Efficient Rare Event Estimation for Branching Random Walks

Frontier Probability Days 2021

Michael Conroy

University of Arizona

December 3, 2021

Background: Exponential tilting for random walks on $\mathbb{R}$

- Let $\{X_i : i \geq 1\}$ be i.i.d. with $E[X_1] < 0$, $S_0 = 0$, and $S_n = \sum_{i=1}^n X_i$.
- If $W = \sup_{n \geq 0} S_n$, then $W < \infty$ a.s. and the events $\{W > t\}$ are rare for $t > 0$.
- Goal: estimate $P(W > t)$ efficiently. $\dfrac{\mathsf{SD}(1(W > t))}{P(W > t)} = \dfrac{\sqrt{P(W > t)P(W \leq t)}}{P(W > t)} \to \infty$.

Background: Exponential tilting for random walks on $\mathbb{R}$

- Let $\{X_i : i \geq 1\}$ be i.i.d. with $E[X_1] < 0$, $S_0 = 0$, and $S_n = \sum_{i=1}^n X_i$.

- If $W = \sup_{n \geq 0} S_n$, then $W < \infty$ a.s. and the events $\{W > t\}$ are rare for $t > 0$.

- Goal: estimate $P(W > t)$ efficiently. $\dfrac{\mathsf{SD}(1(W > t))}{P(W > t)} = \dfrac{\sqrt{P(W > t)P(W \leq t)}}{P(W > t)} \to \infty.$

- Let $m(\theta) = E\left[e^{\theta X_1}\right]$. When X_1 satisfies Cramér-Lundberg conditions

$$m(\alpha) = 1, \qquad m'(\alpha) = E\left[X_1 e^{\alpha X_1}\right] \in (0, \infty) \qquad \text{for some } \alpha > 0,$$

Background: Exponential tilting for random walks on $\mathbb{R}$

- Let $\{X_i : i \geq 1\}$ be i.i.d. with $E[X_1] < 0$, $S_0 = 0$, and $S_n = \sum_{i=1}^n X_i$.
- If $W = \sup_{n \geq 0} S_n$, then $W < \infty$ a.s. and the events $\{W > t\}$ are rare for $t > 0$.
- Goal: estimate $P(W > t)$ efficiently. $\dfrac{\mathsf{SD}(1(W > t))}{P(W > t)} = \dfrac{\sqrt{P(W > t)P(W \leq t)}}{P(W > t)} \to \infty.$
- Let $m(\theta) = E\left[e^{\theta X_1}\right]$. When X_1 satisfies Cramér-Lundberg conditions

$$m(\alpha) = 1, \qquad m'(\alpha) = E\left[X_1 e^{\alpha X_1}\right] \in (0, \infty) \qquad \text{for some } \alpha > 0,$$

- $P(W > t)$ can be sampled by exponential tilting: $\{e^{\alpha S_n} : n \geq 0\}$ is a mean-1 martingale and

$$\widetilde{P}(A) := E\left[1_A e^{\alpha S_n}\right], \qquad A \in \sigma(X_1, \ldots, X_n)$$

induces a probability measure $\widetilde{P}$ on $\sigma(X_i : i \geq 1)$.

Background: Exponential tilting for random walks on $\mathbb{R}$

- Let $\{X_i : i \geq 1\}$ be i.i.d. with $E[X_1] < 0$, $S_0 = 0$, and $S_n = \sum_{i=1}^n X_i$.
- If $W = \sup_{n \geq 0} S_n$, then $W < \infty$ a.s. and the events $\{W > t\}$ are rare for $t > 0$.
- Goal: estimate $P(W > t)$ efficiently. $\frac{\mathsf{SD}(1(W > t))}{P(W > t)} = \frac{\sqrt{P(W > t)P(W \leq t)}}{P(W > t)} \to \infty$.
- Let $m(\theta) = E\left[e^{\theta X_1}\right]$. When X_1 satisfies Cramér-Lundberg conditions

$$m(\alpha) = 1, \qquad m'(\alpha) = E\left[X_1 e^{\alpha X_1}\right] \in (0, \infty) \qquad \text{for some } \alpha > 0,$$

