Large Deviations, Importance Sampling, and Portfolio Credit Risk

Wanmo Kang
Department of Mathematical Science, KAIST, South Korea

Joint work with P. Glasserman (Columbia Univ.) and
the late P. Shahabuddin

Asymptotic Statistics, Risk and Computation in Finance and Insurance 2010
14th-18th December 2010, Tokyo Institute of Technology

December 17, 2010

Outline

- Rare Events, Large Deviations, and Importance Sampling
- Credit Risk, Portfolio Credit Risk
- Portfolio Credit Loss Model and Gaussian Copula Model
- Large Deviations Analysis of Portfolio Credit Risk[1]
- Importance Sampling for Portfolio Credit Risk[2]

1. GKS, Large Deviations in Multifactor Portfolio Credit Risk, *Mathematical Finance* July, 2007.
2. GKS, Fast Simulation of Multifactor Portfolio Credit Risk, *Operations Research* September, 2008.

Rare Events and Large Deviations

- Katrina, Tsunami, Sichuan, Black Monday, LTCM, 9/11, ...
- Very low chance of occurrences.
- Probability measured by the order of magnitude: Large deviations analysis.
- Crude Monte Carlo is impractical.

Crude Monte Carlo

- Crude Monte Carlo simulation of $\mathbb{E}[h(L)]$.
- Sample $h(L)$ independently and use the CLT.
  - Generate $L^{(1)}, \dots, L^{(n)}$;
  - Infer based on $\frac{\sum_{i=1}^{n} h(L^{(i)})}{n}$.
- In general, Monte Carlo is SIMPLE but SLOW!

Rare Event Simulation

- Estimation of $q = \mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ by $Q_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - $\text{Var}(Q_n) = \frac{q-q^2}{n} \approx \frac{q}{n}$ if $q \approx 0$.
- Half Width of CI
  $HW = C \times \sqrt{\text{Var}(Q_n)} = C \times \sqrt{\frac{q}{n}}$
- Relative Error
  $RE = \frac{HW}{q} = \frac{C}{\sqrt{qn}} \to \infty$ as $q \to 0$ for fixed n.

Rare Event Simulation

- Estimation of $q = \mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ by $Q_n = \frac{1}{n}\sum_{i=1}^n X_i$
  - $\mathrm{Var}(Q_n) = \frac{q-q^2}{n} \approx \frac{q}{n}$ if $q \approx 0$.
- Half Width of CI
  $HW = C \times \sqrt{\mathrm{Var}(Q_n)} = C \times \sqrt{\frac{q}{n}}$
- Relative Error
  $RE = \frac{HW}{q} = \frac{C}{\sqrt{qn}} \to \infty$ as $q \to 0$ for fixed n.

Rare Event Simulation

- Estimation of $q = \mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ by $Q_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - $\mathrm{Var}(Q_n) = \frac{q-q^2}{n} \approx \frac{q}{n}$ if $q \approx 0$.
- Half Width of CI
  $HW = C \times \sqrt{\mathrm{Var}(Q_n)} = C \times \sqrt{\frac{q}{n}}$
- Relative Error
  $RE = \frac{HW}{q} = \frac{C}{\sqrt{qn}} \to \infty$ as $q \to 0$ for fixed n.

Rare Event Simulation

- Estimation of $q = \mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ by $Q_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - $\text{Var}(Q_n) = \frac{q-q^2}{n} \approx \frac{q}{n}$ if $q \approx 0$.
- Half Width of CI

  $HW = C \times \sqrt{\text{Var}(Q_n)} = C \times \sqrt{\frac{q}{n}}$
- Relative Error

  $RE = \frac{HW}{q} = \frac{C}{\sqrt{qn}} \to \infty$ as $q \to 0$ for fixed n.

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- Speed-up the Monte Carlo simulation of

$$\mathbb{E}_f[h(L)] = \int h(x)f(x)dx$$

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Importance Sampling (IS):
  1. Choose well a new density $g(\cdot)$ which approximates the importance function $h(\cdot) \times f(\cdot)$;
  2. Generate iid $L^{(i)}$'s from $g(\cdot)$;
  3. Infer from $\frac{1}{n}\sum_{i=1}^{n} h(L^{(i)})f(L^{(i)})/g(L^{(i)})$.

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- Speed-up the Monte Carlo simulation of

$$\mathbb{E}_f[h(L)] = \int h(x)f(x)dx$$

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Importance Sampling (IS):
  1. Choose well a new density $g(\cdot)$ which approximates the importance function $h(\cdot) \times f(\cdot)$;
  2. Generate iid $L^{(i)}$'s from $g(\cdot)$;
  3. Infer from $\frac{1}{n}\sum_{i=1}^{n} h(L^{(i)})f(L^{(i)})/g(L^{(i)})$.

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- Speed-up the Monte Carlo simulation of

$$\mathbb{E}_f[h(L)] = \int h(x)f(x)dx$$

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Importance Sampling (IS):
  1. Choose well a new density $g(\cdot)$ which approximates the importance function $h(\cdot) \times f(\cdot)$;
  2. Generate iid $L^{(i)}$'s from $g(\cdot)$;
  3. Infer from $\frac{1}{n}\sum_{i=1}^{n} h(L^{(i)})f(L^{(i)})/g(L^{(i)})$.

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Second moment of IS estimator

$$\mathbb{E}_f[h(L)^2] \quad \text{vs.} \quad \mathbb{E}_f\left[\frac{h(L)^2 f(L)}{g(L)}\right]$$

- Importance function: If $g(\cdot) \sim h(\cdot) \times f(\cdot)$, then the variance is 0.
- How to choose a GOOD $g(\cdot)$?

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Second moment of IS estimator

$$\mathbb{E}_f[h(L)^2] \quad \text{vs.} \quad \mathbb{E}_f\left[\frac{h(L)^2 f(L)}{g(L)}\right]$$

- Importance function: If $g(\cdot) \sim h(\cdot) \times f(\cdot)$, then the variance is 0.
- How to choose a GOOD $g(\cdot)$?

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Second moment of IS estimator

$$\mathbb{E}_f[h(L)^2] \quad \text{vs.} \quad \mathbb{E}_f\left[\frac{h(L)^2 f(L)}{g(L)}\right]$$

- Importance function: If $g(\cdot) \sim h(\cdot) \times f(\cdot)$, then the variance is 0.
- How to choose a GOOD $g(\cdot)$?

Variance Reduction via Importance Sampling

Assume the existence of a density $f(\cdot)$ of a stochastic loss L.

