Efficient Monte Carlo for Risk Analysis

Jose Blanchet

Columbia Department of IEOR

What is this talk about?

- **General theme of this line of research**

What is this talk about?

- **General theme of this line of research**
- **Design: Light Tails**

What is this talk about?

- **General theme of this line of research**
- **Design: Light Tails**
- **Design: Heavy Tails**

Optimal design of rare event simulation algorithms

- Prof. Varadhan's Abel prize citation on large deviations theory: **"...It has greatly expanded our ability to use computers to analyze rare events."**

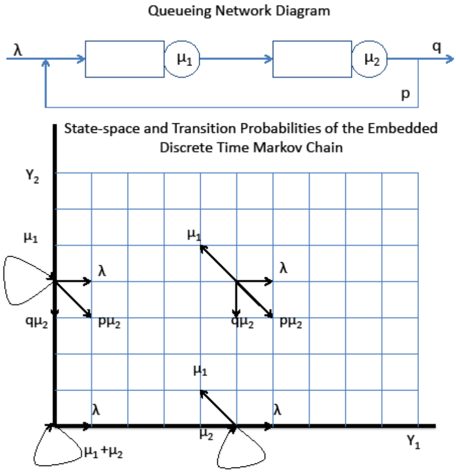
Optimal design of rare event simulation algorithms

- Prof. Varadhan's Abel prize citation on large deviations theory: **"...It has greatly expanded our ability to use computers to analyze rare events."**
- Goal of this line of research: **To investigate exactly HOW?**

Before I Answer How: Why Would Anybody Care?

A fast computational engine
enhances our ability to quantify uncertainty
via sensitivity analysis & stress tests...

Example: A Simple Stochastic Network



Example: A Simple Stochastic Network

Questions of interest:

How would the system perform IF TOTAL POPULATION reaches n inside a busy period?

How likely is this event under different parameters / designs?

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
- Time to estimate with 10% precision $\approx 10^7/1000 = 10^4$ secs ≈ 2.7 hours

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
- Time to estimate with 10% precision $\approx 10^7/1000 = 10^4$ secs ≈ 2.7 hours
- What if you want to do sensitivity analysis for a wide range of parameters?

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
- Time to estimate with 10% precision $\approx 10^7/1000 = 10^4$ secs ≈ 2.7 hours
- What if you want to do sensitivity analysis for a wide range of parameters?
- What about different network designs?

Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Total population overflow $= n = 10$
- Overflow probability $\approx 10^{-\mathbf{5}}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
- Time to estimate with 10% precision $\approx 10^7/1000 = 10^4$ secs ≈ 2.7 hours
- What if you want to do sensitivity analysis for a wide range of parameters?
- What about different network designs?
- Next demo based on **Blanchet '10: Optimal Sampling of Overflow Paths in Jackson Networks.**

Rare Events in Networks

- Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?

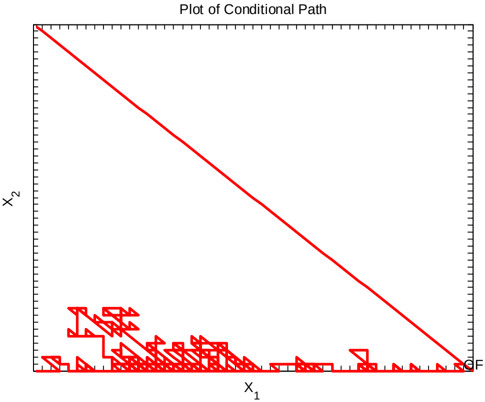
Rare Events in Networks

- Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?
- Say $\lambda = .2$, $\mu_1 = .3$, $\mu_2 = .5$... How can total content reach **50**?

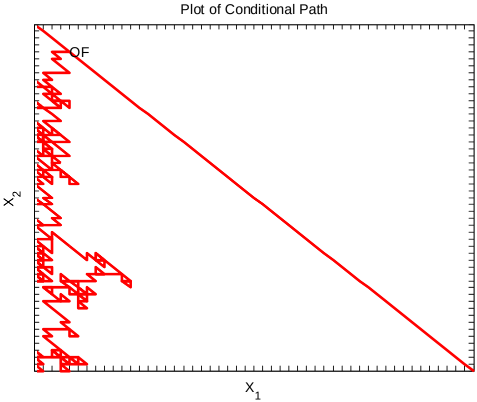
- Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?
- Say $\lambda = .2$, $\mu_1 = .3$, $\mu_2 = .5$... How can total content reach **50**?
- Naive Monte Carlo takes $\approx$ 115 **days**

- Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?
- Say $\lambda = .2$, $\mu_1 = .3$, $\mu_2 = .5$... How can total content reach **50**?
- Naive Monte Carlo takes $\approx$ 115 **days**
- Each picture below took $\approx$ .01 **seconds** (generated with Blanchet (2010) algorithm).

