Efficient Simulation of Rare Events Involving Heavy Tailed Random Walks

Sandeep Juneja, Tata Institute

Jointly with Karthyek R. A. Murthy (Tata Institute)
Some parts
also with
Jose Blanchet (Columbia), Henrik Hult (KTH)

January 7, 2015

The Big Picture

- Importance sampling goes back to Marshall, Ulam, Von Neumann in the late forties

The Big Picture

- Importance sampling goes back to Marshall, Ulam, Von Neumann in the late forties
- Physicists devised many clever variations but with limited analysis

The Big Picture

- Importance sampling goes back to Marshall, Ulam, Von Neumann in the late forties
- Physicists devised many clever variations but with limited analysis
- Siegmund (1976) first provably efficient implementation for rare level crossing probabilities in the light tailed settings.

The Big Picture

- Importance sampling goes back to Marshall, Ulam, Von Neumann in the late forties
- Physicists devised many clever variations but with limited analysis
- Siegmund (1976) first provably efficient implementation for rare level crossing probabilities in the light tailed settings.
- Since then, enormous activity in 80's and 90's on importance sampling for rare event simulation in light tailed settings.

- Asmussen and Binswanger (1997) first to consider estimating $P(S_n > b)$ when increments are heavy-tailed using conditional Monte Carlo.

- Asmussen and Binswanger (1997) first to consider estimating $P(S_n > b)$ when increments are heavy-tailed using conditional Monte Carlo.

- Negative results for state independent methods in importance sampling (2007)

- Asmussen and Binswanger (1997) first to consider estimating $P(S_n > b)$ when increments are heavy-tailed using conditional Monte Carlo.
- Negative results for state independent methods in importance sampling (2007)
- Substantial literature since then focussing on complex state dependent methods

- Asmussen and Binswanger (1997) first to consider estimating $P(S_n > b)$ when increments are heavy-tailed using conditional Monte Carlo.

- Negative results for state independent methods in importance sampling (2007)

- Substantial literature since then focussing on complex state dependent methods

- We propose that a **Divide and Conquer** approach allows simpler state-independent methods to work

Basic assumption

- Let $S_n = \sum_{i=1}^{n} X_i$ where X_i are i.i.d.

Basic assumption

- Let $S_n = \sum_{i=1}^{n} X_i$ where X_i are i.i.d.
- Tail distribution of X_i is regularly varying:

$$P(X_i > x) = \frac{L(x)}{x^\alpha},$$

for $\alpha > 1$, where $L(x)$ is a slowly varying function:

$$\lim_{x \to \infty} L(tx)/L(x) = 1$$

for all $t > 0$.

Basic assumption

- Let $S_n = \sum_{i=1}^{n} X_i$ where X_i are i.i.d.
- Tail distribution of X_i is regularly varying:

$$P(X_i > x) = \frac{L(x)}{x^{\alpha}},$$

for $\alpha > 1$, where $L(x)$ is a slowly varying function:

$$\lim_{x \to \infty} L(tx)/L(x) = 1$$

for all $t > 0$.
- E.g., $L(x)$ is a constant or $L(x) = \log(|x|)^y$ for $y \in \Re$.

The rare event probabilities considered

We develop state independent efficient estimation methodologies for

The rare event probabilities considered

We develop state independent efficient estimation methodologies for

1. Large deviation probabilities $\mathbb{P}\{S_n > na\}$ for $a > 0$ and $EX_i = 0$.

The rare event probabilities considered

We develop state independent efficient estimation methodologies for

1. Large deviation probabilities $\mathbb{P}\{S_n > na\}$ for $a > 0$ and $EX_i = 0$.

2. Level crossing probabilities $\mathbb{P}\{\sup_n S_n > b)$ as $b \nearrow \infty$. where $EX_i < 0$.

The rare event probabilities considered ...

3 Level crossing in a busy cycle $\mathbb{P}\{\sup_{n\leq\tau} S_n > b)$ where $EX_i < 0$, and $\tau = \inf\{n \geq 1 : S_n < 0\}$.

The rare event probabilities considered ...

3 Level crossing in a busy cycle $\mathbb{P}\{\sup_{n\leq\tau} S_n > b)$ where $EX_i < 0$, and $\tau = \inf\{n \geq 1 : S_n < 0\}$.

4 Linear process large exceedance probability

$$P(\sum_{k=1}^{\infty} a_k X_k > b)$$

where $\{X_k\}$ are mean zero, i.i.d., regularly varying, $a_k \geq 0$, $\sum_k a_k^2 < \infty$.

The rare event probabilities considered ...

3 Level crossing in a busy cycle $\mathbb{P}\{\sup_{n\leq\tau} S_n > b)$ where $EX_i < 0$, and $\tau = \inf\{n \geq 1 : S_n < 0\}$.

4 Linear process large exceedance probability

$$P(\sum_{k=1}^{\infty} a_k X_k > b)$$

where $\{X_k\}$ are mean zero, i.i.d., regularly varying, $a_k \geq 0$, $\sum_k a_k^2 < \infty$.

Applications: Ruin probability in insurance

- $\{C_i\}$ - i.i.d. claim value process; $\{A_i\}$ - i.i.d. interarrival times; p - deterministic premium rate and b - initial capital.

Applications: Ruin probability in insurance

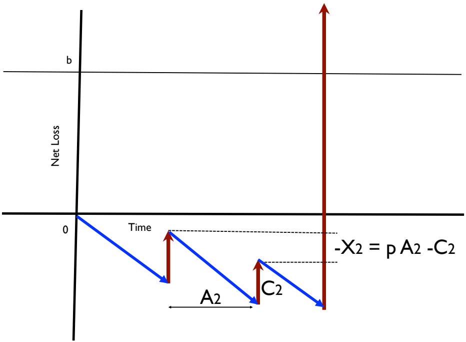
- $\{C_i\}$ - i.i.d. claim value process; $\{A_i\}$ - i.i.d. interarrival times; p - deterministic premium rate and b - initial capital.