- $\mathbb{E}_f[h(L)] = \mathbb{E}_g\left[\frac{h(L)f(L)}{g(L)}\right]$
- Second moment of IS estimator

$$\mathbb{E}_f[h(L)^2] \quad \text{vs.} \quad \mathbb{E}_f\left[\frac{h(L)^2 f(L)}{g(L)}\right]$$

- Importance function: If $g(\cdot) \sim h(\cdot) \times f(\cdot)$, then the variance is 0.
- How to choose a GOOD $g(\cdot)$?

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha \quad \text{or} \quad \mathbb{P}(A_n) = e^{-\alpha n + o(n)}.$$

- For another probability measure $\mathbb{Q}$ and its Radon-Nikodym derivative $Z = \frac{d\mathbb{P}}{d\mathbb{Q}}$,
  - New unbiased estimator via change of measures

$$\mathbb{P}(A_n) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] = \mathbb{E}_{\mathbb{Q}}\left[Z\mathbf{1}_{A_n}\right]$$

  - Second moment of new estimator

$$\mathbb{E}_{\mathbb{Q}}\left[Z^2\mathbf{1}_{A_n}\right]$$

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha \quad \text{or} \quad \mathbb{P}(A_n) = e^{-\alpha n + o(n)}.$$

- For another probability measure $\mathbb{Q}$ and its Radon-Nikodym derivative $Z = \frac{d\mathbb{P}}{d\mathbb{Q}}$,
  - New unbiased estimator via change of measures

$$\mathbb{P}(A_n) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] = \mathbb{E}_{\mathbb{Q}}\left[Z\mathbf{1}_{A_n}\right]$$

  - Second moment of new estimator

$$\mathbb{E}_{\mathbb{Q}}\left[Z^2\mathbf{1}_{A_n}\right]$$

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha \quad \text{or} \quad \mathbb{P}(A_n) = e^{-\alpha n + o(n)}.$$

- For another probability measure $\mathbb{Q}$ and its Radon-Nikodym derivative $Z = \frac{d\mathbb{P}}{d\mathbb{Q}}$,
  - New unbiased estimator via change of measures

$$\mathbb{P}(A_n) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] = \mathbb{E}_{\mathbb{Q}}\left[Z\mathbf{1}_{A_n}\right]$$

  - Second moment of new estimator

$$\mathbb{E}_{\mathbb{Q}}\left[Z^2\mathbf{1}_{A_n}\right]$$

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$, $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha$.
- Assume that we find a sequence $\{\mathbb{Q}_n\}$ and $\{Z_n = \frac{d\mathbb{P}}{d\mathbb{Q}_n}\}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[Z_n^2 \mathbf{1}_{A_n}\right] \le -2\alpha.$$

- Since

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] &\ge \liminf_{n\to\infty} \frac{1}{n} \log \left(\left(\mathbb{E}_{\mathbb{Q}_n}[\mathbf{1}_{A_n} Z_n]\right)^2 \right) \\ &= 2 \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] \\ &= -2\alpha, \end{aligned}$$

- $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] = -2\alpha$

implies the optimality of $\{\mathbb{Q}_n\}$.

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$, $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha$.
- Assume that we find a sequence $\{\mathbb{Q}_n\}$ and $\{Z_n = \frac{d\mathbb{P}}{d\mathbb{Q}_n}\}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[Z_n^2 \mathbf{1}_{A_n}\right] \leq -2\alpha.$$

- Since

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] &\geq \liminf_{n\to\infty} \frac{1}{n} \log\left(\left(\mathbb{E}_{\mathbb{Q}_n}[\mathbf{1}_{A_n} Z_n]\right)^2\right) \\ &= 2 \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] \\ &= -2\alpha, \end{aligned}$$

- $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] = -2\alpha$

implies the optimality of $\{\mathbb{Q}_n\}$.

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$, $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha$.
- Assume that we find a sequence $\{\mathbb{Q}_n\}$ and $\{Z_n = \frac{d\mathbb{P}}{d\mathbb{Q}_n}\}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[Z_n^2 \mathbf{1}_{A_n}\right] \leq -2\alpha.$$

- Since

$$\begin{aligned} & \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] \\ \geq\ & \liminf_{n\to\infty} \frac{1}{n} \log\left(\left(\mathbb{E}_{\mathbb{Q}_n}[\mathbf{1}_{A_n} Z_n]\right)^2\right) \\ =\ & 2 \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] \\ =\ & -2\alpha, \end{aligned}$$

- $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] = -2\alpha$

implies the optimality of $\{\mathbb{Q}_n\}$.

Large Deviations and Optimal Importance Sampling

- For a rare event sequence $\{A_n\}$, $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n) = -\alpha$.
- Assume that we find a sequence $\{\mathbb{Q}_n\}$ and $\{Z_n = \frac{d\mathbb{P}}{d\mathbb{Q}_n}\}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[Z_n^2 \mathbf{1}_{A_n}\right] \leq -2\alpha.$$

- Since

$$\begin{aligned} & \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] \\ \geq\ & \liminf_{n\to\infty} \frac{1}{n} \log \left(\left(\mathbb{E}_{\mathbb{Q}_n}[\mathbf{1}_{A_n} Z_n] \right)^2 \right) \\ =\ & 2 \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n}] \\ =\ & -2\alpha, \end{aligned}$$

- $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}_n}\left[(\mathbf{1}_{A_n} Z_n)^2\right] = -2\alpha$

implies the optimality of $\{\mathbb{Q}_n\}$.

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

  where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

  where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.
- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

  where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

  where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.
- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

  where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

  where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.
- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.
- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

  where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

  where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.

- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

An IS Estimator for Rare Event Simulation

Consider $\mathbb{P}(S_n > n(\mu + \epsilon))$ where $S_n = X_1 + \ldots + X_n$ and X_i IID with mean μ.

- Restrict $g(\cdot)$ (here $\mathbb{P}_\theta$) among an exponential family:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = e^{\theta \cdot S_n - n \cdot \psi(\theta)}$$

  where $\psi(\theta) = \log \mathbb{E}[e^{\theta X_1}]$.

- How to choose θ?

$$\theta : \psi'(\theta) = \mu + \epsilon$$

  where $\psi'(\theta) = \frac{\mathbb{E}[X_1 e^{\theta X_1}]}{\mathbb{E}[e^{\theta X_1}]} = \mathbb{E}_\theta[X_1]$.

- It can be shown that $\mathbb{P}_\theta$ is *logarithmically optimal*.
- Heavy tailed cases: Asmussen, Binswanger, & Hojgaard (2000), Juneja & Shahabuddin (2002)

Importance Sampling: Call Option

Table: Variances of Monte Carlo estimators for the original BM and the drifted BM using Girsanov Theorem. $S_0 = 100$, $r = 0.05$, $\sigma = 0.3$.

