Markov Chain Monte Carlo for rare-event simulation in heavy-tailed settings

Thorbjörn Gudmundsson

Department of Mathematics
KTH Stockholm

Licentiat seminarium, December 2013

- Computational problem. For instance, the probability of default or the expected loss given ruin.

- Computational problem. For instance, the probability of default or the expected loss given ruin.
- Complex system : no analytical solution available

- Computational problem. For instance, the probability of default or the expected loss given ruin.
- Complex system : no analytical solution available
- Simulation techniques
  - i Monte Carlo
  - ii Conditional Monte Carlo
  - iii Splitting methods
  - iv Importance sampling

- Computational problem. For instance, the probability of default or the expected loss given ruin.
- Complex system : no analytical solution available
- Simulation techniques
  - i Monte Carlo
  - ii Conditional Monte Carlo
  - iii Splitting methods
  - iv Importance sampling
  - v Markov chain Monte Carlo (NEW)

Problem

Consider a random variable X with known distribution F and the objective of computing

$$p = \mathbb{P}(X \in A),$$

where $\{X \in A\}$ is thought as rare in the sense that p is small. Event of ruin for instance.

Problem

Consider a random variable X with known distribution F and the objective of computing

$$p = \mathbb{P}(X \in A),$$

where $\{X \in A\}$ is thought as rare in the sense that p is small. Event of ruin for instance.

Example. Random walk $S_n = Y_1 + \cdots + Y_n$ with non-negative steps Y's with known heavy-tailed distribution F_Y and objective of computing

$$p = \mathbb{P}\left(\frac{S_n}{n} > a\right),$$

where a is much larger than $\mathbb{E}[Y]$.

Stochastic Simulation

Want to compute $p = \mathbb{P}(X \in A)$.
In absence of an analytical solution, stochastic simulation offers an alternative.

Stochastic Simulation

Want to compute $p = \mathbb{P}(X \in A)$.
In absence of an analytical solution, stochastic simulation offers an alternative.
Monte Carlo: sample identically distributed and independent copies $X_1, \ldots, X_N$ and compute

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} I\{X_k \in A\}.$$

Shortcomings of Monte Carlo

The relative error of the Monte Carlo estimator is unbounded as $p \to 0$:

$$\frac{\mathbb{V}ar(\hat{p})}{p^2} = \frac{1}{N}\left(\frac{1}{p} - 1\right) \to \infty, \quad \text{as } p \to 0.$$

Example. Standard normal variable X, compute $p = \mathbb{P}(X > a)$ using $N = 10^6$ number of simulations

$$a = 1 \quad : \quad \hat{p} = 0.158, \quad \frac{\text{Stdev}(\hat{p})}{\hat{p}} = 0.002$$

$$a = 3 \quad : \quad \hat{p} = 0.0014, \quad \frac{\text{Stdev}(\hat{p})}{\hat{p}} = 0.027$$

$$a = 5 \quad : \quad \hat{p} = 0, \quad \frac{\text{Stdev}(\hat{p})}{\hat{p}} = \infty$$

Solutions

- Conditional Monte Carlo (Asmussen)
- Splitting methods (Creou et al)
- Importance sampling (Sigmund, Dupuis, Blanchet)

Importance sampling

Goal: construct an efficient estimator $\hat{p}$ of $p = \mathbb{P}(X \in A)$, in the sense that its relative error is bounded.

Importance sampling

Goal: construct an efficient estimator $\hat{p}$ of $p = \mathbb{P}(X \in A)$, in the sense that its relative error is bounded.

The importance sampling approach (Dupuis et al 2007)

- Generate independent copies $X_1, \ldots, X_N$ from a sampling distribution G.
- Compute empirical estimate

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} \frac{dF}{dG}(X_k) \mathbb{I}\{X_k \in A\}.$$

Importance sampling

Goal: construct an efficient estimator $\hat{p}$ of $p = \mathbb{P}(X \in A)$, in the sense that its relative error is bounded.
The importance sampling approach (Dupuis et al 2007)

- Generate independent copies $X_1, \ldots, X_N$ from a sampling distribution G.
- Compute empirical estimate

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} \frac{dF}{dG}(X_k) \mathbb{I}\{X_k \in A\}.$$

$$\mathbb{E}_G[\hat{p}] = \int_A \frac{dF}{dG}(X) dG(X) = F(A) = p.$$

Importance sampling continued

Reduces to finding a suitable sampling distribution G.