- $X_i = C_i - pA_i$ is the net loss between claims $i-1$ and i. Typically, $EX_i < 0$. $\mathbb{P}\{\sup_n S_n > b)$ denotes the ruin probability.

Applications: Ruin probability in insurance

- $\{C_i\}$ - i.i.d. claim value process; $\{A_i\}$ - i.i.d. interarrival times; p - deterministic premium rate and b - initial capital.

- $X_i = C_i - pA_i$ is the net loss between claims $i-1$ and i. Typically, $EX_i < 0$. $\mathbb{P}\{\sup_n S_n > b)$ denotes the ruin probability.



First a brief review

- The problem of rare event simulation

First a brief review

- The problem of rare event simulation
- Notions of performance efficiency

First a brief review

- The problem of rare event simulation
- Notions of performance efficiency
- Importance sampling, zero variance measure

Naive estimation of $\gamma_n = P(S_n > na)$, $a > 0$

- Draw m independent samples $I_1, \ldots, I_m$ of $I(S_n > na)$, then the naive estimator

$$\frac{1}{m}\sum_{i=1}^{m} I_i \to \gamma_n \quad \text{as} \quad m \to \infty.$$

Naive estimation of $\gamma_n = P(S_n > na)$, $a > 0$

- Draw m independent samples $I_1, \ldots, I_m$ of $I(S_n > na)$, then the naive estimator

$$\frac{1}{m}\sum_{i=1}^{m} I_i \to \gamma_n \quad \text{as} \quad m \to \infty.$$

- For ratio of standard deviation to mean of the estimator to be ϵ, need

$$\epsilon = \frac{\sqrt{(1-\gamma_n)\gamma_n/m}}{\gamma_n} \approx 1/\sqrt{\gamma_n m}$$

Naive estimation of $\gamma_n = P(S_n > na)$, $a > 0$

- Draw m independent samples $I_1, \ldots, I_m$ of $I(S_n > na)$, then the naive estimator

$$\frac{1}{m}\sum_{i=1}^{m} I_i \to \gamma_n \quad \text{as} \quad m \to \infty.$$

- For ratio of standard deviation to mean of the estimator to be ϵ, need

$$\epsilon = \frac{\sqrt{(1-\gamma_n)\gamma_n/m}}{\gamma_n} \approx 1/\sqrt{\gamma_n m}$$

  so that

$$m \approx \frac{1}{\gamma_n\ \epsilon^2}$$

  implying that $m \to \infty$ as $\gamma_n \to 0$.

Estimation of $\gamma_n = P(S_n > na)$, $a > 0$

- More generally, for a sequence of unbiased estimators $\{Z_n\}$ of $\{\gamma_n\}$, number of simulation runs needed to get ϵ relative error

$$m = \left(\frac{Var(Z_n)}{\epsilon^2 \gamma_n^2} \right)$$

Efficiency notions of algorithms

- *Weak efficiency:* if $\{Z_n\}$ such that $Var(Z_n) = O(\gamma_n^{2-\epsilon}), \forall \epsilon > 0$

Efficiency notions of algorithms

- *Weak efficiency:* if $\{Z_n\}$ such that $Var(Z_n) = O(\gamma_n^{2-\epsilon}), \forall \epsilon > 0$

  $\Rightarrow m$ is $O(\gamma_n^{-\epsilon})$ $\forall \epsilon > 0$. (e.g., $\gamma_n = e^{-n}$, $Var(Z_n) = n^5 e^{-2n}$).

Efficiency notions of algorithms

- *Weak efficiency:* if $\{Z_n\}$ such that $Var(Z_n) = O(\gamma_n^{2-\epsilon}), \forall \epsilon > 0$

  $\Rightarrow m$ is $O(\gamma_n^{-\epsilon})$ $\forall \epsilon > 0$. (e.g., $\gamma_n = e^{-n}$, $Var(Z_n) = n^5 e^{-2n}$).

- *Strong efficiency:* $Var(Z_n) = O(\gamma_n^2)$, then m remains bounded as $n \nearrow \infty$ (**bounded number of samples no matter how rare the probability**).

Efficiency notions of algorithms

- *Weak efficiency:* if $\{Z_n\}$ such that $Var(Z_n) = O(\gamma_n^{2-\epsilon}), \forall \epsilon > 0$

  $\Rightarrow m$ is $O(\gamma_n^{-\epsilon})$ $\forall \epsilon > 0$. (e.g., $\gamma_n = e^{-n}$, $Var(Z_n) = n^5 e^{-2n}$).

- *Strong efficiency:* $Var(Z_n) = O(\gamma_n^2)$, then m remains bounded as $n \nearrow \infty$ (**bounded number of samples no matter how rare the probability**).

- *Asymptotically vanishing relative error:* $Var(Z_n) = o(\gamma_n^2)$ that is, m vanishes asymptotically (**asymptotically, a single sample gives the desired accuracy**).

Efficiency notions of algorithms

- *Weak efficiency:* if $\{Z_n\}$ such that $Var(Z_n) = O(\gamma_n^{2-\epsilon}), \forall \epsilon > 0$

  $\Rightarrow m$ is $O(\gamma_n^{-\epsilon})$ $\forall \epsilon > 0$. (e.g., $\gamma_n = e^{-n}$, $Var(Z_n) = n^5 e^{-2n}$).

- *Strong efficiency:* $Var(Z_n) = O(\gamma_n^2)$, then m remains bounded as $n \nearrow \infty$ (**bounded number of samples no matter how rare the probability**).

- *Asymptotically vanishing relative error:* $Var(Z_n) = o(\gamma_n^2)$ that is, m vanishes asymptotically (**asymptotically, a single sample gives the desired accuracy**).

