Markov chain Monte Carlo for computing probabilities of rare events in a heavy-tailed random walk

Thorbjorn Gudmundsson and Henrik Hult

Department of Mathematics
KTH Stockholm

KTH, April 2011

Problem and Motivation

Setup

Consider a random walk $S_m = Y_1 + \cdots + Y_m$, increments Y are i.i.d. and distribution known. Compute the probability

$$p_m = \mathbb{P}(S_m > am), \quad \text{for } m \text{ large and } a > \mathbb{E}[Y].$$

- Sometimes no analytical solution known.
- Problems with the most elementary simulation methods: Monte Carlo.

Problem and Motivation

Setup

Consider a random walk $S_m = Y_1 + \cdots + Y_m$, increments Y are i.i.d. and distribution known. Compute the probability

$$p_m = \mathbb{P}(S_m > am), \quad \text{for } m \text{ large and } a > \mathbb{E}[Y].$$

- Sometimes no analytical solution known.
- Problems with the most elementary simulation methods: Monte Carlo.

Problem and Motivation

Problem with Monte Carlo

Monte Carlo:

- Generate $S_m(1), \dots, S_m(n)$ independently.
- Compute empirical estimate $\hat{p}_m = \frac{1}{n} \sum_{i=1}^{n} I\{S_m(i) > am\}$.

Simple to implement, unbiased,

$$\mathbb{E}[\hat{p}_m] = p_m,$$

consistent,

$$\hat{p}_m \to p_m \quad \text{w.p. } 1, \text{ as } n \to \infty.$$

Problem and Motivation

Problem with Monte Carlo

Monte Carlo:

- Generate $S_m(1), \dots, S_m(n)$ independently.
- Compute empirical estimate $\hat{p}_m = \frac{1}{n}\sum_{i=1}^{n} I\{S_m(i) > am\}$.

Simple to implement, unbiased,

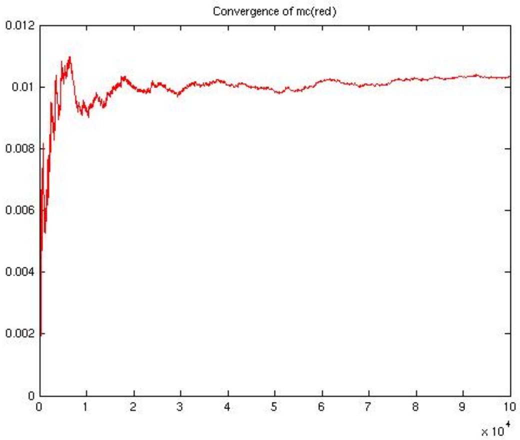
$$\mathbb{E}[\hat{p}_m] = p_m,$$

consistent,

$$\hat{p}_m \to p_m \quad \text{w.p. 1, as } n \to \infty.$$

Problem and Motivation

Convergence of estimator



Problem and Motivation

Problem with Monte Carlo continued

- *What about efficiency?* would like the standard deviation $\text{Std}(\hat{p}_m)$ to be of roughly the same size as p_m.
- For the Monte Carlo estimate

$$\frac{\text{Std}(\hat{p}_m)}{p_m} = \frac{1}{\sqrt{n}} \frac{\sqrt{p_m - p_m^2}}{p_m} \sim \frac{1}{\sqrt{n p_m}}.$$

- *For rare events Monte Carlo requires a large computational cost.*

Problem and Motivation

Problem with Monte Carlo continued

- *What about efficiency?* would like the standard deviation $\mathrm{Std}(\hat{p}_m)$ to be of roughly the same size as p_m.
- For the Monte Carlo estimate

$$\frac{\mathrm{Std}(\hat{p}_m)}{p_m} = \frac{1}{\sqrt{n}} \frac{\sqrt{p_m - p_m^2}}{p_m} \sim \frac{1}{\sqrt{np_m}}.$$

- *For rare events Monte Carlo requires a large computational cost.*

Computing probability using MCMC

Importance sampling

Importance sampling:
Denote the original distribution of S_m by F and density by f.

