Efficient rare-event simulation for sums of dependent random variables

Leonardo Rojas-Nandayapa
joint work with José Blanchet

February 13, 2012
MCQMC
UNSW, Sydney, Australia

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Contents

1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Outline

1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Rare Events

Rare Events

Indexed family of events $\{A_u : u \in \mathbb{R}\}$ with

$$\mathbb{P}(A_u) \to 0, \qquad u \to \infty.$$

Probabilities difficult to estimate.

Algorithm (Estimator)

Indexed set of simulatable random variables $\{Z_u : u \in \mathbb{R}\}$ with

$$\mathbb{E}[Z_u] = \mathbb{P}(A_u)$$

Rare Events

Rare Events

Indexed family of events $\{A_u : u \in \mathbb{R}\}$ with

$$\mathbb{P}(A_u) \to 0, \qquad u \to \infty.$$

Probabilities difficult to estimate.

Algorithm (Estimator)

Indexed set of simulatable random variables $\{Z_u : u \in \mathbb{R}\}$ with

$$\mathbb{E}[Z_u] = \mathbb{P}(A_u)$$

Efficiency

Efficient Algorithms

- Logarithmic Efficiency:

$$\lim_{u \to u_0} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2-\epsilon}(A_u)} = 0, \qquad \forall \epsilon > 0.$$

- Bounded relative error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2}(A_u)} < \infty.$$

- Zero Relative Error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[z_u]}{\mathbb{P}^{2}(A_u)} = 0.$$

Efficiency

Efficient Algorithms

- Logarithmic Efficiency:

$$\lim_{u \to u_0} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2-\epsilon}(A_u)} = 0, \qquad \forall \epsilon > 0.$$

- Bounded relative error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2}(A_u)} < \infty.$$

- Zero Relative Error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[z_u]}{\mathbb{P}^{2}(A_u)} = 0.$$

Efficiency

Efficient Algorithms

- Logarithmic Efficiency:

$$\lim_{u \to u_0} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2-\epsilon}(A_u)} = 0, \qquad \forall \epsilon > 0.$$

- Bounded relative error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[Z_u]}{\mathbb{P}^{2}(A_u)} < \infty.$$

- Zero Relative Error:

$$\lim_{u \to \infty} \frac{\mathbb{V}\mathrm{ar}\,[z_u]}{\mathbb{P}^{2}(A_u)} = 0.$$

Why?

Approximated confidence interval for an MC estimator $\widehat{z}_u$ is

$$\widehat{z}_u \pm \Phi(1-\alpha/2)\sqrt{\frac{\mathbb{V}\mathrm{ar}\, Z_u}{R}}$$

To keep the interval proportional to $\mathbb{P}(A_u)$ we require

$$R \approx \frac{\mathbb{V}\mathrm{ar}\, \widehat{z}_u}{\mathbb{P}^2(A_u)}.$$

How?

Variance reduction techniques.

Why?

Approximated confidence interval for an MC estimator $\widehat{z}_u$ is

$$\widehat{z}_u \pm \Phi(1-\alpha/2)\sqrt{\frac{\mathbb{V}\mathrm{ar}\, Z_u}{R}}$$

To keep the interval proportional to $\mathbb{P}(A_u)$ we require

$$R \approx \frac{\mathbb{V}\mathrm{ar}\, \widehat{z}_u}{\mathbb{P}^2(A_u)}.$$

How?

Variance reduction techniques.

Outline

1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Tail probabilities of sums

Tail probability of a sum

Let $X_1, \ldots, X_n$. Tail probability of the sum

$$\mathbb{P}(X_1 + \cdots + X_n > u), \qquad u \to \infty.$$

Common Fact: i.i.d. case.

An importance sampling algorithm with exponential change of measure

$$F_\theta(dx) := \mathrm{e}^{-\theta x - \kappa(\theta)} F(dx)$$

produces an efficient algorithm if θ is such that $\mathbb{E}_\theta[X] = u/n$.

Tail probabilities of sums

Tail probability of a sum

Let $X_1, \ldots, X_n$. Tail probability of the sum

$$\mathbb{P}(X_1 + \cdots + X_n > u), \qquad u \to \infty.$$

Common Fact: i.i.d. case.

An importance sampling algorithm with exponential change of measure

$$F_\theta(dx) := \mathrm{e}^{-\theta x - \kappa(\theta)} F(dx)$$

produces an efficient algorithm if θ is such that $\mathbb{E}_\theta[X] = u/n$.