Abstract view of importance sampling

- Consider

$$P(A) = \sum_{\omega \in A} P(\omega) = \sum_{\omega \in A} L(\omega) P^*(\omega) = E^*(LI(A))$$

where

$$L(\omega) = \frac{P(\omega)}{P^*(\omega)}.$$

Abstract view of importance sampling

- Consider

$$P(A) = \sum_{\omega \in A} P(\omega) = \sum_{\omega \in A} L(\omega) P^*(\omega) = E^*(LI(A))$$

where

$$L(\omega) = \frac{P(\omega)}{P^*(\omega)}.$$

- Average of independent samples of $L \times I(A)$ under P^* give an unbiased estimator of $P(A)$.

- Variance equals

$$E^*[L^2 I(A)] - P(A)^2 = \sum_{\omega \in A} \frac{P(\omega)^2}{P^*(\omega)} - P(A)^2$$

- Variance equals

$$E^*[L^2 I(A)] - P(A)^2 = \sum_{\omega \in A} \frac{P(\omega)^2}{P^*(\omega)} - P(A)^2$$

- Zero variance under the conditional measure

$$P^*(\omega) = \frac{P(\omega)}{P(A)}$$

for $\omega \in A$ and zero otherwise.

Importance sampling to estimate $\gamma_n := \mathbb{P}\{S_n > na\}$

- Let f denote the pdf of X_i

Importance sampling to estimate $\gamma_n := \mathbb{P}\{S_n > na\}$

- Let f denote the pdf of X_i
- Draw samples of X_i from suitably chosen density functions $\tilde{f}_{X_i|\mathbf{X_{i-1}}}$.

Importance sampling to estimate $\gamma_n := \mathbb{P}\{S_n > na\}$

- Let f denote the pdf of X_i
- Draw samples of X_i from suitably chosen density functions $\tilde{f}_{X_i|\mathbf{X_{i-1}}}$.
- *Unbiased* estimator of γ_n along sample $(x_1, x_2, \ldots, x_n)$ is given by

$$Z_n = \prod_{i=1}^{n} \frac{f(x_i)}{\tilde{f}_{X_i|\mathbf{X_{i-1}}=\mathbf{x_{i-1}}}(x_i)} I(\sum_{i=1}^{n} x_i > na).$$

Importance sampling to estimate $\gamma_n := \mathbb{P}\{S_n > na\}$

- Let f denote the pdf of X_i
- Draw samples of X_i from suitably chosen density functions $\tilde{f}_{X_i|\mathbf{X_{i-1}}}$.
- *Unbiased* estimator of γ_n along sample $(x_1, x_2, \ldots, x_n)$ is given by

$$Z_n = \prod_{i=1}^{n} \frac{f(x_i)}{\tilde{f}_{X_i|\mathbf{X_{i-1}}=\mathbf{x_{i-1}}}(x_i)} I(\sum_{i=1}^{n} x_i > na).$$

- Take sample average of m independent samples of Z_n.

State-dependent and independent importance sampling

- If the distribution of X_i depends on $\{X_1, \ldots, X_{i-1}\}$, the method is state-dependent

State-dependent and independent importance sampling

- If the distribution of X_i depends on $\{X_1, \dots, X_{i-1}\}$, the method is state-dependent
- If no such dependence exists and the samples of $\{X_i : 1 \leq i \leq n\}$ can be drawn independently, then we call it state-independent

State-dependent and independent importance sampling

- If the distribution of X_i depends on $\{X_1, \ldots, X_{i-1}\}$, the method is state-dependent
- If no such dependence exists and the samples of $\{X_i : 1 \leq i \leq n\}$ can be drawn independently, then we call it state-independent
- Typically, easier to generate samples using state-independent methods

Zero Variance Estimator is State Dependent

- Under conditional measure

$$\tilde{f}(x_1, \ldots, x_n) = \frac{\prod_{i=1}^{n} f(x_i)}{P(S_n > na)} I(\sum_{i=1}^{n} x_i > na).$$

- Set

$$\tilde{f}_{X_i|\mathbf{X_{i-1}}=\mathbf{x_{i-1}}}(x) = \frac{f(x)P(S_n > na|S_i = s_{i-1} + x)}{P(S_n > na|S_{i-1} = s_{i-1})}$$

Zero Variance Estimator is State Dependent

- Under conditional measure

$$\tilde{f}(x_1, \ldots, x_n) = \frac{\prod_{i=1}^n f(x_i)}{P(S_n > na)} I(\sum_{i=1}^n x_i > na).$$

- Set

$$\tilde{f}_{X_i|\mathbf{X_{i-1}}=\mathbf{x_{i-1}}}(x) = \frac{f(x)P(S_n > na|S_i = s_{i-1} + x)}{P(S_n > na|S_{i-1} = s_{i-1})}$$

- This is a zero variance estimator as along $\{\sum_{i=1}^n x_i > na\}$

$$\prod_{i=1}^n \frac{f(x_i)}{\tilde{f}_{X_i|\mathbf{X_{i-1}}=\mathbf{x_{i-1}}}(x_i)} = P(S_n > na).$$

It is state-dependent.

Estimating $\gamma_n := \mathbb{P}\{S_n > na\}$ for light tailed increments

- When these X_i are light-tailed, exponential twisting based importance sampling methods are provably successful. X_i remain independent under the importance sampling measure (Sadowsky and Bucklew 1990, 91)

Estimating $\gamma_n := \mathbb{P}\{S_n > na\}$ for light tailed increments

- When these X_i are light-tailed, exponential twisting based importance sampling methods are provably successful. X_i remain independent under the importance sampling measure (Sadowsky and Bucklew 1990, 91)

- Samples $(x_1, \ldots, x_n)$ drawn for increments from the exponentially twisted density:

$$\hat{f}(x) = e^{\theta_a x - \Lambda(\theta_a)} f(x)$$

for appropriately chosen $\theta_a > 0$, where $\Lambda(\theta) = \log E[e^{\theta X}]$.

- Simulation output equals

$$\begin{aligned} & \exp(-\theta_a \sum_{i=1}^{n} x_i + n\Lambda(\theta_a)) I(\sum_{i=1}^{n} x_i > na) \\ \leq \; & \exp(-n(\theta_a a - \Lambda(\theta_a))) \end{aligned}$$

and the estimator is shown to be weakly efficient.