- $P(W > t)$ can be sampled by exponential tilting: $\{e^{\alpha S_n} : n \geq 0\}$ is a mean-1 martingale and

$$\widetilde{P}(A) := E\left[1_A e^{\alpha S_n}\right], \qquad A \in \sigma(X_1, \ldots, X_n)$$

  induces a probability measure $\widetilde{P}$ on $\sigma(X_i : i \geq 1)$.
- Since $\widetilde{E}[X_1] = E\left[X_1 e^{\alpha X_1}\right] = m'(\alpha) > 0$, $\widetilde{P}(W > t) = 1$.
- If $\tau(t) = \inf\{n > 0 : S_n > t\}$, $P(W > t) = \widetilde{E}\left[e^{-\alpha S_{\tau(t)}}\right]$.

Can we do this for branching random walks?

- Motivation: The stationary waiting time W of a multi-server queue with certain synchronization requirements solves

$$W \stackrel{D}{=} \left(\max_{1 \leq i \leq N} (X_i + W_i) \right)^+, \qquad \{W_i\} \sim_{\text{iid}} W, \text{ indep. of } (N, \{X_i\}).$$

Can we do this for branching random walks?

- Motivation: The stationary waiting time W of a multi-server queue with certain synchronization requirements solves

$$W \stackrel{D}{=} \left(\max_{1 \le i \le N} (X_i + W_i) \right)^+, \qquad \{W_i\} \sim_{\text{iid}} W, \text{ indep. of } (N, \{X_i\}).$$

- W has the law of the all-time maximum of a branching random walk with offspring distribution N and increment distributions $X_1, \ldots, X_N$.

Can we do this for branching random walks?

- Motivation: The stationary waiting time W of a multi-server queue with certain synchronization requirements solves

$$W \stackrel{D}{=} \left(\max_{1 \le i \le N} (X_i + W_i) \right)^+, \qquad \{W_i\} \sim_{\text{iid}} W, \text{ indep. of } (N, \{X_i\}).$$

- W has the law of the all-time maximum of a branching random walk with offspring distribution N and increment distributions $X_1, \ldots, X_N$.
- Let

$$\rho(\theta) = E\left[\sum_{i=1}^{N} e^{\theta X_i} \right],$$

and suppose $(N, \{X_i\})$ satisfies the Cramér-Lundberg-type conditions

$$\rho(\alpha) = 1, \quad \rho'(\alpha) = E\left[\sum_{i=1}^{N} X_i e^{\alpha X_i} \right] \in (0, \infty) \quad \text{for some } \alpha > 0,$$

and $\rho(\beta) < 1$ for some $\beta \in (0, \alpha)$.Also, $P(N \ge 1) = 1$.

Can we do this for branching random walks?

- Motivation: The stationary waiting time W of a multi-server queue with certain synchronization requirements solves

$$W \stackrel{D}{=} \left(\max_{1 \le i \le N} (X_i + W_i) \right)^+, \qquad \{W_i\} \sim_{\text{iid}} W, \text{ indep. of } (N, \{X_i\}).$$

- W has the law of the all-time maximum of a branching random walk with offspring distribution N and increment distributions $X_1, \dots, X_N$.
- Let

$$\rho(\theta) = E\left[\sum_{i=1}^{N} e^{\theta X_i} \right],$$

and suppose $(N, \{X_i\})$ satisfies the Cramér-Lundberg-type conditions

$$\rho(\alpha) = 1, \quad \rho'(\alpha) = E\left[\sum_{i=1}^{N} X_i e^{\alpha X_i} \right] \in (0, \infty) \quad \text{for some } \alpha > 0,$$

and $\rho(\beta) < 1$ for some $\beta \in (0, \alpha)$.Also, $P(N \ge 1) = 1$.
- By Jensen's inequality,

$$E\left[\max_{1 \le i \le N} X_i \right] \le \frac{1}{\beta} \log E\left[\max_{1 \le i \le N} e^{\beta X_i} \right] \le \frac{1}{\beta} \log \rho(\beta) < 0.$$

