Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

USING REGENERATIVE SIMULATION TO CALIBRATE EXPONENTIAL APPROXIMATIONS TO RISK MEASURES OF HITTING TIMES TO RARELY VISITED SETS

Peter W. Glynn

Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, USA

Marvin K. Nakayama

Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102, USA

Bruno Tuffin

Inria
Campus de Beaulieu, 263 Avenue Général Leclerc
35042 Rennes, FRANCE

ABSTRACT

We develop simulation estimators of risk measures associated with the distribution of the hitting time to a rarely visited set of states of a regenerative process. In various settings, the distribution of the hitting time divided by its expectation converges weakly to an exponential as the rare set becomes rarer. This motivates approximating the hitting-time distribution by an exponential whose mean is the expected hitting time. As the mean is unknown, we estimate it via simulation. We then obtain estimators of a quantile and conditional tail expectation of the hitting time by computing these values for the exponential approximation calibrated with the estimated mean. Similarly, the distribution of the sum of lengths of cycles before the one hitting the rare set is often well-approximated by an exponential, and we analogously exploit this to estimate the two risk measures of the hitting time. Numerical results demonstrate the effectiveness of our estimators.

1 INTRODUCTION

Many stochastic processes possess a regenerative structure, and such a process "probabilistically restarts" at an increasing sequence of regeneration times; e.g., see Shedler (1993) and Kalashnikov (1994). Suppose the process rarely visits some set $\mathscr{A}$ of states, and we are interested in estimating (performance or risk) measures associated with the distribution of the hitting time T to $\mathscr{A}$. For example, in a stable GI/G/1 queue, the set $\mathscr{A}$ may correspond to a large number of customers in the system (e.g., buffer overflow), so $\mathscr{A}$ is rarely hit. In a highly reliable Markovian system consisting of a collection of components that fail and get repaired, system failures occur when certain combinations of components are down; in this case, the set $\mathscr{A}$ corresponds to the failed states, which are rarely visited.

Under various assumptions and asymptotic regimes, as visits to $\mathscr{A}$ become rarer, the distribution of the ratio of the hitting time T to $\mathscr{A}$ divided by its expectation μ converges weakly to an exponential; see Chapter 3 of Kalashnikov (1997). These results generalize a classical result by Rényi (Proposition 1.1.2 of Kalashnikov 1997), which establishes that as $p \to 0$, the product of p times the sum of a geometrically distributed number (with parameter p) of independent and identically distributed (i.i.d.) nonnegative random variables with finite mean converges weakly to an exponential. For our regenerative setting, the weak convergence motivates approximating the distribution F of T by an exponential with mean μ. As μ is unknown, we estimate it via simulation to calibrate the approximation. We then obtain estimators of the q-quantile (for a fixed $0 < q < 1$) and the conditional tail expectation (CTE) of T by computing these

978-1-5386-6572-5/18/$31.00 ©2018 IEEE

values for the calibrated exponential approximation, where the CTE is the conditional expectation of T given that it exceeds its q-quantile. (In Glynn et al. 2017 we provide a theoretical comparison of different estimators of the mean μ of T.)

We also extend the idea by exploiting similar exponential approximations to the distribution G of the sum S of the lengths of the cycles before the one in which the rare set $\mathscr{A}$ is hit. The exponential approximation depends on the unknown mean η of S, and we use simulation to estimate η, which provides us with an estimator for G. We further simulate to estimate the distribution H of the time V to hit $\mathscr{A}$ in the cycle in which $\mathscr{A}$ is visited. We can then express the hitting time T to $\mathscr{A}$ as the sum of S and V. The regenerative property guarantees that $S \sim G$ and $V \sim H$ are independent, so the distribution F of T is the convolution of G and H. Taking the convolution of our simulation estimators of G and H thus leads to an estimator of F, and we then compute the q-quantile and CTE of the estimated F. We present numerical results showing the effectiveness of our methods.

The rest of the paper unfolds as follows. Section 2 describes the problem mathematically and develops the notation. Section 3 explains the asymptotic regimes under which hitting the set $\mathscr{A}$ is a rare event, and discusses the weak convergence of T/μ and S/η to exponentials. Section 4 (resp., 5) exploits the resulting exponential approximation to T (resp., S) to develop our estimators of the q-quantile and the CTE of the hitting time. We give numerical results in Section 6, and concluding remarks appear in Section 7.

2 PROBLEM DESCRIPTION AND NOTATION

Consider a continuous-time stochastic process $X = [X(t) : t \geq 0]$ evolving on a state space $\mathscr{S}$. We assume that X is (classically) regenerative, with $0 = \Gamma_0 < \Gamma_1 < \Gamma_2 < \cdots$ as the sequence of regeneration times of X, so the process "probabilistically restarts" at each Γ_i; see p. 19 of Kalashnikov (1994). For example, an irreducible continuous-time Markov chain (CTMC) on a finite state space is regenerative, with successive hits to a fixed state forming a sequence of regeneration times.

For $\mathscr{A} \subset \mathscr{S}$ a subset of states (e.g., "failed states"), define $T = \inf\{t \geq 0 : X(t) \in \mathscr{A}\}$ as the *hitting time* (or first passage time) to $\mathscr{A}$. Let F be the cumulative distribution function (CDF) of T. For fixed $0 < q < 1$, our goal is to estimate the q-quantile

$$\xi = F^{-1}(q) \equiv \inf\{t : F(t) \geq q\} \tag{1}$$

of F and the *conditional tail expectation* (CTE)

$$\gamma = E[T \mid T > \xi]. \tag{2}$$

In the finance context, where F instead denotes the CDF of the loss in a portfolio of investments (e.g., Section 2.2 of McNeil et al. 2005 and Hong et al. 2014), a quantile is often called a *value-at-risk* (VaR), and the CTE is also known as the *expected shortfall* or the *conditional value-at-risk* (CVaR).