Importance sampling continued

Reduces to finding a suitable sampling distribution G.
The zero-variance distribution

$$F_A(x) = \mathbb{P}(X \leq x | X \in A).$$

If we can choose $G = F_A$, then $\frac{dF}{dF_A}(X)\mathbb{I}\{X \in A\} = p$, so

$$\hat{p} = \frac{1}{N}\sum_{k=1}^{N} \frac{dF}{dF_A}(X_k)\mathbb{I}\{X_k \in A\} = p,$$

with zero variance!

Importance sampling continued

Reduces to finding a suitable sampling distribution G.
The zero-variance distribution

$$F_A(x) = \mathbb{P}(X \leq x | X \in A).$$

If we can choose $G = F_A$, then $\frac{dF}{dF_A}(X)\mathbb{I}\{X \in A\} = p$, so

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} \frac{dF}{dF_A}(X_k)\mathbb{I}\{X_k \in A\} = p,$$

with zero variance!
Requires knowledge of $\mathbb{P}(X \in A)$ ...

The idea

Want: sample from $F_A(x) = \mathbb{P}(X \leq x | X \in A)$.
Assuming the existence of a density, it takes the form

$$f_A(x) = \frac{f(x)\mathbb{I}\{x \in A\}}{\mathbb{P}(X \in A)}.$$

The idea

Want: sample from $F_A(x) = \mathbb{P}(X \leq x | X \in A)$.
Assuming the existence of a density, it takes the form

$$f_A(x) = \frac{f(x)\mathbb{I}\{x \in A\}}{\mathbb{P}(X \in A)}.$$

The main idea is to construct a Markov chain $(X_k)_{k \geq 1}$ for which f_A is the invariant density via MCMC. Then *extract* information about the normalising constant from the sample.

Estimator

- Construct a Markov chain $(X_k)_{k \geq 1}$ via MCMC sampler, with the zero-variance distribution F_A as its invariant distribution.

Estimator

- Construct a Markov chain $(X_k)_{k\geq 1}$ via MCMC sampler, with the zero-variance distribution F_A as its invariant distribution.
- For any $v \geq 0$ such that $\int_A v(x)dx = 1$, consider

$$u((X_k)_{k\geq 1}) = \frac{1}{N}\sum_{k=1}^{N} \frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}.$$

Estimator continued

- For $\int_A v(x)dx = 1$ it holds

$$\mathbb{E}_{F_A}\Big[\frac{1}{N}\sum_{k=1}^{N}\frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}\Big] = \int_A \frac{v(x)}{f(x)}\frac{f(x)}{p}dx = \frac{1}{p}\int_A v(x)dx = \frac{1}{p}.$$

Estimator continued

- For $\int_A v(x)dx = 1$ it holds

$$\mathbb{E}_{F_A}\left[\frac{1}{N}\sum_{k=1}^{N}\frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}\right] = \int_A \frac{v(x)}{f(x)}\frac{f(x)}{p}dx = \frac{1}{p}\int_A v(x)dx = \frac{1}{p}.$$

- Define $\hat{q} = \frac{1}{N}\sum_{k=1}^{N}\frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}$ estimator of $1/p$.

Design issues

Estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} \frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}$ of $1/p$.

- Choice of the MCMC sampler: crucial to control the dependence of the Markov chain, to ensure the large sample efficiency

$$\mathbb{V}ar(\hat{q}) \to 0, \quad \text{as } N \to \infty.$$

Design issues

Estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} \frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}$ of $1/p$.

- Choice of the MCMC sampler: crucial to control the dependence of the Markov chain, to ensure the large sample efficiency

$$\mathbb{V}ar(\hat{q}) \to 0, \quad \text{as } N \to \infty.$$

- Choice of v: controls the variance, set to ensure rare-event efficiency

$$\frac{\text{Std}(\hat{q})}{1/p} = p\,\text{Std}(\hat{q}) \to 0, \quad \text{as } p \to 0.$$

Controlling the variance

Estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} u(X_k)$, with $u(X_k) = \frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}$.
Goal is to show $p\,\mathrm{Std}(\hat{q})$ tends to zero as $p \to 0$.

