Regenerative-Simulation-Based Estimators of Risk Measures for Hitting Times to Rarely Visited Sets

Peter W. Glynn, Marvin K. Nakayama, Bruno Tuffin

Stanford University, NJIT, Inria

RESIM 2021

Context and goals

- Common in the literature to estimate $\mu = \mathbb{E}[T]$ with $T = \inf\{t \geq 0 : X(t) \in \mathcal{A}\}$, hitting time of $\mathcal{A}$
- But what about the distribution F of T, in a rare event context?
  - May be required to estimate q-quantiles ($0 < q < 1$) of hitting times:
  $$\xi = F^{-1}(q) \equiv \inf\{t : F(t) \geq q\}$$
  - and the *conditional tail expectation* (CTE)
  $$\gamma = E[T \mid T > \xi].$$
- We designed two (*exponential and convolution*) estimators in a regenerative context — Glynn, Nakayama & T., WSC 2018
- And even 3 variations not particularly producing improvements — Glynn, Nakayama & T., WSC 2020

Our Goals:

Context and goals

- Common in the literature to estimate $\mu = \mathbb{E}[T]$ with $T = \inf\{t \geq 0 : X(t) \in \mathcal{A}\}$, hitting time of $\mathcal{A}$
- But what about the distribution F of T, in a rare event context?
  - May be required to estimate q-quantiles ($0 < q < 1$) of hitting times:
  $$\xi = F^{-1}(q) \equiv \inf\{t : F(t) \geq q\}$$
  - and the *conditional tail expectation* (CTE)
  $$\gamma = E[T \mid T > \xi].$$
- We designed two (*exponential and convolution*) estimators in a regenerative context — Glynn, Nakayama & T., WSC 2018
- And even 3 variations not particularly producing improvements — Glynn, Nakayama & T., WSC 2020

Our Goals:

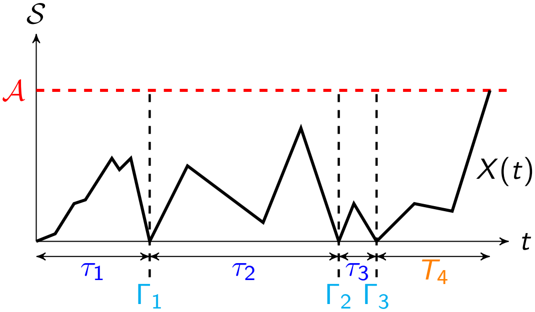
1/ Recall the exponential and convolution estimators and their efficiency
2/ Illustrate their respective power on a simple example.

Outline

1. Model: regenerative process
2. Estimators
   - Exponential Approximation Estimator
   - Convolution Estimator
3. Numerical efficiency
4. Analysis on a toy example
5. Conclusions

Regenerative system

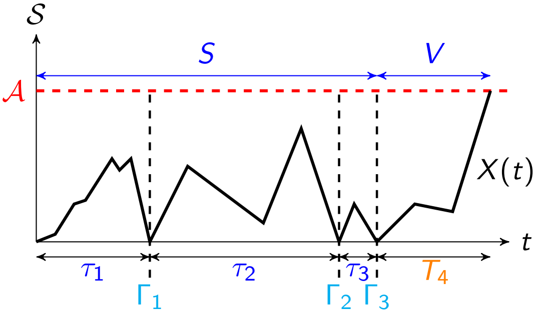
- Regeneration times $0 = \Gamma_0 < \Gamma_1 < \cdots$,
  with iid cycles $((\tau_k, (X(\Gamma_{k-1} + s) : 0 \leq s < \tau_k) : k \geq 1)$



- $\tau_k = \Gamma_k - \Gamma_{k-1}$, length of the kth regenerative cycle
- $T_k = \inf\{t \geq 0 : X(\Gamma_{k-1} + t) \in \mathcal{A}\}$ first hitting to $\mathcal{A}$ after regeneration Γ_{k-1}
- $M = \sup\{i > 0 : T_i > \tau_i\}$ (# cycles before first hitting $\mathcal{A}$)

Regenerative system

- Regeneration times $0 = \Gamma_0 < \Gamma_1 < \cdots$,
  with iid cycles $((\tau_k, (X(\Gamma_{k-1} + s) : 0 \leq s < \tau_k) : k \geq 1)$



- We can express

$$T = S + V \equiv \sum_{i=1}^{M} \tau_i + T_{M+1},$$

where the *geometric sum* S is independent of V.