Strike Price K	80	100	120	140	160	180	200
Option Value	26.6	14.3	7.0	3.2	1.4	0.6	0.2
STD for Crude MC	27.8	22.6	16.6	9.3	7.5	4.9	3.1
STD for MC with IS	97.4	27.8	10.0	2.6	1.7	0.7	0.3

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Credit Risk Models

- Single name vs. Portfolio
- Static vs. Dynamic
- Pricing vs. Risk Management
- Structural vs. Reduced-form
- Light-tailed vs. Heavy-tailed
  (Bassamboo, Juneja, Zeevi (2008), Chan, Kroese (2010))
- Bottom-up vs. Top-down
  (Longstaff, Rajan (2008), Errais, Giesecke, Goldberg (2009))

Modelling of Portfolio Credit Risk

	Structural Model	Reduced-form Model
Dependence	Asset Values	Intensities

- Copula function *couples* the uniform marginal distributions
- Copula function separates the dependence structure from the marginals

Copula Function

- Copula function, $C(\cdot)$, *couples* the uniform marginal distributions.

$$C(u_1, \ldots, u_n) = \mathbb{P}(U_1 \le u_1, \ldots, U_n \le u_n).$$

  - $C : [0,1]^n \to [0,1]$ and U_k's are uniform r.v.'s on $[0,1]$.
  - Desire to find appropriate $C_F(\cdot)$ such that
    $F(x_1, \ldots, x_n) = C_F(F_1(x_1), \ldots, F_n(x_n))$
    since $F_k(X_k)$ is a standard uniform if $X_k \sim F_k(\cdot)$.

- For any $F(\cdot)$,
  $F(x_1, \ldots, x_n) = F(F_1^{-1}(F_1(x_1)), \ldots, F_n^{-1}(F_n(x_n)))$. So

$$C_F(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)).$$

Gaussian Copula Function

- Gaussian Copula

$$C_{Gauss}(u_1, \ldots, u_n) = \Phi_\Sigma(\Phi_1^{-1}(u_1), \ldots, \Phi_n^{-1}(u_n))$$

where Φ_Σ stands for a multivariate normal CDF with zero mean and Σ correlation.

- Gaussian Copula as the Dependece Structure of Default Times, T_k with Gaussian marginals i.e. $F_i(T_i) = \Phi_i(X_i)$ where $(X_1, \ldots, X_n)$ is multivariate normal.
  - $F(t_1, \ldots, t_n) = C_{Gauss}(F_1(t_1), \ldots, F_n(t_n)) = \Phi_\Sigma(x_1, \ldots, x_n)$

Gaussian Copula Function

- Gaussian Copula

$$C_{Gauss}(u_1, \ldots, u_n) = \Phi_\Sigma(\Phi_1^{-1}(u_1), \ldots, \Phi_n^{-1}(u_n))$$

where Φ_Σ stands for a multivariate normal CDF with zero mean and Σ correlation.

- Gaussian Copula as the Dependece Structure of Default Times, T_k with Gaussian marginals i.e. $F_i(T_i) = \Phi_i(X_i)$ where $(X_1, \ldots, X_n)$ is multivariate normal.
  - $F(t_1, \ldots, t_n) = C_{Gauss}(F_1(t_1), \ldots, F_n(t_n)) = \Phi_\Sigma(x_1, \ldots, x_n)$

Single-Period Portfolio Credit Loss Model

$L_m = \ell_1 Y_1 + \cdots + \ell_m Y_m$: Total loss from defaults

- ℓ_k : Loss resulting from default of k-th obligor
- Y_k : Default indicator ($= 0$ or 1) for k-th obligor
- p_k : Marginal probability that k-th obligor defaults
- m : The number of obligors

For some l, what is $\mathbb{P}(L_m > l)$?

For the portfolio credit risk, characterization of dependence structure among defaults is very important.

Single-Period Portfolio Credit Loss Model

$L_m = \ell_1 Y_1 + \cdots + \ell_m Y_m$: Total loss from defaults

- ℓ_k : Loss resulting from default of k-th obligor
- Y_k : Default indicator ($= 0$ or 1) for k-th obligor
- p_k : Marginal probability that k-th obligor defaults
- m : The number of obligors

For some l, what is $\mathbb{P}(L_m > l)$?

For the portfolio credit risk, characterization of dependence structure among defaults is very important.

Modeling via Asset Values

- Recall $L_m = \ell_1 Y_1 + \cdots + \ell_m Y_m$: total loss from defaults
- Let $(X_1, \ldots, X_m)$ be $N(0,1)$ variables (called latent variables representing relative *asset values*) such that for some x_k,

$$Y_k = \mathbf{1}\{X_k > x_k\} = \begin{cases} 1 & \text{if } X_k > x_k \\ 0 & \text{otherwise.} \end{cases}$$

- Select x_k such that

$$\mathbb{P}(Y_k = 1) = \mathbb{P}(X_k > x_k) = p_k \quad \text{or} \quad x_k = \Phi^{-1}(1 - p_k) \; .$$

Dependence Structure: Gaussian Copula Model

- $X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k \quad , k = 1, \ldots, m.$
  - $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_d)$: Systematic risk factors,

    e.g. Macro economic indices, Country factors, Industrial Sectors.
  - $\varepsilon_k \sim N(0,1)$, independent of $\mathbf{Z}$ and other $\varepsilon_{k'}$, $k' \neq k$: Idiosyncratic risks.
- Note that X_k's are independent given $\mathbf{Z}$: *Conditional Independence.*
- Industry standard for portfolio credit risk.
- Neither analytical nor numerical results exist for $\mathbb{P}(L_m > l)$.
- Large deviations and rare event simulation are viable alternatives.

Dependence Structure: Gaussian Copula Model

- $X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k \quad , k = 1, \ldots, m.$
  - $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_d)$: Systematic risk factors,

    e.g. Macro economic indices, Country factors, Industrial Sectors.
  - $\varepsilon_k \sim N(0,1)$, independent of $\mathbf{Z}$ and other $\varepsilon_{k'}$, $k' \neq k$: Idiosyncratic risks.
- Note that X_k's are independent given $\mathbf{Z}$: *Conditional Independence.*
- Industry standard for portfolio credit risk.
- Neither analytical nor numerical results exist for $\mathbb{P}(L_m > l)$.
- Large deviations and rare event simulation are viable alternatives.

Dependence Structure: Gaussian Copula Model

- $X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k \quad , k = 1, \ldots, m.$
  - $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_d)$: Systematic risk factors,

    e.g. Macro economic indices, Country factors, Industrial Sectors.
  - $\varepsilon_k \sim N(0,1)$, independent of $\mathbf{Z}$ and other $\varepsilon_{k'}$, $k' \neq k$: Idiosyncratic risks.
- Note that X_k's are independent given $\mathbf{Z}$: *Conditional Independence.*
- Industry standard for portfolio credit risk.
- Neither analytical nor numerical results exist for $\mathbb{P}(L_m > l)$.
- Large deviations and rare event simulation are viable alternatives.