$\lambda = .1$, $\mu_1 = .3$, $\mu_2 = .5$

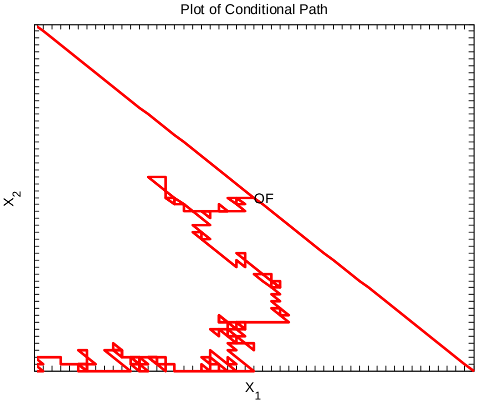


$\lambda = .1,\ \mu_1 = .5,\ \mu_2 = .2$

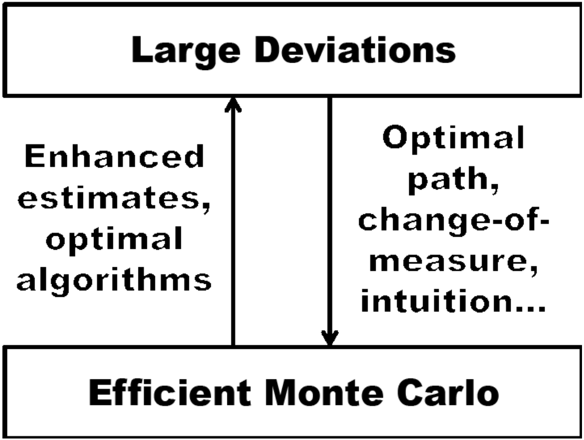


Rare Events in Networks

$\lambda = .1,\ \mu_1 = .4,\ \mu_2 = .4$



Large Deviations and Monte Carlo: A Conceptual Diagram



Performance analysis of rare event simulation algorithms

- $P(A_n) = \exp(-nI + o(n))$ as $n \nearrow \infty$ for $I > 0$.

Performance analysis of rare event simulation algorithms

- $P(A_n) = \exp(-nI + o(n))$ as $n \nearrow \infty$ for $I > 0$.
- **Asymptotic weak optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \mathbf{Com}(n) = \exp(o(n)).$$

Performance analysis of rare event simulation algorithms

- $P(A_n) = \exp(-nI + o(n))$ as $n \nearrow \infty$ for $I > 0$.
- **Asymptotic weak optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \mathbf{Com}(n) = \exp(o(n)).$$

- **Asymptotic strong optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \mathbf{Com}(n) = O(1).$$

Performance analysis of rare event simulation algorithms

- $P(A_n) = \exp(-nI + o(n))$ as $n \nearrow \infty$ for $I > 0$.
- **Asymptotic weak optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \textbf{Com}(n) = \exp(o(n)).$$

- **Asymptotic strong optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \textbf{Com}(n) = O(1).$$

- Sufficient number of replications of Z_n to get ε relative error with $1-\delta$ confidence:

$$\varepsilon^{-2}\delta^{-1}\textbf{Com}(n)$$

Performance analysis of rare event simulation algorithms

- $P(A_n) = \exp(-nI + o(n))$ as $n \nearrow \infty$ for $I > 0$.
- **Asymptotic weak optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

$$\frac{EZ_n^2}{P(A_n)^2} = \mathbf{Com}(n) = \exp(o(n)).$$

- **Asymptotic strong optimality:** Z_n satisfies $EZ_n = P(A_n)$ and

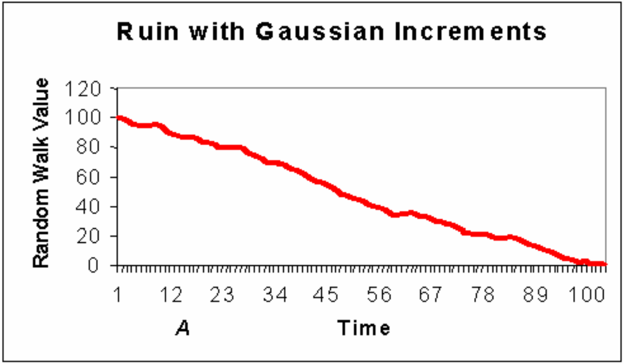
$$\frac{EZ_n^2}{P(A_n)^2} = \mathbf{Com}(n) = O(1).$$

- Sufficient number of replications of Z_n to get ε relative error with $1 - \delta$ confidence:

$$\varepsilon^{-2}\delta^{-1}\mathbf{Com}(n)$$

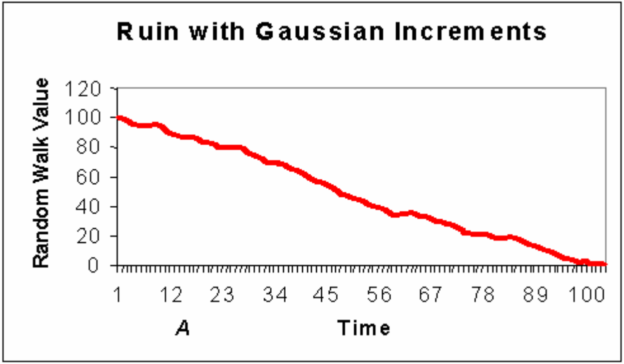
- Total cost $\mathbf{TC}(n) = \mathbf{Com}(n) \times$ *Cost per replication*

