Estimating Joint Default Probability by Efficient Importance Sampling with Applications from Bottom Up

Chuan-Hsiang Han

Dept. of Quantitative Finance

Natl. Tsing-Hua University, Taiwan

MCQMC, Warsaw

August, 17, 2010

Outline

- Credit Risk Modeling: Classical Models

- Joint Default Probability, Importance Sampling, and Large Deviation

- Homogenization by Singular Perturbation and Effect of (Stochastic) Correlation

Modeling Default Times: Bottom Up Approach

Notation: τ_i: default time of firm i.

- Intensity-Based (Reduced Form)

View firm's default as exogenous.

$$\mathbb{P}(\tau_i \leq t) = F_i(t) := 1 - \exp\left\{-\int_0^t h_i(s)ds\right\}.$$

- Asset Value-Based (Structural Form)

Firm asset values follow correlated processes, say geometric Brownian motions:

$$dS_{i\,t} = \mu_i\, S_{i\,t}\, dt + \sigma_i\, S_{i\,t} dW_{i\,t}, d < W_i, W_j >_t = \rho_{ij} dt.$$

A default event $\{\tau_i \leq T\} := \mathbf{I}(S_{i\,T} \leq B_i)$.

Application I: Loss Density Function

The loss random variable $L(T)$ is defined by

$$L(T) = \sum_{i=1}^{N} c_i \, \mathbf{I}(\tau_i \leq T).$$

If the density of $L(T)$ is known, one can investigate credit portfolio risk management, pricing credit derivatives, etc.

Application II: Evaluation of Credit Swaps

$$premium = \frac{\mathbb{E}\left\{(1-R) \times B(0,\tau) \times \mathbf{I}(\tau < T)\right\}}{\mathbb{E}\left\{\sum_{j=1}^{N} \triangle_{j-1,\,j} \times B(0,t_j) \times \mathbf{I}(\tau > t_j)\right\}}$$

Notations: τ: default time, R: recovery rate, $B(0,t)$: discount factor, $\triangle_{j-1,\,j}$: time increment.

CDS: τ is the time to default of an asset.

BDS: τ is an order statistics of $\tau_1, \tau_2, \cdots, \tau_n$.

Then we ask a question

$$\text{JDP} = E\left\{\Pi_{i=1}^{n}\mathbf{I}(\tau_i \leq T)\right\}?$$

And hope this leads to the estimation of
(1) $P(L(T) = i) = p_i$,
(2) $P(\tau \leq T)$, where τ is the k-th order statistics of $\{\tau_i\}_{i=1}^{n}$.

JDP = joint default probability

Correlation under Reduced Form Model: Copula Method*

1. Default Time: $\left\{\tau_i = F_i^{-1}(U_i)\right\}_{i=1}^{n}$, U's are [0,1]-uniform random variables.
2. Copula is a distribution function on $[0,1]^n$ with uniform marginal distributions.
3. Through a **copula function**, one can build up correlations between default times.

*Cherubini, Luciano, Vecchiato (2004), Nelson(2006).

Characterization of Default Events

Gaussian Copula Factor Model (Laurent and Gregory (2003))

$$\{\tau_i = F_i^{-1}(\Phi(W_i)) \le T\}$$

$$= \left\{ W_i := \rho_i Z_0 + \sqrt{1-\rho_i^2} Z_i \le \Phi^{-1}(F_i(T)) \right\}$$

$$= \left\{ Z_0 \le \frac{\Phi^{-1}(F_i(T)) - \sqrt{1-\rho_i^2} z_i}{\rho_i} \right\} \text{when} Z_i = z_i$$

$$= \left\{ Z_i \le \frac{\Phi^{-1}(F_i(T)) - \rho_i z_0}{\sqrt{1-\rho_i^2}} \right\} \text{when} Z_0 = z_0.$$

(Conditional) Importance Sampling

Estimate the JDP $\mathbb{E}\left\{\prod_{i=1}^{n}\mathbf{I}(\tau_i \le T)\right\}$ by