For $i \geq 1$, let $\tau_i = \Gamma_i - \Gamma_{i-1}$, and the process $[X(\Gamma_{i-1}+s) : 0 \leq s < \tau_i]$ is called the ith (regenerative) *cycle* of X, which has length τ_i. As X is regenerative, $(\tau_i, [X(\Gamma_{i-1}+s) : 0 \leq s < \tau_i])$, $i \geq 1$, is a sequence of i.i.d. pairs of cycle lengths and cycles. Let τ be a generic copy of τ_i. For $i \geq 1$, let $T_i = \inf\{t \geq 0 : X(\Gamma_{i-1}+t) \in \mathscr{A}\}$ be the time elapsing after Γ_{i-1} until the next hit to $\mathscr{A}$. For $x, y \in \Re$, let $x \wedge y = \min(x,y)$ and $x \vee y = \max(x,y)$. Let $\mathscr{I}(\cdot)$ be the indicator function, which takes value 1 (resp., 0) when its argument is true (resp., false). The regenerative property implies that $(\tau_i, T_i \wedge \tau_i, \mathscr{I}(T_i < \tau_i))$, $i \geq 1$, form an i.i.d. sequence of triples. Let $N(0) = 0$, and for $j \geq 1$, let $N(j) = \inf\{i > N(j-1) : T_i < \tau_i\}$ be the index i of the cycle corresponding to the jth cycle in which $\mathscr{A}$ is hit. For $j \geq 1$, let $M(j) = N(j) - N(j-1) - 1$, which is the number of cycles that do not hit $\mathscr{A}$ between cycles $N(j-1)$ and $N(j)$. We can then express the hitting time to $\mathscr{A}$ as

$$T = \sum_{i=1}^{M(1)} \tau_i + T_{M(1)+1}, \tag{3}$$

where $M(1)$ may depend on $\tau_1, \tau_2, \ldots, \tau_{M(1)}$ and $T_{M(1)+1}$.

As in Section 7.3 of Kalashnikov (1997), we next give a stochastically equivalent representation of T in (3) in terms of *independent* random variables. Let W be a random variable having CDF G_W with

$$G_W(x) = P(\tau \leq x \mid \tau < T), \tag{4}$$

so the CDF of W is the conditional CDF of τ given $\tau < T$. Let V be a random variable with CDF H, where

$$H(y) = P(T \leq y \mid T < \tau); \tag{5}$$

i.e., the distribution of V is the conditional distribution of T given $T < \tau$. Let M be a geometric random variable with $P(M = k) = p(1-p)^k$ for each $k \geq 0$, where

$$p = P(T < \tau). \tag{6}$$

Let $W_1, W_2, \ldots$ be i.i.d. copies of W, which are independent of V and M, where V and M are also independent. Define $S = \sum_{i=1}^{M} W_i$. Let G be the CDF of S, and let $\eta = E[S]$. The regenerative property of X ensures that

$$T \stackrel{\mathscr{D}}{=} S + V, \quad \text{with } S \sim G \text{ independent of } V \sim H, \tag{7}$$

where $\stackrel{\mathscr{D}}{=}$ denotes equality in distribution.

Define $\mu = E[T]$, which is the expected hitting time to the set $\mathscr{A}$. As is well known (e.g., see Goyal et al. 1992 and Glynn et al. 2017), the regenerative structure of X allows us to express μ as a ratio

$$\mu = \frac{E[T \wedge \tau]}{p} \equiv \frac{\zeta}{p} \tag{8}$$

with p from (6), and both the numerator and denominator in (8) are expectations of cycle-based quantities.

3 ASYMPTOTIC REGIMES

To develop estimators of the q-quantile $\xi = F^{-1}(q)$ of the CDF F of $T \stackrel{\mathscr{D}}{=} \sum_{i=1}^{M} W_i + V$ and the CTE γ, we consider some approximations that require p in (6) to be small. For a theoretical framework to accommodate this, we parameterize the problem by introducing a *rarity parameter* $\varepsilon > 0$ and examine the behavior of $F \equiv F_\varepsilon$ as $\varepsilon \to 0$, where we assume that

$$p \equiv p_\varepsilon \to 0 \quad \text{as } \varepsilon \to 0. \tag{9}$$

We now provide examples in which such parameterizations arise. In the first example the rarity comes from a receding set $\mathscr{A} \equiv \mathscr{A}_\varepsilon$ of failed states, with step-wise probability distributions independent of the parameterization. In the second example, it is the opposite: the transitions of the discrete-event system depend on the parameterization, but the set $\mathscr{A}$ of failed states does not.

Example 1 For a stable GI/G/1 queue with first-in, first-out discipline, let $X(t)$ denote the number of customers in the system at time $t \geq 0$, where the first customer arrives at time $t = 0$ to an empty system. The process X with the state space $\mathscr{S} = \{0, 1, 2, \ldots\}$ is regenerative with the beginnings of busy periods as regeneration times; e.g., see p. 16 of Kalashnikov (1994). We are interested in the distribution of the time when X first hits a high level $b_\varepsilon \equiv \lceil 1/\varepsilon \rceil$, where the interarrival- and service-time distributions do not vary with ε. Thus, we let the set $\mathscr{A} \equiv \mathscr{A}_\varepsilon = \{b_\varepsilon, b_\varepsilon + 1, b_\varepsilon + 2, \ldots\}$, and $T \equiv T_\varepsilon = \inf\{t \geq 0 : X(t) \in \mathscr{A}_\varepsilon\}$ is the first time that the queue length hits $b_\varepsilon - 1$. Theorem 1 of Sadowsky (1991) shows that (9) holds.

Example 2 Consider a *highly reliable Markovian system* (HRMS), as studied in Shahabuddin (1994), Nicola et al. (2001), and Rubino and Tuffin (2009). The system consists of a finite set of components, each of which has exponentially distributed lifetimes and repair times. The components may be of different types, which may require different classes of repairmen. The system is failed when certain combinations of components are failed. At time $t = 0$, all components are operational. We can model the evolution of the system as a CTMC $X = [X(t) : t \geq 0]$ on a finite state space $\mathscr{S}$, where each state in $\mathscr{S}$ specifies the number of failed components of each type, along with any necessary information about the queueing of failed components for each repairman class. Thus, X is regenerative, with hits to the state with all components operational as regeneration times. Assume that the system is highly reliable in the sense that the components' failure rates are much smaller than their repair rates. We then parameterize component failure rates as positive powers of the rarity parameter ε, and assume that repair rates and the set $\mathscr{A}$ of failed states do not change as $\varepsilon \to 0$. Under appropriate assumptions (see Shahabuddin 1994), we have that (9) holds. In this setting, $\mu_\varepsilon = E_\varepsilon[T_\varepsilon]$ is often called the *mean time to failure* (MTTF).

Because we now actually are considering a family of models indexed by the rarity parameter $\varepsilon > 0$, we should thus write $T = T_\varepsilon$, $W = W_\varepsilon$, $M = M_\varepsilon$, $V = V_\varepsilon$, $p = p_\varepsilon$, $F = F_\varepsilon$, $G = G_\varepsilon$, $G_W = G_{W,\varepsilon}$, $H = H_\varepsilon$, $P = P_\varepsilon$, $E = E_\varepsilon$, $\mu = \mu_\varepsilon$, etc. However, we often omit the subscript ε to simplify notation.