Insurance Reserves (Random Walk): Light Tails (e.g. Car Insurance)



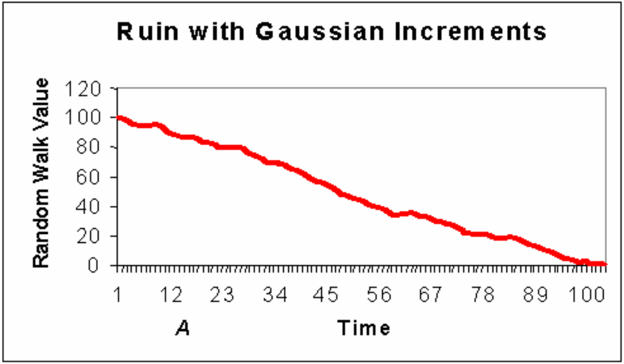
- ORIGINAL increments are *Gaussian* drift $+1$ and variance $+1$

Insurance Reserves (Random Walk): Light Tails (e.g. Car Insurance)



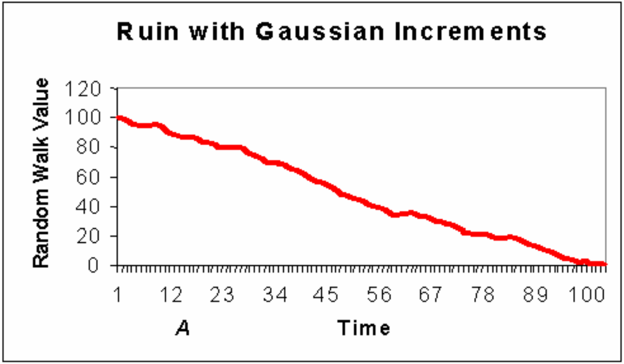
- ORIGINAL increments are *Gaussian* drift $+1$ and variance $+1$
- Picture represents the reserve CONDITIONAL ON RUIN!

Insurance Reserves (Random Walk): Light Tails (e.g. Car Insurance)



- ORIGINAL increments are *Gaussian* drift $+1$ and variance $+1$
- Picture represents the reserve CONDITIONAL ON RUIN!
- Light tails: Exponential, Gamma, Gaussian, mixtures of these, etc.

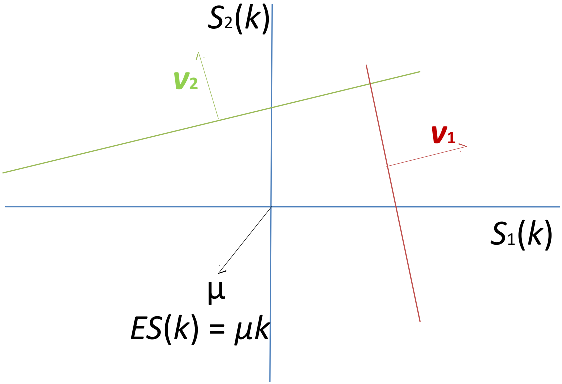
Insurance Reserves (Random Walk): Light Tails (e.g. Car Insurance)



- ORIGINAL increments are *Gaussian* drift $+1$ and variance $+1$
- Picture represents the reserve CONDITIONAL ON RUIN!
- Light tails: Exponential, Gamma, Gaussian, mixtures of these, etc.
- Generated with Siegmund's 76 algorithm

Stylized Example: Two Dimensional Ruin Problem

- Two dimensional random walk
- $A = \{s : v_2^T s \geq 1\}$ and $B = \{s : v_1^T s \geq 1\}$



- **Efficiently estimate as** $n \nearrow \infty$

$$u_n(0) = P_0[S_k / n \textbf{ hits } A \textbf{ OR } B \textbf{ Eventually}]$$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}, \quad Y_k$'s are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$
- Note $EZ_k^{(1)} = v_1^T \mu < 0$ and $EZ_k^{(2)} = v_2^T \mu < 0$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + \ldots + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$
- Note $EZ_k^{(1)} = v_1^T \mu < 0$ and $EZ_k^{(2)} = v_2^T \mu < 0$
- Assume there are $\theta_1^*, \theta_2^* > 0$ such that

$$
\begin{aligned}
E\exp(\theta_1^* Z_k^{(1)}) &= 1, \quad E\exp(\theta_2^* Z_k^{(2)}) = 1 \\
E[\exp(\theta_1^* Z_k^{(1)}) Z_k^{(1)}] &< \infty, \quad E[\exp(\theta_2^* Z_k^{(2)}) Z_k^{(2)}] < \infty
\end{aligned}
$$