(1) Condition on marginal factors (Chiang, Yuah, Hsieh (2007))

$$\mathbb{E}\left\{\tilde{\mathbb{E}}\left\{\Pi_{i=1}^{n}\mathbf{I}\left(Z_0 \le \frac{c_i - \sqrt{1-\rho_i^2}Z_i}{\rho_i}\right)L(Z_0;u)|Z_1,\cdots,Z_n\right\}\right\}$$

(2) Condition on common factor

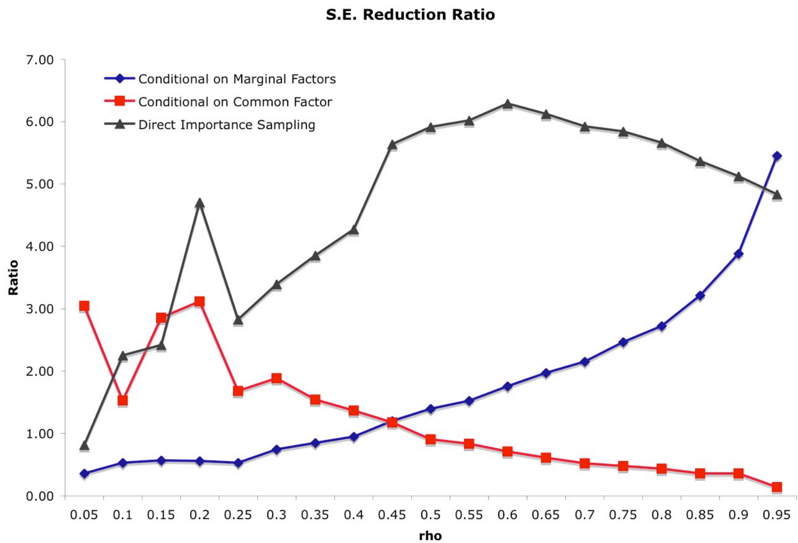
$$\mathbb{E}\left\{\tilde{\mathbb{E}}^{(u_1,\cdots,u_n)}\left\{\Pi_{i=1}^{n}\mathbf{I}(Z_i \le \frac{c_i-\rho_i Z_0}{\sqrt{1-\rho_i^2}})\Pi_{i=1}^{n}L(Z_i;u_i) \mid Z_0\right\}\right\}$$

(3) Direct Change of Measure

$$\tilde{\mathbb{E}}\left\{\prod_{i=1}^{n}\mathbf{I}(W_i \le c_i)\Pi_{i=1}^{n}L(W_i;w_i)\right\}$$

Notations: $c_i = \Phi^{-1}(F_i(T))$ and $L(\cdot,\cdot)$ the Likelihood ratio.

Variance Reduction Comparison of IS Estimators: Guassian Distribution



Asymptotic Optimality of Direct Change of Measure

Let W be a centered multivariate normal

$JDP = E\{\mathbf{I}(W < c)\} = E_\mu \{\mathbf{I}(W < c)\Pi_{i=1}^n L(W_i; \mu_i)\}$,
where $L(w; \mu) = \exp(-\mu w + \mu^2/2)$ is the likelihood function.

Theorem The variance of $\mathbf{I}_{\{W<c\}}\Pi_{i=1}^n L(W_i; \mu_i)$ is optimally minimized at $\mu = c$ when each component in the vector $-c$ is sufficiently large.
Proof: By Cramer's Theorem in large deviation theory.

Generalizations

- Computation of tail probability for multivariate normals. (versus a Matlab program **mvncdf.m**, based on Genz and Bretz ('99))

- Equivalent to Black-Scholes's structural-form model model in high dimension to estimate $\mathbb{E}\left\{\Pi_{i=1}^{n}\mathbf{I}(S_{iT} \leq B_i)\right\}$.