Rare event and exponential limit

- Context: $p = \mathbb{P}(T < \tau)$ is small (rare event)
  - Model indexed by ϵ, rarity parameter, such that $p \equiv p_\epsilon \to 0$
- Ex: GI/G/1 queue
  - buffer size $b \equiv b_\epsilon = \lceil 1/\epsilon \rceil$
  - $\mathcal{A} = \mathcal{A}_\epsilon = \{b_\epsilon, b_\epsilon + 1, \dots\}$
- Ex: Highly Reliable System (HRS; HRMS in the Markovian case)
  - Multicomponent system with component j failure rate $\lambda_j = c_j \epsilon^{d_j}$ $(d_j > 0)$
  - Repair distributions independent of ϵ
  - Set $\mathcal{A}$: states with combinations of components down.

Theorem (Renyi's)

In the above contexts, if $p_\epsilon \to 0$ as $\epsilon \to 0$, T_ϵ/μ_ϵ converges weakly to an exponential

$$\lim_{\epsilon \to 0} \mathbb{P}_\epsilon(T_\epsilon/\mu_\epsilon \leq t) = 1 - e^{-t}, \quad \forall t \geq 0.$$

Outline

Exponential Approximation Estimator

Glynn, Nakayama and T., WSC'18

- $F(t) = \mathbb{P}_\epsilon(T_\epsilon < t) = \mathbb{P}_\epsilon(T_\epsilon/\mu_\epsilon < t/\mu_\epsilon) \approx 1 - e^{-t/\mu_\epsilon}$
- "Just" estimate the mean μ_ϵ
- Use the expression

$$\mu_\epsilon = \frac{\mathbb{E}_\epsilon[T_\epsilon \wedge \tau_\epsilon]}{\mathbb{P}_\epsilon(T_\epsilon < \tau_\epsilon)} \equiv \frac{\zeta_\epsilon}{p_\epsilon}$$

- Measure-specific IS (MSIS): simulate n cycles Shahabuddin et al. (1988)
  - $n_{CS} \equiv \gamma n$ cycles to estimate by crude simulation (CS): $\widehat{\zeta}_n \approx \zeta_\epsilon$
  - $n_{IS} \equiv (1-\gamma)n$ cycles to estimate by importance sampling (IS): $\widehat{p}_n \approx p_\epsilon$
- Resulting estimator $\widehat{\mu}_n = \dfrac{\widehat{\zeta}_n}{\widehat{p}_n}$.

Estimator

The exponential estimator of the cdf $F(t)$ of T is

$$\widehat{F}_{exp,n}(t) = 1 - e^{-t/\widehat{\mu}_n}$$

Exponential estimators

- Remarkably, estimating the distribution is reduced to estimating its mean.
- From

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

we get

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

Using an efficient estimator $\widehat{\mu}$ of μ from the literature

L'Ecuyer & T., Annals of OR 2011

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

Convolution Estimator

Glynn, Nakayama and T., WSC'18

- From $T = S + V$, the cdf F can be expressed as the convolution

$$F = G \star H \quad \text{with } S \sim G \text{ and } V \sim H.$$

Convolution Estimator

Glynn, Nakayama and T., WSC'18

- From $T = S + V$, the cdf F can be expressed as the convolution
$$F = G \star H \quad \text{with } S \sim G \text{ and } V \sim H.$$

- Exponential approximation for G when for $p \approx 0$: Kalashnikov (1997)
For each $t \geq 0$, $G(t) \approx \widetilde{G}_{\text{exp}}(t) = 1 - e^{-t/\eta}$ where
$\eta = \mathbb{E}[S] = \mathbb{E}[M] \cdot \mathbb{E}[\tau \mid \tau < T]$.

Convolution Estimator

Glynn, Nakayama and T., WSC'18

- From $T = S + V$, the cdf F can be expressed as the convolution
$$F = G \star H \quad \text{with } S \sim G \text{ and } V \sim H.$$
- Exponential approximation for G when for $p \approx 0$: Kalashnikov (1997)
For each $t \geq 0$, $G(t) \approx \widetilde{G}_{\text{exp}}(t) = 1 - e^{-t/\eta}$ where $\eta = \mathbb{E}[S] = \mathbb{E}[M] \cdot \mathbb{E}[\tau \mid \tau < T]$.
- Estimator (CS): $\widehat{G}_{\text{exp},n}(t) = 1 - e^{-t/\widehat{\eta}_n}$ with $\widehat{\eta}_n = \frac{1}{\widehat{p}_n \, n_{CS}} \sum_{i=1}^{n_{CS}} \tau_i \mathcal{I}(\tau_i < T_i)$.