Large Deviations for the Stylized Example

- Then

$$
\begin{aligned}
u_n(x) &= P_x[W_n(t) \ \textbf{hits } A \textbf{ OR } B] \\
&\sim c_1 \exp(-n\theta_1^*(1 - v_1^T x)) + c_2 \exp(-n\theta_2^*(1 - v_2^T x)) \\
&= \exp(-nh(x) + o(n))
\end{aligned}
$$

as $n \nearrow \infty$, where

$$
h(x) = \min[\theta_1^*(1 - v_1^T x), \theta_2^*(1 - v_2^T x)].
$$

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda [\lambda^T x - \psi(\lambda)]$

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda[\lambda^T x - \psi(\lambda)]$
- $f_\lambda(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ $\longleftarrow$ Controls

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda[\lambda^T x - \psi(\lambda)]$
- $f_\lambda(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <— Controls
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_\lambda E[e^{-\lambda^T Y + \psi(\lambda)} C_n(w + Y/n)]$$

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_{\lambda}[\lambda^T x - \psi(\lambda)]$
- $f_{\lambda}(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <— Controls
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_{\lambda} E[e^{-\lambda^T Y + \psi(\lambda)} C_n(w + Y/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - n[g(w+Y/n) - g(w)]}] \\ &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - \partial g(w)}] = \min_{\lambda}[\psi(\lambda) + \psi(-\lambda - \partial g(w))] \end{aligned}$$

Control-theoretic Approach by Dupuis-Wang

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_{\lambda}[\lambda^T x - \psi(\lambda)]$
- $f_{\lambda}(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <— Controls
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_{\lambda} E[e^{-\lambda^T Y + \psi(\lambda)} C_n(w + Y/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - n[g(w+Y/n) - g(w)]}] \\ &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - \partial g(w)}] = \min_{\lambda}[\psi(\lambda) + \psi(-\lambda - \partial g(w))] \end{aligned}$$

- GET so-called Isaacs equation:

$$\begin{aligned} \psi(-\partial g(w)/2) &= 0, \ \lambda^*(w) = -\partial g(w)/2 \\ \text{Subject to } g(w) &= 0 \text{ on } A \cup B. \end{aligned}$$

Harmonic Functions & Doob's h-transform

- $u_n(x) = 1$ on $A \cup B$ and

$$u_n(x) = P_x(T_{A \cup B} < \infty) = E[u_n(x + Y_1/n)]$$

Harmonic Functions & Doob's h-transform

- $u_n(x) = 1$ on $A \cup B$ and

$$u_n(x) = P_x(T_{A\cup B} < \infty) = E[u_n(x + Y_1/n)]$$

- Zero-variance sampler is:

$$\begin{aligned} & P(Y_{k+1} \in dy | k < T_{A\cup B} < \infty, S_k = nx) \\ = \; & P^*(Y_{k+1} \in dy | S_k = nx) = f(y) \frac{u_n(x + y/n)}{u_n(x)} dy \end{aligned}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^*\left(Y_{k+1} \in dy \middle| S_k = nx \right) = f\left(y\right) \frac{u_n\left(x + y/n\right)}{u_n\left(x\right)} dy$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^*\left(\left. Y_{k+1} \in dy \right| S_k = nx\right) = f\left(y\right) \frac{u_n\left(x + y/n\right)}{u_n\left(x\right)} dy$$

- Recall: $u_n\left(x\right) = \exp\left(-nh\left(x\right) + o\left(n\right)\right)$

$$\begin{aligned} & P^*\left(\left. Y_{k+1} \in dy \right| S_k = nx\right) \\ = \quad & f\left(y\right) \exp\left(-n[h\left(x + y/n\right) - h\left(x\right)] + o\left(1\right)\right) dy \\ = \quad & f\left(y\right) \exp\left(-\partial h\left(x\right) \cdot y + o\left(1\right)\right) dy \end{aligned}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^*\left(Y_{k+1} \in dy \middle| S_k = nx\right) = f\left(y\right) \frac{u_n\left(x + y/n\right)}{u_n\left(x\right)} dy$$

- Recall: $u_n\left(x\right) = \exp\left(-nh\left(x\right) + o\left(n\right)\right)$

$$\begin{aligned} & P^*\left(Y_{k+1} \in dy \middle| S_k = nx\right) \\ = \;& f\left(y\right)\exp\left(-n[h\left(x + y/n\right) - h\left(x\right)] + o\left(1\right)\right) dy \\ = \;& f\left(y\right)\exp\left(-\partial h\left(x\right) \cdot y + o\left(1\right)\right) dy \end{aligned}$$

- But

$$1 = \int P^*\left(Y_{k+1} \in dy \middle| S_k = nx\right) \Longrightarrow \psi\left(-\partial h\left(x\right)\right) = 0$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(Y_{k+1}\in dy\right| S_{k}=nx)=f\left(y\right) \frac{u_{n}\left(x+y/n\right) }{u_{n}\left(x\right) }dy$$

- Recall: $u_n(x) = \exp(-nh(x) + o(n))$

$$\begin{aligned} & P^{*}\left(Y_{k+1}\in dy\right| S_{k}=nx) \\ = \quad & f\left(y\right) \exp\left(-n[h\left(x+y/n\right) -h\left(x\right)]+o\left(1\right) \right) dy \\ = \quad & f\left(y\right) \exp\left(-\partial h\left(x\right) \cdot y+o\left(1\right) \right) dy \end{aligned}$$