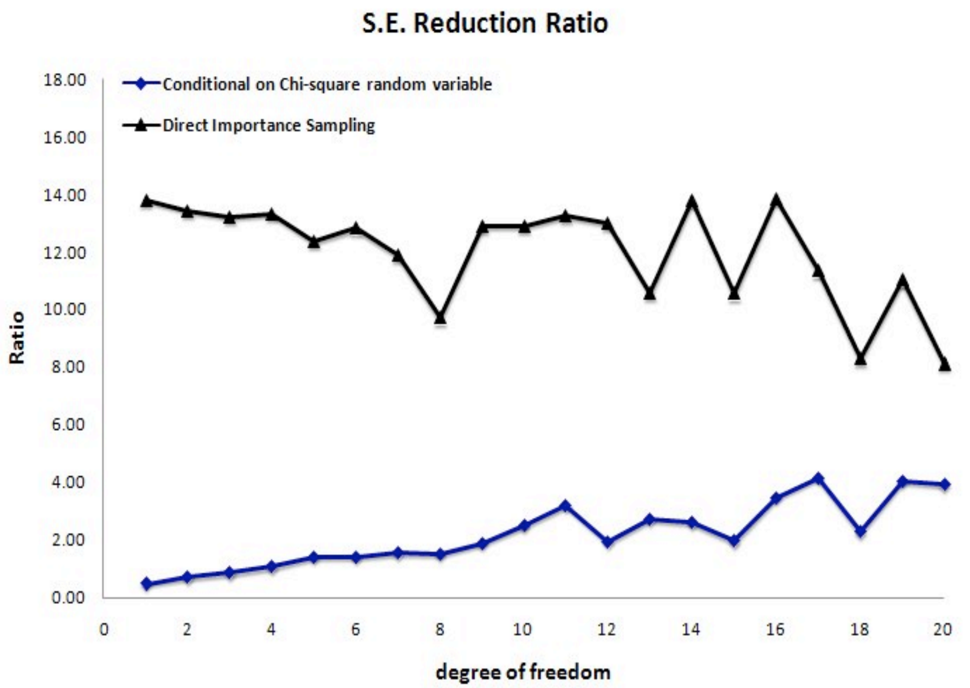
- Based on Glasserman et al. (2002), compute tail probability of multivariate Student T dist. (versus **mvtcdf.m**).

Gaussian Tail Probability Estimation: IS vs. mvncdf.m

	Basic MC		Importance Sampling		Quasi MC	
n	$Mean$	SE	$Mean$	SE	$Value$	Error
5	4E-05	4E-05	1.41E-05	5.62E-07	1.40E-05	1.59E-07
10	-	-	2.10E-07	1.96E-08	1.99E-07	1.33E-08
15	-	-	1.42E-08	2.20E-09	1.58E-08	2.92E-09
20	-	-	1.99E-09	5.36E-10	2.48E-09	5.13E-10
25	-	-	5.48E-10	1.28E-10	6.98E-10	5.20E-10
30	-	-	1.71E-10	6.81E-11	-	-
50	-	-	4.06E-12	2.17E-12	-	-

$c = -2, \rho = 0.5$, and the total number of simulations=25000. Averaged CPU time: 4.29E-02, 9.64E-02, 2.16E-01, respectively, without dimensions of 30 and 50.

Variance Reduction Comparison of IS Estimators: Student T Distribution



Student T Tail Probability Estimation: IS vs. mvtcdf.m

	Basic MC		IS		Quasi MC	
n	$Mean$	SE	$Mean$	SE	$Value$	Error
5	2.4E-04	9.8E-05	2.00E-04	6.41E-06	1.94E-04	1.19E-05
10	-	-	1.30E-05	9.18E-07	1.31E-05	4.17E-06
15	-	-	3.25E-06	3.42E-07	2.40E-06	9.26E-07
20	-	-	8.89E-07	1.85E-07	1.03E-06	8.93E-07
25	-	-	2.51E-07	5.23E-08	1.70E-07	1.25E-07
30	-	-	1.21E-07	2.54E-08	-	-
50	-	-	7.53E-09	4.53E-09	-	-

$c = -2$, $\rho = 0.5$, degree of freedom is 10, and the total number of simulation is 25000.
Averaged CPU time: 4.39E-02, 1.27E-01, 2.39E-01, respectively.