Constructing W

- Construct W as follows: Let $U = \{\emptyset\} \cup \{\boldsymbol{i} = (i_1, i_2, \ldots, i_k) : i_j \in \mathbb{N}_+, k \geq 1\}$ be strings of positive integers, endowed with length-lexicographic order, and let

$$\left\{(N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}_{j\geq 1}) : \boldsymbol{i} \in U\right\}$$

  be i.i.d. copies of $(N, \{X_j\})$, where $(\boldsymbol{i}, j) = (i_1, \ldots, i_k, j)$ when $|\boldsymbol{i}| = k$.

Constructing W

- Construct W as follows: Let $U = \{\emptyset\} \cup \{\boldsymbol{i} = (i_1, i_2, \ldots, i_k) : i_j \in \mathbb{N}_+, k \geq 1\}$ be strings of positive integers, endowed with length-lexicographic order, and let

$$\left\{(N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}_{j\geq 1}) : \boldsymbol{i} \in U\right\}$$

  be i.i.d. copies of $(N, \{X_j\})$, where $(\boldsymbol{i}, j) = (i_1, \ldots, i_k, j)$ when $|\boldsymbol{i}| = k$.

- $\{N_{\boldsymbol{i}}\}$ determines the structure of a random tree $\mathcal{T}$.

Constructing W

- Construct W as follows: Let $U = \{\emptyset\} \cup \{\boldsymbol{i} = (i_1, i_2, \ldots, i_k) : i_j \in \mathbb{N}_+, k \geq 1\}$ be strings of positive integers, endowed with length-lexicographic order, and let

$$\left\{(N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}_{j\geq 1}) : \boldsymbol{i} \in U\right\}$$

  be i.i.d. copies of $(N, \{X_j\})$, where $(\boldsymbol{i}, j) = (i_1, \ldots, i_k, j)$ when $|\boldsymbol{i}| = k$.

- $\{N_{\boldsymbol{i}}\}$ determines the structure of a random tree $\mathcal{T}$.

- For each $\boldsymbol{i} = (i_1, \ldots, i_k) \in \mathcal{T}$, let

$$S_\emptyset = 0, \qquad S_{\boldsymbol{i}} = \sum_{j=1}^{k} X_{(i_1,\ldots,i_j)}.$$

- Then, $W = \sup_{\boldsymbol{i}\in\mathcal{T}} S_{\boldsymbol{i}}$.

Constructing W

- Construct W as follows: Let $U = \{\emptyset\} \cup \{\boldsymbol{i} = (i_1, i_2, \ldots, i_k) : i_j \in \mathbb{N}_+, k \geq 1\}$ be strings of positive integers, endowed with length-lexicographic order, and let

$$\left\{(N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}_{j\geq 1}) : \boldsymbol{i} \in U\right\}$$

  be i.i.d. copies of $(N, \{X_j\})$, where $(\boldsymbol{i}, j) = (i_1, \ldots, i_k, j)$ when $|\boldsymbol{i}| = k$.

- $\{N_{\boldsymbol{i}}\}$ determines the structure of a random tree $\mathcal{T}$.

- For each $\boldsymbol{i} = (i_1, \ldots, i_k) \in \mathcal{T}$, let

$$S_\emptyset = 0, \qquad S_{\boldsymbol{i}} = \sum_{j=1}^{k} X_{(i_1,\ldots,i_j)}.$$

- Then, $W = \sup_{\boldsymbol{i} \in \mathcal{T}} S_{\boldsymbol{i}}$.

- We want to estimate $P(W > t)$.