- Simulation output equals

$$\exp(-\theta_a \sum_{i=1}^{n} x_i + n\Lambda(\theta_a)) I(\sum_{i=1}^{n} x_i > na)$$
$$\leq \exp(-n(\theta_a a - \Lambda(\theta_a)))$$

and the estimator is shown to be weakly efficient.

- Can show that as $n \to \infty$, the zero variance density

$$\tilde{f}_{X_i|S_{i-1}=s}(x) = \frac{f(x)P(S_n > na|S_i = s + x)}{P(S_n > na|S_{i-1} = s)} \to \frac{e^{\theta_a x}}{\mathbb{E}[e^{\theta_a X}]} f(x).$$

- Simulation output equals

$$\exp(-\theta_a \sum_{i=1}^{n} x_i + n\Lambda(\theta_a)) I(\sum_{i=1}^{n} x_i > na)$$
$$\leq \exp(-n(\theta_a a - \Lambda(\theta_a)))$$

and the estimator is shown to be weakly efficient.

- Can show that as $n \to \infty$, the zero variance density

$$\tilde{f}_{X_i|S_{i-1}=s}(x) = \frac{f(x)P(S_n > na|S_i = s + x)}{P(S_n > na|S_{i-1} = s)} \to \frac{e^{\theta_a x}}{\mathbb{E}[e^{\theta_a X}]} f(x).$$

- **For heavy tailed rv, exponential twisting no longer feasible as $E[e^{\theta X}] = \infty$ for $\theta > 0$.**

Negative result for $\mathbb{P}\{\sup_{n \leq \tau} S_n > b)$, Bassamboo, J, Zeevi 07

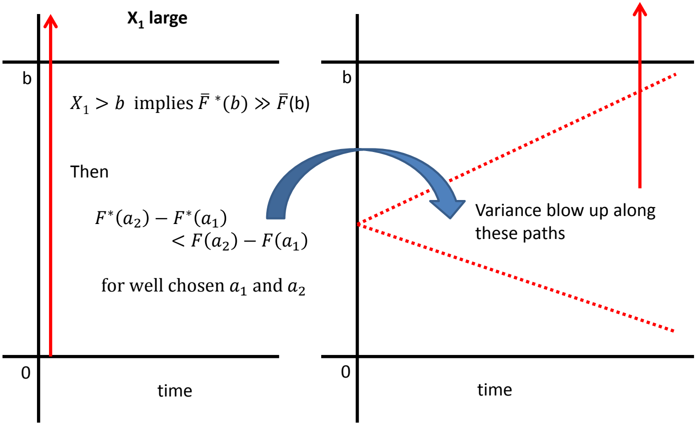
- *Algorithm* - Under importance sampling $(X_1, \ldots, X_\tau)$ are iid.

Negative result for $\mathbb{P}\{\sup_{n\leq\tau} S_n > b)$, Bassamboo, J, Zeevi 07

- *Algorithm* - Under importance sampling $(X_1, \ldots, X_\tau)$ are iid.
- *Result* - Weakly efficient state independent measure does not exist.

Negative result for $\mathbb{P}\{\sup_{n\leq\tau} S_n > b)$, Bassamboo, J, Zeevi 07

- *Algorithm* - Under importance sampling $(X_1, \ldots, X_\tau)$ are iid.
- *Result* - Weakly efficient state independent measure does not exist.



Remaining talk overview

- Motivated by this and the form of the zero variance estimator, there is large evolving literature that develops complex *state dependent* importance sampling methods for efficient simulation of these probabilities. (Blanchet and Liu 08, 12, Dupuis, Leder, Wang 07, Blanchet and Glynn 08, Chan, Deng, Lai 12)

Remaining talk overview

- Motivated by this and the form of the zero variance estimator, there is large evolving literature that develops complex *state dependent* importance sampling methods for efficient simulation of these probabilities. (Blanchet and Liu 08, 12, Dupuis, Leder, Wang 07, Blanchet and Glynn 08, Chan, Deng, Lai 12)

- We propose that by suitably decomposing these probabilities into a dominant and further residual components, simpler state-independent importance sampling algorithms can be devised with a desirable vanishing relative error property.

Remaining talk overview

- Motivated by this and the form of the zero variance estimator, there is large evolving literature that develops complex *state dependent* importance sampling methods for efficient simulation of these probabilities. (Blanchet and Liu 08, 12, Dupuis, Leder, Wang 07, Blanchet and Glynn 08, Chan, Deng, Lai 12)

- We propose that by suitably decomposing these probabilities into a dominant and further residual components, simpler state-independent importance sampling algorithms can be devised with a desirable vanishing relative error property.

- When the increments have infinite variance, there is an added complexity in estimating the level crossing probability as even the well known zero variance estimator has an infinite expected termination time.

- We adapt our algorithms so that this expectation is finite while the estimators remain strongly efficient.

- We adapt our algorithms so that this expectation is finite while the estimators remain strongly efficient.

- We show how our approach may be applied to estimate rare probabilities such as level crossing in a busy cycle as well as large exceedance probability of a linear process.

- We adapt our algorithms so that this expectation is finite while the estimators remain strongly efficient.
- We show how our approach may be applied to estimate rare probabilities such as level crossing in a busy cycle as well as large exceedance probability of a linear process.
- Numerically, the proposed estimators perform at least as well, and sometimes substantially better than the existing state-dependent estimators in the literature.