Controlling the variance

Estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} u(X_k)$, with $u(X_k) = \frac{v(X_k)\mathbb{I}\{X_k \in A\}}{f(X_k)}$.
Goal is to show $p\,\mathrm{Std}(\hat{q})$ tends to zero as $p \to 0$.

- Consider the term

$$\begin{aligned} p^2 \mathbb{V}ar\big(u(X)\big) &= p^2 \big(\mathbb{E}[u(X)^2] - \mathbb{E}[u(X)]^2\big) \\ &= p^2 \Big(\int_A \frac{v^2(x)}{f^2(x)} \frac{f(x)}{p} dx - 1 \Big) \\ &= p \int_A \frac{v^2(x)}{f(x)} dx - 1. \end{aligned}$$

Controlling the variance continued

Choosing $v(x) = f_A(x) = \frac{f(x)\mathbb{I}\{x \in A\}}{p}$ implies

$$p^2 \mathbb{V}ar(u(X)) = p \int_A \frac{f^2(x)/p^2}{f(x)} dx - 1 = \frac{1}{p} \int_A f(x) dx - 1 = 0.$$

Controlling the variance continued

Choosing $v(x) = f_A(x) = \frac{f(x)\mathbb{I}\{x \in A\}}{p}$ implies

$$p^2 \mathbb{V}ar(u(X)) = p \int_A \frac{f^2(x)/p^2}{f(x)} dx - 1 = \frac{1}{p} \int_A f(x) dx - 1 = 0.$$

Choose v as an approximation of the zero-variance density!

Recipe

- Sample $(X_k)_{k \geq 1}$ under F_A via some MCMC sampler

Recipe

- Sample $(X_k)_{k \geq 1}$ under F_A via some MCMC sampler
- Show $p^2 \mathbb{V}ar\left(u(X)\right) \to 0$ as $p \to 0$

Recipe

- Sample $(X_k)_{k\geq 1}$ under F_A via some MCMC sampler
- Show $p^2 \mathbb{V}ar\left(u(X)\right) \to 0$ as $p \to 0$
- Show $(X_k)_{k\geq 1}$ is geometric ergodic

Setup

- Consider a random walk $S_n = Y_1 + \cdots + Y_n$ with non-negative steps Y's with known heavy-tailed distribution F_Y and objective of computing

$$p = \mathbb{P}\left(\frac{S_n}{n} > a\right),$$

where a is much larger than $\mathbb{E}[Y]$.

Setup

- Consider a random walk $S_n = Y_1 + \cdots + Y_n$ with non-negative steps Y's with known heavy-tailed distribution F_Y and objective of computing

$$p = \mathbb{P}\Big(\frac{S_n}{n} > a\Big),$$

where a is much larger than $\mathbb{E}[Y]$.

- Construct $(\mathbf{Y}_k)_{k\geq 1}$ via MCMC with invariant density

$$f_A(\mathbf{y}) = \frac{f_{\mathbf{Y}}(\mathbf{y})\mathbb{I}\{y_1 + \cdots + y_n > an\}}{\mathbb{P}(S_n > an)}.$$

Setup

- Consider a random walk $S_n = Y_1 + \cdots + Y_n$ with non-negative steps Y's with known heavy-tailed distribution F_Y and objective of computing

$$p = \mathbb{P}\left(\frac{S_n}{n} > a\right),$$

where a is much larger than $\mathbb{E}[Y]$.
- Construct $(\mathbf{Y}_k)_{k \geq 1}$ via MCMC with invariant density

$$f_A(\mathbf{y}) = \frac{f_{\mathbf{Y}}(\mathbf{y})\mathbb{I}\{y_1 + \cdots + y_n > an\}}{\mathbb{P}(S_n > an)}.$$

- A typical such a random walk has a $n - 1$ number of "small" steps and one "large" step.