Spine change of measure

- Define a random path in $\mathcal{T}$: let $\boldsymbol{J}_0 = \emptyset$, and for each $k \geq 0$,

$$\boldsymbol{J}_{k+1} = (\boldsymbol{J}_k, i) \quad \text{w.p.} \quad \frac{e^{\alpha X_{(\boldsymbol{J}_k, i)}}}{\sum_{j=1}^{N_{\boldsymbol{J}_k}} e^{\alpha X_{(\boldsymbol{J}_k, j)}}}.$$

Spine change of measure

- Define a random path in $\mathcal{T}$: let $\boldsymbol{J}_0 = \emptyset$, and for each $k \geq 0$,

$$\boldsymbol{J}_{k+1} = (\boldsymbol{J}_k, i) \quad \text{w.p.} \quad \frac{e^{\alpha X_{(\boldsymbol{J}_k, i)}}}{\sum_{j=1}^{N_{\boldsymbol{J}_k}} e^{\alpha X_{(\boldsymbol{J}_k, j)}}}.$$

- Then define $L_0 = 1$, and for $n \geq 1$,

$$L_n = \prod_{k=0}^{n-1} \sum_{i=1}^{N_{\boldsymbol{J}_k}} e^{\alpha X_{(\boldsymbol{J}_k, i)}},$$

  which is a mean-1 martingale with respect to

$$\mathcal{G}_n = \sigma\left(\{(N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}) : \boldsymbol{i} \in \mathcal{T}, |\boldsymbol{i}| < n\} \cup \{\boldsymbol{J}_k : k < n\}\right).$$

- This induces the measure

$$\widetilde{P}(A) = E\left[1_A L_n\right], \qquad A \in \mathcal{G}_n,$$

  which extends to all of $\sigma\left(\cup_{n=0}^{\infty} \mathcal{G}_n\right)$.

Spine change of measure

Theorem

For $\boldsymbol{i} \in \mathcal{T}$ *with* $|\boldsymbol{i}| = k$,

$$\widetilde{P}\left((N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}) \in \cdot \,\middle|\, \boldsymbol{i} \neq \boldsymbol{J}_k\right) = P\left((N, \{X_j\}) \in \cdot\right), \quad \textit{and}$$

$$\widetilde{P}\left((N_{\boldsymbol{i}}, \{X_{(\boldsymbol{i},j)}\}) \in \cdot \,\middle|\, \boldsymbol{i} = \boldsymbol{J}_k\right) = E\left[1\left((N, \{X_j\}) \in \cdot\right) \sum_{i=1}^{N} e^{\alpha X_i}\right].$$

If $\hat{X}_k = X_{\boldsymbol{J}_k}$ *for each* k, *then* $\{\hat{X}_i : i \geq 0\}$ *are i.i.d. with CDF*

$$G(x) = E\left[\sum_{i=1}^{N} e^{\alpha X_i} 1(X_i \leq x)\right].$$

In particular,

$$\mu := \widetilde{E}\left[\hat{X}_1\right] = E\left[\sum_{i=1}^{N} X_i e^{\alpha X_i}\right] = \rho'(\alpha) > 0.$$

Spine change of measure

- Let $V_k = S_{\boldsymbol{J}_k} = \hat{X}_1 + \cdots + \hat{X}_k$, and define

$$\gamma(t) = \inf\{\boldsymbol{i} \in \mathcal{T} : S_{\boldsymbol{i}} > t\}, \qquad \tau(t) = \inf\{n > 0 : V_n > t\}.$$

  Then,

$$P(W > t) = P\left(|\gamma(t)| < \infty\right) = \widetilde{E}\left[1(\gamma(t) = \boldsymbol{J}_{\tau(t)})e^{-\alpha V_{\tau(t)}}\right].$$

Spine change of measure

- Let $V_k = S_{\boldsymbol{J}_k} = \hat{X}_1 + \cdots + \hat{X}_k$, and define

$$\gamma(t) = \inf\{\boldsymbol{i} \in \mathcal{T} : S_{\boldsymbol{i}} > t\}, \qquad \tau(t) = \inf\{n > 0 : V_n > t\}.$$