- But

$$1=\int P^{*}\left(Y_{k+1}\in dy\right| S_{k}=nx)\Longrightarrow \psi\left(-\partial h\left(x\right) \right) =0$$

- Equivalent to Isaacs equation with $g(x) = 2h(x)$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^*\left(Y_{k+1} \in dy \middle| S_k = nx \right) = f\left(y\right) \frac{u_n\left(x + y/n\right)}{u_n\left(x\right)} dy$$

- Recall: $u_n\left(x\right) = \exp\left(-nh\left(x\right) + o\left(n\right)\right)$

$$\begin{aligned} & P^*\left(Y_{k+1} \in dy \middle| S_k = nx \right) \\ = \; & f\left(y\right) \exp\left(-n[h\left(x + y/n\right) - h\left(x\right)] + o\left(1\right)\right) dy \\ = \; & f\left(y\right) \exp\left(-\partial h\left(x\right) \cdot y + o\left(1\right)\right) dy \end{aligned}$$

- But

$$1 = \int P^*\left(Y_{k+1} \in dy \middle| S_k = nx \right) \Longrightarrow \psi\left(-\partial h\left(x\right)\right) = 0$$

- Equivalent to Isaacs equation with $g\left(x\right) = 2h\left(x\right)$
- CONCLUSION (Dupuis-Wang 04): Sampler (mollified) is *(weakly)* asymptotically optimal... *BUT*

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^Q\left(Y_{k+1} \in dy \mid S_k = nx\right) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^{Q}\left(Y_{k+1} \in dy \right| S_k = nx) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

- Likelihood ratio

$$r\left(W_n(0), W_n\left(1/n\right)\right) ... r(W_n(T_{A\cup B} - 1), W_n(T_{A\cup B}))$$

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^Q\left(Y_{k+1} \in dy \middle| S_k = nx\right) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

- Likelihood ratio

$$r\left(W_n(0), W_n\left(1/n\right)\right) \ldots r\left(W_n\left(T_{A\cup B} - 1\right), W_n\left(T_{A\cup B}\right)\right)$$

- Second moment of estimator

$$s(x) = E_x[r(x, x + Y/n)s(x + Y/n)]$$

subject to $s(x) = 1$ for $x \in A \cup B$.

The Lyapunov Inequality

Lemma

Blanchet & Glynn '08: Lyapunov inequality

$$v(x) \geq E_x[r(x, x + Y/n)v(x + Y/n)]$$

subject to $v(x) \geq 1$ *for* $x \in A \cup B$. *Then,* $v(x) \geq s(x)$.

- **How to use the result?**

The Lyapunov Inequality

Lemma

Blanchet & Glynn '08: Lyapunov inequality

$$v(x) \geq E_x[r(x, x + Y/n)v(x + Y/n)]$$

subject to $v(x) \geq 1$ *for* $x \in A \cup B$. *Then,* $v(x) \geq s(x)$.

- **How to use the result?**
- 1) Identify a change-of-measure,

The Lyapunov Inequality

Lemma

Blanchet & Glynn '08: Lyapunov inequality

$$v(x) \geq E_x[r(x, x + Y/n)v(x + Y/n)]$$

subject to $v(x) \geq 1$ *for* $x \in A \cup B$. *Then,* $v(x) \geq s(x)$.

- **How to use the result?**
- 1) Identify a change-of-measure,
- 2) Use Large Deviations to force $v(x) \approx u_n(x)^2$.

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-n\gamma(x))$ & $\lambda = -\partial\gamma(x)/2$

$$1 \geq E[\exp(-\lambda^T Y + \psi(\lambda)) \exp(-n[\gamma(x + Y/n) - \gamma(x)])]$$

subject to $\gamma(x) \leq 0$ for $x \in A \cup B$. Then, $v(x) \geq s(x)$.

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-n\gamma(x))$ & $\lambda = -\partial\gamma(x)/2$

$$1 \geq E[\exp(-\lambda^T Y + \psi(\lambda)) \exp(-n[\gamma(x + Y/n) - \gamma(x)])]$$

subject to $\gamma(x) \leq 0$ for $x \in A \cup B$. Then, $v(x) \geq s(x)$.

- Expanding as $n \nearrow \infty$ we get

$$1 + O(1/n) \geq \exp[2\psi(-\partial\gamma(x)/2)]$$

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-n\gamma(x))$ & $\lambda = -\partial\gamma(x)/2$

$$1 \geq E[\exp(-\lambda^T Y + \psi(\lambda)) \exp(-n[\gamma(x + Y/n) - \gamma(x)])]$$

  subject to $\gamma(x) \leq 0$ for $x \in A \cup B$. Then, $v(x) \geq s(x)$.