Gibbs sampler

Initial state $\mathbf{Y}_0 = (Y_{0,1}, \ldots, Y_{0,n})$ such that $Y_{0,1} > an$ and $Y_{0,i} = 0$ for other indices. Given $\mathbf{Y}_k = (Y_{k,1}, \ldots, Y_{k,n})$, $k = 0, 1, \ldots$ the next state $\mathbf{Y}_{k+1}$ is sampled as follows

- Take a copy of the current state, let $Y_{k+1,i} = Y_{k,i}$,

Gibbs sampler

Initial state $\mathbf{Y}_0 = (Y_{0,1}, \ldots, Y_{0,n})$ such that $Y_{0,1} > an$ and $Y_{0,i} = 0$ for other indices. Given $\mathbf{Y}_k = (Y_{k,1}, \ldots, Y_{k,n})$, $k = 0, 1, \ldots$ the next state $\mathbf{Y}_{k+1}$ is sampled as follows

- Take a copy of the current state, let $Y_{k+1,i} = Y_{k,i}$,
- Draw a random index $j \in \{1, \ldots, n\}$,

Gibbs sampler

Initial state $\mathbf{Y}_0 = (Y_{0,1}, \ldots, Y_{0,n})$ such that $Y_{0,1} > an$ and $Y_{0,i} = 0$ for other indices. Given $\mathbf{Y}_k = (Y_{k,1}, \ldots, Y_{k,n})$, $k = 0, 1, \ldots$ the next state $\mathbf{Y}_{k+1}$ is sampled as follows

- Take a copy of the current state, let $Y_{k+1,i} = Y_{k,i}$,
- Draw a random index $j \in \{1, \ldots, n\}$,
- Sample $Y_{k+1,j}$ from the conditional distribution of Y given that the sum exceeds the threshold,

$$\mathbb{P}(Y_{k+1,j} \in \cdot) = \mathbb{P}\Big(Y \in \cdot \mid Y + \sum_{i \neq j} Y_{k,i} > an\Big).$$

Gibbs sampler

Initial state $\mathbf{Y}_0 = (Y_{0,1}, \ldots, Y_{0,n})$ such that $Y_{0,1} > an$ and $Y_{0,i} = 0$ for other indices. Given $\mathbf{Y}_k = (Y_{k,1}, \ldots, Y_{k,n})$, $k = 0, 1, \ldots$ the next state $\mathbf{Y}_{k+1}$ is sampled as follows

- Take a copy of the current state, let $Y_{k+1,i} = Y_{k,i}$,
- Draw a random index $j \in \{1, \ldots, n\}$,
- Sample $Y_{k+1,j}$ from the conditional distribution of Y given that the sum exceeds the threshold,

$$\mathbb{P}(Y_{k+1,j} \in \cdot) = \mathbb{P}\Big(Y \in \cdot \mid Y + \sum_{i \neq j} Y_{k,i} > an\Big).$$

- Permutate the steps in $\mathbf{Y}_{k+1}$.

Gibbs sampler continued

Proposition

The Markov chain $(\mathbf{Y}_k)_{k\geq 1}$ constructed using the proposed Gibbs sampler has the conditional distribution F_A as its invariant distribution.

MCMC estimator

- The MCMC estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} \frac{v(\mathbf{y}_k)\mathbb{I}\{S_n > an\}}{f(\mathbf{y}_k)}$. The steps are heavy-tailed in the sense that

$$\frac{\mathbb{P}(M_n > an)}{\mathbb{P}(S_n > an)} \to 1,$$

where $M_n = \max_i\{y_{k,i}\}$.

MCMC estimator

- The MCMC estimator $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} \frac{v(\mathbf{y}_k)\mathbb{I}\{S_n > an\}}{f(\mathbf{y}_k)}$. The steps are heavy-tailed in the sense that

$$\frac{\mathbb{P}(M_n > an)}{\mathbb{P}(S_n > an)} \to 1,$$

where $M_n = \max_i\{y_{k,i}\}$.

- Therefore seems smart to use

$$\mathbb{P}(\mathbf{Y} \in \cdot \mid M_n > an) \text{ as a proxy for } \mathbb{P}(\mathbf{Y} \in \cdot \mid S_n > an).$$

Propose

$$v(\mathbf{y}_k) = \frac{f(\mathbf{y}_k)\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}.$$

MCMC estimator continued

Choosing $v(\mathbf{y}) = \frac{f(\mathbf{y})\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$ yields

$$u(\mathbf{y}) = \frac{v(\mathbf{y})\mathbb{I}\{S_n > an\}}{f(\mathbf{y})} = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}.$$

