Importance Sampling for Portfolio Credit Risk

Glasserman et al. (2005)

Presenter: Masoud M. Supervisor: Prof. Steve Golbeck

December 22, 2015

Outline

Introduction

- A primary problem in risk management is the estimation of the profit-and-loss distribution over a specific horizon.
- Estimating probabilities exceeding a large threshold x, the size (quantile) of loss associated with a particular loss probability (VaR) and the expected shortfall are important in guiding many risk decisions

$$P(L > x) = p \qquad , \qquad P(L > x_\alpha) = \alpha$$

- Large multi-asset portfolios with complex securities (including derivatives) require huge computational budget for pricing and estimating risk measures
- Monte Carlo simulation is a favorable computational tool in risk measurement because it is very general but considerably slow (probabilistic error of $O(1/\sqrt{N})$)

- (Challenges) portfolio credit risk settings
  - Requires accurate estimation of low-probability (rare) events of large losses (defaults)
  - Capturing dependency among credit risk sources, adds complexity, particularly in the context of large losses (rare) - default events.
- Sampling procedures for rare-event simulation are introduced to tackle these challenges through:
  - Normal Copula model to capture dependency between default (rare) events
  - Two-step Importance sampling for variance reduction (tighter confidence interval - less computational budget)

Normal Copula Model

For a portfolio with m credits:

- Unconditional marginal default probability p_k is assumed known from credit ratings or market prices of corporate bonds over the specified horizon
- k^{th} obligor defaults if:

$$X_k > x_k$$

X_k : "Latent" variable of k^{th} credit, assumed to be standard normal (demeaned and standardized).

$$X_k \sim N(0, 1)$$

x_k : default threshold for credit k (equivalent to p_k)

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

setup

- Total portfolio loss from defaults:

$$L = \sum_{k=1}^{m} c_k Y_k$$

$$\begin{aligned} c_k &: \text{ Known loss from } k^{th} \text{ obligor default} \\ Y_k &: \text{ Default indicator for } k^{th} \text{ obligor} \end{aligned}$$

$$Y_k = \mathbb{1}\{X_k > x_k\}$$

- Dependence among obligors default events is presented as dependence among defaults indicators $(Y_1, ..., Y_m)$
- In the normal copula framework, this dependence is introduced through the MVN vector $(X_1, ..., X_m)$

- Correlation among X is specified through factor model

$$X_k = a_k^T Z + b_k \epsilon_k$$

Z : the (uncorrelated) systematic *common* default risk factors vector. $Z_i \sim N(0,1)$

a_k : the factors loadings vector for the k^{th} obligor

  - assumed non-negative to ensure default events have positive correlation

ϵ_k : idiosyncratic risk associated with k^{th} obligor $\epsilon_k \sim N(0,1)$

- Conditional on Z, default indicators Y become independent with conditional default probabilities:

$$\begin{aligned} p_k(Z) = P(Y_k = 1|Z) &= P(X_k > x_k|Z) \\ &= P(a_k^T Z + b_k \epsilon_k > \Phi^{-1}(1 - p_k)|Z) \\ &= \Phi(\frac{a_k^T Z + \Phi^{-1}(p_k)}{b_k}) \end{aligned}$$

Exponential Twisting

- To better estimate probabilities $P(L > x)$, we would like to increase sampling the rare-events (defaults).
- Why? because we need the default events to be more accurately represented in estimating loss distribution.
- Supposedly, we can achieve that by shifting the loss distribution closer to those events, OR, we can increase the default probabilities.
- exponential twisting helps to achieve both by robustly increasing the conditional default probabilities that results in twisting and shifting the loss distribution closer to the default events and in correspondence closer to the event $\{L > x\}$

Exponential Twisting - Setup

- Assuming independent Y, replace the default probability p_k (under $\mathbb{P}$ measure) with probability $p_{k,\theta}$ (under θ measure)

$$p_{k,\theta} = \frac{p_k e^{c_k \theta}}{(1 + p_k(e^{ck\theta} - 1))} \qquad , p_{k,\theta} \geq p_k \quad if \quad \theta \geq 0$$

  Larger exposure c_k results in a greater increase in default probability (more significance on the portfolio loss)