- Expanding as $n \nearrow \infty$ we get

$$1 + O(1/n) \geq \exp[2\psi(-\partial\gamma(x)/2)]$$

- Yields subsolution to Isaacs equation (Dupuis-Wang '07)

$$\psi(-\partial\gamma(x)/2) \leq 0 \text{ s.t. } \gamma(x) \leq 0,\ x \in A \cup B$$

A Lyapunov Inequality

- Select

$$
\begin{aligned}
v(x) &= (w_1(x) + w_2(x))^2 \quad \text{<– square of LD approx} \\
w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\
w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x))
\end{aligned}
$$

A Lyapunov Inequality

- Select

$$
\begin{aligned}
v(x) &= (w_1(x) + w_2(x))^2 \quad \text{<– square of LD approx} \\
w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\
w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x))
\end{aligned}
$$

- Mixture sampler from density $\widetilde{f}(y)$

$$
\frac{\widetilde{f}(y)}{f(y)} = \frac{w_1(x)}{w_1(x) + w_2(x)} \exp\left(\theta_1^* v_1^T y\right) + \frac{w_2(x)}{w_1(x) + w_2(x)} \exp\left(\theta_2^* v_2^T y\right)
$$

A Lyapunov Inequality

- Select

$$\begin{aligned} v(x) &= (w_1(x) + w_2(x))^2 \quad \text{<– square of LD approx} \\ w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\ w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x)) \end{aligned}$$

- Mixture sampler from density $\widetilde{f}(y)$

$$\frac{\widetilde{f}(y)}{f(y)} = \frac{w_1(x)}{w_1(x) + w_2(x)} \exp\left(\theta_1^* v_1^T y\right) + \frac{w_2(x)}{w_1(x) + w_2(x)} \exp\left(\theta_2^* v_2^T y\right)$$

- Boundary condition on $A \cup B$

$$v(x) = (w_1(x) + w_2(x))^2 \geq 1$$

for $v_1^T x \geq 1$ OR $v_2^T x \geq 1$... OK!

A Lyapunov Inequality

- $$\begin{aligned} w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\ w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x)) \end{aligned}$$

A Lyapunov Inequality

- $$\begin{aligned} w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\ w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x)) \end{aligned}$$

- $$\begin{aligned} v(x) &= (w_1(x) + w_2(x))^2 \\ w_1(x + Y/n) &= w_1(x)\, e^{\theta^* v_1^T Y} \\ w_2(x + Y/n) &= w_2(x)\, e^{\theta^* v_2^T Y} \end{aligned}$$

A Lyapunov Inequality

- $$\begin{aligned} w_1(x) &= \exp(-n\theta_1^*(1 - v_1^T x)) \\ w_2(x) &= \exp(-n\theta_2^*(1 - v_2^T x)) \end{aligned}$$

- $$\begin{aligned} v(x) &= (w_1(x) + w_2(x))^2 \\ w_1(x + Y/n) &= w_1(x)\, e^{\theta^* v_1^T Y} \\ w_2(x + Y/n) &= w_2(x)\, e^{\theta^* v_2^T Y} \end{aligned}$$

- $$\begin{aligned} & E\frac{v(x+Y/n)}{v(x)} \frac{1}{\frac{w_1(x)}{w_1(x)+w_2(x)} e^{\theta_1^* v_1^T Y} + \frac{w_2(x)}{w_1(x)+w_2(x)} e^{\theta_2^* v_2^T Y}} \\ = \quad & E\frac{w_1(x)\exp\left(\theta_1^* v_1^T Y\right) + w_2(x)\exp\left(\theta_2^* v_2^T Y\right)}{w_1(x) + w_2(x)} = 1. \end{aligned}$$

Conclusion of the Example

- By Lyapunov inequality

  $$2nd \text{ Moment of estimator} \leq v(0) = (w_1(0) + w_2(0))^2 \leq O(u_n(0)^2)$$

Conclusion of the Example

- By Lyapunov inequality

  $2nd$ Moment of estimator $\leq v(0) = (w_1(0) + w_2(0))^2 \leq O(u_n(0)^2)$

- So, sampler is STRONGLY OPTIMAL.

Notes and Summary: Light Tails

- **Smoothness of solution to Isaacs equation plays an important role in the performance**

Notes and Summary: Light Tails

- **Smoothness of solution to Isaacs equation plays an important role in the performance**
- **Dupuis-Wang 04, 07 introduced subsolutions approach for LIGHT-TAILED problems**

Notes and Summary: Light Tails

- **Smoothness of solution to Isaacs equation plays an important role in the performance**
- **Dupuis-Wang 04, 07 introduced subsolutions approach for LIGHT-TAILED problems**
- **Blanchet, Glynn and Leder 10 study sharper complexity via Lyapunov inequalities**

Heavy-tailed Case

- Rich large deviations theory for random walks with *subexponential* increments:

$$P\left(X_1 + X_2 > b\right) \sim P\left(\max\left(X_1, X_2\right) > b\right)$$

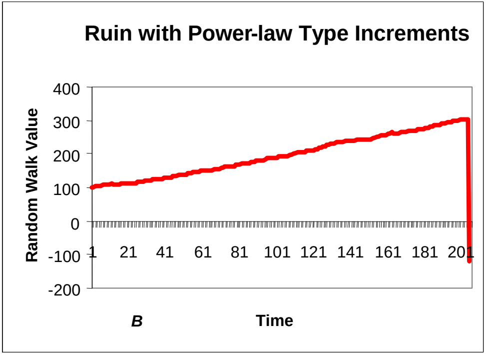
as $b \longrightarrow \infty$.