  Then,

$$P(W > t) = P\left(|\gamma(t)| < \infty\right) = \widetilde{E}\left[1(\gamma(t) = \boldsymbol{J}_{\tau(t)})e^{-\alpha V_{\tau(t)}}\right].$$

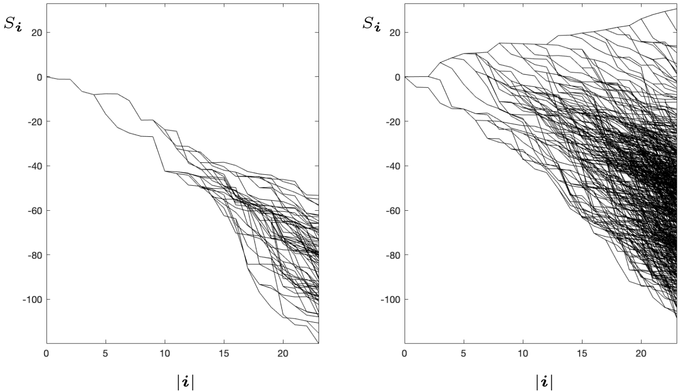

Figure: A branching random walk simulated under both P (left) and $\widetilde{P}$ (right).

Properties of the algorithm

- $P(W > t)$ can be estimated unbiasedly by sampling

$$Z(t) = 1(\gamma(t) = \boldsymbol{J}_{\tau(t)})e^{-\alpha V_{\tau(t)}}$$

  under $\widetilde{P}$:

  1. Generate a branching random walk and $\{\boldsymbol{J}_k\}$ until the first node $\boldsymbol{i} = \gamma(t)$ where $S_{\boldsymbol{i}} > t$.
  2. If $\boldsymbol{i} \in \{\boldsymbol{J}_k\}$, set $Z(t) = e^{-\alpha V_{\tau(t)}}$. Else, set $Z(t) = 0$.

- When $N \equiv 1$, $Z(t) = e^{-\alpha V_{\tau(t)}}$ is just the estimator from exponential tilting.

Properties of the algorithm

- $P(W > t)$ can be estimated unbiasedly by sampling

$$Z(t) = 1(\gamma(t) = \boldsymbol{J}_{\tau(t)})e^{-\alpha V_{\tau(t)}}$$

under $\widetilde{P}$:

1. Generate a branching random walk and $\{\boldsymbol{J}_k\}$ until the first node $\boldsymbol{i} = \gamma(t)$ where $S_{\boldsymbol{i}} > t$.
2. If $\boldsymbol{i} \in \{\boldsymbol{J}_k\}$, set $Z(t) = e^{-\alpha V_{\tau(t)}}$. Else, set $Z(t) = 0$.

- When $N \equiv 1$, $Z(t) = e^{-\alpha V_{\tau(t)}}$ is just the estimator from exponential tilting.

- $\tau(t) \sim t/\mu$ as $t \to \infty$ $\widetilde{P}$-a.s.

Properties of the algorithm

- $P(W > t)$ can be estimated unbiasedly by sampling

$$Z(t) = 1(\gamma(t) = \boldsymbol{J}_{\tau(t)})e^{-\alpha V_{\tau(t)}}$$

under $\widetilde{P}$:

1. Generate a branching random walk and $\{\boldsymbol{J}_k\}$ until the first node $\boldsymbol{i} = \gamma(t)$ where $S_{\boldsymbol{i}} > t$.
2. If $\boldsymbol{i} \in \{\boldsymbol{J}_k\}$, set $Z(t) = e^{-\alpha V_{\tau(t)}}$. Else, set $Z(t) = 0$.

