MCMC Technique for Rare-Event Simulation for Heavy-Tailed SRE

Thorbjörn Gudmundsson

Department of Mathematics
KTH Royal Institute of Technology
Sweden

RESIM, Amsterdam, August 27–29, 2014

based on joint work with H. Hult

Outline

Setup

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

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

  where $\{X \in C\}$ is thought as rare in the sense that p is small.
- Assume that no analytical solution is known.
- The event is rare so standard Monte Carlo simulation is ineffective.

The idea

- Construct a Markov chain $(X_t)_{t\geq 0}$ having

$$F_C(\cdot) = \mathbb{P}(X \in \cdot \mid X \in C)$$

  as its invariant distribution.
- Then extract information about the normalising constant from the sample.

The idea

- Construct a Markov chain $(X_t)_{t\geq 0}$ having

$$F_C(\cdot) = \mathbb{P}(X \in \cdot \mid X \in C)$$

  as its invariant distribution.
- Then extract information about the normalising constant from the sample.

Definition

- Construct a Markov chain $(X_t)_{t \geq 0}$ via MCMC sampler having F_C as its invariant distribution.
- For any distribution V such that $V \ll F_C$ consider

$$u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- Then

$$\mathbb{E}_{F_C}[u(X)] = \int_C \frac{dV}{dF} dF_C = \frac{1}{p} \int_C dV = \frac{1}{p}.$$

- Motivates the following expression as an estimate for p

$$\left(\frac{1}{T} \sum_{t=0}^{T-1} u(X_t) \right)^{-1}.$$

Definition

- Construct a Markov chain $(X_t)_{t\geq 0}$ via MCMC sampler having F_C as its invariant distribution.
- For any distribution V such that $V \ll F_C$ consider

$$u(X) = \frac{dV}{dF}(X)I\{X \in C\}.$$

- Then

$$\mathbb{E}_{F_C}\left[u(X)\right] = \int_C \frac{dV}{dF}dF_C = \frac{1}{p}\int_C dV = \frac{1}{p}.$$

- Motivates the following expression as an estimate for p

$$\left(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\right)^{-1}.$$

Definition

- Construct a Markov chain $(X_t)_{t \geq 0}$ via MCMC sampler having F_C as its invariant distribution.
- For any distribution V such that $V \ll F_C$ consider

$$u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- Then

$$\mathbb{E}_{F_C}\left[u(X)\right] = \int_C \frac{dV}{dF} dF_C = \frac{1}{p} \int_C dV = \frac{1}{p}.$$

- Motivates the following expression as an estimate for p

$$\Big(\frac{1}{T} \sum_{t=0}^{T-1} u(X_t)\Big)^{-1}.$$

Design issues

For a sample $(X_t)_{t \geq 0}$ from a MCMC sampler, then

$$\hat{p} = \Big(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\Big)^{-1},$$

$$u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- Design of the MCMC sampler: crucial to control the dependence of the Markov chain.
- Choice of V: controls the variance, set to ensure rare-event efficiency of the algorithm.

$$p^2 \mathbb{V}ar_{F_C}(u(X)) = \ldots = p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1.$$

Design issues

For a sample $(X_t)_{t\geq 0}$ from a MCMC sampler, then

$$\hat{p} = \left(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\right)^{-1},$$
$$u(X) = \frac{dV}{dF}(X)I\{X \in C\}.$$

- Design of the MCMC sampler: crucial to control the dependence of the Markov chain.
- Choice of V: controls the variance, set to ensure rare-event efficiency of the algorithm.

$$p^2 \mathbb{V}ar_{F_C}(u(X)) = \ldots = p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1.$$

Design issues

For a sample $(X_t)_{t\geq 0}$ from a MCMC sampler, then

$$\hat{p} = \Big(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\Big)^{-1},$$
$$u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- Design of the MCMC sampler: crucial to control the dependence of the Markov chain.
- Choice of V: controls the variance, set to ensure rare-event efficiency of the algorithm.

$$p^2 \mathbb{V}ar_{F_C}(u(X)) = \ldots = p\mathbb{E}_V\Big[\frac{dV}{dF}\Big] - 1.$$