- We adapt our algorithms so that this expectation is finite while the estimators remain strongly efficient.
- We show how our approach may be applied to estimate rare probabilities such as level crossing in a busy cycle as well as large exceedance probability of a linear process.
- Numerically, the proposed estimators perform at least as well, and sometimes substantially better than the existing state-dependent estimators in the literature.
- Our key contribution thus is to **question the prevailing view that one needs to resort to state-dependent methods for efficient computation of rare event probabilities involving large number of heavy-tailed random variables.**

Proposed method for $P(S_n > na)$

Asymptotics for heavy tailed sums

- Let $M_n = \max\{X_1, \ldots, X_n\}$. For heavy tails, as $n \nearrow \infty$,

$$\mathbb{P}\{S_n > na\} \sim \mathbb{P}\{M_n > na\} \sim nP(X_1 > na),$$

Asymptotics for heavy tailed sums

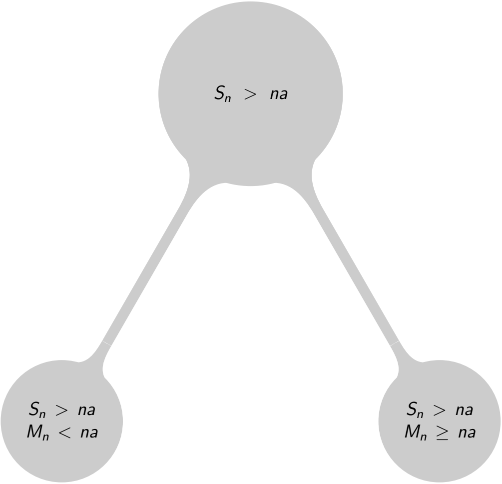
- Let $M_n = \max\{X_1, \ldots, X_n\}$. For heavy tails, as $n \nearrow \infty$,

$$\mathbb{P}\{S_n > na\} \sim \mathbb{P}\{M_n > na\} \sim nP(X_1 > na),$$

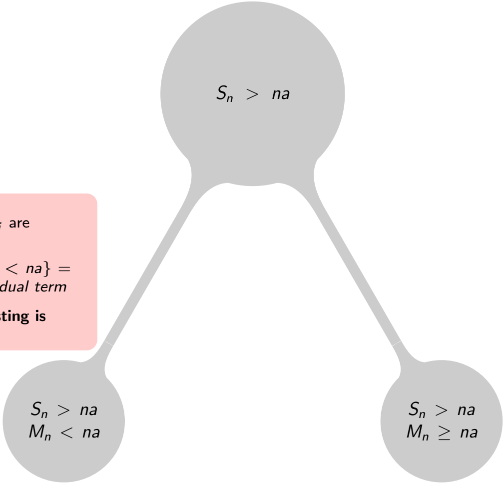
so that

$$\mathbb{P}\{S_n > na, M_n \leq na\} = o(nP(X_1 > na)).$$

Solving $\mathbb{P}\{S_n > na\}$ as $n \nearrow \infty$, through decomposition

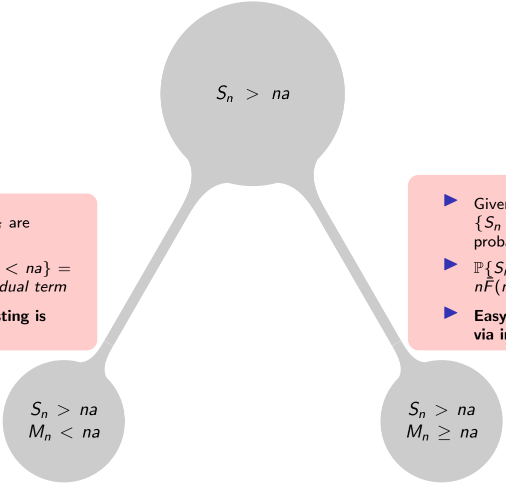


Solving $\mathbb{P}\{S_n > na\}$ as $n \nearrow \infty$, through decomposition



- In this set, all X_i are bounded by na!
- $\mathbb{P}\{S_n > na, M_n < na\} = o(n\bar{F}(na))$. *Residual term*
- **exponential twisting is applicable!**

Solving $\mathbb{P}\{S_n > na\}$ as $n \nearrow \infty$, through decomposition



- In this set, all X_i are bounded by na!
- $\mathbb{P}\{S_n > na, M_n < na\} = o(n\bar{F}(na))$. *Residual term*
- **exponential twisting is applicable!**

- Given $\{M_n \geq na\}$, $\{S_n > na\}$ occurs with probability 1 as $n \to \infty$!
- $\mathbb{P}\{S_n > na, M_n \geq na\} \sim n\bar{F}(na)$. *Dominant term*
- **Easy to efficiently simulate via importance sampling**

Algorithm to estimate dominant $\mathbb{P}\{S_n > na, M_n \geq na\}$

- (Chan, Deng, Lai 2012)
  1. Choose an index I uniformly at random from $\{1, \ldots, n\}$

Algorithm to estimate dominant $\mathbb{P}\{S_n > na, M_n \geq na\}$

- (Chan, Deng, Lai 2012)
  1. Choose an index I uniformly at random from $\{1, \ldots, n\}$
  2. For $k = 1, \ldots, n$, generate a realization of X_k from $F(\cdot|X_k \geq na)$ if $k = I$; otherwise, generate X_k from $F(\cdot)$.

To estimate residual $\mathbb{P}\{S_n > na, M_n < na\}$

- On the set $\{M_n < na\}$, each X_i is bounded for fixed n; we can apply exponential twisting!

To estimate residual $\mathbb{P}\{S_n > na, M_n < na\}$

- On the set $\{M_n < na\}$, each X_i is bounded for fixed n; we can apply exponential twisting!
- *IS distribution:* $\tilde{f}(x) = c_n e^{\theta_n x} f(x) \mathbf{1}(x \leq na)$
- The estimator:

$$Z_{res}(n) = \frac{1}{c_n^n} e^{-\theta_n S_n} \mathbb{I}_{\{S_n > na, M_n < na\}}$$