Credit Risk Modeling:
Structural Form Approach

Multi-Names Dynamics: for $1 \leq i \leq n$

$$
\begin{aligned}
dS_{it} &= \mu_i S_{it} dt + \sigma_i S_{it} dW_{it}, \\
d\left\langle W_{it}, W_{jt} \right\rangle &= \rho_{ij} dt.
\end{aligned}
$$

Each default time τ_i for the i^{th} name is defined as $\tau_i = \text{inf}\{t \geq 0 : S_{it} \leq B_i\}$, where B_i denotes the i^{th} debt level.
The i^{th} default event is defined as $\{\tau_i \leq T\}$.

Joint Default Probability: First Passage Time Problem in High Dim.

Q: How to compute $JDP = \mathbb{E}\left\{\Pi_{i=1}^{n}\mathbf{I}(\tau_i \leq T)\right\}$ under structural-form models?
Explicit Formulas exist for 1-name case (Black and Cox '76) and 2-name case (Zhou '01).

Note that Carmona, Fouque, and Vestal ('09) dealed with a similar problem by means of Interacting Particle Systems.

Multi-Dimensional Girsanov Theorem

Given the Radon-Nikodym derivative

$$\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} = Q_T^h = e^{\left(\int_0^T h(s,S_s)\cdot d\tilde{W}_s - \frac{1}{2}\int_0^T ||h(s,S_s)||^2 ds\right)},$$

$\tilde{W}_t = W_t + \int_0^t h(s, S_s)ds$ is a vector of Brownian motions under $\tilde{\mathbb{P}}$. Thus

$$DP = \tilde{\mathbb{E}}\left\{\prod_{i=1}^n \mathbf{I}(\tau_i \leq T) Q_T^h\right\}.$$

Monte Carlo Simulations: Importance Sampling

An importance sampling method is developed to satisfy

$$\tilde{E}\left\{S_{iT}|\mathcal{F}_0\right\} = B_i, i = 1, \cdots, n.$$

The new measure is characterized by solving **the linear system** $\Sigma_{j=1}^{i}\rho_{ij}h_j = \frac{\mu_i}{\sigma_i} - \frac{\ln B_i/S_{i0}}{\sigma_i T}$ so that by Girsanov Theorem

$$JDP = \tilde{E}\left\{\Pi_{i=1}^{n}\mathbf{I}(\tau_i \leq T)\, Q_T\right\}.$$

Single Name Default Probability

B	BMC	Exact Sol	Importance Sampling
50	0.0886 (0.0028)	0.0945	0.0890 (0.0016)
20	0 (0)	$7.7 * 10^{-5}$	$7.2 * 10^{-5}(2.3 * 10^{-6})$
1	0 (0)	$1.3 * 10^{-30}$	$1.8 * 10^{-30}(3.4 * 10^{-31})$

The number of simulations is 10^4 and the Euler discretization takes time step size $T/400$, where T is one year.

Other parameters are $S_0 = 100, \mu = 0.05$ and $\sigma = 0.4$. Standard errors are shown in parenthesis.

Asymptotic Optimality
Efficient Importance Sampling

Assume $\ln(B_i/S_{i0}) = \frac{-1}{\varepsilon}$ for each $1 \leq i \leq N$. Denote by P_ε JDP and $M_{2\varepsilon}$ the second moment under a new measure:

$$\begin{aligned} P_\varepsilon &= I\!\!E\left[\Pi_{i=1}^N \mathbf{I}\left(\inf_{0\leq t\leq T} S_{it} \leq B_i\right)\right] \\ M_{2\varepsilon} &= \tilde{I\!\!E}\left[\Pi_{i=1}^N \mathbf{I}\left(\inf_{0\leq t\leq T} S_{it} \leq B_i\right) Q_T\right] \end{aligned}$$

Theorem: By $M_{2\varepsilon} \approx (P_\varepsilon)^2$ for **small** ε (spatial scale) we observe the optimality of chosen measure.