Dependence Structure: Gaussian Copula Model

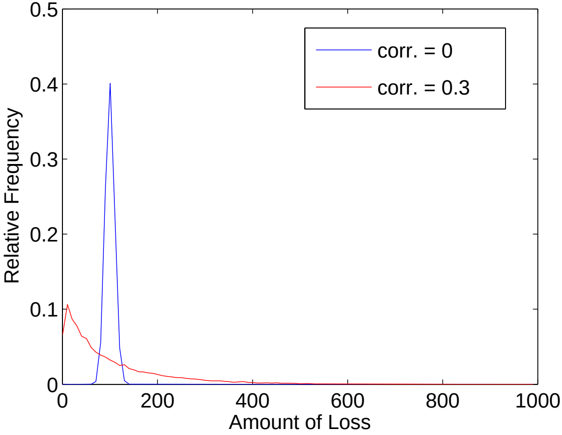
- $X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k \quad , k = 1, \ldots, m.$
  - $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_d)$: Systematic risk factors,
    e.g. Macro economic indices, Country factors, Industrial Sectors.
  - $\varepsilon_k \sim N(0,1)$, independent of $\mathbf{Z}$ and other $\varepsilon_{k'}$, $k' \neq k$: Idiosyncratic risks.
- Note that X_k's are independent given $\mathbf{Z}$: *Conditional Independence.*
- Industry standard for portfolio credit risk.
- Neither analytical nor numerical results exist for $\mathbb{P}(L_m > l)$.
- Large deviations and rare event simulation are viable alternatives.

Dependence Structure: Gaussian Copula Model

- $X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k \quad , k = 1, \ldots, m.$
  - $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_d)$: Systematic risk factors,

    e.g. Macro economic indices, Country factors, Industrial Sectors.
  - $\varepsilon_k \sim N(0, 1)$, independent of $\mathbf{Z}$ and other $\varepsilon_{k'}$, $k' \neq k$: Idiosyncratic risks.
- Note that X_k's are independent given $\mathbf{Z}$: *Conditional Independence.*
- Industry standard for portfolio credit risk.
- Neither analytical nor numerical results exist for $\mathbb{P}(L_m > l)$.
- Large deviations and rare event simulation are viable alternatives.

Effects of Dependence



$$L = \mathbf{1}\{X_1 > x\} + \ldots + \mathbf{1}\{X_{1000} > x\}, \quad X_k = \rho N + \sqrt{1-\rho^2} N_k, \quad p_k = 0.1$$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model $(p_k \equiv p,\ c_k \equiv c,\ \mathbf{a}_j \equiv \rho \in \mathbb{R})$ i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L > x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\boldsymbol{\mu} = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L>x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}),\mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}),\mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\boldsymbol{\mu} = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}),\mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L > x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\boldsymbol{\mu} = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L>x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\mu = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L>x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\boldsymbol{\mu} = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L>x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x|\mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\mu = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L > x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\mu = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

Glasserman and Li (2005)

- Large deviations results for homogeneous single factor model ($p_k \equiv p$, $c_k \equiv c$, $\mathbf{a}_j \equiv \rho \in \mathbb{R}$) i.e. $X_k = \rho Z + \sqrt{1-\rho^2}\varepsilon_k$
- Provably efficient importance sampling algorithm for homogeneous single factor model.
- Importance sampling procedure for heterogeneous multifactor cases.
- $\psi(\theta, \mathbf{z}) = \mathbb{E}[e^{\theta L} | \mathbf{Z} = \mathbf{z}]$
- $\theta_x(\mathbf{z}) : \frac{\partial \psi(\theta, \mathbf{z})}{\partial \theta} = x$
- $\mathbb{P}(L > x | \mathbf{Z}) = \mathbb{E}[\mathbf{1}_{\{L>x\}} | \mathbf{Z}] \leq \mathbb{E}[e^{\theta_x(\mathbf{Z})(L-x)} | \mathbf{Z}] = e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}$
- $\mathbb{P}(L > x) = \mathbb{E}[\mathbb{P}(L > x | \mathbf{Z})] \leq \mathbb{E}[e^{\psi(\theta_x(\mathbf{Z}), \mathbf{Z}) - x\theta_x(\mathbf{Z})}]$
- $\boldsymbol{\mu} = \mathrm{argmax}_{\mathbf{z}}\, e^{\psi(\theta_x(\mathbf{z}), \mathbf{z}) - x\theta_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z}}$

GKS 2007, 2008

- For heterogeneous multifactor Gaussian copula model,
  - Large deviations results;
  - Provably efficient importance sampling procedure.
- For the t-Copula model,
  - Heuristic importance sampling procedure.

Finite Types Assumption

- Recall the dependence structure,

$$X_k = \mathbf{a}_k^\top \mathbf{Z} + b_k \varepsilon_k.$$

- We partition the m obligors into t types, t: fixed.
  - $\{1, \ldots, m\} = \bigcup_{j=1}^t \mathcal{I}_j^{(m)} \quad (\mathcal{I}_j^{(m)}$: disjoint $)$
  - Obligors belonging to the same type have the same $\mathbf{a}_k$ and b_k.
  - $r_j = \lim_{m\to\infty} \frac{\left|\mathcal{I}_j^{(m)}\right|}{m}$ and $C_j = \lim_{m\to\infty} \frac{1}{m} \sum_{k\in\mathcal{I}_j^{(m)}} \mathbb{E}[\ell_k]$

Two Rare Event Regimes

$$\mathbb{P}(L_m > l) = \mathbb{P}\left(\sum_{k=1}^{m} c_k Y_k > l\right) = \mathbb{P}\left(\sum_{k=1}^{m} c_k \mathbf{1}_{\{X_k > x_k\}} > l\right)$$

(Recall $x_k = \Phi^{-1}(1 - p_k)$)

- Large Loss Threshold (LLT) : e.g. 70% loss in one year. l is large.
- Small Default Probability (SDP) : e.g. 5% loss in one week. x_k is large, i.e. p_k is small.

Two Rare Event Regimes

$$\mathbb{P}(L_m > l) = \mathbb{P}\left(\sum_{k=1}^m c_k Y_k > l\right) = \mathbb{P}\left(\sum_{k=1}^m c_k \mathbf{1}_{\{X_k > x_k\}} > l\right)$$

(Recall $x_k = \Phi^{-1}(1 - p_k)$)

Make l or p_k depend on m, i.e. $p_k^{(m)}$ and l_m. Then increase m to ∞.