- Under the new probability θ measure:

$$P(L > x) = E_\theta[\mathbb{1}\{L > x\} \prod_{k=1}^{m} (\frac{p_k}{p_{k,\theta}})^{Y_k} (\frac{1 - p_k}{1 - p_{k,\theta}})^{1-Y_k}]$$
$$= E_\theta[\mathbb{1}\{L > x\} e^{-\theta L + \psi(\theta)}]$$

The Likelihood Ratio (LR) $= e^{-\theta L + \psi(\theta)}$
The Cumulant Generating Function (CGF)

$$\psi(\theta) = log(E[e^{\theta L}]) = \sum_{k=1}^{m} log(1 + p_k(e^{\theta c_k} - 1))$$

- (Which value of θ?) For more confidence in estimation (tighter C.I), choose θ that minimizes the variance of the estimator, or the second moment

$$\begin{aligned} M_2(x,\theta) &= E_\theta[\mathbb{1}\{L > x\} e^{-2\theta L + 2\psi(\theta)}] \\ &\le e^{-2\theta x + 2\psi(\theta)} \end{aligned}$$

- Minimizing the second moment is difficult; minimize the upper bound instead over the interval $\theta \geq 0$

$$min_\theta\{e^{-2\theta x+2\psi(\theta)}\} \rightarrow (\psi'(\theta) = x)$$

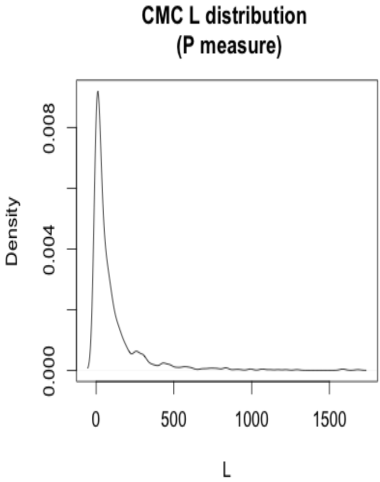
- $\psi(\theta)$ is strictly convex and passes through the origin

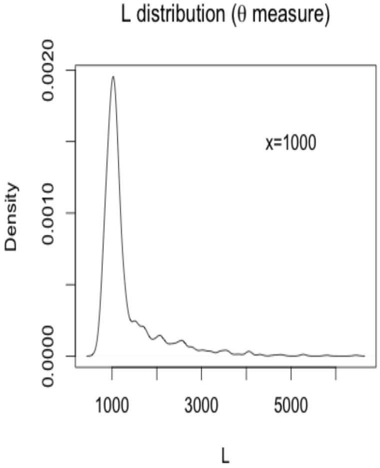
$$\begin{aligned} \theta_x &= \textit{unique solution}(\psi'(\theta) = x) \quad & \textit{if} \quad x > \psi'(0) \\ &= 0 & \textit{if} \quad x \leq \psi'(0) \end{aligned}$$

- It can be shown by direct differentiation of $\psi(\theta)$, that (under θ measure):

$$E_\theta[L] = \psi'(\theta) \rightarrow (E_\theta[L] = x)$$

- The exponential twisting shifted the distribution of portfolio loss to x, such that the event $\{L > x\}$ is no longer a rare-event.


CMC L distribution
(P measure)
Density
0.000
0.004
0.008
0
500
1000
1500
L


L distribution (θ measure)
x=1000
Density
0.0000
0.0010
0.0020
1000
3000
5000
L

One-Step Algorithm

- Generate the d-dimensional common risk factors (marginal $Z_i \sim N(0,1)$)
- Given Z vector, calculate the conditional default probabilities

$$p_k(Z) = P(Y_k = 1|Z) = \frac{\Phi(a_k^T Z + \Phi^{-1}(p_k))}{b_k}$$

- Calculate the conditional portfolio Loss expectation

$$E[L|Z] = \psi'(0,Z) = \sum_{k=1}^{m} p_k(Z) c_k$$

- Solve for θ_x :

$$\begin{aligned}\theta_x(Z) &= unique\ solution(\psi'(\theta,Z)=x) \quad if \ \ x > \psi'(0,Z)\\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad if \ \ x \le \psi'(0,Z)\end{aligned}$$

$$\psi'(\theta, Z) = \frac{\partial}{\partial\theta} \sum_{k=1}^{m} log(1 + p_k(Z)(e^{\theta_x(Z)c_k} + 1))$$