Right tail probabilities of sums

Fact

Severe difficulties occur in the construction of efficient algorithms in presence of heavy tails (Asmussen et al., 2000).

Heavy Tails

The Laplace transform (mgf) is not defined for a heavy-tailed random variable X.

Right tail probabilities of sums

Fact

Severe difficulties occur in the construction of efficient algorithms in presence of heavy tails (Asmussen et al., 2000).

Heavy Tails

The Laplace transform (mgf) is not defined for a heavy-tailed random variable X.

Literature Review

Heavy-tailed independent random variables

- Asmussen and Binswanger (1997). First efficient algorithm for regularly varying distribution.
- Asmussen and Kroese (2006). A refined version which is proved to be efficient in the Lognormal and Weibull case.
- Boots and Shahabuddin (2001) y Juneja and Shahabuddin (2002). Importance sampling (hazard rate).
- Dupuis et al. (2006). Importance sampling algorithm for regularly varying distributions and based on mixtures.
- Blanchet and Li (2011). Adaptive method for subexponential random variables.

Literature Review

Our contribution: Non-independent case (this talk)

- Blanchet et al. (2008). Strategies for Log-elliptical random variables.
- Asmussen et al. (2009). Considers correlated lognormal random variables.
- Blanchet and Rojas-Nandayapa (2012) Improved algorithm for tail probabilities of log-elliptics. An additional more general algorithm for sums.

Related papers

Klöppel et al. (2010) y Chan and Kroese (2009). Similar strategies.

Literature Review

Our contribution: Non-independent case (this talk)

- Blanchet et al. (2008). Strategies for Log-elliptical random variables.
- Asmussen et al. (2009). Considers correlated lognormal random variables.
- Blanchet and Rojas-Nandayapa (2012) Improved algorithm for tail probabilities of log-elliptics. An additional more general algorithm for sums.

Related papers

Klöppel et al. (2010) y Chan and Kroese (2009). Similar strategies.

Another interesting problem

Current work with S. Asmussen and J.L. Jensen.

Efficient estimation of

$$\mathbb{P}(X_1 + \cdots + X_n < nx), \qquad x \to 0.$$

- Exponential Twisting.
- Main difficulty: approximate the Laplace transform.
- Bounded relative error.