- Large Loss Threshold (LLT): Large l_m and moderate p_k. We use $l_m = \Phi(s\sqrt{\log m}) \sum_{k=1}^m \mathbb{E}[\ell_k]$ where $0 < s < 1$ and p_k independent of m.
- Small Default Probability (SDP): Small $p_k^{(m)}$ and moderate l_m. $\ell_k = c_k$. We use $p_k^{(m)} = \Phi(-s_j\sqrt{m})$ where $s_j > 0$, $l_m = q \sum_{k=1}^m c_k$ where $0 < q < 1$.

Two Rare Event Regimes

- $\mathbb{P}(L_m > l) = \mathbb{P}\left(\sum_{k=1}^m c_k Y_k > l\right) = \mathbb{P}\left(\sum_{k=1}^m c_k \mathbf{1}_{\{X_k > x_k\}} > l\right)$
  (Recall $x_k = \Phi^{-1}(1 - p_k)$)
- Make l or p_k depend on m, i.e. $p_k^{(m)}$ and l_m and increase m to ∞.
  - Large Loss Threshold (LLT)
    - e.g. 70% loss in one year. l is large.
    - Large l_m and moderate p_k: $l_m = \Phi(s\sqrt{\log m}) \sum_{k=1}^m \mathbb{E}[\ell_k]$ where $0 < s < 1$ and p_k independent of m.
  - Small Default Probability (SDP)
    - e.g. 5% loss in one week. x_k is large, i.e. p_k is small.
    - Small $p_k^{(m)}$ and moderate l_m: $\ell_k = c_k$, $p_k^{(m)} = \Phi(-s_j\sqrt{m})$ where $s_j > 0$, $l_m = q\sum_{k=1}^m c_k$ where $0 < q < 1$.

Large Deviations Analysis

in the Gaussian Copula Model

Large Deviations Results for SDP

Theorem (GKS 2007)

Under finite types assumption, $0 < \ell_k = c_k \le \overline{c} < \infty$, $p_k^{(m)} = \Phi(-s_j\sqrt{m})$ *where* $s_j > 0$ *and* $l_m = q\sum_{k=1}^m c_k$ *with* $0 < q < 1$,

$$\lim_{m\to\infty} \frac{1}{m} \log \mathbb{P}(L_m > l_m) = -\frac{1}{2}\|\gamma_*\|^2$$

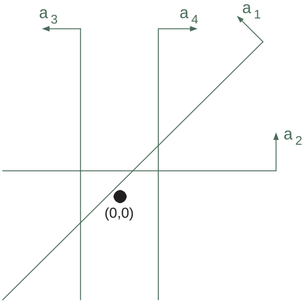
where

$$\begin{aligned}
\mathcal{M}_q &= \Big\{\mathcal{J} \in \{1,\ldots,t\} : \max_{\mathcal{J}' \subsetneq \mathcal{J}} \sum_{j\in\mathcal{J}'} C_j < qC < \sum_{j\in\mathcal{J}} C_j\Big\}, \\
G_{\mathcal{J}} &= \{\mathbf{z} \ : \ \mathbf{a}_j^\top \mathbf{z} \ge s_j \ , j \in \mathcal{J}\} \quad \textit{for } \mathcal{J} \in \mathcal{M}_q, \\
\gamma_{\mathcal{J}} &= \begin{cases} \operatorname{argmin}\{\|\mathbf{z}\| \ : \ \mathbf{z} \in G_{\mathcal{J}}\} & \textit{if } G_{\mathcal{J}} \ne \emptyset \\ (\infty,\ldots,\infty)^\top & \textit{if } G_{\mathcal{J}} = \emptyset, \end{cases} \\
\|\gamma_*\| &= \min_{\mathcal{J}\in\mathcal{M}_q} \|\gamma_{\mathcal{J}}\|.
\end{aligned}$$

Example: Two Risk Factors, 4 Obligor Types

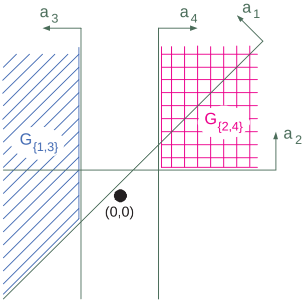
- $C_1 = 2,\ C_2 = 2,\ C_3 = 3,\ C_4 = 3$: Maximum average loss for each type.
- $C = 10$: Maximum loss of portfolio.
- $q = 0.45$: 45% loss $qC = 4.5$: Threshold
- $$\begin{aligned} \mathcal{M}_q &= \left\{ \mathcal{J} : \max_{\mathcal{J}' \subsetneq \mathcal{J}} \sum_{j \in \mathcal{J}'} C_j < qC < \sum_{j \in \mathcal{J}} C_j \right\} \\ &= \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}. \end{aligned}$$

Example: Two Risk Factors, 4 Obligor Types



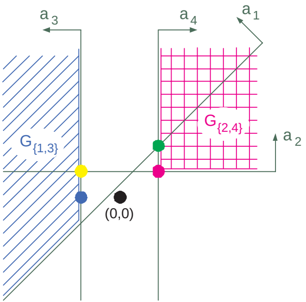
- $C_1 = 2,\ C_2 = 2,\ C_3 = 3,\ C_4 = 3.$
- $C = 10.$
- $q = 0.45, \quad qC = 4.5.$
- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}.$
- $G_{\mathcal{J}} = \{\mathbf{z} \ : \ \mathbf{a}_j^\top \mathbf{z} \geq s_j \ , j \in \mathcal{J}\}.$

Example: Two Risk Factors, 4 Obligor Types



- $C_1 = 2,\ C_2 = 2,\ C_3 = 3,\ C_4 = 3.$
- $C = 10.$
- $q = 0.45, \quad qC = 4.5.$
- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}.$
- $G_{\{1,3\}} = ///,\ G_{\{2,3\}} \subset G_{\{1,3\}},\ G_{\{2,4\}} = \boxplus,$ $G_{\{1,4\}} \subset G_{\{2,4\}},\ G_{\{3,4\}} = \emptyset$
- $\gamma_{\mathcal{J}} = \begin{cases} \operatorname{argmin}\{\|\mathbf{z}\| \ : \ \mathbf{z} \in G_{\mathcal{J}}\} & \text{if } G_{\mathcal{J}} \neq \emptyset \\ (\infty, \ldots, \infty)^\top & \text{if } G_{\mathcal{J}} = \emptyset, \end{cases}$

Example: Two Risk Factors, 4 Obligor Types



- $C_1 = 2,\ C_2 = 2,\ C_3 = 3,\ C_4 = 3.$
- $C = 10.$
- $q = 0.45, \quad qC = 4.5.$
- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}.$
- $G_{\{1,3\}} = ///$, $G_{\{2,3\}} \subset G_{\{1,3\}}$, $G_{\{2,4\}} = \boxplus$, $G_{\{1,4\}} \subset G_{\{2,4\}}$, $G_{\{3,4\}} = \emptyset$
- $\gamma_{\{1,3\}} = \bullet$, $\gamma_{\{1,4\}} = \bullet$, $\gamma_{\{2,3\}} = \bullet$, $\gamma_{\{2,4\}} = \bullet$, $\gamma_{\{3,4\}} = (\infty, \infty).$
- $\gamma_* = \bullet.$