- Generate $S_m(1), \ldots, S_m(n)$ independently from a sampling distribution G.
- Compute empirical estimate

$$\hat{p}_m = \frac{1}{n} \sum_{i=1}^{n} \frac{dF}{dG} I\{S_m(i) > am\}.$$

Both unbiased and consistent.

Computing probability using MCMC

Zero variance sampling distribution

There exists a **best choice** for G that gives zero variance. The best sampling distribution G is the conditional distribution given the event itself

$$\mathbb{P}(S_m \in \cdot | S_m > am).$$

The density

$$g(x) = \frac{f(x)I\{x > am\}}{\mathbb{P}(S_m > am)}.$$

Problem: This distribution requires to know $p_m = \mathbb{P}(S_m > am)$ - the very probability we are trying to compute.

Computing probability using MCMC

Zero variance sampling distribution

There exists a **best choice** for G that gives zero variance. The best sampling distribution G is the conditional distribution given the event itself

$$\mathbb{P}(S_m \in \cdot | S_m > am).$$

The density

$$g(x) = \frac{f(x)I\{x > am\}}{\mathbb{P}(S_m > am)}.$$

Problem: This distribution requires to know $p_m = \mathbb{P}(S_m > am)$ - the very probability we are trying to compute.

Computing probability using MCMC

MCMC Algorithm

- An MCMC algorithm is a tool to sample a random variable despite only knowing its density up to a normalising constant.
- The density of S_m under G is precisely of that nature

$$g(x) = \frac{f(x)I\{x > am\}}{p_m}.$$

We can generate via MCMC a sample of random variables with g as density - but they are dependent!

Computing probability using MCMC
MCMC Algorithm

- An MCMC algorithm is a tool to sample a random variable despite only knowing its density up to a normalising constant.
- The density of S_m under G is precisely of that nature

$$g(x) = \frac{f(x)I\{x > am\}}{p_m}.$$

We can generate via MCMC a sample of random variables with g as density - but they are dependent!

Computing probability using MCMC

Execute MCMC and extract data

Suppose sampling $S_m(1), \ldots, S_m(n)$ via MCMC (dependent) from the zero variance distribution G.

$$S_m(i) \sim g(\cdot) = \frac{f(\cdot) I\{\cdot > am\}}{p_m}.$$

How to extract the information about the normalising constant?

Computing probability using MCMC

Execute MCMC and extract data continued

$$\mathbb{E}[u(S_m)] = \int u(x)g(x)dx = \int_{x>am} u(x)\frac{f(x)}{p_m}dx.$$

Setting $u(x) = \frac{v(x)}{f(x)} I\{x > am\}$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}\int_{x>am} v(x)dx.$$

So choosing v is such that $\int_{x>am} v(x)dx = 1$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}.$$

Computing probability using MCMC

Execute MCMC and extract data continued

$$\mathbb{E}[u(S_m)] = \int u(x)g(x)dx = \int_{x>am} u(x)\frac{f(x)}{p_m}dx.$$

Setting $u(x) = \frac{v(x)}{f(x)} I\{x > am\}$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}\int_{x>am} v(x)dx.$$

So choosing v is such that $\int_{x>am} v(x)dx = 1$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}.$$

Computing probability using MCMC

Execute MCMC and extract data continued

$$\mathbb{E}[u(S_m)] = \int u(x)g(x)dx = \int_{x>am} u(x)\frac{f(x)}{p_m}dx.$$

Setting $u(x) = \frac{v(x)}{f(x)} I\{x > am\}$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}\int_{x>am} v(x)dx.$$

So choosing v is such that $\int_{x>am} v(x)dx = 1$

$$\mathbb{E}[u(S_m)] = \frac{1}{p_m}.$$

Computing probability using MCMC

Estimator

- Consistent estimator based on MCMC:

$$\hat{p}_m = \left(\frac{1}{n}\sum_{i=1}^{n} u(S_m(i))\right)^{-1}$$

- Control efficiency by choosing a v.