Choice of V

$$p^2 \mathbb{V}ar(u(X)) = \ldots = p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1.$$

- For any $R \subseteq C$ for which $r = \mathbb{P}(X \in R)$ can be computed explicitly, a candidate for V is

$$V(\cdot) = \mathbb{P}(X \in \cdot \mid X \in R).$$

- Such a choice is a good one if r is close to p since

$$p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1 = p\mathbb{E}_V\left[\frac{dF/r}{dF}\right] - 1 = \frac{p}{r} - 1 \to 0.$$

Choice of V

$$p^2 \mathbb{V}ar(u(X)) = \ldots = p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1.$$

- For any $R \subseteq C$ for which $r = \mathbb{P}(X \in R)$ can be computed explicitly, a candidate for V is

$$V(\cdot) = \mathbb{P}(X \in \cdot \mid X \in R).$$

- Such a choice is a good one if r is close to p since

$$p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1 = p\mathbb{E}_V\left[\frac{dF/r}{dF}\right] - 1 = \frac{p}{r} - 1 \to 0.$$

Choice of V

$$p^2 \mathbb{V}ar(u(X)) = \ldots = p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1.$$

- For any $R \subseteq C$ for which $r = \mathbb{P}(X \in R)$ can be computed explicitly, a candidate for V is

$$V(\cdot) = \mathbb{P}(X \in \cdot \mid X \in R).$$

- Such a choice is a good one if r is close to p since

$$p\mathbb{E}_V\left[\frac{dV}{dF}\right] - 1 = p\mathbb{E}_V\left[\frac{dF/r}{dF}\right] - 1 = \frac{p}{r} - 1 \to 0.$$

Setup

- Let $\mathbf{A} = (A_1, \ldots, A_m)$ and $\mathbf{B} = (B_1, \ldots, B_m)$ be independent sequences of i.i.d. random variables. Consider the solution X_m to the SRE

$$\begin{aligned} X_k &= A_k X_{k-1} + B_k, \quad \text{for } k = 1, \ldots, m, \\ X_0 &= 0, \end{aligned}$$

and the problem of computing

$$p = \mathbb{P}(X_m > c).$$

Design of the MCMC sampler

First task is to construct a Markov chain $(\mathbf{A}_t, \mathbf{B}_t)_{t \geq 0}$ having

$$F_C(\cdot) = \mathbb{P}(\mathbf{A}, \mathbf{B} \in \cdot \mid X_m > c),$$

as its invariant distribution.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \ldots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \ldots, A_{t,m}, B_{t,1}, \ldots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \ldots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \ldots, A_{t,m}, B_{t,1}, \ldots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \ldots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \ldots, A_{t,m}, B_{t,1}, \ldots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \ldots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \ldots, A_{t,m}, B_{t,1}, \ldots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \dots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \dots, A_{t,m}, B_{t,1}, \dots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Initial state $(\mathbf{A}_0, \mathbf{B}_0)$ such that $X_m > c$. Given $(\mathbf{A}_t, \mathbf{B}_t)$, $t = 0, 1, \ldots$ the next state $(\mathbf{A}_{t+1}, \mathbf{B}_{t+1})$ is sampled as follows

- Randomly pick one of the variables $A_{t,1}, \ldots, A_{t,m}, B_{t,1}, \ldots, B_{t,m}$,
- If $A_{t,k}$ is to be updated, sample A' from

$$\mathbb{P}(A' \in \cdot \mid A' > s)$$

  where $\{A' > s\}$ ensures that $\{X_m > c\}$ when $A_{t,k}$ is replaced with A'.
- Set $A_{t+1,i} = A_{t,i}$ and $B_{t+1,i} = B_{t,i}$ for all i except $A_{t+1,k} = A'$.
- Similar if $B_{t,k}$ is to be updated.

Gibbs sampler

Proposition

The Markov chain $(\boldsymbol{A}_t, \boldsymbol{B}_t)_{t\geq 0}$ *constructed using the proposed Gibbs sampler has the conditional distribution* F_C *as its invariant distribution.*

Model assumptions

Assume heavy-tailed innovations.