Fast Simulation in the Monte Carlo Simulation of the Gaussian Copula Model

Importance Sampling for Gaussian Copula

- Given $\mathbf{Z}$, Y_k's are independent binary random variables.
- There exist standard importance sampling(IS) procedures, involving exponential twisting, to estimate the probability of the sum of independent random variables exceeding a given threshold, i.e. $\sum_{k=1}^{m} c_k Y_k > l$.
- Hence one procedure is: Generate $\mathbf{Z}$ and use conditional (on $\mathbf{Z}$) IS.
- However we also need to change the measure of $\mathbf{Z}$, so that there is greater chance of $\sum_{k=1}^{m} c_k Y_k > l$ (given $\mathbf{Z}$).
- This is accomplished by shifting the mean of $\mathbf{Z}$.
- We derive an appropriate shift and prove that it is asymptotically optimal.

Conditional IS on Defaults

- CGF of LGD: $\Lambda_k(\theta) = \log \mathbb{E}\left[e^{\theta \ell_k}\right]$.
- Exponential twisting: $p_{k,\theta}(\mathbf{Z}) = \frac{p_k(\mathbf{Z})e^{\Lambda_k(\theta)}}{1+p_k(\mathbf{Z})\left(e^{\Lambda_k(\theta)}-1\right)}$.
- Likelihood ratio:

$$\prod_{k=1}^{m} \left(\frac{p_k(\mathbf{Z})}{p_{k,\theta}(\mathbf{Z})}\right)^{Y_k} \left(\frac{1-p_k(\mathbf{Z})}{1-p_{k,\theta}(\mathbf{Z})}\right)^{1-Y_k} = e^{-\sum_{k=1}^{m} Y_k \Lambda_k(\theta) + m\psi_m(\theta,\mathbf{Z})}$$

- Conditional CGF of Portfolio Loss:

$$\begin{aligned} \psi_m(\theta, \mathbf{z}) &= \frac{1}{m} \log \mathbb{E}\left[e^{\theta L_m} \middle| \mathbf{Z} = \mathbf{z}\right] \\ &= \frac{1}{m} \sum_{k=1}^{m} \log\left(1 + p_k(\mathbf{z})\left(e^{\Lambda_k(\theta)} - 1\right)\right). \end{aligned}$$

Conditional IS on Loss Given Default

- Exponential twisting: $f_{\ell_k,\theta}(\ell) = f_{\ell_k}(\ell)e^{\theta Y_k \ell - \Lambda_k(\theta Y_k)}$.
- Likelihood ratio:

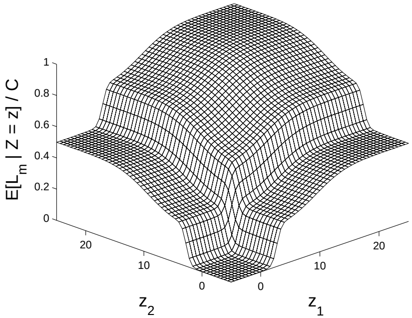
$$\prod_{k=1}^{m} \frac{f_{\ell_k}(\ell_k)}{f_{\ell_k,\theta}(\ell_k)} = \prod_{k=1}^{m} e^{-\theta Y_k \ell_k + \Lambda_k(Y_k\theta)} = e^{-\theta \sum_{k=1}^{m} Y_k \ell_k + \sum_{k=1}^{m} \Lambda_k(Y_k\theta)}.$$

- Combined likelihood ratio:

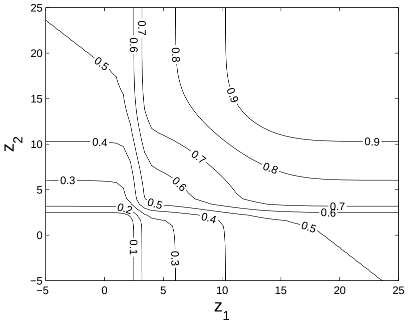
$$e^{-\theta L_m + m\psi_m(\theta, \mathbf{Z})}.$$

Conditional Importance Function

$$\mathbb{E}[\ell_k] \equiv 1,\ p_k \equiv 0.01,\ \mathbf{a}_1^\top = (0.85, 0),\ \mathbf{a}_2^\top = (0, 0.25),\ \mathbf{a}_3^\top = (0, 0.85),\ \mathbf{a}_4^\top = (0, 0.25).$$



3D plot of $\frac{1}{C}\mathbb{E}[L_m|\mathbf{Z} = \mathbf{z}]$



Contours of $\frac{1}{C}\mathbb{E}[L_m|\mathbf{Z} = \mathbf{z}]$

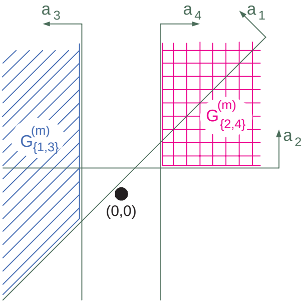
Mean Shifting of Common Factors

- $G_j^{(m)} \triangleq \left\{ \mathbf{z} \in \mathbb{R}^d : \mathbf{a}_j^\top \mathbf{z} \geq \alpha_1^{(m)} \Phi^{-1}(1 - \overline{p}_j) + \alpha_2^{(m)} b_j \Phi^{-1}(q) \right\}$.
- $G_{\mathcal{J}}^{(m)} \triangleq \bigcap_{j \in \mathcal{J}} G_j^{(m)}$ for $\mathcal{J} \in \mathcal{M}_q$
- $G^{(m)} \triangleq \bigcup_{\mathcal{J} \in \mathcal{M}_q} G_{\mathcal{J}}^{(m)}$.
- Sufficient subfamily, $\mathcal{S}_q$:
  - Feasibility: For each $\mathcal{J} \in \mathcal{S}_q$, $G_{\mathcal{J}}^{(m)} \neq \emptyset$ for all m;
  - Covering property: $\bigcup_{\mathcal{J} \in \mathcal{S}_q} G_{\mathcal{J}}^{(m)} = G^{(m)}$ for all m.
- $\mu_{\mathcal{J}}^{(m)} \triangleq \operatorname{argmin} \left\{ \|\mathbf{z}\| \ : \ \mathbf{z} \in G_{\mathcal{J}}^{(m)} \right\}$.
- Sample $\mathbf{Z}$ from a mixture of $N(\mu_{\mathcal{J}}^{(m)}, \mathbf{I})$.