- When $N \equiv 1$, $Z(t) = e^{-\alpha V_{\tau(t)}}$ is just the estimator from exponential tilting.
- $\tau(t) \sim t/\mu$ as $t \to \infty$ $\widetilde{P}$-a.s.
- $Z(t)$ has bounded relative error:

$$\limsup_{t\to\infty} \frac{\widetilde{\mathrm{Var}}\,(Z(t))}{P(W > t)^2} < \infty,$$

where $\widetilde{\mathrm{Var}}$ denotes variance under $\widetilde{P}$.

Remarks

1. *Computational complexity: requires*

$$\approx n(E[N])^{t/\mu}$$

   copies of $(N, \{X_i\})$ *to produce a sample of* $Z(t)$ *of size* n.
2. *The algorithm is virtually guaranteed to terminate on the spine.*

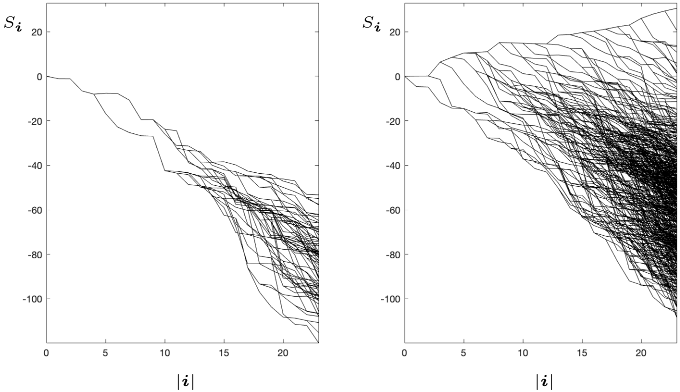

Figure: A branching random walk simulated under both P (left) and $\widetilde{P}$ (right).

Efficient Hybrid Estimation

- For each k, let

$$F_k(x) = P\left(W^{(k)} \le x\right), \qquad W^{(k)} = \max_{|\boldsymbol{i}| \le k} S_{\boldsymbol{i}}.$$

Efficient Hybrid Estimation

- For each k, let

$$F_k(x) = P\left(W^{(k)} \leq x\right), \qquad W^{(k)} = \max_{|\boldsymbol{i}| \leq k} S_{\boldsymbol{i}}.$$

Theorem

For any $t > 0$,

$$P(W > t) = \widetilde{E}\left[e^{-\alpha V_{\tau(t)}} \prod_{k=1}^{\tau(t)} \prod_{\boldsymbol{i} \in B_k^{\prec}} F_{\tau(t)-k}(t - S_{\boldsymbol{i}}) \prod_{\boldsymbol{j} \in B_k^{\succ}} F_{\tau(t)-k-1}(t - S_{\boldsymbol{j}})\right],$$

where

$$B_k^{\prec} = \{\boldsymbol{i} \in \mathcal{T} : \boldsymbol{i} \text{ has parent } \boldsymbol{J}_{k-1} \text{ and } \boldsymbol{i} \prec \boldsymbol{J}_k\}$$
$$B_k^{\succ} = \{\boldsymbol{i} \in \mathcal{T} : \boldsymbol{i} \text{ has parent } \boldsymbol{J}_{k-1} \text{ and } \boldsymbol{i} \succ \boldsymbol{J}_k\},$$

and with the conventions that $F_{-1}(x) \equiv 1$ and $\prod_{i=1}^{0} x_i \equiv 1$.

Efficient Hybrid Estimation

- For each k, let

$$F_k(x) = P\left(W^{(k)} \le x\right), \qquad W^{(k)} = \max_{|\boldsymbol{i}| \le k} S_{\boldsymbol{i}}.$$