Under a variety of different sets of assumptions under which (9) is true, the scaled hitting time $T_\varepsilon/\mu_\varepsilon$ converges weakly to an exponential: for each $x \geq 0$,

$$P_\varepsilon(T_\varepsilon/\mu_\varepsilon \leq x) \to 1 - e^{-x} \quad \text{as} \quad \varepsilon \to 0, \tag{10}$$

where we recall that $\mu_\varepsilon = E_\varepsilon[T_\varepsilon]$. To handle settings as in Example 1 of the GI/G/1 queue with receding sets $\mathscr{A}_\varepsilon$, we can apply Theorem 3.4.1 of Kalashnikov (1997), which provides conditions to ensure (10) holds for a Harris-recurrent Markov chain, where the set $\mathscr{A}_\varepsilon$ varies with ε so that (9) remains valid. For the HRMS in Example 2, the system dynamics (transition probabilities and holding-time distributions) change with ε, which requires a "triangular array" formulation; Theorem 3.2.5 of Kalashnikov (1997) gives general conditions (not only for HRMSs) guaranteeing the validity of (10) when p_ε, $G_{W,\varepsilon}$, and H_ε are allowed to depend on ε (see also Nakayama and Shahabuddin 2004 for the specific HRMS context).

It is often the case that when (10) is true, we further have that the sum $S_\varepsilon = \sum_{i=1}^{M_\varepsilon} W_{\varepsilon,i}$ satisfies

$$P_\varepsilon(S_\varepsilon/\eta_\varepsilon \leq y) \to 1 - e^{-y} \quad \text{as} \quad \varepsilon \to 0 \tag{11}$$

for each $y \geq 0$, where we recall that $\eta_\varepsilon = E_\varepsilon[S_\varepsilon]$. For example, if we assume that $V = V_\varepsilon \equiv 0$ in (7), the conditions in Theorem 3.2.5 of Kalashnikov (1997) also ensure that (11) holds.

Throughout the rest of the paper, we apply non-simulation approximations based on the asymptotic results in (10) and (11), and then calibrate the approximations by estimating the unknown parameters μ_ε and η_ε via simulation. To distinguish the estimators and approximations for the different methods, we adopt the following notational convention. For an unknown quantity α, such as a CDF or parameter, we let $\widetilde{\alpha}$ denote a non-simulation approximation to α. Also, we let $\widehat{\alpha}$ denote an estimator of α constructed from simulation-generated data.

4 APPROXIMATING THE CDF F OF T BY AN EXPONENTIAL

The limiting result (10) suggests that for small ε (which we now drop to simplify the notation),

$$F(t) = P(T \leq t) = P(T/\mu \leq t/\mu) \approx 1 - e^{-t/\mu} \equiv \widetilde{F}_{\exp}(t) \tag{12}$$

for each $t \geq 0$, where we recall that $\mu = E[T]$. We will next use the approximation $\widetilde{F}_{\exp}$ to F to obtain approximations to the q-quantile $\xi = F^{-1}(q)$ in (1) for a fixed $0 < q < 1$ and the CTE γ in (2).

The exponential approximation in (12) motivates approximating ξ by

$$\widetilde{\xi}_{\exp} = \widetilde{F}_{\exp}^{-1}(q) = -\mu \ln(1-q). \tag{13}$$

For the CTE, if T has exactly CDF $\widetilde{F}_{\exp}$, then its CTE is

$$\widetilde{\gamma}_{\exp} = \widetilde{\xi}_{\exp} + \mu = \mu[1 - \ln(1-q)] \tag{14}$$

by the memoryless property of $\widetilde{F}_{\mathrm{exp}}$. But in (12), (13), and (14), the parameter μ is unknown, so we next calibrate our approximations by estimating μ via simulation.

4.1 Simulation Estimator of μ

As we saw in (8), the expected hitting time can be represented as a ratio $\mu = \zeta/p$. Because the asymptotic result (10) needs (9) to hold, the denominator $p = P(T < \tau)$ is a rare-event probability, so we will estimate it with *importance sampling* (IS); see Chapters V and VI of Asmussen and Glynn (2007) for an overview of this variance-reduction technique. But Shahabuddin et al. (1988) note that the numerator $\zeta = E[T \wedge \tau]$ can often be more efficiently handled by *crude simulation* (i.e., without IS), so we will independently estimate ζ and p. Shahabuddin et al. (1988) call this approach *measure-specific importance sampling*, which we implement as follows.

To estimate the numerator $\zeta = E[T \wedge \tau]$ in (8), we generate $T_i \wedge \tau_i$, $i = 1,2,\ldots,s$, as s i.i.d. copies of $T \wedge \tau$ sampled using crude simulation. Generating each $T_i \wedge \tau_i$ entails simulating a cycle until either it ends or $\mathscr{A}$ is hit, whichever occurs first. An estimator of ζ is then

$$\widehat{\zeta} \equiv \frac{1}{s}\sum_{i=1}^{s} T_i \wedge \tau_i. \tag{15}$$

Independently of the simulation runs employed to construct $\widehat{\zeta}$ in (15), we use IS to estimate the denominator $p = P(T < \tau)$ in (8) as follows. Applying a change of measure, write

$$p = E[\mathscr{I}(T < \tau)] = E'[\mathscr{I}(T < \tau)L], \tag{16}$$

where E' denotes expectation under the IS measure, and L is the corresponding likelihood ratio. The representation (16) motivates the following approach to estimate p. Let $(\mathscr{I}(T_i' < \tau_i'), T_i' \wedge \tau_i', L_i')$, $i = 1,2,\ldots,r$, be i.i.d. copies of $(\mathscr{I}(T < \tau), T \wedge \tau, L)$ generated via IS. We then estimate p by

$$\widehat{p} = \frac{1}{r}\sum_{i=1}^{r} \mathscr{I}(T_i' < \tau_i')L_i'. \tag{17}$$

We combine the two estimators $\widehat{\zeta}$ from (15) and $\widehat{p}$ from (17) to obtain

$$\widehat{\mu} = \widehat{\zeta}/\widehat{p} \tag{18}$$

as the simulation estimator of μ in (8).

4.2 Simulation Estimators of F, ξ, and γ

We next employ the simulation estimator $\widehat{\mu}$ from (18) to calibrate the approximate CDF $\widetilde{F}_{\mathrm{exp}}$ in (12) of T. Specifically, we replace the unknown μ in (12) by $\widehat{\mu}$ to obtain

$$\widehat{F}_{\mathrm{exp}}(t) = 1 - e^{-t/\widehat{\mu}} \tag{19}$$

as a simulation estimator of $F(t)$ for each $t \geq 0$.