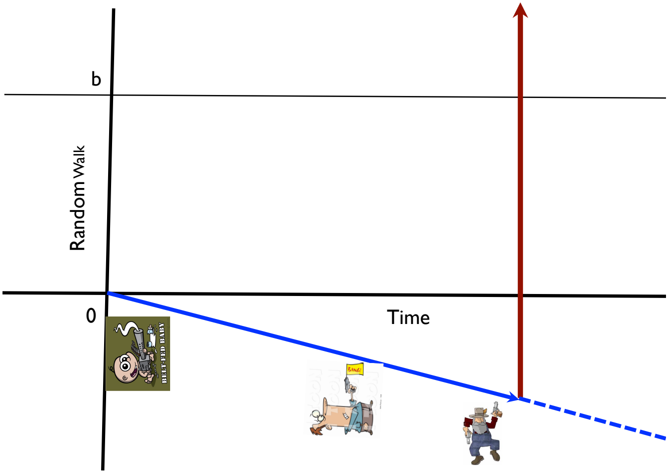
$$\theta_n = \frac{-\log n\bar{F}(na)}{na}$$

Theorem

Above algorithm offers asymptotically vanishing relative error in the estimation of $\mathbb{P}\{S_n \geq na\}$ *as* $n \nearrow \infty$. *The results generalize to* na *replaced by* $b_n \geq \tilde{c} n^{\frac{1}{2}+\epsilon}$ *for any positive constant* $\tilde{c}$ *and* ϵ.

Proposed method for $P(\sup_{n\geq 1} S_n > b)$

Level crossing probability: Big jump principle on fluid scale



$$P(\sup_{n \geq 1} S_n > b) = P(\tau_b < \infty) \sim \sum_{n=1}^{\infty} P(X_n > b + (n-1)\mu).$$

Estimating level crossing probability

- Naive simulation no longer feasible

Estimating level crossing probability

- Naive simulation no longer feasible
- We partition and divide

$$
\begin{aligned}
P(\tau_b < \infty) &= \sum_{k \geq 1} P(n_{k-1} < \tau_b \leq n_k) \\
&= \sum_{k \geq 1} \frac{P(n_{k-1} < \tau_b \leq n_k)}{p_k} p_k \\
&= E\left(\frac{P(n_{K-1} < \tau_b \leq n_K)}{p_K} \right)
\end{aligned}
$$

Approximate zero variance randomization

$$P(\tau_b < \infty) = E\left(\frac{P(n_{K-1} < \tau_b \leq n_K)}{p_K}\right)$$

Approximate zero variance randomization

$$P(\tau_b < \infty) = E\left(\frac{P(n_{K-1} < \tau_b \le n_K)}{p_K}\right)$$

- If we know $P(n_{k-1} < \tau_b \le n_k)$ and select

$$p_k = \frac{P(n_{k-1} < \tau_b \le n_k)}{P(\tau_b < \infty)}$$

then the estimator has zero variance.

Approximate zero variance randomization

$$P(\tau_b < \infty) = E\left(\frac{P(n_{K-1} < \tau_b \leq n_K)}{p_K}\right)$$

- If we know $P(n_{k-1} < \tau_b \leq n_k)$ and select

$$p_k = \frac{P(n_{k-1} < \tau_b \leq n_k)}{P(\tau_b < \infty)}$$

  then the estimator has zero variance.

- We know $P(n_{k-1} < \tau_b \leq n_k) \sim \sum_{i=n_{k-1}+1}^{n_k} P(X_i > b + (i-1)\mu)$.

Approximate zero variance randomization

$$P(\tau_b < \infty) = E\left(\frac{P(n_{K-1} < \tau_b \leq n_K)}{p_K}\right)$$

- If we know $P(n_{k-1} < \tau_b \leq n_k)$ and select

$$p_k = \frac{P(n_{k-1} < \tau_b \leq n_k)}{P(\tau_b < \infty)}$$

  then the estimator has zero variance.

- We know
  $P(n_{k-1} < \tau_b \leq n_k) \sim \sum_{i=n_{k-1}+1}^{n_k} P(X_i > b + (i-1)\mu)$.

- Use these approximations to generate K.

For $P(n_{k-1} < \tau_b \leq n_k)$ further divide and ...

- The dominant event $A_k \cap \{n_{k-1} < \tau_b \leq n_k\}$

$$A_k = \bigcup_{n_{k-1}+1}^{n_k} \{X_i > b + (i-1)\mu\}$$

For $P(n_{k-1} < \tau_b \leq n_k)$ further divide and ...

- The dominant event $A_k \cap \{n_{k-1} < \tau_b \leq n_k\}$

$$A_k = \bigcup_{n_{k-1}+1}^{n_k} \{X_i > b + (i-1)\mu\}$$

- Residual events $B_k \cap \{n_{k-1} < \tau_b \leq n_k\}$ where B_k denotes all $X_i's$ truncated.

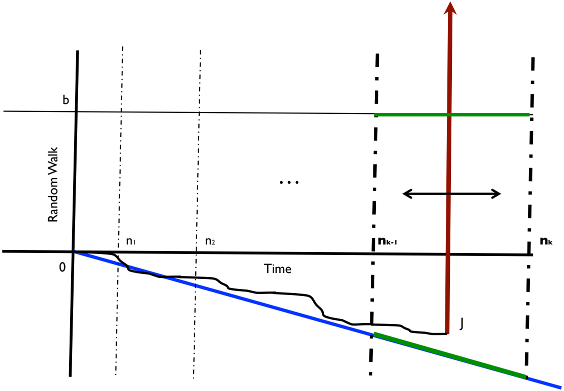
For $P(n_{k-1} < \tau_b \leq n_k)$ further divide and ...

- The dominant event $A_k \cap \{n_{k-1} < \tau_b \leq n_k\}$

$$A_k = \bigcup_{n_{k-1}+1}^{n_k} \{X_i > b + (i-1)\mu\}$$

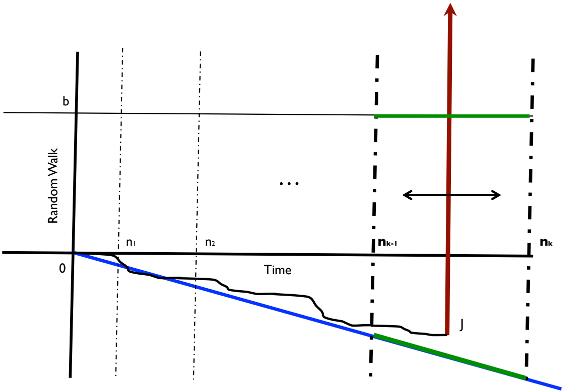
- Residual events $B_k \cap \{n_{k-1} < \tau_b \leq n_k\}$ where B_k denotes all $X_i^\prime s$ truncated.
- and $C_k \cap \{n_{k-1} < \tau_b \leq n_k\}$ where C_k denotes unsuccessful big jumps early on.