- *Focus on an important class* of a subexponential distributions: *regularly varying distributions* (basically *power-law* type)

$$P\left(X_1 > t\right) = t^{-\alpha} L\left(t\right)$$

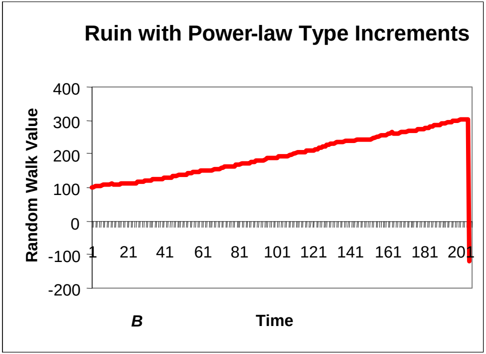
for $\alpha > 1$ and $L\left(t\beta\right)/L\left(t\right) \longrightarrow 1$ as $t \nearrow \infty$ for each $\beta > 0$.

How Does Ruin Occur with Heavy-tails (e.g. Catastrophic Insurance)?



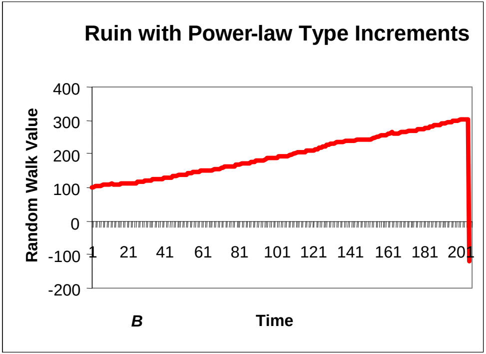
- Increments t-distributed (power law density) drift $+1$ variance $+1$

How Does Ruin Occur with Heavy-tails (e.g. Catastrophic Insurance)?



- Increments t-distributed (power law density) drift $+1$ variance $+1$
- Picture represents reserve CONDITIONAL ON RUIN!

How Does Ruin Occur with Heavy-tails (e.g. Catastrophic Insurance)?



- Increments t-distributed (power law density) drift $+1$ variance $+1$
- Picture represents reserve CONDITIONAL ON RUIN!
- Generated with Blanchet and Glynn (2008)'s algorithm

Problem Formulation

- Let X_1, X_2,... are i.i.d. regularly varying
- $EX_i = \eta < 0$
- $S_n = X_1 + ... + X_n$, $(S_0 = 0)$.
- $T_b = \inf\{n \geq 0 : S_n > b\}$.
- Object of interest:

$$u_b(s) = P_s(T_b < \infty).$$

Interpretation of the Picture

- **Interpretation**: Prior to ruin, the random walk has drift $\eta < 0$ and a large jump of size b occurs suddenly...

Interpretation of the Picture

- **Interpretation**: Prior to ruin, the random walk has drift $\eta < 0$ and a large jump of size b occurs suddenly...
- So, at time k, $S_k \approx \eta k$ and chance of reaching b in the next step *given* $(T_b < \infty)$ is

$$\frac{P(X > b - \eta k)}{\sum_{k=1}^{\infty} P(X > b - \eta k)} \approx \frac{P(X > b - \eta k)}{\int_0^{\infty} P(X > b - \eta u)\, du}$$

$$\approx \frac{-\eta P(X > b - \eta k)}{\int_b^{\infty} P(X > s)\, du} = O\left(\frac{1}{b}\right).$$

Interpretation of the Picture

- **Interpretation**: Prior to ruin, the random walk has drift $\eta < 0$ and a large jump of size b occurs suddenly...
- So, at time k, $S_k \approx \eta k$ and chance of reaching b in the next step *given* $(T_b < \infty)$ is

$$\frac{P\left(X > b - \eta k\right)}{\sum_{k=1}^{\infty} P\left(X > b - \eta k\right)} \approx \frac{P\left(X > b - \eta k\right)}{\int_0^{\infty} P\left(X > b - \eta u\right) du}$$

$$\approx \frac{-\eta P\left(X > b - \eta k\right)}{\int_b^{\infty} P\left(X > s\right) du} = O\left(\frac{1}{b}\right).$$

- The analysis also gives

$$P_0\left(T_b < \infty\right) \approx -\frac{1}{\eta} \int_b^{\infty} P\left(X > s\right) ds$$

as $b \rightarrow \infty$. (Pakes-Veraberbeke Thm.)