Similarly, for the q-quantile $\xi = F^{-1}(q)$ of F, we approximated it in (13) by $\widetilde{\xi}_{\mathrm{exp}} = -\mu \ln(1-q)$. Replacing μ by its simulation estimator $\widehat{\mu}$ from (18) leads to the estimator

$$\widehat{\xi}_{\mathrm{exp}} = \widehat{F}_{\mathrm{exp}}^{-1}(q) = -\widehat{\mu}\ln(1-q) \tag{20}$$

of ξ. Finally, for the CTE, we replace $\widetilde{\xi}_{\mathrm{exp}}$ and μ in (14) by their simulation estimators $\widehat{\xi}_{\mathrm{exp}}$ and $\widehat{\mu}$ to obtain

$$\widehat{\gamma}_{\mathrm{exp}} = \widehat{\xi}_{\mathrm{exp}} + \widehat{\mu} = \widehat{\mu}[1 - \ln(1-q)] \tag{21}$$

as a simulation estimator of the CTE γ.

The simulation estimators $\widehat{F}_{\text{exp}}$ in (19) of F, $\widehat{\xi}_{\text{exp}}$ in (20) of ξ, and $\widehat{\gamma}_{\text{exp}}$ in (21) of γ are based on the approximation (12), which becomes more accurate as the rarity parameter $\varepsilon \to 0$ in (10). But for an actual physical system, we have a fixed value of $\varepsilon > 0$, so the exponential approximation $\widetilde{F}_{\text{exp}}(t)$ in (12) typically will not exactly equal $F(t)$. Thus, the simulation estimators $\widehat{F}_{\text{exp}}$, $\widehat{\xi}_{\text{exp}}$, and $\widehat{\gamma}_{\text{exp}}$ will often be biased, and the bias does not vanish as we increase the sample sizes s and r used to construct $\widehat{\mu}$ in (18).

By (20) and (21), the estimators $\widehat{\xi}_{\text{exp}}$ and $\widehat{\gamma}_{\text{exp}}$ are simply constant multiples of $\widehat{\mu}$. Because $\widehat{\mu}$ satisfies a central limit theorem (CLT; Shahabuddin 1994), $\widehat{\xi}_{\text{exp}}$ and $\widehat{\gamma}_{\text{exp}}$ also obey CLTs, with centering constants $\widetilde{\xi}_{\text{exp}}$ from (13) and $\widetilde{\gamma}_{\text{exp}}$ from (14), respectively, because of the bias.

5 APPROXIMATING THE CDF G OF S BY AN EXPONENTIAL

Recall that $T \stackrel{\mathscr{D}}{=} S+V$, where $S \sim G$ is independent of $V \sim H$ by (7). We next devise methods that separately estimate G and H to estimate the CDF F of T, its q-quantile ξ, and the CTE γ. To do this, we will approximate G by an exponential CDF, which is motivated by the asymptotic result in (11).

By (7), we can write the CDF F of T as a convolution

$$F(t) = G \star H(t) = \int H(t-x)\, dG(x), \tag{22}$$

where $\star$ denotes the convolution operator. The convergence in (11) suggests that for small ε (which is now dropped to simplify the notation), we have the parametric approximation

$$G(x) = P(S \leq x) = P(S/\eta \leq x/\eta) \approx 1 - e^{-x/\eta} \equiv \widetilde{G}_{\text{exp}}(x). \tag{23}$$

As $\eta = E[S]$ is unknown, we will use simulation to estimate it, as will be discussed in Section 5.1, along with the estimation of the CDF H of V.

Let $\xi = F^{-1}(q)$ be the q-quantile of F. We next obtain another representation for the CTE γ, which requires the regenerative property of X but not the limiting result (11) nor the approximation (23).

Theorem 1 If F is continuous at ξ, then the CTE satisfies

$$\gamma = \frac{1}{1-q}\left[\int x\,[1-H(\xi-x)]\, dG(x) + \int y\,[1-G(\xi-y)]\, dH(y)\right]. \tag{24}$$

Proof. First write the CTE as

$$\gamma = E[T \mid T > \xi] = \frac{E[T\,\mathscr{I}(T>\xi)]}{P(T>\xi)} = \frac{E[T\,\mathscr{I}(T>\xi)]}{1-q}, \tag{25}$$

where the last step follows from the continuity of F at ξ. Express $T \stackrel{\mathscr{D}}{=} S+V$ using (7), so the numerator in the right side of (25) then satisfies

$$\begin{aligned}
E[T\,\mathscr{I}(T>\xi)] &= E[(S+V)\,\mathscr{I}(S+V>\xi)] \\
&= E[S\,\mathscr{I}(S+V>\xi)] + E[V\,\mathscr{I}(S+V>\xi)] \\
&= E\big[E[S\,\mathscr{I}(V>\xi-S) \mid S]\big] + E\big[E[V\,\mathscr{I}(S>\xi-V) \mid V]\big] \\
&= E[S\,P(V>\xi-S \mid S)] + E[V\,P(S>\xi-V \mid V)] \\
&= E\big[S\,[1-H(\xi-S)]\big] + E\big[V\,[1-G(\xi-V)]\big]
\end{aligned}$$

because of the independence of S and V by (7). Thus, (24) immediately follows. □

5.1 Simulation Estimators of η, G, and H

As the approximate CDF $\widetilde{G}_{\text{exp}}$ in (23) and the CTE representation in (24) depend on the unknown η, G, H, and ξ, we next describe simulation estimators for them. We first explain how to handle $\eta = E[S] = E[\sum_{i=1}^{M} W_i]$. Writing $S = \sum_{i=1}^{\infty} W_i \mathscr{I}(M \geq i)$, we see that

$$\eta = \sum_{i=1}^{\infty} E[W_i \mathscr{I}(M \geq i)] = E[W] \sum_{i=1}^{\infty} E[\mathscr{I}(M \geq i)] = E[W] \sum_{i=1}^{\infty} P(M \geq i) = \frac{1-p}{p} \nu \tag{26}$$

