cv^2 - \sqrt{K(1 + cv^2) - K^2cv^2} \right), \quad \lambda = \frac{K - \alpha}{\tau}.$$ 

The $h_2b$ and MECon are convenient choices, but any Ph (or other) distribution with the appropriate properties can be chosen, and might provide more desirable sample paths in some applications. For instance, the $h_2b$ will mix exponentials with very different means when $cv^2$ is large.