Mean Shifting of Common Factors

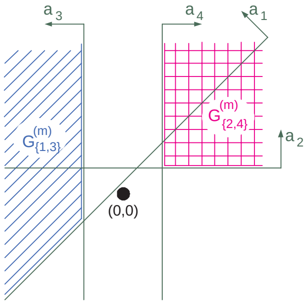
- $G_j^{(m)} \triangleq \left\{ \mathbf{z} \in \mathbb{R}^d : \mathbf{a}_j^\top \mathbf{z} \geq \alpha_1^{(m)} \Phi^{-1}(1 - \bar{p}_j) + \alpha_2^{(m)} b_j \Phi^{-1}(q) \right\}$.
- $G_{\mathcal{J}}^{(m)} \triangleq \bigcap_{j \in \mathcal{J}} G_j^{(m)}$ for $\mathcal{J} \in \mathcal{M}_q$
- $G^{(m)} \triangleq \bigcup_{\mathcal{J} \in \mathcal{M}_q} G_{\mathcal{J}}^{(m)}$.
- Sufficient subfamily, $\mathcal{S}_q$:
  - Feasibility: For each $\mathcal{J} \in \mathcal{S}_q$, $G_{\mathcal{J}}^{(m)} \neq \emptyset$ for all m;
  - Covering property: $\bigcup_{\mathcal{J} \in \mathcal{S}_q} G_{\mathcal{J}}^{(m)} = G^{(m)}$ for all m.
- $\boldsymbol{\mu}_{\mathcal{J}}^{(m)} \triangleq \operatorname{argmin} \left\{ \|\mathbf{z}\| \ : \ \mathbf{z} \in G_{\mathcal{J}}^{(m)} \right\}$.
- Sample $\mathbf{Z}$ from a mixture of $N(\boldsymbol{\mu}_{\mathcal{J}}^{(m)}, \mathbf{I})$.

Example: Two Risk Factors, 4 Obligor Types



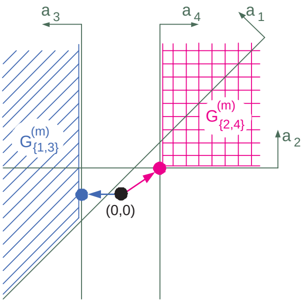
- $C_1 = 2, \ C_2 = 2, \ C_3 = 3, \ C_4 = 3.$
- $C = 10.$
- $q = 0.45, \qquad qC = 4.5.$
- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}.$
- $G^{(m)}_{\{1,3\}} = ///, \ G^{(m)}_{\{2,3\}} \subset G^{(m)}_{\{1,3\}}, \ G^{(m)}_{\{2,4\}} = \boxplus,$
  $G^{(m)}_{\{1,4\}} \subset G^{(m)}_{\{2,4\}}, \ G^{(m)}_{\{3,4\}} = \emptyset$

Example: Two Risk Factors, 4 Obligor Types



- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}$.
- Reduction of $\mathcal{M}_q$ as $\mathcal{S}_q = \{\{1,3\},\{2,4\}\}$.
  $G^{(m)} = G^{(m)}_{\{1,3\}} \cup G^{(m)}_{\{2,4\}}$
  $= G^{(m)}_{\{1,3\}} \cup G^{(m)}_{\{1,4\}} \cup G^{(m)}_{\{2,3\}} \cup G^{(m)}_{\{2,4\}} \cup G^{(m)}_{\{3,4\}}$
- $\boldsymbol{\mu}^{(m)}_{\{1,3\}} = \bullet$, $\boldsymbol{\mu}^{(m)}_{\{2,4\}} = \bullet$.
- To sample **Z**, use the mixture of two bivariate normal distributions with mean vectors • and •.

Example: Two Risk Factors, 4 Obligor Types



- $\mathcal{M}_q = \{\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}$.
- Reduction of $\mathcal{M}_q$ as $\mathcal{S}_q = \{\{1,3\},\{2,4\}\}$.
  $G^{(m)} = G^{(m)}_{\{1,3\}} \cup G^{(m)}_{\{2,4\}}$
  $= G^{(m)}_{\{1,3\}} \cup G^{(m)}_{\{1,4\}} \cup G^{(m)}_{\{2,3\}} \cup G^{(m)}_{\{2,4\}} \cup G^{(m)}_{\{3,4\}}$
- $\boldsymbol{\mu}^{(m)}_{\{1,3\}} = \bullet$, $\boldsymbol{\mu}^{(m)}_{\{2,4\}} = \bullet$.
- To sample **Z**, use the mixture of two bivariate normal distributions with mean vectors • and •.

Mixed Importance Sampling (MIS) Procedure

- Factor shifting direction: choose the mean vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ to shift the common factors $\mathbf{Z}$ and their weights $\lambda_1, \ldots, \lambda_k$ for $k \geq 1$.
- Main Loop: repeat for replications $i = 1, \ldots, \lambda_j \cdot n$, and for each type $j = 1, \ldots, t$
  1. Sample $\mathbf{Z}$ from $N(\boldsymbol{\mu}_j, \mathbf{I})$.
  2. Find $\theta_m(\mathbf{Z})$ by $\mathrm{argmin}_{\theta \geq 0} \{-\theta x + m\psi_m(\theta, \mathbf{Z})\}$.
  3. Compute the twisted conditional default probabilities $p_{k,\theta_m(\mathbf{Z})}(\mathbf{Z})$, $k = 1, \ldots, m$ and generate Y_k, $k = 1, \ldots, m$.
  4. For k with $Y_k = 1$, generate the loss ℓ_k under the twisted conditional distribution. If the loss is deterministic, set $\ell_k = c_k$.
  5. Calculate $I_i^{(j)} = 1\{L_m > x\}$ $\times e^{-\theta_m(\mathbf{Z})L_m + m\psi_m(\theta_m(\mathbf{Z}),\mathbf{Z})} \left(\sum_{i=1}^{k} \lambda_i \exp\left(\boldsymbol{\mu}_i^\top \mathbf{Z} - \frac{1}{2}\boldsymbol{\mu}_i^\top \boldsymbol{\mu}_i \right) \right)^{-1}$.
- Return the estimate $\frac{1}{n} \sum_{j=1}^{t} \sum_{i=1}^{\lambda_j \cdot n} I_i^{(j)}$