Theorem

For any $t > 0$,

$$P(W > t) = \widetilde{E}\left[e^{-\alpha V_{\tau(t)}} \prod_{k=1}^{\tau(t)} \prod_{\boldsymbol{i} \in B_k^{\prec}} F_{\tau(t)-k}(t - S_{\boldsymbol{i}}) \prod_{\boldsymbol{j} \in B_k^{\succ}} F_{\tau(t)-k-1}(t - S_{\boldsymbol{j}})\right],$$

where

$$B_k^{\prec} = \{\boldsymbol{i} \in \mathcal{T} : \boldsymbol{i} \text{ has parent } \boldsymbol{J}_{k-1} \text{ and } \boldsymbol{i} \prec \boldsymbol{J}_k\}$$
$$B_k^{\succ} = \{\boldsymbol{i} \in \mathcal{T} : \boldsymbol{i} \text{ has parent } \boldsymbol{J}_{k-1} \text{ and } \boldsymbol{i} \succ \boldsymbol{J}_k\},$$

and with the conventions that $F_{-1}(x) \equiv 1$ and $\prod_{i=1}^{0} x_i \equiv 1$.

- If we had estimators $\{\hat{F}_k\}$ for $\{F_k\}$, this suggests the estimator

$$\hat{Z}(t) = e^{-\alpha V_{\tau(t)}} \prod_{k=1}^{\tau(t)} \prod_{\boldsymbol{i} \in B_k^{\prec}} \hat{F}_{\tau(t)-k}(t - S_{\boldsymbol{i}}) \prod_{\boldsymbol{j} \in B_k^{\succ}} \hat{F}_{\tau(t)-k-1}(t - S_{\boldsymbol{j}})$$

sampled under $\widetilde{P}$.

Estimating $\{F_k : k \geq 0\}$

- The *population dynamics algorithm* uses bootstrapping to generate approximate samples of $W^{(k)}$ under P of a given sample size.
- For $m, K \in \mathbb{N}_+$, the samples

$$\left\{\hat{W}_1^{(j,m)}, \ldots, \hat{W}_m^{(j,m)}\right\}, \qquad j \leq K$$

approximately from the laws of $W^{(j)}$, $j \leq K$, can be generated from $m \cdot K$ copies of $(N, \{X_i\})$.

Estimating $\{F_k : k \geq 0\}$

- The *population dynamics algorithm* uses bootstrapping to generate approximate samples of $W^{(k)}$ under P of a given sample size.
- For $m, K \in \mathbb{N}_+$, the samples

$$\left\{\hat{W}_1^{(j,m)}, \dots, \hat{W}_m^{(j,m)}\right\}, \qquad j \leq K$$

  approximately from the laws of $W^{(j)}$, $j \leq K$, can be generated from $m \cdot K$ copies of $(N, \{X_i\})$.
- To approximate $\{F_k\}$, we can use

$$F_k(x) \approx \hat{F}_{k \wedge K, m}(x) = \frac{1}{m} \sum_{i=1}^{m} 1\left(\hat{W}_i^{(k \wedge K, m)} \leq x\right).$$

Estimating $\{F_k : k \geq 0\}$

- The *population dynamics algorithm* uses bootstrapping to generate approximate samples of $W^{(k)}$ under P of a given sample size.
- For $m, K \in \mathbb{N}_+$, the samples

$$\left\{\hat{W}_1^{(j,m)}, \ldots, \hat{W}_m^{(j,m)}\right\}, \qquad j \leq K$$

  approximately from the laws of $W^{(j)}$, $j \leq K$, can be generated from $m \cdot K$ copies of $(N, \{X_i\})$.
- To approximate $\{F_k\}$, we can use

$$F_k(x) \approx \hat{F}_{k\wedge K,m}(x) = \frac{1}{m}\sum_{i=1}^{m} 1\left(\hat{W}_i^{(k\wedge K,m)} \leq x\right).$$

Remark

Since $\tau(t) \sim t/\mu$ $\widetilde{P}$-a.s., generating $\{\hat{F}_k : k \leq K\}$ then producing a sample of size n of $\hat{Z}(t)$ requires about

$$mK + \frac{nt}{\mu}$$

copies of $(N, \{X_i\})$. In particular, there is no dependence on $E[N]$!