by the independence of M and W_i, $i \geq 1$, which are i.i.d. with mean $\nu = E[W]$. Thus, (26) shows that η can be expressed as a function of expectations of cycle-based quantities, where we recall that $W \sim G_W$, with G_W the conditional CDF of τ given that $\tau < T$, as defined in (4). Hence, $\nu = E[\tau \mid \tau < T] = E[\tau \mathscr{I}(\tau < T)]/(1-p)$ by (6), and we estimate $E[\tau \mathscr{I}(\tau < T)]$ by $(1/s)\sum_{i=1}^{s} \tau_i \mathscr{I}(\tau_i < T_i)$, where $(\tau_i, \mathscr{I}(\tau_i < T_i))$, $i = 1, 2, \ldots, s$, are from the same i.i.d. crude-simulation data used to construct $\widehat{\zeta}$ in (15). Then we estimate ν via

$$\widehat{\nu} = \frac{1}{(1-\widehat{p})s} \sum_{i=1}^{s} \tau_i \mathscr{I}(\tau_i < T_i), \tag{27}$$

where $\widehat{p}$ is from (17). Therefore, we arrive at our estimator of η based on (26) as

$$\widehat{\eta} = \frac{1-\widehat{p}}{\widehat{p}} \widehat{\nu}. \tag{28}$$

Replacing η in (23) by its estimator $\widehat{\eta}$ from (28), we obtain a parametric estimator of $\widetilde{G}_{\text{exp}}(x)$ in (23) as

$$\widehat{G}_{\text{exp}}(x) = 1 - e^{-x/\widehat{\eta}}. \tag{29}$$

We next discuss how to use IS to estimate the CDF H of V, where we recall that H is the conditional CDF of T given that $T < \tau$, as in (5). Note that

$$\begin{aligned} H(y) = P(V \leq y) = P(\min(T,\tau) \leq y \mid T < \tau) &= \frac{P(\min(T,\tau) \leq y, T < \tau)}{p} \\ = \frac{E[\mathscr{I}(\min(T,\tau) \leq y, T < \tau)]}{p} &= \frac{E'[\mathscr{I}(\min(T,\tau) \leq y, T < \tau) L]}{p}, \end{aligned} \tag{30}$$

where, as in (16), E' is the expectation operator under IS, and L is the corresponding likelihood ratio. We previously obtained the IS estimator $\widehat{p}$ in (17) to handle the denominator p of (30). For the numerator, let $(\mathscr{I}(T_i' < \tau_i'), T_i' \wedge \tau_i', L_i')$, $i = 1, 2, \ldots, r$, be the same i.i.d. copies of $(\mathscr{I}(T < \tau), T \wedge \tau, L)$ obtained through IS that we employed to construct the estimator $\widehat{p}$. We then build a nonparametric estimator $\widehat{H}$ of H as

$$\widehat{H}(y) = \frac{1}{\widehat{p} \cdot r} \sum_{i=1}^{r} \mathscr{I}(T_i' \wedge \tau_i' \leq y, T_i' < \tau_i') L_i'. \tag{31}$$

While other nonparametric estimators of H can also be constructed, e.g., by interpolating $\widehat{H}$ in (31), or by using kernel methods, etc., we will only work with $\widehat{H}$ in (31).

5.2 Simulation Estimators of F, ξ, and γ

Now that (29) and (31) provide simulation estimators for the CDFs G and H of S and V, respectively, we use them to build an estimator for the CDF F of T. The representation of F in (22) as a convolution of G and H suggests estimating $F(t)$ by

$$\widehat{F}_\star(t) \equiv \widehat{G}_{\text{exp}} \star \widehat{H}(t) = \int \widehat{H}(t-x)\, d\widehat{G}_{\text{exp}}(x). \tag{32}$$

The following result works out an expression for $\widehat{F}_\star(t)$.

Proposition 1 The estimator $\widehat{F}_\star$ in (32) of the CDF F of T satisfies

$$\widehat{F}_\star(t) = 1 - \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(T_i' < \tau_i')\, L_i' e^{-(t-A_i')^+/\widehat{\eta}}, \tag{33}$$

where $A_i' = (T_i' \wedge \tau_i')$, $\widehat{\eta}$ is defined in (28), and $x^+ = \max(x,0)$.

Proof. Put (29) and (31) into (32) and use the fact that $\widehat{G}_{\exp}(x) = 0$ for $x < 0$ to get

$$\begin{aligned}
\widehat{F}_\star(t) &= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(T_i' < \tau_i')\, L_i' \int_0^t \mathscr{I}(A_i' \le t-x)\, d\widehat{G}_{\exp}(x) \\
&= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(T_i' < \tau_i')\, L_i' \int_0^{(t-A_i')^+} d\widehat{G}_{\exp}(x) \\
&= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(T_i' < \tau_i')\, L_i' \widehat{G}_{\exp}((t-A_i')^+),
\end{aligned}$$

which equals (33) by (29) and (17). □

The corresponding estimator of the q-quantile $\xi = F^{-1}(q)$ is

$$\widehat{\xi}_\star = \widehat{F}_\star^{-1}(q). \tag{34}$$

Computing $\widehat{\xi}_\star$ typically requires applying a numerical root-finding method, such as the bisection method or Newton's method, incurring additional computational cost.

We next give a simulation estimator for the CTE γ.

Proposition 2 When ξ, G, and H in (24) are replaced by their respective estimators $\widehat{\xi}_\star$ in (34), $\widehat{G}_{\exp}$ in (29), and $\widehat{H}$ in (31), the resulting CTE estimator is

$$\widehat{\gamma}_\star = \frac{1}{1-q}\left[\frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i' \mathscr{I}(T_i' < \tau_i')\left[(\widehat{\xi}_\star \vee A_i') + \widehat{\eta}\right] e^{-(\widehat{\xi}_\star - A_i')^+/\widehat{\eta}}\right]. \tag{35}$$

Proof. The first term in the outer square brackets in (24) is expressed in terms of $1-H(\cdot)$, and by (31),

$$\begin{aligned}
1-\widehat{H}(y) &= \frac{\widehat{p}}{\widehat{p}} - \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(A_i' \le y)\, \mathscr{I}(T_i' < \tau_i')\, L_i' \\
&= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} [1-\mathscr{I}(A_i' \le y)]\, \mathscr{I}(T_i' < \tau_i')\, L_i' = \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} \mathscr{I}(A_i' > y, T_i' < \tau_i') L_i'.
\end{aligned}$$

where the second equality holds by (17), and $A_i' = (T_i' \wedge \tau_i')$, as in Proposition 1. Thus, replacing ξ, G, and $1-H$ in the first term inside the outer square brackets of (24) by their respective estimators leads to

$$\begin{aligned}
\int x[1-\widehat{H}(\widehat{\xi}_\star - x)]\, d\widehat{G}_{\exp}(x) &= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i' \mathscr{I}(T_i' < \tau_i') \int_{x=0}^{\infty} x\, \mathscr{I}(A_i' > \widehat{\xi}_\star - x)\, d\widehat{G}_{\exp}(x) \\
&= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i' \mathscr{I}(T_i' < \tau_i') \int_{x=(\widehat{\xi}_\star - A_i')^+}^{\infty} x\, d\widehat{G}_{\exp}(x) \\
&= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i' \mathscr{I}(T_i' < \tau_i')\left[(\widehat{\xi}_\star - A_i')^+ + \widehat{\eta}\right] e^{-(\widehat{\xi}_\star - A_i')^+/\widehat{\eta}},
\end{aligned} \tag{36}$$

where we recall $\widehat{\eta}$ is defined in (28).