Convolution Estimator

Glynn, Nakayama and T., WSC'18

- From $T = S + V$, the cdf F can be expressed as the convolution
$$F = G \star H \quad \text{with } S \sim G \text{ and } V \sim H.$$
- Exponential approximation for G when for $p \approx 0$: Kalashnikov (1997)
For each $t \geq 0$, $G(t) \approx \widetilde{G}_{\exp}(t) = 1 - e^{-t/\eta}$ where
$\eta = \mathbb{E}[S] = \mathbb{E}[M] \cdot \mathbb{E}[\tau \mid \tau < T]$.
- Estimator (CS): $\widehat{G}_{\exp,n}(t) = 1 - e^{-t/\widehat{\eta}_n}$ with $\widehat{\eta}_n = \dfrac{1}{\widehat{p}_n\, n_{CS}} \sum_{i=1}^{n_{CS}} \tau_i \mathcal{I}(\tau_i < T_i)$.
- Estimator (IS) of H: $\widehat{H}_n(x) = \dfrac{1}{\widehat{p}_n\, n_{IS}} \sum_{i=1}^{n_{IS}} \mathcal{I}(T_i' \wedge \tau_i' \leq x, T_i' < \tau_i') L_i'$

Convolution Estimator

Glynn, Nakayama and T., WSC'18

- From $T = S + V$, the cdf F can be expressed as the convolution
$$F = G \star H \quad \text{with } S \sim G \text{ and } V \sim H.$$

- Exponential approximation for G when for $p \approx 0$: Kalashnikov (1997)
For each $t \geq 0$, $G(t) \approx \widetilde{G}_{\exp}(t) = 1 - e^{-t/\eta}$ where
$\eta = \mathbb{E}[S] = \mathbb{E}[M] \cdot \mathbb{E}[\tau \mid \tau < T]$.

- Estimator (CS): $\widehat{G}_{\exp,n}(t) = 1 - e^{-t/\widehat{\eta}_n}$ with $\widehat{\eta}_n = \frac{1}{\widehat{p}_n \, n_{CS}} \sum_{i=1}^{n_{CS}} \tau_i \mathcal{I}(\tau_i < T_i)$.

- Estimator (IS) of H: $\widehat{H}_n(x) = \frac{1}{\widehat{p}_n \, n_{IS}} \sum_{i=1}^{n_{IS}} \mathcal{I}(T'_i \wedge \tau'_i \leq x, T'_i < \tau'_i) L'_i$

Estimator (For cdf $F(t)$; estimators of quantile and CTE deduced)

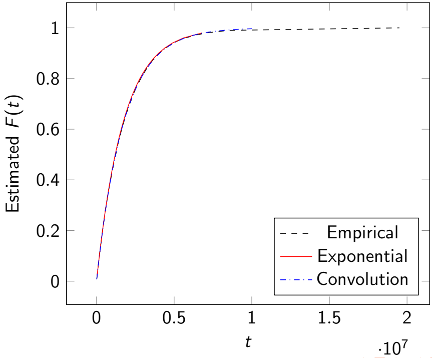
$$\widehat{F}_{conv,n}(t) = (\widehat{G}_{exp,n} \star \widehat{H}_n)(t) = 1 - \frac{1}{\widehat{p}_n \cdot n_{IS}} \sum_{i=1}^{n_{IS}} \mathcal{I}(T'_i < \tau'_i)\, L'_i\, e^{-(t-(T'_i \wedge \tau'_i))^+/\widehat{\eta}_n}.$$

Outline

Numerical example

- Highly reliable Markovian system with three component types
- five components of each type
- 15 repairmen
- system up whenever at least two components of each type work
- Each component has failure rate ϵ and repair rate 1.

With $\epsilon = 10^{-2}$



Numerical results

- Quantile estimators (CPU in sec.)

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

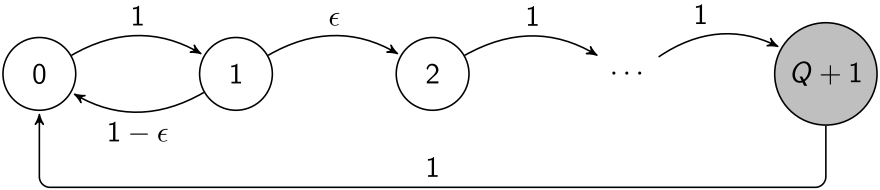
- CTE estimators (CPU in sec.)

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

  - Very efficient
  - But biased.... for small ϵ, it does not *seem* a problem in practice.