MCMC estimator continued

Choosing $v(\mathbf{y}) = \frac{f(\mathbf{y})\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$ yields

$$u(\mathbf{y}) = \frac{v(\mathbf{y})\mathbb{I}\{S_n > an\}}{f(\mathbf{y})} = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}.$$

$\hat{q} = \mathbb{P}(M_n > an)^{-1} \frac{1}{N} \sum_{k=1}^{N} \mathbb{I}\{M_n(k) > an\}$

Efficiency

Since $u(\mathbf{y}) = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$, we have:

$$p^2 \mathbb{V}ar_{F_A}(u(\mathbf{Y})) = \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \mathbb{V}ar_{F_A}(\mathbb{I}\{M_n > an\})$$

Efficiency

Since $u(\mathbf{y}) = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$, we have:

$$p^2 \mathbb{V}ar_{F_A}(u(\mathbf{Y})) = \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \mathbb{V}ar_{F_A}(\mathbb{I}\{M_n > an\})$$

$$= \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \left(\mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}] - \mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}]^2 \right)$$

Efficiency

Since $u(\mathbf{y}) = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$, we have:

$$p^2 \mathbb{V}ar_{F_A}(u(\mathbf{Y})) = \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \mathbb{V}ar_{F_A}(\mathbb{I}\{M_n > an\})$$

$$= \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \left(\mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}] - \mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}]^2 \right)$$

$$= \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \left(\frac{\mathbb{P}(M_n > an)}{\mathbb{P}(S_n > an)} - \frac{\mathbb{P}(M_n > an)^2}{\mathbb{P}(S_n > an)^2} \right)$$

Efficiency

Since $u(\mathbf{y}) = \frac{\mathbb{I}\{M_n > an\}}{\mathbb{P}(M_n > an)}$, we have:

$$p^2 \mathbb{V}ar_{F_A}(u(\mathbf{Y})) = \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \mathbb{V}ar_{F_A}(\mathbb{I}\{M_n > an\})$$

$$= \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \left(\mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}] - \mathbb{E}_{F_A}[\mathbb{I}\{M_n > an\}]^2 \right)$$

$$= \frac{\mathbb{P}(S_n > an)^2}{\mathbb{P}(M_n > an)^2} \left(\frac{\mathbb{P}(M_n > an)}{\mathbb{P}(S_n > an)} - \frac{\mathbb{P}(M_n > an)^2}{\mathbb{P}(S_n > an)^2} \right)$$

$$= \frac{\mathbb{P}(S_n > an)}{\mathbb{P}(M_n > an)} - 1 \to 0 \quad \text{as } p \to 0.$$

Geometric ergodicity

- The design of the Gibbs sampler ensures that the Markov chain $(\mathbf{Y}_k)_{k\geq 1}$ is (uniformly) ergodic.
- This guarantees that the chain mixes sufficiently and hence that $\mathbb{V}ar(\hat{p}) \to 0$ as $N \to \infty$ at same speed as $1/N$.

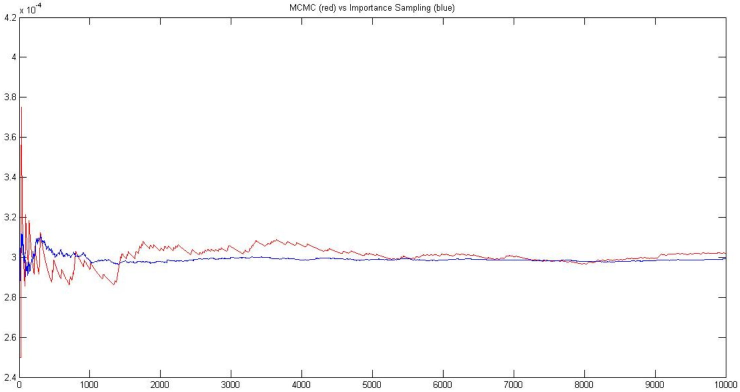
Numerical experiments

- The MCMC estimator $\hat{q}^{-1}$ of the probability p tested against importance sampling and standard Monte Carlo.
- Steps are Pareto(2) distributed.
- Number of batches: 25, simulations per batch: $10,000$.