Simulating $\{n_{k-1} < \tau_b \leq n_k, A_k\}$



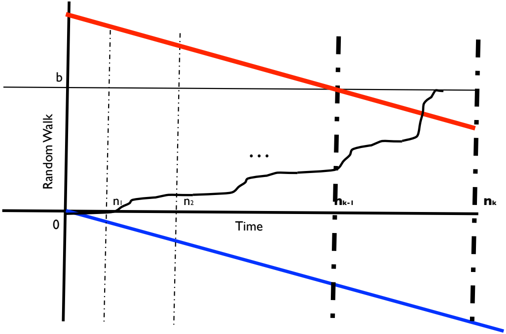
1. Select an index J: $\Pr\{J = n\} = \frac{\bar{F}(b+(n-1)\mu)}{\sum_{i=n_{k-1}+1}^{n_k} \bar{F}(b+(i-1)\mu)}$, for $n_{k-1} < n \leq n_k$.

Simulating $\{n_{k-1} < \tau_b \leq n_k, A_k\}$



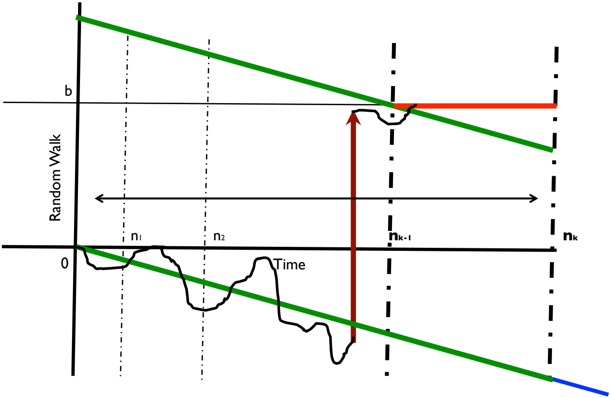
1. Select an index J: $\Pr\{J = n\} = \frac{\bar{F}(b+(n-1)\mu)}{\sum_{i=n_{k-1}+1}^{n_k} \bar{F}(b+(i-1)\mu)}$, for $n_{k-1} < n \leq n_k$.
2. Simulate the increment X_n from $F(\cdot|X_n \geq b + (n-1)\mu)$, if $n = J$; otherwise, simulate X_n from $F(\cdot)$, for any $n \leq n_k$.

Simulating $\{n_{k-1} < \tau_b \leq n_k, B_k\}$



- Sample $X_1, \ldots, X_{\tau_b \wedge n_k}$ independently from appropriate exponentially twisted distribution.

Simulating $\{n_{k-1} < \tau_b \leq n_k, C_k\}$



Some results

- Theorem
  The family of unbiased estimators $(Z(b) : b > 0)$ *achieves asymptotically vanishing relative error for the computation of* $\mathbb{P}\{\tau_b < \infty\}$, *as* $b \nearrow \infty$; *that is:*

$$\overline{\lim_{b\to\infty}} \frac{\mathrm{Var}\left[Z(b)\right]}{\mathbb{P}\{\tau_b < \infty\}^2} = 0.$$

Some results

- Theorem
  The family of unbiased estimators $(Z(b) : b > 0)$ *achieves asymptotically vanishing relative error for the computation of* $\mathbb{P}\{\tau_b < \infty\}$*, as* $b \nearrow \infty$*; that is:*

$$\overline{\lim_{b\to\infty}} \frac{\mathrm{Var}\,[Z(b)]}{\mathbb{P}\{\tau_b < \infty\}^2} = 0.$$

- Theorem
  If $\bar{F}(\cdot)$ *is regularly varying with index* $\alpha > 2$*, under the proposed algorithm, the computational effort* ν_b*:*

$$E[\nu_b] \leq \frac{r}{\mu(\alpha - 2)} b, \text{ as } b \nearrow \infty.$$

Infinite Variance $1 < \alpha < 2$

- For tails $\bar{F}(\cdot)$ with regularly varying index $1 < \alpha < 2$, we have that $\mathbb{E}[\tau_b|\tau_b < \infty] = \infty$; that is, the zero-variance measure has infinite expected termination time!

Infinite Variance $1 < \alpha < 2$

- For tails $\bar{F}(\cdot)$ with regularly varying index $1 < \alpha < 2$, we have that $\mathbb{E}[\tau_b | \tau_b < \infty] = \infty$; that is, the zero-variance measure has infinite expected termination time!
- The proposed $\{p_k\}$ asymptotically match the zero variance measure

Infinite Variance $1 < \alpha < 2$

- For tails $\bar{F}(\cdot)$ with regularly varying index $1 < \alpha < 2$, we have that $\mathbb{E}[\tau_b | \tau_b < \infty] = \infty$; that is, the zero-variance measure has infinite expected termination time!
- The proposed $\{p_k\}$ asymptotically match the zero variance measure
- Can see that infinite expected termination time for the proposed algorithm for $1 < \alpha < 2$.

For $1.5 < \alpha < 2$

- Let $\beta \in [2, 2\alpha - 1)$.

For $1.5 < \alpha < 2$

- Let $\beta \in [2, 2\alpha - 1)$.