Necessity of Mixture

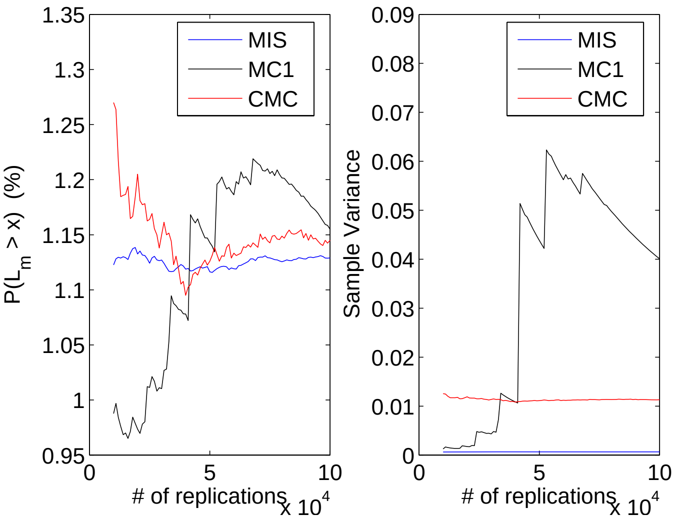
- A single shift method suggested by Glasserman & Li:
  - $\operatorname{argmax}_{\mathbf{z}} \mathbb{P}(L > x | \mathbf{Z} = \mathbf{z}) e^{-\mathbf{z}^\top \mathbf{z}/2}$.
  - $\operatorname{argmax}_{\mathbf{z}} \left\{ F_x(\mathbf{z}) - \frac{1}{2}\mathbf{z}^\top \mathbf{z} \right\}$
    where $F_x(\mathbf{z}) \triangleq -\theta_x(\mathbf{z})x + m\psi(\theta_x(\mathbf{z}), \mathbf{z})$.
- An example

$$
\begin{aligned}
X_{2k-1} &= 0.7Z_1 && +\sqrt{0.51}\ \varepsilon_{2k-1} \\
X_{2k} &= \qquad 0.65Z_2 && +\sqrt{0.5775}\ \varepsilon_{2k}
\end{aligned}
$$

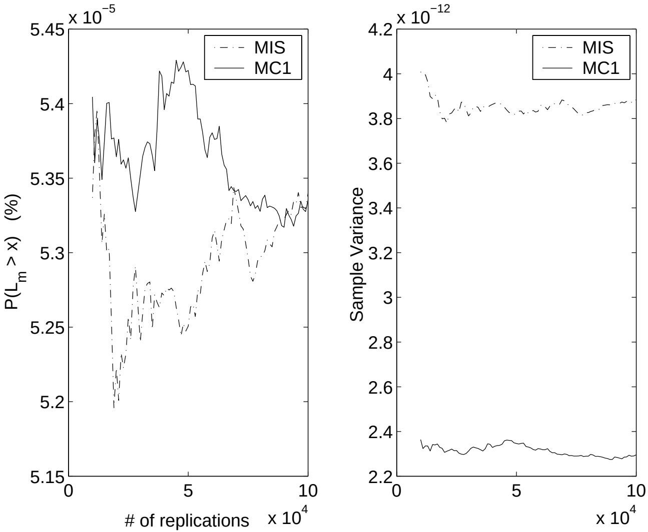
for $k = 1, 2, \ldots, 1000$.

  - $p_k \equiv 5\%$, $\ell_k \equiv 1$, and $m = 1000$.
  - $\mathbb{P}\left(\sum_{1 \le k \le 1000} \mathbf{1}\left\{ X_k > \Phi^{-1}(0.95) \right\} > 0.3 \cdot 1000 \right)$?

MIS vs. Single Shift ($q = 0.3$)

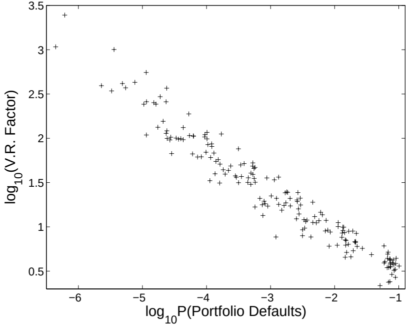


MIS vs. Single Shift ($q = 0.8$)



Small Random Examples

- 30 instances of 25 types and 5 common factors.
  - 60% of the factor-loading coefficients are non-zero. Each coefficient comes from $U(-0.2, 1)$. Then $\|\mathbf{a}_j\| \in U(0.1, 0.7)$.
  - $c_k \in U\{1, 2, \ldots, 30\}$.
  - $p_k = 0.0255 + 0.0245 \times \sin(16\pi k/m)$. $p_k \in (0.1\%, 5\%)$.
  - $q = 0.05, 0.075, \ldots, 0.175$.



Large Sparse Example

- 8 instances of 100 types and 21–22 common factors.
  - $p_k = 0.01 \cdot (1 + \sin(16\pi k/m)), \quad k = 1, \ldots, 1000.$
  - $c_k = 1 + \frac{99}{999}(k-1), \quad k = 1, \ldots, 1000.$

$$A = \left(\begin{array}{c|ccc|c} & F & & & G \\ R & & \ddots & & \vdots \\ & & & F & G \end{array} \right), \quad G = \begin{pmatrix} c_G & & \\ & \ddots & \\ & & c_G \end{pmatrix}.$$

- Approximate Importance Sampling by PCA.

# of Dominating Factors	$(\alpha_R, \alpha_F, \alpha_G)$			
	(0.8,0.4,0.4)	(0.5,0.4,0.4)	(0.2,0.4,0.4)	(0.25, 0.15,0.05)
Single Factor in $\mathbb{R}^{21}$	79%	60%	25%	74%
Two Factors in $\mathbb{R}^{22}$	80%	64%	31%	77%

Large Sparse Example

Loss $x(q)$	$\mathbb{P}(L_m > x)$	V.R. Est.
10000 (20%)	0.0114	25
15000 (30%)	0.0056	44
20000 (40%)	0.0027	75
25000 (50%)	0.0013	127
30000 (60%)	0.0007	224
35000 (70%)	0.0002	441
40000 (80%)	7.4×10^{-5}	1081

Loss $x(q)$	$\mathbb{P}(L_m > x)$	V.R. Est.
10000 (20%)	0.0077	16
15000 (30%)	0.0031	60
20000 (40%)	0.0012	120
25000 (50%)	0.0004	245
30000 (60%)	0.0001	584

Tractability of Sizes

Types	d	Bound	$\|\mathcal{V}\|$	$n_{0.1}$	$n_{0.3}$	$n_{0.5}$	$n_{0.7}$	$n_{0.9}$
20	4	6195	574.6	16.9	48.5	44.5	14.3	0.2
20	5	21699	932.2	25.0	78.8	69.0	19.5	0.4
25	4	15275	1224.9	33.5	90.5	74.6	16.0	0.2
25	5	68405	2036.5	39.7	138.4	137.7	28.2	0.0

Conclusion

- Large deviations bounds on the tail of heterogeneous credit portfolio
- Efficient simulation procedure for credit portfolio
- Sometimes, we really need to consider the multiple factors!