Table

n	a	MCMC	IS	MC	
5	10	3.40e-3 (0.81e-4) [4.1]	2.91e-3 (1.77e-4) [3.4]	2.83e-3 (4.74e-4) [0.7]	Avg. est. (Std. dev.) [Avg. time (ms)]
10	20	3.34e-4 (5.83e-6)	3.02e-4 (2.02e-6)	2.68e-4 (162.58e-6)	Avg. est. (Std. dev.)

10, 000 simulations for $m = 10$ and $a = 20$



Setup

Consider a random walk $S_{N_n} = Y_1 + \cdots + Y_{N_n}$ with non-negative heavy-tailed steps Y, discrete random variable N_n and the objective of computing

$$p = \mathbb{P}(S_{N_n} > a\mathbb{E}[N_n]),$$

where a is much larger than $\mathbb{E}[Y]$.

The challenge

How to design a Gibbs sampler to construct a Markov chain with the following invariant distribution

$$F_A(\cdot) = \mathbb{P}\big((N, Y_1, \ldots, Y_N) \in \cdot \mid S_{N_n} > a_n\big).$$

The challenge

How to design a Gibbs sampler to construct a Markov chain with the following invariant distribution

$$F_A(\cdot) = \mathbb{P}\big((N, Y_1, \ldots, Y_N) \in \cdot \mid S_{N_n} > a_n\big).$$

The trick was to sample N from $\mathbb{P}(N = k \mid N \geq k^*)$ where $k^* = min\{k : Y_1 + \ldots + Y_k > a_n\}$.

Numerical experiments

- The MCMC estimator $\hat{q}^{-1}$ of the probability p tested against importance sampling and standard Monte Carlo.
- Steps are Pareto(1) distributed.
- Number of steps is Geometric(0.2) distributed
- Number of batches: 25, simulations per batch: $10,000$.

Numerical experiments

a	MCMC	IS	MC	
100	1.149e-2 (4e-5) [25]	1.087e-2 (6e-5) [11]	1.089-2 (35e-5) [1.2]	Avg. est. (Std. dev.) [Avg. time (ms)]
$5 \cdot 10^7$	2.000003e-8 (6e-14)	1.999325e-8 (1114e-14)		Avg. est. (Std. dev.)

Setup

Consider the following setup for the risk reserve U_k, for positive claim size B:

$$\begin{aligned} U_k &= R_k(U_{k-1} - B_k), \quad \text{for } k \geq 1, \\ U_0 &= u. \end{aligned}$$

Setup

Consider the following setup for the risk reserve U_k, for positive claim size B:

$$\begin{aligned} U_k &= R_k(U_{k-1} - B_k), \quad \text{for } k \geq 1, \\ U_0 &= u. \end{aligned}$$

Iteration gives: $U_n = R_n \cdots R_1 u - \left(R_n \cdots R_1 B_1 + \cdots + R_N B_n\right)$.

Setup

Consider the following setup for the risk reserve U_k, for positive claim size B:

$$\begin{aligned} U_k &= R_k(U_{k-1} - B_k), \quad \text{for } k \geq 1, \\ U_0 &= u. \end{aligned}$$

Iteration gives: $U_n = R_n \cdots R_1 u - \left(R_n \cdots R_1 B_1 + \cdots + R_N B_n\right)$.
Writing $A_k = 1/R_k$ then

$$\begin{aligned} A_1 \cdots A_n U_n &= u - W_n, \quad \text{where} \\ W_n &= B_1 + A_1 B_2 + \cdots + A_1 \cdots A_{n-1} B_n. \end{aligned}$$

Problem

Thus the event of ruin can be expressed as follows

$$\{\inf_k U_k < 0\} = \{\sup_k W_k > u\}.$$

Problem

Thus the event of ruin can be expressed as follows

$$\{\inf_k U_k < 0\} = \{\sup_k W_k > u\}.$$

Goal: Construct an MCMC estimator for computing

$$p = \mathbb{P}\big(\sup_k W_k > u\big).$$

Gibbs sampler

Construct a Markov chain $(\mathbf{A}_t, \mathbf{B}_t)_{t \geq 0}$ with the invariant distribution

$$\mathbb{P}\big((\mathbf{A}, \mathbf{B}) \in \cdot \mid \sup_k W_k > u\big).$$

Gibbs sampler

Construct a Markov chain $(\mathbf{A}_t, \mathbf{B}_t)_{t \geq 0}$ with the invariant distribution

$$\mathbb{P}\big((\mathbf{A}, \mathbf{B}) \in \cdot \mid \sup_k W_k > u\big).$$

Carried out by updating one of $(A_1, \dots, A_n, B_1, \dots, B_n)$ at a time, conditioned so that

$$\max_{1 \leq k \leq n} W_k = \max_{1 \leq k \leq n} B_1 + A_1 B_2 + \cdots + A_1 \cdots A_{k-1} B_k > u.$$