- B has regularly varying tail distribution with index $-\alpha < 0$.
- A fullfills the Breiman condition,

$$\mathbb{E}[A^{\alpha+\varepsilon}] < \infty, \quad \text{for some } \varepsilon > 0.$$

- Then the following heavy-tail asymptotics holds

$$p \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{E}[A^{\alpha}]^k, \quad \text{as } c \to \infty.$$

Model assumptions

Assume heavy-tailed innovations.

- B has regularly varying tail distribution with index $-\alpha < 0$.
- A fullfills the Breiman condition,

$$\mathbb{E}[A^{\alpha+\varepsilon}] < \infty, \quad \text{for some } \varepsilon > 0.$$

- Then the following heavy-tail asymptotics holds

$$p \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{E}[A^{\alpha}]^{k}, \quad \text{as } c \to \infty.$$

Choice of V

Event of interest: $X_m = B_m + A_m B_{m-1} + \cdots + A_m \cdots A_2 B_1 > c$.

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

$$\{\mathbf{A}, \mathbf{B} \in R\} = \{\exists k : A_m \cdots A_{k+1} B_k > c, A_m, \dots, A_{k+1} > a\}.$$

- Then $r = \mathbb{P}(\mathbf{A}, \mathbf{B} \in R)$ can be computed explicitly and asymptotically

$$r \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}.$$

Choice of V

Event of interest: $X_m = B_m + A_m B_{m-1} + \cdots + A_m \cdots A_2 B_1 > c$.

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

$$\{\mathbf{A}, \mathbf{B} \in R\} = \{\exists k : A_m \cdots A_{k+1} B_k > c, A_m, \ldots, A_{k+1} > a\}.$$

- Then $r = \mathbb{P}(\mathbf{A}, \mathbf{B} \in R)$ can be computed explicitly and asymptotically

$$r \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}.$$

Choice of V

Event of interest: $X_m = B_m + A_m B_{m-1} + \cdots + A_m \cdots A_2 B_1 > c$.

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

$$\{\mathbf{A}, \mathbf{B} \in R\} = \{\exists k : A_m \cdots A_{k+1} B_k > c, A_m, \ldots, A_{k+1} > a\}.$$

- Then $r = \mathbb{P}(\mathbf{A}, \mathbf{B} \in R)$ can be computed explicitly and asymptotically

$$r \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}.$$

Choice of V

Event of interest: $X_m = B_m + A_m B_{m-1} + \cdots + A_m \cdots A_2 B_1 > c$.

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

$$\{\mathbf{A}, \mathbf{B} \in R\} = \{\exists k : A_m \cdots A_{k+1} B_k > c, A_m, \dots, A_{k+1} > a\}.$$

- Then $r = \mathbb{P}(\mathbf{A}, \mathbf{B} \in R)$ can be computed explicitly and asymptotically

$$r \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}.$$

Choice of a

$$
\begin{aligned}
p &\sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{E}[A^\alpha]^k \\
r &\sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}
\end{aligned}
$$

- The free parameter a is set so that $\lim_{c \to \infty} p/r = 1$, that is

$$
\mathbb{E}[A^\alpha] = \mathbb{P}(A > a) a^\alpha .
$$

Choice of a

$$p \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{E}[A^\alpha]^k$$

$$r \sim \mathbb{P}(B > c) \sum_{k=1}^{m} \mathbb{P}(A > a)^k a^{\alpha k}$$

- The free parameter a is set so that $\lim_{c \to \infty} p/r = 1$, that is

$$\mathbb{E}[A^\alpha] = \mathbb{P}(A > a) a^\alpha .$$

The MCMC estimator

- The MCMC estimator is defined by

$$\hat{p} = \Big(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\Big)^{-1}, \quad u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- For our choice of $V(\cdot) = \mathbb{P}(\mathbf{A}, \mathbf{B} \in \cdot \mid \mathbf{A}, \mathbf{B} \in R)$ then

$$u(\mathbf{A}, \mathbf{B}) = \frac{1}{r} I\{\mathbf{A}, \mathbf{B} \in R\}.$$

- Then the MCMC estimator is

$$\hat{p} = r\Big(\frac{1}{T}\sum_{t=0}^{T-1} I\{\mathbf{A}_t, \mathbf{B}_t \in R\}\Big)^{-1}.$$