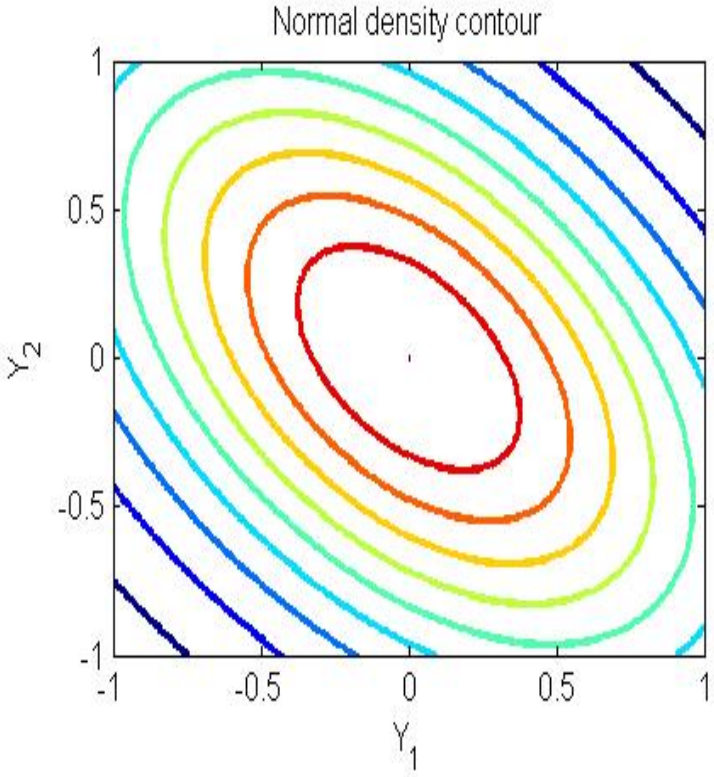
Outline

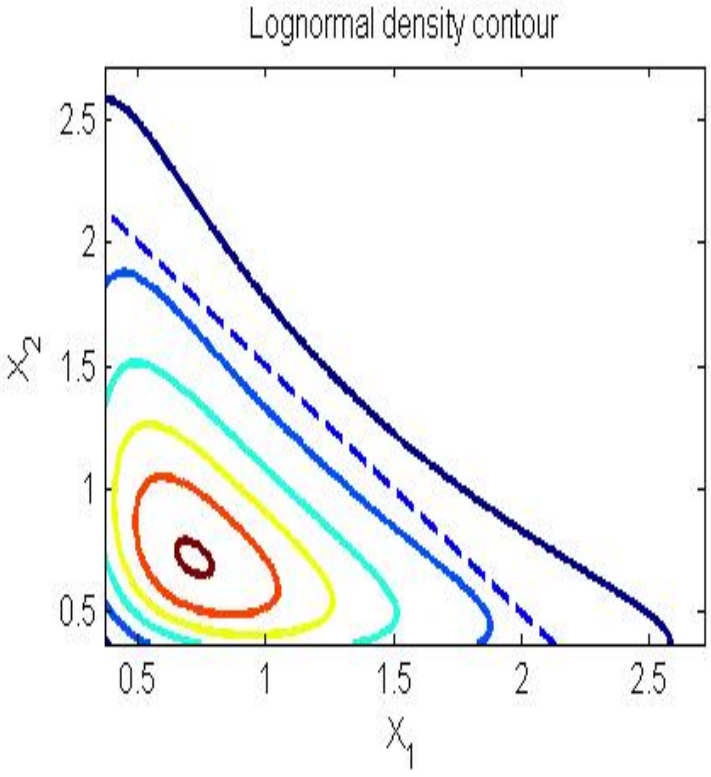
1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Exploratory analysis

Correlated Lognormals





Scaled variance algorithm

Asmussen et al. (2009)

Use the importance sampling distribution

$$\mathrm{LN}(\mu, \delta(x)\Sigma)$$

where $\delta(x)$ is the scaling function.

1. Under very mild conditions of $\delta(x)$: logarithmically efficient.
2. Cross-entropy selection: excellent numerical results

Scaled variance algorithm

Asmussen et al. (2009)

Use the importance sampling distribution

$$\mathrm{LN}(\mu, \delta(x)\Sigma)$$

where $\delta(x)$ is the scaling function.

1. Under very mild conditions of $\delta(x)$: logarithmically efficient.
2. Cross-entropy selection: excellent numerical results

Outline

1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Elliptical Distributions

Definition

$\mathbf{X}$ is elliptical, denoted $E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, if

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R\,\mathbf{C}\,\Theta.$$

- **Location**: $\boldsymbol{\mu} \in \mathbb{R}^n$. **Dispersion**: $\boldsymbol{\Sigma} = \mathbf{C}^t\mathbf{C}$ with $\mathbf{C} \in \mathbb{R}^{n\times k}$.
- **Spherical**: Θ uniform random vector on the unit spheroid. **Radial**: R positive random variable.
- R and Θ independent of each other.

Elliptical Distributions

Definition

$\mathbf{X}$ is elliptical, denoted $E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, if

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R\,\mathbf{C}\,\Theta.$$

- **Location**: $\boldsymbol{\mu} \in \mathbb{R}^n$. **Dispersion**: $\boldsymbol{\Sigma} = \mathbf{C}^t\mathbf{C}$ with $\mathbf{C} \in \mathbb{R}^{n\times k}$.
- **Spherical**: Θ uniform random vector on the unit spheroid. **Radial**: R positive random variable.
- R and Θ independent of each other.

Elliptical Distributions

Example

Logelliptical distributions

- Multivariate Normal.
- Normal Mixtures.
- Symmetric Generalized Hyperbolic Distributions: Hyperbolic Distributions, Multivariate Normal Inverse Gaussian (NIG), Generalized Laplace, Bessel or *Symmetric Variance-Gamma*, Multivariate t.

The *Symmetric Generalized Hyperbolic Distributions* offer better adjustments than the multivariate normal distributions in financial applications (McNeil et al., 2005).

Elliptical Distributions

Example

Logelliptical distributions

- Multivariate Normal.
- Normal Mixtures.
- Symmetric Generalized Hyperbolic Distributions: Hyperbolic Distributions, Multivariate Normal Inverse Gaussian (NIG), Generalized Laplace, Bessel or *Symmetric Variance-Gamma*, Multivariate t.

The *Symmetric Generalized Hyperbolic Distributions* offer better adjustments than the multivariate normal distributions in financial applications (McNeil et al., 2005).

Heavy Tails and Log-elliptical distributions

Log-elliptical Distributions

The exponential transformation (component-wise) of an elliptical random vector is known as *logelliptical*. Commonly the marginals are dependent heavy-tailed random variables.