Efficiency

Assume that

- The claim size *B* is Pareto(α) distributed
- The stochastic return *R* fulfills $\mathbb{E}[R^{-\alpha-\epsilon}] < \infty$ for some $\epsilon > 0$

Efficiency

Assume that

- The claim size B is Pareto(α) distributed
- The stochastic return R fulfills $\mathbb{E}[R^{-\alpha-\epsilon}] < \infty$ for some $\epsilon > 0$

Then we have the asymptotic result

$$\frac{\mathbb{P}(\sup_{1 \leq k \leq n} W_k > u)}{\mathbb{P}(B > u) \sum_{k=0}^{n-1} \mathbb{E}[A^\alpha]^k} \to 1, \quad \text{as } n \to \infty.$$

Efficiency continued

Now $W_n = B_1 + A_1 B_2 + \cdots + A_1 \cdots A_{n-1} B_n$.
Based on the existing asymptotic results we propose the following choice for V

$$V(\cdot) = \mathbb{P}\big((\mathbf{A}, \mathbf{B}) \in \cdot \mid (\mathbf{A}, \mathbf{B}) \in R\big),$$

where

$$R = \{B_1 > u\}$$

Efficiency continued

Now $W_n = B_1 + A_1 B_2 + \cdots + A_1 \cdots A_{n-1} B_n$.
Based on the existing asymptotic results we propose the following choice for V

$$V(\cdot) = \mathbb{P}\big((\mathbf{A}, \mathbf{B}) \in \cdot \mid (\mathbf{A}, \mathbf{B}) \in R\big),$$

where

$$R = \{B_1 > u\} \cup \{A_1 > a, B_2 > u/a\}$$

Efficiency continued

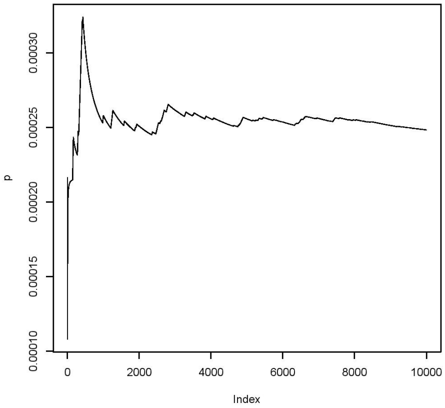
Now $W_n = B_1 + A_1 B_2 + \cdots + A_1 \cdots A_{n-1} B_n$.
Based on the existing asymptotic results we propose the following choice for V

$$V(\cdot) = \mathbb{P}\big((\mathbf{A}, \mathbf{B}) \in \cdot \mid (\mathbf{A}, \mathbf{B}) \in R\big),$$

where

$$\begin{aligned} R &= \{B_1 > u\} \cup \{A_1 > a, B_2 > u/a\} \cup \dots \\ &\cup \{A_1 > a, \dots, A_{n-1} > a, B_n > u/a^{n-1}\}. \end{aligned}$$

$10,000$ simulations for $n = 10$ and $u = 10^5$



Conclusion

Established a framework for new and simple method within stochastic simulation: Markov chain Monte Carlo methodology.

Conclusion

Established a framework for new and simple method within stochastic simulation: Markov chain Monte Carlo methodology.
Applied the framework and proved efficiency on four concrete examples:

- Random walk with heavy-tails
- Random sum with heavy-tails
- Solution to stochastic recurrent equations with heavy-tailed innovations
- Insurance model with risky investments and Pareto distributed claim size

Conclusion

Possibilities for future work:

- Extension to random walk with light-tails
- Perfect simulation / coupling form the past
- Solution to stochastic recurrent equations where the ruin event is controlled by the stochastic returns rather than the claim size

Thank you for your attention!