Proof: by Freidlin-Wentzellthm or a PDE argument.

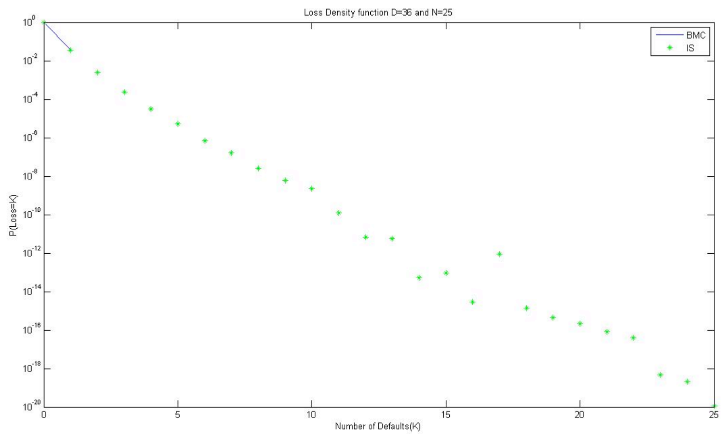
Tail Probability Estimation : the first passage time problem *

	Basic MC		Importance Sampling	
$Names$	$Mean$	SE	$Mean$	SE
2	1.1E-03	3.31E-04	1.04E-03	2.83E-05
5	-	-	6.36E-06	3.72E-07
10	-	-	2.90E-07	2.66E-08
15	-	-	9.45E-09	1.16E-09
20	-	-	1.15E-09	1.98E-10
25	-	-	2.06E-10	3.84E-11
30	-	-	6.76E-11	2.36E-11
35	-	-	1.35E-11	2.89E-12
40	-	-	6.59E-12	1.58E-12
45	-	-	3.25E-12	1.08E-12
50	-	-	6.76E-13	2.26E-13

Parameters are $S_0 = 100, \mu = 0.05$, $\sigma = 0.3$, $\rho = 0.3$, and $B = 50$.

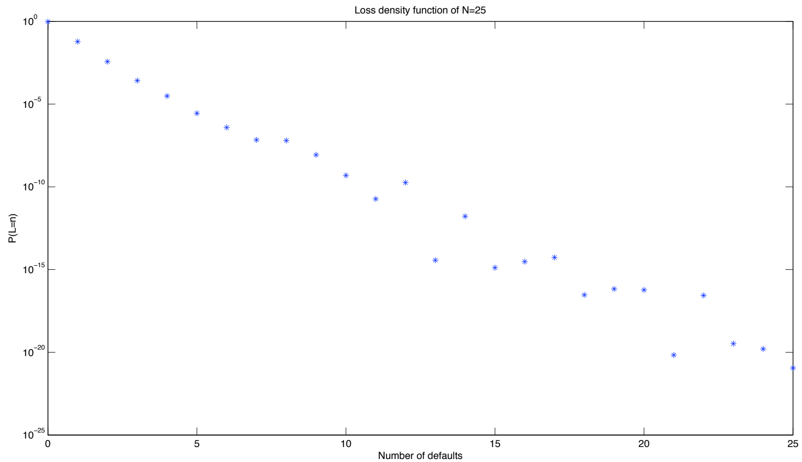
*H. (2010)

Loss Density Function: 25 Names
Diffusion Model



Note: Consider both survival and default probabilities.
Applications: Pricing CDOs, Risk Management of credit portfolios, etc.

Loss Density Function: 25 Names
Jump Diffusion Model



Use the compound poisson jump as a common factor.
But optimal efficiency can not be obtained.