How to choose v?

Computing probability using MCMC

Estimator

- Consistent estimator based on MCMC:

$$\hat{p}_m = \left(\frac{1}{n}\sum_{i=1}^{n} u(S_m(i))\right)^{-1}$$

- Control efficiency by choosing a v.

How to choose v?

Computing probability using MCMC

Estimator's variance

Consider the variance of $\hat{p}_m = \left(\frac{1}{n}\sum_{i=1}^n u(S_m(i))\right)^{-1}$.

- Taylor: $h(x) \approx h(x_0) + h'(x_0)(x - x_0)$ so

$$\mathbb{V}ar(h(x)) \approx \left(h'(x_0)\right)^2 \mathbb{V}ar(x).$$

- Applied on $h(x) = 1/x$ for $x = \frac{1}{n}\sum_{i=1}^n u(S_m(i))$ and $x_0 = \mathbb{E}[x] = 1/p_m$

$$\mathbb{V}ar(\hat{p}_m) \approx \left(\frac{-1}{x_0^2}\right)^2 \mathbb{V}ar(x) = \frac{p_m^4}{n} \mathbb{V}ar(u(S_m)).$$

Computing probability using MCMC

Estimator's variance

Consider the variance of $\hat{p}_m = \left(\frac{1}{n}\sum_{i=1}^{n} u(S_m(i))\right)^{-1}$.

- Taylor: $h(x) \approx h(x_0) + h'(x_0)(x - x_0)$ so

$$\mathbb{V}ar(h(x)) \approx \left(h'(x_0)\right)^2 \mathbb{V}ar(x).$$

- Applied on $h(x) = 1/x$ for $x = \frac{1}{n}\sum_{i=1}^{n} u(S_m(i))$ and $x_0 = \mathbb{E}[x] = 1/p_m$

$$\mathbb{V}ar(\hat{p}_m) \approx \left(\frac{-1}{x_0^2}\right)^2 \mathbb{V}ar(x) = \frac{p_m^4}{n} \mathbb{V}ar(u(S_m)).$$

Computing probability using MCMC

Estimator's variance

Consider the variance of $\hat{p}_m = \left(\frac{1}{n}\sum_{i=1}^{n} u(S_m(i))\right)^{-1}$.

- Taylor: $h(x) \approx h(x_0) + h'(x_0)(x - x_0)$ so

$$\mathbb{V}ar(h(x)) \approx \left(h'(x_0)\right)^2 \mathbb{V}ar(x).$$

- Applied on $h(x) = 1/x$ for $x = \frac{1}{n}\sum_{i=1}^{n} u(S_m(i))$ and $x_0 = \mathbb{E}[x] = 1/p_m$

$$\mathbb{V}ar(\hat{p}_m) \approx \left(\frac{-1}{x_0^2}\right)^2 \mathbb{V}ar(x) = \frac{p_m^4}{n} \mathbb{V}ar\left(u(S_m)\right).$$

Computing probability using MCMC

Estimator's variance continued

For MCMC estimator

$$\begin{aligned}
\frac{\mathbb{V}ar(\hat{p}_m)}{p_m^2} &\approx \frac{p_m^2}{n}\mathbb{V}ar\left(u(S_m)\right) \\
&= \frac{p_m^2}{n}\left(\mathbb{E}\left[u(S_m)^2\right] - \left(\mathbb{E}[u(S_m)]\right)^2\right) \\
&= \frac{p_m^2}{n}\left(\mathbb{E}\left[u(S_m)^2\right] - \frac{1}{p_m^2}\right) \\
&= \frac{1}{n}\left(p_m^2 \int_{x>am} \frac{v(x)^2}{f(x)^2} - 1\right)
\end{aligned}$$