Example

Sum of Logellipticals

$$G(r,\theta) = \sum_{i=1}^{d} \exp\left(\mu_i + r\left\langle \mathbf{A}_i, \theta \right\rangle\right).$$

Conditional Monte Carlo

Estimator

An unbiased estimator of $\mathbb{P}(G(R,\Theta) > u)$ is

$$\mathbb{P}(G(R,\Theta) > u|\Theta).$$

Algorithm

- Simulate Θ.
- Determine $B_\Theta := \{r > 0 : G(r,\Theta) > u\}$.
- Return $\mathbb{P}(R \in B_\Theta)$.

Conditional Monte Carlo

Estimator

An unbiased estimator of $\mathbb{P}(G(R,\Theta) > u)$ is

$$\mathbb{P}(G(R,\Theta) > u|\Theta).$$

Algorithm

- Simulate Θ.
- Determine $B_\Theta := \{r > 0 : G(r,\Theta) > u\}$.
- Return $\mathbb{P}(R \in B_\Theta)$.

Conditional Monte Carlo

Case 1: Sums of logelliptical random variables

Logarithmic efficient if

$$\lim_{r\to\infty} \frac{r\, f_R(r)}{\mathbb{P}(R > r)^{1-\varepsilon}} = 0 \qquad \forall \varepsilon > 0.$$

f_R is the density of F

Conditional Monte Carlo

Efficient for more general functions

Conditions

- G is continuous in the two variables and differentiable in r.
- There exists $\delta_0 > 0$, $s_* \in \mathcal{S}_d$, $r_0 > 0$ and $v > 0$ such that for all $0 < \delta \leq \delta_0$ and all $r > r_0$ it holds

$$\sup_{\theta \in \mathcal{S}} G(r,\theta)^{1-v\delta} \leq \inf_{\theta \in \mathcal{D}(\delta, s^*)} G(r,\theta)$$

$$\sup_{\theta \in \mathcal{D}(\delta_0, s^*)} G(r,\theta) = \sup_{\theta \in \mathcal{S}_d} G(r,\theta),$$

  where $\mathcal{D}(\delta, s^*) = \{\theta \in \mathcal{S}_d : \|\theta - s^*\| < \delta\}$.
- $0 < \delta_1 < 1$ chosen such that for all $r > r_0$ and $\theta \in \mathcal{D}(\delta, s^*)$ it holds

$$\delta_1 \leq \frac{d \log G(r,\theta)}{dr} \leq \frac{1}{\delta_1}.$$

Conditional Monte Carlo

Interesting cases where we have proved efficiency

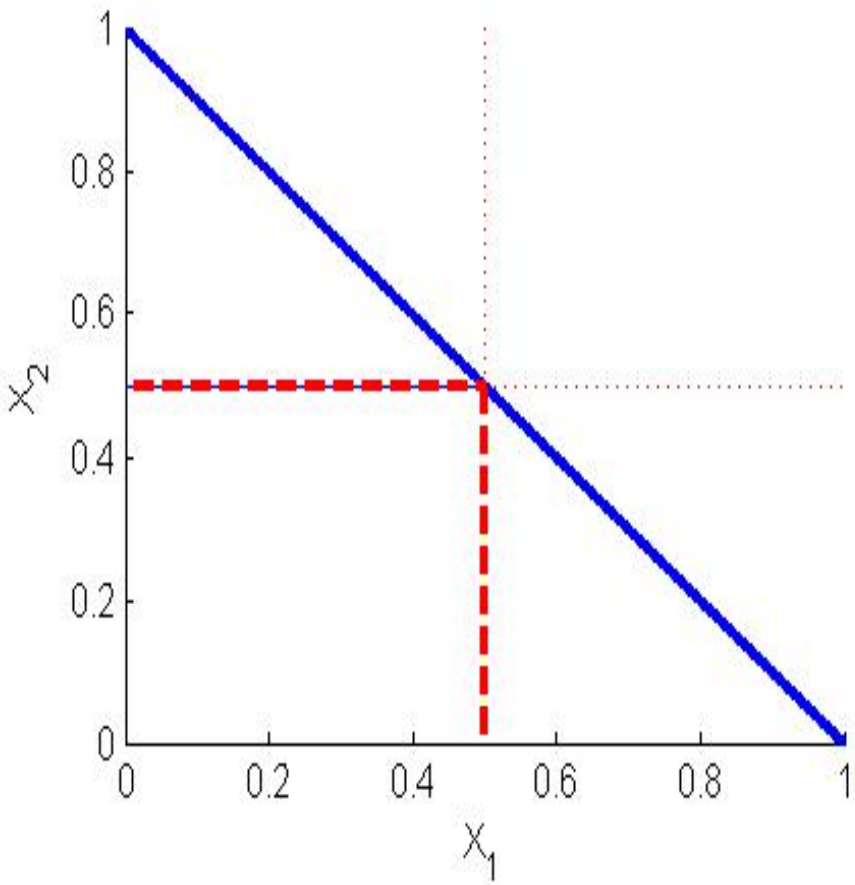
Sums, maxima, norms and portfolios of options with Symmetric Generalized Hyperbolic Distributions.