Outline

Simple $(Q+2)$-states example (discrete time)



- With $Q = \lfloor \epsilon^{-w} \rfloor$ $(w \geq 0)$ increasing as $\epsilon \to 0$
- Mixes both asymptotic regimes; what about the convergence to an exponential?
- Only two possibilities for paths of $T \wedge \tau$.
- We have $p = \mathbb{P}(T < \tau) = \epsilon$ and

$$T = 2M + Q + 1,$$

with $S = 2M$ and $V = Q + 1$ with M number of cycles between 0 and 1,

$$\mathbb{P}(M = m) = (1-\epsilon)^m \epsilon \quad (m = 0, 1, \dots)$$

On the weak convergence of $S_\epsilon/\mathbb{E}_\epsilon[S_\epsilon]$ to an exponential

M geometric with starting value 0; $\mathbb{E}[S_\epsilon] = 2(1-\epsilon)/\epsilon$

$S_\epsilon/\mathbb{E}_\epsilon[S_\epsilon]$ converges to an exponential whatever $w \geq 0$.

$$\begin{aligned}
P\left(\frac{S}{\mathbb{E}[S]} \leq y\right) &= P\left(M \leq \frac{y(1-\epsilon)}{\epsilon}\right) \\
&= P\left(M+1 \leq \frac{y(1-\epsilon)}{\epsilon}+1\right) \\
&= 1-P\left(M+1 > \frac{y(1-\epsilon)}{\epsilon}+1\right) \\
&= 1-(1-\epsilon)^{\lceil \frac{y(1-\epsilon)}{\epsilon} \rceil + 1} \\
&= 1-e^{\log(1-\epsilon)\left(\lceil \frac{y(1-\epsilon)}{\epsilon} \rceil + 1\right)} \\
&= 1-e^{-(\epsilon+o(\epsilon))\left(\lceil \frac{y(1-\epsilon)}{\epsilon} \rceil + 1\right)} \\
&= 1-e^{-y+o(1)}.
\end{aligned}$$

On the weak convergence of $T_\epsilon/\mathbb{E}_\epsilon[T_\epsilon]$ to an exponential

Recall that $T = 2M + Q + 1$ and $\mathbb{E}[T] = 2(1-\epsilon)/\epsilon + 1/\epsilon^w + 1$.

$$\begin{aligned} P\left(\frac{T}{\mathbb{E}[T]} \leq x\right) &= P\left(M + 1 \leq x\left(\frac{(1-\epsilon)}{\epsilon} + \frac{Q+1}{2}\right) - \frac{Q+1}{2} + 1\right) \\ &= 1 - (1-\epsilon)^{\lceil x\left(\frac{(1-\epsilon)}{\epsilon} + \frac{Q+1}{2}\right) - \frac{Q+1}{2} + 1\rceil} \\ &= 1 - e^{\log(1-\epsilon)\lceil x\left(\frac{(1-\epsilon)}{\epsilon} + \frac{\epsilon^{-w}+1}{2}\right) - \frac{\epsilon^{-w}+1}{2} + 1\rceil}. \end{aligned}$$

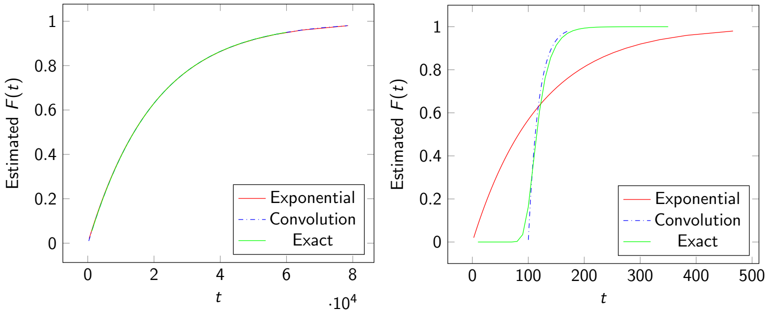
$T_\epsilon/\mathbb{E}_\epsilon[T_\epsilon]$ converges to an exponential if and only is $0 \leq w < 1$.

Basically, the exponential approximation to $T = S + V$ is valid when V is negligible compared to S.

Plots of CDFs

IS: replacing transition probability ϵ from state 0 to state 1 by $\alpha = 0.5$.

With $\epsilon = 0.0001$ and $w = 0.5$ (left) and $\epsilon = 0.1$ and $w = 2$ (right)



The convolution estimator is the only one always matching the true distribution.

Outline

Conclusions

- Rare-event estimation often focuses on determining a mean
- But other measures are of interest: quantiles, conditional tail expectation, etc.
- For rare events, not so many existing techniques
- We have described two estimators making use of an exponential limit for regenerative systems.
- The convolution estimator is
  - more robust to asymptotic settings
  - and expected to be less biased.

Thank you!