The estimator of the second term inside the outer square brackets of (24) is

$$\int y\,[1-\widehat{G}_{\exp}(\widehat{\xi}_\star - y)]\,d\widehat{H}(y) = \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i'\,\mathscr{I}(T_i' < \tau_i')\int_{y=0}^{\infty} y\,[1-\widehat{G}_{\exp}(\widehat{\xi}_\star - y)]\,\mathscr{I}(A_i' \in dy)$$
$$= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i'\,\mathscr{I}(T_i' < \tau_i')\,A_i'\,[1-\widehat{G}_{\exp}((\widehat{\xi}_\star - A_i')^+)]$$
$$= \frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i'\,\mathscr{I}(T_i' < \tau_i')\,A_i'\,e^{-(\widehat{\xi}_\star - A_i')^+/\widehat{\eta}}. \tag{37}$$

By replacing the two terms inside the outer square brackets of (24) by their estimators (36) and (37), we obtain the estimator of (24) as

$$\widehat{\gamma}_\star = \frac{1}{1-q}\left[\frac{1}{\widehat{p}\cdot r}\sum_{i=1}^{r} L_i'\,\mathscr{I}(T_i' < \tau_i')\left[(\widehat{\xi}_\star - A_i')^+ + \widehat{\eta} + A_i'\right] e^{-(\widehat{\xi}_\star - A_i')^+/\widehat{\eta}}\right],$$

which equals (35) because $(\widehat{\xi}_\star - A_i')^+ + A_i' = (\widehat{\xi}_\star - A_i' + A_i') \vee (0 + A_i') = \widehat{\xi}_\star \vee A_i'$. □

6 NUMERICAL RESULTS

We now present numerical results for our estimators, from Sections 4.2 and 5.2, of the CDF F of the hitting time T, its q-quantile ξ, and the CTE γ. Due to space limitations, we focus on only one simple model of an HRMS, as in Example 2 of Section 3. Specifically, the HRMS has three component types, with five components of each type. There are 15 repairmen, so failed components never queue for repair. The system is up if and only if at least two components of each type work, so the failure set $\mathscr{A}$ comprises states having at least four components down of one type. Each component has failure rate ε and repair rate 1. We consider two versions of the model: one with $\varepsilon = 10^{-2}$, and the other has $\varepsilon = 10^{-4}$.

To implement our methods, we need to specify the IS distribution employed to construct the estimators $\widehat{p}$ in (17) and $\widehat{H}$ in (31), which subsequently are used in the estimators of F, ξ, and γ. Rubino and Tuffin (2009) provide an overview of IS techniques designed to simulate HRMSs, where the basic idea is to increase the probability of failure transitions. In our experiments, we applied the IS approach known as zero-variance approximation (ZVA) of L'Ecuyer and Tuffin (2012). ZVA produces estimators of the MTTF μ in (8) having the (desirable) bounded relative error (BRE) property, with the (even better) vanishing relative error (VRE) holding under certain conditions.

For the methods in Section 4, computing $\widehat{F}_{\exp}$, $\widehat{\xi}_{\exp}$, and $\widehat{\gamma}_{\exp}$ in (19)–(21) requires simulation to estimate only $\mu = \zeta/p$. To do this, we simulated for each ε a total of 10^4 cycles, of which we allocated s to construct the estimator $\widehat{\zeta}$ in (15) using crude simulation, and we sampled the remaining $r = 10^4 - s$ cycles with IS to build the estimator $\widehat{p}$ in (17). Goyal et al. (1992) derive the optimal allocation of s and r for a fixed total budget to minimize the work-normalized asymptotic variance of the MTTF ratio estimator $\widehat{\mu}$ in (18). We ran pilot simulations with 10^3 cycles without IS (resp., with IS) for the numerator (resp., denominator) to estimate the parameters determining the optimal allocation. Adding constraints that at least 10% of the 10^4 second-stage cycles are for simulating the numerator and the same for the denominator, we got the optimal $r = 9000$ and $s = 1000$ for both values of ε in our example. To build $\widehat{F}_\star$ in (33) of Section 5, we used the same $s + r = 10^4$ simulated second-stage cycles to compute the estimators $\widehat{\nu}$ in (27) and $\widehat{H}$ in (31).

In the discussion below, we refer to $\widehat{F}_{\exp}$ (resp., $\widehat{F}_\star$) and its corresponding estimators of ξ and γ in (20) and (21) (resp., (34) and (35)) as *exponential* (resp., *convolution*) *estimators*. For comparison, we also built an *empirical* distribution $\acute{F}$ from $n = 10^4$ i.i.d. observations $\acute{T}_1, \acute{T}_2, \ldots, \acute{T}_n$, of T for $\varepsilon = 10^{-2}$ generated via crude simulation, where $\acute{F}(t) = (1/n)\sum_{i=1}^{n} \mathscr{I}(\acute{T}_i \le t)$. (We could not obtain an empirical distribution for $\varepsilon = 10^{-4}$, as we will explain later.) To define the corresponding estimators of ξ and γ, let