- Set p_k proportional to

$$\sum_{n=n_{k-1}+1}^{n_k} \frac{P(X_n > b + (n-1)\mu)}{(b + (n-1)\mu)^{\beta-\alpha}}.$$

For $1.5 < \alpha < 2$

- Let $\beta \in [2, 2\alpha - 1)$.
- Set p_k proportional to

$$\sum_{n=n_{k-1}+1}^{n_k} \frac{P(X_n > b + (n-1)\mu)}{(b + (n-1)\mu)^{\beta - \alpha}}.$$

- Theorem
  With above chosen randomization probabilities
  1. *strong efficiency:* $\overline{\lim}_{b \to \infty} \frac{\mathrm{Var}[Z(b)]}{\mathbb{P}\{\tau_b < \infty\}^2} < \infty$, *and*
  2. *finite expected termination time:* $\mathbb{E}[\nu_b] \leq \frac{r + o(1)}{\mu(\beta - 2)} b$, *as* $b \nearrow \infty$.

For $1 < \alpha < 1.5$, an impossibility result

- Theorem
  If the tail index $\alpha < 1.5$, there does not exist an assignment of $(p_k, n_k : k \geq 1)$ such that both $\mathbb{E}^Q[Z^2(b)]$ and $\mathbb{E}^Q[\nu_b]$ are simultaneously finite.

  - Similar result by Blanchet and Liu 2012, in a different state-dependent setting

Level Crossing in a Busy Cycle

.

$$P(\sup_{n \leq \tau} S_n > b)$$

Divide and conquer!

- $(X_i : i \geq 1)$ are i.i.d, negative mean, random variables with regularly varying tail. $S_n = \sum_{i=1}^{n} X_i$.
- $$\tau = \inf\{n \geq 1 : S_n < 0\}.$$
  $$\tau_b = \inf\{n \geq 1 : S_n \geq b\}.$$
- Probability of interest
  $$P(\max_{k \leq \tau} S_k > b) = P(\tau_b < \tau).$$
- The following decomposition is easily seen
  $$P(\tau_b < \tau) = \bar{F}(b)E(\tau_b \wedge \tau) + P(\tau_b < \tau, \max_{k \leq \tau_b} X_k < b).$$

Estimating large deviations probability for linear processes

.

$$P(\sum_k a_k X_k > b)$$

- Probability of interest $P(\sum_k a_k X_k > b)$ for large b, where a_k are non-negative and $\sum_k a_k^2 < \infty$.

- Probability of interest $P(\sum_k a_k X_k > b)$ for large b, where a_k are non-negative and $\sum_k a_k^2 < \infty$.
- Easy to see that $P(\sum_k a_k X_k > b)$

$$
\begin{aligned}
&= \lim_{n\to\infty} P(\sum_{k=1}^{n} a_k X_k > b) \\
&= \sum_{n=1}^{\infty} \left(P(\sum_{k=1}^{n} a_k X_k > b) - P(\sum_{k=1}^{n-1} a_k X_k > b) \right) \\
&= \sum_{n=1}^{\infty} \left(\frac{P(\sum_{k=1}^{n} a_k X_k > b) - P(\sum_{k=1}^{n-1} a_k X_k > b)}{p_n} \right) p_n
\end{aligned}
$$

- Well known that

$$P(\sum_{k=1}^{n} a_k X_k > b) \sim \sum_{k=1}^{n} P(a_k X_k > b)$$

- Well known that

$$P(\sum_{k=1}^{n} a_k X_k > b) \sim \sum_{k=1}^{n} P(a_k X_k > b)$$

- Can set $\{p_n\}$ and develop fast simulation methods for

$$P(\sum_{k=1}^{n} a_k X_k > b) - P(\sum_{k=1}^{n-1} a_k X_k > b) =$$

to achieve vanishing relative error.

Numerical experiment: $P(S_n > n)$

$X = \Lambda R$, $\mathbb{P}\{\Lambda > x\} = \min(1, \frac{1}{x^4})$, $R \sim Laplace(1)$

n	True value	Point estimate	CV* of Prop. Algo	CV of Algo BL
100	2.21×10^{-5}	2.17×10^{-5}	1.9	4.7
500	1.04×10^{-7}	1.05×10^{-7}	0.7	4.1
1000	1.25×10^{-8}	1.29×10^{-8}	0.6	3.8

Table: Comparing proposed algorithm with that of Blanchet and Liu (2008). Sample average of 10,000 samples

*Coefficient of variation, $\text{CV} = \frac{\text{Standard deviation of the estimator}}{\text{Mean of the estimator}}$

In addition to variance reduction, the proposed algorithm requires much less computational effort in generating samples (due to state independence). For common range of input parameters, it runs at least 100 times faster than the existing state-dependent methods

Conclusion

- Revisited the problem of efficient simulation of the rare large deviations as well as the level crossing probability of a random walk.

Conclusion

- Revisited the problem of efficient simulation of the rare large deviations as well as the level crossing probability of a random walk.
- Showed that simple state-independent importance sampling methods, that are at least as efficient as the existing state-dependent methods, can be devised.

Conclusion

- Revisited the problem of efficient simulation of the rare large deviations as well as the level crossing probability of a random walk.
- Showed that simple state-independent importance sampling methods, that are at least as efficient as the existing state-dependent methods, can be devised.
- Devised efficient schemes for large exceedance for a linear process as well as for random walk in a busy cycle.

Conclusion

- Revisited the problem of efficient simulation of the rare large deviations as well as the level crossing probability of a random walk.
- Showed that simple state-independent importance sampling methods, that are at least as efficient as the existing state-dependent methods, can be devised.
- Devised efficient schemes for large exceedance for a linear process as well as for random walk in a busy cycle.
- Our approach relied on partitioning the rare event of interest into elementary events that were amenable to straight forward state-independent importance sampling methods.

Conclusion

- Revisited the problem of efficient simulation of the rare large deviations as well as the level crossing probability of a random walk.
- Showed that simple state-independent importance sampling methods, that are at least as efficient as the existing state-dependent methods, can be devised.
- Devised efficient schemes for large exceedance for a linear process as well as for random walk in a busy cycle.
- Our approach relied on partitioning the rare event of interest into elementary events that were amenable to straight forward state-independent importance sampling methods.
- We expect this approach will generalize to more complex, multi-dimensional problems, and for similar problems involving Weibull-type sub-exponential tail distributions.