Computing probability using MCMC

Bounded Relative Error Criteria

Choosing

$$v(x) = g(x) = \frac{f(x)I\{x > am\}}{p_m}.$$

Gives

$$\frac{\mathbb{V}ar(\hat{p}_m)}{p_m^2} \approx \frac{1}{n}\left(p_m^2 \int_{x>am} \frac{v(x)^2}{f(x)^2} - 1\right) = 0.$$

Result:

v* is chosen as an approximation of the zero variance density *g

Random Walk with Heavy-tailed Increments

Setup

- Random walk $S_m = Y_1 + \cdots + Y_m$. Compute $\mathbb{P}(S_m > am)$.
- Zero variance distribution

$$\mathbb{P}(S_m \leq x | S_m > am),$$

- Say Y are heavy-tailed if following holds:

$$\frac{\mathbb{P}(S_m > am)}{\mathbb{P}(M_m > am)} \to 1 \quad \text{as } m \to \infty,$$

$M_m = \max\{Y_1, \dots, Y_m\}$, e.g. Cauchy, regularly varying, subexponential.

- Choose v as the density of

$$\mathbb{P}(S_m \leq x | M_m > am) = \frac{\mathbb{P}(S_m \leq x, M_m > am)}{\mathbb{P}(M_m > am)}.$$

Random Walk with Heavy-tailed Increments

Setup

- Random walk $S_m = Y_1 + \cdots + Y_m$. Compute $\mathbb{P}(S_m > am)$.
- Zero variance distribution

$$\mathbb{P}(S_m \leq x | S_m > am),$$

- Say Y are heavy-tailed if following holds:

$$\frac{\mathbb{P}(S_m > am)}{\mathbb{P}(M_m > am)} \to 1 \quad \text{as } m \to \infty,$$

  $M_m = \max\{Y_1, \dots, Y_m\}$, e.g. Cauchy, regularly varying, subexponential.
- Choose v as the density of

$$\mathbb{P}(S_m \leq x | M_m > am) = \frac{\mathbb{P}(S_m \leq x, M_m > am)}{\mathbb{P}(M_m > am)}.$$

Random Walk with Heavy-tailed Increments

MCMC estimator

This choice of v gives MCMC estimator:

$$\begin{aligned}
\hat{p}_m &= \Big(\frac{1}{n}\sum_{i=1}^{n}\frac{v(S_m(i))}{f(S_m(i))}I\{S_m(i) > am\}\Big)^{-1} \\
&= \Big(\frac{1}{n}\sum_{i=1}^{n}\frac{f(S_m(i))I\{M_m(i) > am\}/p_{\max}}{f(S_m(i))}I\{S_m(i) > am\}\Big)^{-1} \\
&= p_{\max}\Big(\frac{1}{n}\sum_{i=1}^{n}I\{M_m(i) > am\}\Big)^{-1},
\end{aligned}$$

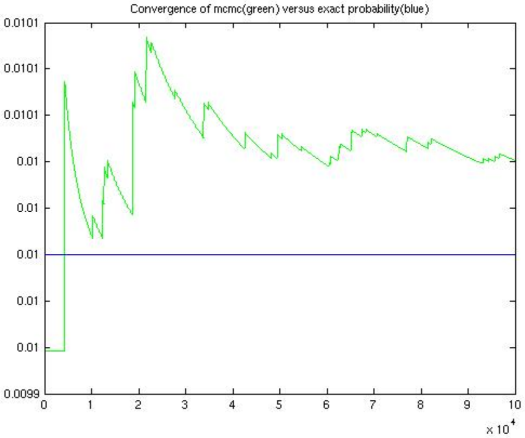
where

$$p_{\max} = \mathbb{P}(M_m > am) = 1 - F_Y(am)^m,$$

is easily calculated.

Random Walk with Heavy-tailed Increments

Cauchy: MCMC estimate vs true probability



Random Walk with Heavy-tailed Increments

Cauchy: MCMC estimate vs Monte Carlo