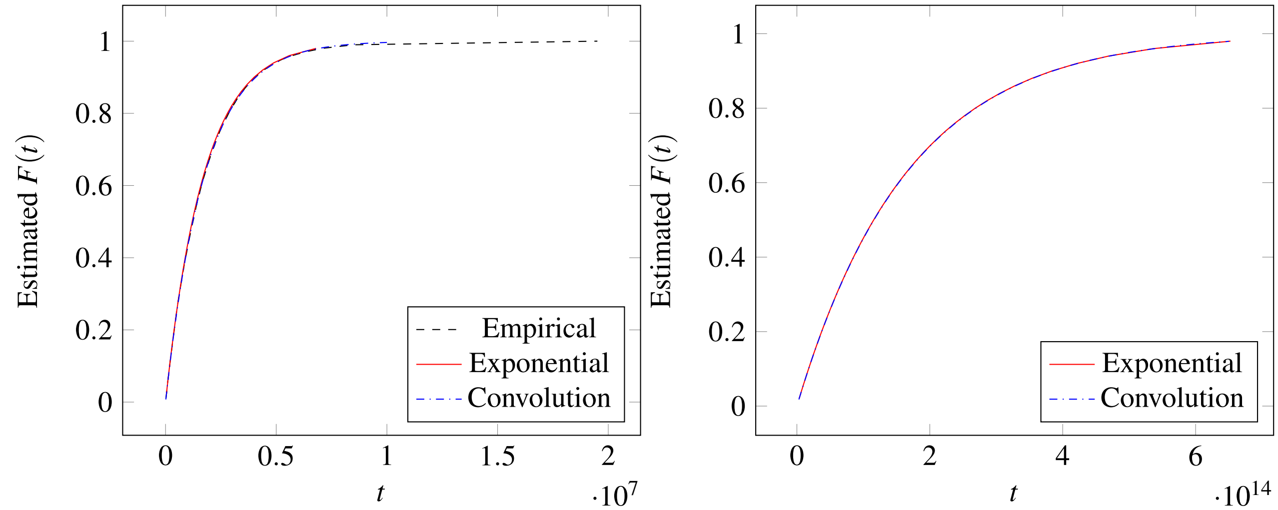

Figure 1: Plots of the empirical estimator $\acute{F}(t)$ of $F(t)$, the exponential estimator $\widehat{F}_{\text{exp}}(t)$ from (19), and the convolution estimator $\widehat{F}_{\star}(t)$ from (32) for $\varepsilon = 10^{-2}$ (left) and $\varepsilon = 10^{-4}$ (right).

$\acute{T}_{1:n} \le \acute{T}_{2:n} \le \cdots \le \acute{T}_{n:n}$ be the sorted $\acute{T}_i$ values. Then the empirical estimators of the q-quantile and CTE are $\acute{\xi} = \acute{F}^{-1}(q) = \acute{T}_{\lceil nq \rceil:n}$ and $\acute{\gamma} = [1/((1-q)n)] \sum_{i=\lceil nq \rceil}^{n} \acute{T}_{i:n}$, respectively, where $\lceil \cdot \rceil$ is the ceiling function.

The left side of Figure 1 plots $\acute{F}(t)$, $\widehat{F}_{\text{exp}}(t)$, and $\widehat{F}_{\star}(t)$ for $\varepsilon = 10^{-2}$, which shows that the three curves closely align. But the CPU times are not of the same order of magnitude, as seen in Table 1 for the estimation of quantiles (for each particular method, computing the CDF and quantile estimators requires roughly the same time). The right side of Figure 1 plots the exponential and convolution estimators of $F(t)$ for $\varepsilon = 10^{-4}$, but we were not able to obtain the empirical distribution. Indeed, it would have required a CPU time of more than one year to sample 10^4 observations of T for ε^{-4} using crude simulation; on average, a single run of T takes more than one hour. But as ε shrinks, the approximations in (12) and (23) become more accurate by virtue of the limiting results in (10) and (11), so the CDF estimators $\widehat{F}_{\text{exp}}$ and $\widehat{F}_{\star}$ should be close to the true F for $\varepsilon = 10^{-4}$. Moreover, Table 1 shows that for this model, the CPU times for the exponential and convolution estimators do not increase when ε decreases.

Table 1 contains results for the empirical, exponential, the convolution estimators of the q-quantile, for three values of q. We used the bisection method to numerically compute the inverse in (34) for the convolution estimator. For $\varepsilon = 0.01$, the exponential and convolution estimators of ξ are close to the

Table 1: Quantile estimators.

ε	q	Empirical 95% CI	CPU	Expon. Est.	Expon. 95% CI	CPU	Convol. Est.	CPU
0.01	0.1	(1.701e+05, 1.971e+05)	890 sec	1.830e+05	(1.764e+05, 1.896e+05)	0.3 sec	1.865e+05	0.4 sec
0.01	0.5	(1.206e+06, 1.271e+06)	890 sec	1.204e+06	(1.161e+06, 1.247e+06)	0.3 sec	1.227e+06	0.4 sec
0.01	0.9	(3.958e+06, 4.135e+06)	890 sec	4.000e+06	(3.856e+06, 4.143e+06)	0.3 sec	4.075e+06	0.4 sec
10^{-4}	0.1	N/A	N/A	1.757e+13	(1.756e+13, 1.758e+13)	0.3 sec	1.762e+13	0.4 sec
10^{-4}	0.5	N/A	N/A	1.155e+14	(1.154e+14, 1.157e+14)	0.3 sec	1.159e+14	0.4 sec
10^{-4}	0.9	N/A	N/A	3.840e+14	(3.838e+14, 3.842e+14)	0.3 sec	3.850e+14	0.4 sec

empirical estimator, but with much less computational effort expended. For example, sampling $T \wedge \tau$ (resp., $\mathscr{I}(T < \tau)L$) with crude simulation (resp., IS) to compute (15) (resp., (17)), (18), and (19), requires simulating, on average, only 2.30 (resp., 4.25) transitions for $\varepsilon = 0.01$. For the exponential estimator, we include a *biased* 95% confidence interval (CI) based on the CI for μ; see the paragraph before Section 5. (For ε fixed, the approximation in (12) leads to estimators $\widehat{F}_{\text{exp}}$ and $\widehat{\xi}_{\text{exp}}$ having bias, which do not go away as the sample sizes grow large. In contrast, the bias of the ratio estimator (18) of μ vanishes as sample sizes increase.) For $\varepsilon = 10^{-4}$, we are not able to provide an empirical estimator of ξ, but the estimator (18) of μ is so good that the (biased) CI of ξ based on the exponential estimator has a relative width of only about 0.1%. But the exponential and convolution quantile estimators differ by about 0.3%, so the (biased) exponential CIs do not include $\widehat{\xi}_\star$. This may indicate that $\widehat{\xi}_\star$ and $\widehat{\xi}_{\exp}$ have different levels of bias.

Table 2 provides results for the CTE estimators. These numbers exhibit similar behavior to what Table 1 shows for the quantile estimators: much smaller CPU times for the exponential and convolution estimators than for the empirical, and very narrow (biased) exponential CIs (see the paragraph before Section 5), with the convolution estimators for $\varepsilon = 10^{-4}$ just above the exponential CIs.

Table 2: CTE estimators.