The MCMC estimator

- The MCMC estimator is defined by

$$\hat{p} = \Big(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\Big)^{-1}, \quad u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- For our choice of $V(\cdot) = \mathbb{P}(\mathbf{A}, \mathbf{B} \in \cdot \mid \mathbf{A}, \mathbf{B} \in R)$ then

$$u(\mathbf{A}, \mathbf{B}) = \frac{1}{r} I\{\mathbf{A}, \mathbf{B} \in R\}.$$

- Then the MCMC estimator is

$$\hat{p} = r\Big(\frac{1}{T}\sum_{t=0}^{T-1} I\{\mathbf{A}_t, \mathbf{B}_t \in R\}\Big)^{-1}.$$

The MCMC estimator

- The MCMC estimator is defined by

$$\hat{p} = \Big(\frac{1}{T}\sum_{t=0}^{T-1} u(X_t)\Big)^{-1}, \quad u(X) = \frac{dV}{dF}(X) I\{X \in C\}.$$

- For our choice of $V(\cdot) = \mathbb{P}(\mathbf{A}, \mathbf{B} \in \cdot \mid \mathbf{A}, \mathbf{B} \in R)$ then

$$u(\mathbf{A}, \mathbf{B}) = \frac{1}{r} I\{\mathbf{A}, \mathbf{B} \in R\}.$$

- Then the MCMC estimator is

$$\hat{p} = r\Big(\frac{1}{T}\sum_{t=0}^{T-1} I\{\mathbf{A}_t, \mathbf{B}_t \in R\}\Big)^{-1}.$$

Efficiency

Asymptotically $\lim_{c\to\infty} \frac{p}{r} = 1$.

$$
\begin{aligned}
p^2 \mathbb{V}ar_{F_c}(u(\mathbf{A},\mathbf{B})) &= \frac{p^2}{r^2}\left(\mathbb{E}_{F_c}[I\{\mathbf{A},\mathbf{B}\in R\}] - \mathbb{E}_{F_c}[I\{\mathbf{A},\mathbf{B}\in R\}]^2\right) \\
&= \frac{p^2}{r^2}\left(\frac{r}{p} - \frac{r^2}{p^2}\right) \\
&= \frac{p}{r} - 1 \to 0, \quad \text{as } c \to \infty.
\end{aligned}
$$

Rare-event efficiency in 3 lines!

Efficiency

Asymptotically $\lim_{c\to\infty} \frac{p}{r} = 1$.

$$\begin{aligned} p^2 \mathbb{V}ar_{F_C}(u(\mathbf{A},\mathbf{B})) &= \frac{p^2}{r^2}\left(\mathbb{E}_{F_C}\left[I\{\mathbf{A},\mathbf{B}\in R\}\right] - \mathbb{E}_{F_C}\left[I\{\mathbf{A},\mathbf{B}\in R\}\right]^2\right) \\ &= \frac{p^2}{r^2}\left(\frac{r}{p} - \frac{r^2}{p^2}\right) \\ &= \frac{p}{r} - 1 \to 0, \quad \text{as } c \to \infty. \end{aligned}$$

Rare-event efficiency in 3 lines!

Efficiency

Asymptotically $\lim_{c\to\infty} \frac{p}{r} = 1$.

$$
\begin{aligned}
p^2 \mathbb{V}ar_{F_c}(u(\mathbf{A},\mathbf{B})) &= \frac{p^2}{r^2}\left(\mathbb{E}_{F_c}[I\{\mathbf{A},\mathbf{B} \in R\}] - \mathbb{E}_{F_c}[I\{\mathbf{A},\mathbf{B} \in R\}]^2\right) \\
&= \frac{p^2}{r^2}\left(\frac{r}{p} - \frac{r^2}{p^2}\right) \\
&= \frac{p}{r} - 1 \to 0, \quad \text{as } c \to \infty.
\end{aligned}
$$

Rare-event efficiency in 3 lines!