Outline

1. Introduction
   - Rare Event Simulation
   - Efficient Simulation for Sums of Random Variables
2. Efficient simulation of sums of dependent random variables
   - Scaled variance for sums of correlated lognormals
   - Conditional Monte Carlo for functions of logellipticals
   - Sums with general dependent heavy tails

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Exploratory analysis



Importance Sampling Distribution

Auxiliary IS distribution

- Set $b = \log(x/n)$.
- Take the distribution G_b of any efficient IS algorithm for

$$\mathbb{E}\left[\sum_{i=1}^{d} I(Y_i > b)\right].$$

- Use G_b as an IS distribution for estimating $\mathbb{P}(\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_n} > x)$.

Efficiency

Efficiency

The last algorithm is efficient if

$$\lim_{b\to\infty} \frac{\log \mathbb{P}\left(Y_i > b - c\right)}{\log \mathbb{P}\left(Y_i > b\right)} = 1.$$

We say Y_i are *Logarithmically Long Tailed*.

Observation

This condition includes the most practical heavy-tailed distributions.

Efficiency

Efficiency

The last algorithm is efficient if

$$\lim_{b\to\infty} \frac{\log \mathbb{P}\left(Y_i > b - c\right)}{\log \mathbb{P}\left(Y_i > b\right)} = 1.$$

We say Y_i are *Logarithmically Long Tailed*.

Observation

This condition includes the most practical heavy-tailed distributions.

Thanks!

Bibliography

Asmussen, S. and K. Binswanger (1997). Simulation of ruin probabilities for subexponential claims. *ASTIN Bulletin 27*, 297–318.

Asmussen, S., K. Binswanger, and B. Højgaard (2000). Rare events simulation for heavy-tailed distributions. *Bernoulli 6*, 303–322.

Asmussen, S., J. Blanchet, S. Juneja, and L. Rojas-Nandayapa (2009). Efficient simulation of tail probabilities of sums of correlated lognormals. *Annals of Operations Research*. To appear.

Asmussen, S. and D. P. Kroese (2006). Improved algorithms for rare event simulation with heavy tails. *Advances in Applied Probability 38*, 545–558.

Blanchet, J., S. Juneja, and L. Rojas-Nandayapa (2008). Efficient tail estimation for sums of correlated lognormals. In *Proceedings of the 2008 Winter Simulation Conference*, Miami, FL., USA, pp. 607–614. IEEE.

Blanchet, J. and C. Li (2011). Efficient rare-event simulation for heavy-tailed compound sums. *ACM TOMACS 21*. Forthcoming.

Blanchet, J. and L. Rojas-Nandayapa (2012). Efficient simulation of tail probabilities of sums of dependent random variables. *Journal of Applied Probability 48A*, 147–164.

Boots, N. K. and P. Shahabuddin (2001). Simulating ruin probabilities in insurance risk processes with subexponential claims. In *Proceedings of the 2001 Winter Simulation Conference*, Arlington, VA., USA, pp. 468–476. IEEE.

Chan, J. C. C. and D. Kroese (2009). Rare-event probability estimation with conditional monte carlo. *Annals of Operations Research*. To appear.

Dupuis, P., K. Leder, and H. Wang (2006). Importance sampling for sums of random variables with regularly varying tails. *ACM TOMACS 17*, 1–21.

Juneja, S. and P. Shahabuddin (2002). Simulating heavy-tailed processes using delayed hazard rate twisting. *ACM TOMACS 12*, 94–118.

Klöppel, S., R. Reda, and W. Scharchermayer (2010). A rotationally invariant technique for rare event simulation. *Risk Magazine 22*, 90–94.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. New Jersey: Princeton University Press.