ε	q	Empir. Est.	CPU	Expon. Est.	Expon. 95% CI	CPU	Convol. Est.	CPU
0.01	0.1	1.964e+06	890 sec	1.920e+06	(1.851e+06, 1.989e+06)	0.3 sec	1.956e+06	0.4 sec
0.01	0.5	3.011e+06	890 sec	2.941e+06	(2.836e+06, 3.046e+06)	0.3 sec	2.996e+06	0.4 sec
0.01	0.9	5.915e+06	890 sec	5.737e+06	(5.531e+06, 5.942e+06)	0.3 sec	5.844e+06	0.4 sec
10^{-4}	0.1	N/A	N/A	1.839e+14	(1.834e+14, 1.845e+14)	0.3 sec	1.848e+14	0.4 sec
10^{-4}	0.5	N/A	N/A	2.817e+14	(2.809e+14, 2.826e+14)	0.3 sec	2.831e+14	0.4 sec
10^{-4}	0.9	N/A	N/A	5.495e+14	(5.479e+14, 5.512e+14)	0.3 sec	5.523e+14	0.4 sec

7 CONCLUDING REMARKS

We used simulation to calibrate approximations to the distribution F of the hitting time T to a rarely visited set $\mathscr{A}$ of states for a regenerative process. The exponential approximations in (12) and (23) to F and G, respectively, require the rarity parameter $\varepsilon \to 0$. But for an actual physical system, we have a fixed $\varepsilon > 0$, which introduces bias in both exponential approximations. Our numerical results may indicate that the two approximations produce different levels of bias, but this requires further study. Also, we are currently working on CLTs for the convolution estimators, to complement those for the exponential estimators (see the paragraph before Section 5). Chapter 3 of Kalashnikov (1997) provides upper and lower bounds to the true CDFs F and G, and we are investigating extending them to the q-quantile and CTE of F.

ACKNOWLEDGMENTS

This work has been supported in part by the National Science Foundation under Grant No. CMMI-1537322. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York: Springer.

Glynn, P. W., M. K. Nakayama, and B. Tuffin. 2017. "On the Estimation of the Mean Time to Failure by Simulation". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 1844–1855. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Goyal, A., P. Shahabuddin, P. Heidelberger, V. Nicola, and P. W. Glynn. 1992. "A Unified Framework for Simulating Markovian Models of Highly Dependable Systems". *IEEE Transactions on Computers* C-41(1):36–51.

Hong, L. J., Z. Hu, and G. Liu. 2014. "Monte Carlo Methods for Value-at-Risk and Conditional Value-at-Risk: A Review". *ACM Transactions on Modeling and Computer Simulataion* 24(4):22:1–22:37.

Kalashnikov, V. 1994. *Topics on Regenerative Processes*. Boca Raton: CRC Press.

Kalashnikov, V. 1997. *Geometric Sums: Bounds for Rare Events with Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

L'Ecuyer, P., and B. Tuffin. 2012. "Approximating Zero-Variance Importance Sampling in a Reliability Setting". *Annals of Operations Research* 189(1):277–297.

McNeil, A. J., R. Frey, and P. Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton, NJ: Princeton University Press.

Nakayama, M. K., and P. Shahabuddin. 2004. "Quick Simulation Techniques for Estimating the Unreliability of Large Regenerative Models of Highly Reliable Systems". *Probability in the Engineering and Informational Sciences* 18(3):339–368.

Nicola, V. F., P. Shahabuddin, and M. K. Nakayama. 2001. "Techniques for Fast Simulation of Models of Highly Dependable Systems". *IEEE Transactions on Reliability* 50(3):246–264.

Rubino, G., and B. Tuffin. 2009. "Markovian Models for Dependability Analysis". In *Rare Event Simulation using Monte Carlo Methods*, edited by G. Rubino and B. Tuffin, 125–144. Chichester, UK: John Wiley & Sons.

Sadowsky, J. S. 1991. "Large Deviations Theory and Efficient Simulation of Excessive Backlogs in a GI/GI/m Queue". *IEEE Transactions on Automatic Control* 36(12):1383–1394.

Shahabuddin, P. 1994. "Importance Sampling for Highly Reliable Markovian Systems". *Management Science* 40(3):333–352.

Shahabuddin, P., V. F. Nicola, P. Heidelberger, A. Goyal, and P. W. Glynn. 1988. "Variance Reduction in Mean Time to Failure Simulations". In *Proceedings of the 1988 Winter Simulation Conference*, edited by M. Abrams et al., 491–499. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Shedler, G. S. 1993. *Regenerative Stochastic Simulation*. San Diego: Academic Press.

AUTHOR BIOGRAPHIES

PETER W. GLYNN is the Thomas Ford Professor in the Department of Management Science and Engineering (MS&E) at Stanford University. He is a Fellow of INFORMS and of the Institute of Mathematical Statistics, has been co-winner of Best Publication Awards from the INFORMS Simulation Society in 1993, 2008, and 2016, and was the co-winner of the John von Neumann Theory Prize from INFORMS in 2010. In 2012, he was elected to the National Academy of Engineering. His research interests lie in stochastic simulation, queueing theory, and statistical inference for stochastic processes. His email address is glynn@stanford.edu.

MARVIN K. NAKAYAMA is a professor in the Department of Computer Science at the New Jersey Institute of Technology. He received a Ph.D. in operations research from Stanford University. He is a recipient of a CAREER Award from the National Science Foundation, and a paper he co-authored received the Best Theoretical Paper Award for the 2014 Winter Simulation Conference. He served as the simulation area editor for the *INFORMS Journal on Computing* from 2007–2016, and is an associate editor for *ACM Transactions on Modeling and Computer Simulation*. His research interests include simulation, modeling, statistics, risk analysis, and energy. His email address is marvin@njit.edu.

BRUNO TUFFIN received his PhD degree in applied mathematics from the University of Rennes 1 (France) in 1997. Since then, he has been with Inria in Rennes. His research interests include developing Monte Carlo and quasi-Monte Carlo simulation techniques for the performance evaluation of telecommunication systems and telecommunication-related economical models. He is currently Area Editor for *INFORMS Journal on Computing* and Associate Editor for *ACM Transactions on Modeling and Computer Simulation*. He has written or co-written three books (two devoted to simulation): *Rare event simulation using Monte Carlo methods* published by John Wiley & Sons in 2009, *La simulation de Monte Carlo* (in French), published by Hermes Editions in 2010, and *Telecommunication Network Economics: From Theory to Applications*, published by Cambridge University Press in 2014. His email address is bruno.tuffin@inria.fr.