The One Dimensional Case

- **Family of changes-of-measure:** Given s current position of the walk (NOTE $p(s)$ and $a \in (0,1)$)

$$
\begin{aligned}
f_{X|s}(x|\, s) &= p(s) \frac{f_X(x) I(x > a(b-s))}{P(X > a(b-s))} \\
&\quad + (1-p(s)) \frac{f_X(x) I(x \leq a(b-s))}{P(X > a(b-s))}
\end{aligned}
$$

- In other words, $s_0 = s$ and $s_1 = s_0 + x$

$$
\begin{aligned}
\frac{f_{X|s}(x|s)}{f(x)} &:= r(s_0, s_1)^{-1} = p(s_0) \frac{I(s_1 - s_0 > a(b-s_0))}{P(X > a(b-s_0))} \\
&\quad + (1-p(s_1)) \frac{I(s_1 - s_0 \leq a(b-s_0))}{P(X \leq a(b-s_0))}
\end{aligned}
$$

Recall Lyapunov Inequality

- **Lyapunov Inequalities for Variance Control:**

Lemma (Blanchet & Glynn '08)

Suppose that there is a positive function $g(\cdot)$ *such that*

$$E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \leq 1$$

for all $s \leq b$ *and* $g(s) \geq 1$ *for* $s > b$. *Then,* $g(s)$ *bounds second moment of importance sampling estimator, that is*

$$E_s\left(\prod_{j=1}^{T_b - 1} r(S_j, S_{j+1})\, I(T_b < \infty)\right) \leq g(s).$$

Choosing the Candidate Lyapunov Function

- Want *strong efficiency*, so pick (by Pakes-Veraberbeke Thm)

$$g(s) = \min\left(\kappa\left(\int_{b-s}^{\infty} P(X > u)\, du\right)^2, 1\right).$$

- Pick

$$p(s) = \theta \frac{P(X > b - s)}{\int_{b-s}^{\infty} P(X > s)\, du}$$

- Just select θ, $\kappa > 0$ to force Lyapunov inequality!

Testing the Lyapunov Inequality

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
= \; & \frac{E(g(s+X); X > a(b-s))\, P(X > a(b-s))}{p(s)\, g(s)} \\
& + \frac{E(g(s+X); X \leq a(b-s))\, P(X \leq a(b-s))}{(1-p(s))\, g(s)}
\end{aligned}
$$

Testing the Lyapunov Inequality

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
= \; & \frac{E(g(s+X); X > a(b-s))\, P(X > a(b-s))}{p(s)\, g(s)} \\
& + \frac{E(g(s+X); X \leq a(b-s))\, P(X \leq a(b-s))}{(1-p(s))\, g(s)} \\
\leq \; & \frac{P(X > a(b-s))^2}{p(s)\, g(s)} + \frac{E(g(s+X); X \leq a(b-s))}{(1-p(s))\, g(s)}
\end{aligned}
$$

Choosing the Parameters and Testing the Inequality

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
= \; & \frac{E(g(s+X); X > a(b-s))\, P(X > a(b-s))}{p(s)\, g(s)} \\
& + \frac{E(g(s+X); X \leq a(b-s))\, P(X \leq a(b-s))}{(1-p(s))\, g(s)} \\
\leq \; & \frac{P(X > a(b-s))^2}{p(s)\, g(s)} + \frac{E(g(s+X); X \leq a(b-s))}{(1-p(s))\, g(s)} \\
\approx \; & \frac{a^{-\alpha} P(X > a(b-s))}{\theta \kappa \int_{b-s}^{\infty} P(X > u)\, du} + 1 + 2(\eta + \theta) \frac{P(X > (b-s))}{\left(\int_{b-s}^{\infty} P(X > u)\, du\right)} \leq 1
\end{aligned}
$$

The One Dimensional Case

- **Corresponding Algorithm:**
- Select $a \in (0,1)$, then choose θ and κ based on Lyapunov inequality
- AT EACH TIME STEP TEST
    -
        - IF $g(s) < 1$ apply Imp. Sampling according $p(s)$ - mixture
        - ELSE do NOT apply I.S. and continue until hitting.
    - OUTPUT PRODUCT OF LOCAL LIKELIHOOD RATIOS Z

Conclusion of Example: Blanchet and Glynn '08

$$2nd \text{ moment } \leq g(0) \sim \kappa \left(\int_b^\infty P(X > s)\, ds \right)^2 = O\left(P(T_b < \infty)^2 \right)$$

so **strong optimality holds.**

Notes and Summary: Heavy Tails

- **Change-of-measure should be chosen depending on tail environment**

Notes and Summary: Heavy Tails

- **Change-of-measure should be chosen depending on tail environment**
- **Use large deviations to select your Lyapunov inequality, similar to light-tailed case**

References

1. Blanchet (2010). http://www.columbia.edu/~jb2814/
2. Blanchet and Glynn (2008). Annals of Applied Probability.
3. Blanchet, Glynn and Leder (2010). http://www.columbia.edu/~jb2814/
4. **Blanchet and Lam (2011) Surveys in Oper. Res. and Mgmt. Sc. (To appear) http://www.columbia.edu/~jb2814/**
5. Dupuis and Wang (2004). Annals of Applied Probability
6. Dupuis and Wang (2007) Math. of Operations Research