Efficiency

Asymptotically $\lim_{c\to\infty} \frac{p}{r} = 1$.

$$
\begin{aligned}
p^2 \mathbb{V}ar_{F_c}(u(\mathbf{A},\mathbf{B})) &= \frac{p^2}{r^2}\Big(\mathbb{E}_{F_c}\left[I\{\mathbf{A},\mathbf{B} \in R\}\right] - \mathbb{E}_{F_c}\left[I\{\mathbf{A},\mathbf{B} \in R\}\right]^2\Big) \\
&= \frac{p^2}{r^2}\Big(\frac{r}{p} - \frac{r^2}{p^2}\Big) \\
&= \frac{p}{r} - 1 \to 0, \quad \text{as } c \to \infty.
\end{aligned}
$$

Rare-event efficiency in 3 lines!

Geometric ergodicity

- Guarantees that the chain $(\mathbf{A}_t, \mathbf{B}_t)_{t \geq 0}$ mixes sufficiently and thus that $\mathbb{V}ar(\hat{p}) \to 0$ as $T \to \infty$ at same speed as $1/T$.
- Problem!

$$X_m = B_m + A_m B_{m-1} + A_m A_{m-1} B_{m-2} + \cdots + A_m \cdots A_2 B_1$$

The chain tends to get stuck with large value for B_m and low for any other B's...

- Causes bias is the estimate.

Geometric ergodicity

- Guarantees that the chain $(\mathbf{A}_t, \mathbf{B}_t)_{t\geq 0}$ mixes sufficiently and thus that $\mathbb{V}ar(\hat{p}) \to 0$ as $T \to \infty$ at same speed as $1/T$.
- Problem!

$$X_m = B_m + A_m B_{m-1} + A_m A_{m-1} B_{m-2} + \cdots + A_m \cdots A_2 B_1$$

The chain tends to get stuck with large value for B_m and low for any other B's...

- Causes bias is the estimate.

Geometric ergodicity

- Guarantees that the chain $(\mathbf{A}_t, \mathbf{B}_t)_{t\geq 0}$ mixes sufficiently and thus that $\mathbb{V}ar(\hat{p}) \to 0$ as $T \to \infty$ at same speed as $1/T$.
- Problem!

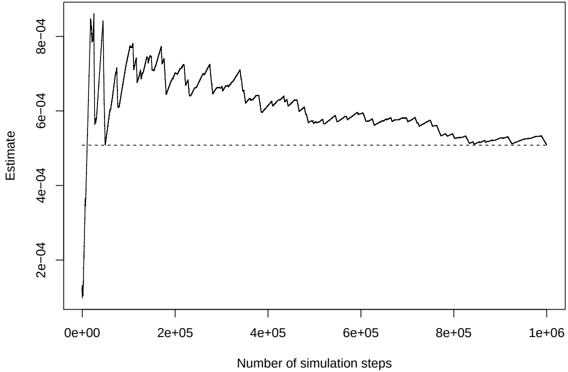
$$X_m = B_m + A_m B_{m-1} + A_m A_{m-1} B_{m-2} + \cdots + A_m \cdots A_2 B_1$$

The chain tends to get stuck with large value for B_m and low for any other B's...

- Causes bias is the estimate.

Figure

The point estimate of $\mathbb{P}(X_4 > 25)$ as a function of simulations.



Table

- Innovations B are Pareto(2)-distributed.
- Returns A are Exponentially(4)-distributed.

Table: Numerical comparison of computing $\mathbb{P}(X_4 > c)$.

$c = 10$	MCMC	IS
Estimate	1.043671e-02	1.041979e-02
Std. deviation	2.476812e-04	1.837578e-04
Rel. error	2.373174e-02	1.763545e-02
$c = 1,000$	MCMC	IS
Estimate	1.044860e-06	1.140318e-06
Std. deviation	8.879878e-08	1.459354e-08
Rel. error	8.498627e-02	1.279778e-02

References

- T. Gudmundsson and H. Hult.
  MCMC for computing rare-event probabilities for heavy-tailed random walk.
  Journal of Applied Probability, **51**(2), June 2014.
- T. Gudmundsson and H. Hult.
  MCMC for computing rare-event probabilities for SRE with heavy-tailed innovations.
  Preprint.
- H. Hult, Blanchet. J., and K. Leder.
  IS for computing rare-event probabilities for SRE with heavy-tailed innovations.
  ACM T. Model. Comput. S., 2013.