- Given $\theta_x(Z)$, calculate the θ measure conditional default probabilities

$$p_{k,\theta}(Z) = \frac{p_k e^{c_k \theta_x(Z)}}{(1 + p_k(Z)(e^{ck\theta_x(Z)} + 1))}$$

- Generate independent $Y_k \sim Bernouli(p_{k,\theta}(Z))$
- Calculate portfolio Loss $L = \sum_{k=1}^{m} c_k Y_k$
- Calculate Likelihood Ratio (LR) $= e^{-\theta_x(Z)L + \psi(\theta_x(Z), Z)}$
- Repeat n.sim times, $P(L > y) = E_\theta[\mathbb{1}\{L > y\} LR]$
  (y is any value, even much larger than x)

Factors Mean Shifting

- The estimator $\hat{p}_x$ of $P(L > x)$ has a variance:

$$Var[\hat{p}_x] = E[Var[\hat{p}_x|Z]] + Var[E[\hat{p}_x|Z]]$$

- The $Var[\hat{p}_x|Z]$ is minimized through the One-Step Importance Sampling.
- Since $E[\hat{p}_x|Z] = P(L > x|Z)$, The second term can be minimized by choosing a distribution of Z that would reduce the variance in estimating the integral of $P(L > x|Z = z)$ against the density of Z

$$\text{Choose } Z \sim N(\mu, I) \text{ that:}$$
$$max_z\{P(L > x|Z = z)e^{-z^Tz/2}\}$$

Finding μ - Tail Bound Approximation

- Maximize

$$P(L > x|Z = z)e^{-z^T z/2} = E_\theta[\mathbb{1}\{L > x\}e^{F_L}]e^{-z^T z/2}$$

$$F_L = -\theta_x(z)L + \psi(\theta_x(z), z)$$

- The upper bound of the estimator

$$P(L > x|Z = z)e^{-z^T z/2} \leq e^{F_x - z^T z/2}$$

$$F_x = -\theta_x(z)x + \psi(\theta_x(z), z)$$

- Maximize the upperbound

$$max_z\{J(x, \mu) = F_x - z^T z/2\}$$

The Two-Step Algorithm

Two-Step Algorithm

1. Step 1: Find Risk Factors μ vector:
   - Numerically maximize $J(x, \mu)$ (note that the objective function has stochastic component)
2. Step 2: Same as the importance sampling One-Step algorithm (above) except
   - Generate the d-dimensional common risk factors (marginal $Z_i \sim N(\mu_i, 1)$)
   - Calculate Likelihood Ratio $LR = e^{-\theta_x(Z)L+\psi(\theta_x(Z),Z)} e^{-\mu^T z+\mu^T \mu}$

Results

Example 1

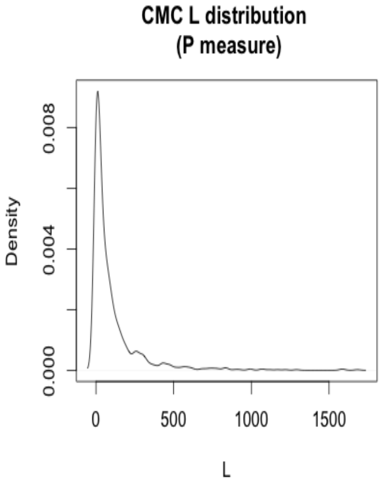
- m=1000 obligors , d=10 factors
- $p_k = 0.01(1 + sin(16\pi k/m))$
- $c_k = \lceil 5k/m \rceil^2$
- $a_{kj} \sim U(0, 1/\sqrt{d})$

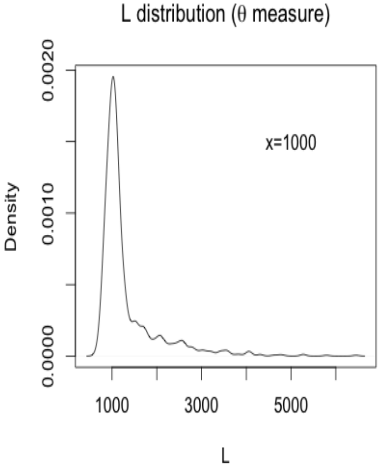
1. Step 1: μ vector:

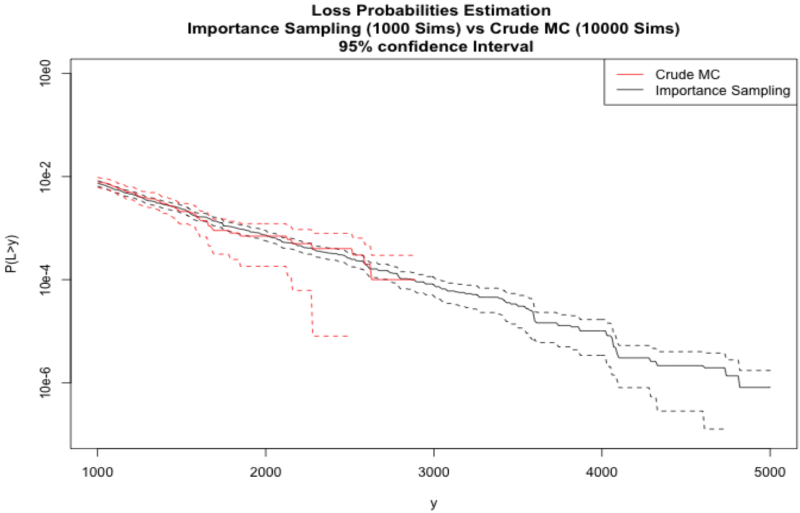
0.53	0.64	0.68	0.35	0.8
0.59	0.87	0.58	0.8	0.78

Glasserman reported a mean vector of 0.8

Note: the optimization objective function is not convex and it has stochastic components. hard to arrive to same local optima


CMC L distribution
(P measure)
Density
0.000
0.004
0.008
0
500
1000
1500
L


L distribution (θ measure)
x=1000
Density
0.0000
0.0010
0.0020
1000
3000
5000
L


Loss Probabilities Estimation
Importance Sampling (1000 Sims) vs Crude MC (10000 Sims)
95% confidence Interval
Crude MC
Importance Sampling
P(L>y)
10e0
10e-2
10e-4
10e-6
1000
2000
3000
4000
5000
y

Example 2

- m=1000 obligors , d=21 factors, p_k (same as Ex1)
- c_k increases linearly from 1 to 100
- a_{kj}:

  First market-wide factor affecting all obligors $a_{k1} = 0.8$
  20 factors are 10 industry and 10 geographical factors loading of 0.4 (see structure in Glasserman et al (2005))

1. Step 1: μ vector:

2.169	0.312	0.316	0.241	0.079	0.466	0.255
0.004	0.437	0.137	0.413	0.372	0.231	0.189
0.471	0.466	0.236	0.031	0.274	0.377	0.305

Glasserman reported a mean vector of **2.46** for first factor and around 0.2 for the other 20 factors

- Using the Step 1 μ as an input to step 2

y	P($L > y$) Glasserman	VRR Glasserman	P($L > y$)	VRR
10000	0.0114	33	0.011	12
14000	0.0065	53	0.0064	14
18000	0.0037	83	0.0036	26
22000	0.0021	125	0.002	42
30000	0.0006	278	0.0007	72
40000	0.0001	977	0.0001	1391

$$VRR = \frac{Variance\ CMC}{Variance\ Imp.\ Sampling}$$

- Using Glasserman reported vector c(2.46 , rep (0.2 , 20) μ as an input to step 2

y	P($L > y$) Glasserman	VRR Glasserman	P($L > y$)	VRR
10000	0.0114	33	0.011	28
14000	0.0065	53	0.006	49
18000	0.0037	83	0.0034	84
22000	0.0021	125	0.002	126
30000	0.0006	278	0.0006	236
40000	0.0001	977	0.0001	3779

Summary

- Dependency between obligors adds complexity to estimating portfolio credit risk measurements (rare-events dependency)
- Normal copula model provides good framework to capture this dependency
- The two-step Importance Sampling algorithm reduces the simulation computational budget through variance reduction:
  - Exponential twisting results in mean shifting that increases the probabilities of the credit default rare-events (shifted Loss distribution)
  - Shifting the underlying risk factors mean vector further reduce the variance
- The estimator variance (and VRR) is considerably more sensitive to risk factors mean vector compared to the estimator accuracy.

References

- Glasserman P., Li J. (2005), Importance Sampling for Portfolio Credit Risk Mgmt. Sci 46, 1643-1656
- Glasserman P. (2003), Monte Carlo Methods in Financial Engineering, Springer Publisher, New York.