Consistency of $\hat{Z}(t)$

Proposition

For $K, m \in \mathbb{N}$, *let*

$$\hat{Z}(t) = e^{-\alpha V_{\tau(t)}} \prod_{k=1}^{\tau(t)} \prod_{\boldsymbol{i} \in B_k^{\prec}} \hat{F}_{(\tau(t)-k)\wedge K, m}(t - S_{\boldsymbol{i}}) \prod_{\boldsymbol{j} \in B_k^{\succ}} \hat{F}_{(\tau(t)-k-1)\wedge K, m}(t - S_{\boldsymbol{j}}),$$

and suppose that $\rho(\beta) < 1$ *and*

$$\widetilde{E}[N] = E\left[N \sum_{i=1}^{N} e^{\alpha X_i} \right] < \infty.$$

Then,

$$\limsup_{t\to\infty} \left| \frac{\widetilde{E}\left[\hat{Z}(t)\right] - P(W > t)}{P(W > t)} \right| \leq C \left(\rho(\beta)^{K/2} + m^{-1/4} \right)$$

for $C \in (0, \infty)$.

Numerical examples ($K = 20$, $m = 5000$, $n = 5000$)

1. $\{X_i\} \sim_{\text{iid}}$ Exp(5) − Exp$(1/4)$, $N \sim$ Ber$(1/2) + 2$, N independent of $\{X_i\}$.

t	$Z(t)$	Time (sec.)	$\hat{Z}(t)$	Time	Total time	Rel. bias
1	1.8304e-03	5.95	1.8719e-03	2.13	6.20	0.0227
2	2.4038e-05	76.93	2.4425e-05	2.19	6.25	0.0161
3	3.1800e-07	133.84	3.3133e-07	2.81	6.88	0.0419
4	4.4749e-09	700.01	4.3597e-09	3.14	7.21	0.0257
5	5.7474e-11	1643.12	5.6500e-11	3.60	7.67	0.0170
6	7.3680e-13	5761.24	7.6879e-13	4.06	8.13	0.0434
7	1.0007e-14	46447.66	1.0447e-14	4.85	8.92	0.0440

Numerical examples ($K = 20$, $m = 5000$, $n = 5000$)

1. $\{X_i\} \sim_{\text{iid}}$ Exp(5) − Exp$(1/4)$, $N \sim$ Ber$(1/2) + 2$, N independent of $\{X_i\}$.

t	$Z(t)$	Time (sec.)	$\hat{Z}(t)$	Time	Total time	Rel. bias
1	1.8304e-03	5.95	1.8719e-03	2.13	6.20	0.0227
2	2.4038e-05	76.93	2.4425e-05	2.19	6.25	0.0161
3	3.1800e-07	133.84	3.3133e-07	2.81	6.88	0.0419
4	4.4749e-09	700.01	4.3597e-09	3.14	7.21	0.0257
5	5.7474e-11	1643.12	5.6500e-11	3.60	7.67	0.0170
6	7.3680e-13	5761.24	7.6879e-13	4.06	8.13	0.0434
7	1.0007e-14	46447.66	1.0447e-14	4.85	8.92	0.0440

2. $\{X_i\} \sim_{\text{iid}}$ N$(-5, 1)$, $N \sim$ Unif$\{1, 2, \ldots, 99\}$, N independent of $\{X_i\}$. $E[N] = 50$.

t	$\hat{Z}(t)$	Time (sec.)	Total time
1	3.6419e-08	17.25	36.46
2	6.1537e-11	15.80	35.01
3	3.1472e-14	23.89	43.10
4	5.4582e-18	33.05	52.26
5	4.0702e-22	33.89	53.10

References

- B. Basrak, M. Conroy, M. Olvera-Cravioto, and Z. Palmowski. Importance sampling for maxima on trees. *arXiv:2004.08966*, 2020.

- M. Conroy and M. Olvera-Cravioto. Efficient hybrid estimation for maxima on trees. In preparaton.