A Modification: Stochastic Correlation*

$$\begin{cases} dS_t^1 = rS_t^1 dt + \sigma_1 S_t^1 dW_t^1 \\ dS_t^2 = rS_t^2 dt + \sigma_2 S_t^2 (\rho(Y_t) dW_t^1 + \sqrt{1-\rho^2(Y_t)} dW_t^2) \\ dY_t = \frac{1}{\varepsilon}(m - Y_t)dt + \frac{\sqrt{2}\beta}{\sqrt{\varepsilon}} dZ_t \ \textbf{(Scaling in Time)} \end{cases}$$

Joint default probability

$$P^\varepsilon(t, x_1, x_2, y) := \mathbb{E}_{x_1,x_2,y}\left\{ \Pi_{i=1}^2 \mathbf{I}(\min_{t \le u \le T} S_u^i \le B_i) \right\}$$

*Hull, Presescu, White (2005)

Full Expansion of P^ε

Theorem

$$P^\varepsilon(t, x_1, x_2, y) = \sum_{i=0}^{\infty} \varepsilon^i P_i(t, x_1, x_2, y),$$

where P_i's can be obtained recursively by solving a seq. of Poisson eqns.

Proof: by means of Singular Perturbation Techniques.
Accuracy results are ensured given smoothness of terminal condition.

Leading Order Term

$P_0(t, x_1, x_2)$ solves the **homogenized** PDE (y-independent).

$$\left(\mathcal{L}_{1,0} + \bar{\rho}\,\mathcal{L}_{1,1}\right) P_0(t, x_1, x_2) = 0$$

$\bar{\rho} =< \rho(y) >$, average taken wrt the **invartiant measure** of Y.

Differential operators are

$$\mathcal{L}_{1,0} = \frac{\partial}{\partial t} + \sum_{i=1}^{2} \frac{\sigma_i^2 x_i^2}{2} \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^{2} \mu_i x_i \frac{\partial}{\partial x_i}$$

$$\mathcal{L}_{1,1} = \sigma_1 \sigma_2 x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2}.$$

Other Terms

$$P_{n+1}(t, x_1, x_2, y) = \sum_{i \geq 0, j \geq 1}^{i+j=n+1} \varphi_{i,j}^{(n+1)}(y) \, \mathcal{L}_{1,0}^{i} \, \mathcal{L}_{1,1}^{j} \, P_n$$

where a seq. of Poisson eqns to be solved:

$$\mathcal{L}_0 \, \varphi_{i+1,j}^{(n+1)}(y) = \left(\varphi_{i,j}^{(n)}(y) - < \varphi_{i,j}^{(n)}(y) > \right)$$

$$\mathcal{L}_0 \, \varphi_{i,j+1}^{(n+1)}(y) = \left(\rho(y) \, \varphi_{i,j}^{(n)}(y) - < \rho \, \varphi_{i,j}^{(n)} > \right),$$

where $\mathcal{L}_0 = \beta^2 \frac{\partial^2}{\partial y^2} + (m - y) \frac{\partial}{\partial y}$.

Stochastic Correlation I

$\alpha = \frac{1}{\varepsilon}$	BMC	Importance Sampling
0.1	$0.0037(6 * 10^{-4})$	$0.0032(1 * 10^{-4})$
1	$0.0074(9 * 10^{-4})$	$0.0065(2 * 10^{-4})$
10	$0.0112(1 * 10^{-3})$	$0.0116(4 * 10^{-4})$
50	$0.0163(1 * 10^{-3})$	$0.0137(5 * 10^{-4})$
100	$0.016(1 * 10^{-3})$	$0.0132(4 * 10^{-4})$

Parameters are $S_{10} = S_{20} = 100, B_1 = 50, B_2 = 40, m = \pi/4, \nu = 0.5, \rho(y) = |sin(y)|$.

Using the homogenized term in IS, note the effect of correlation.

Stochastic Correlation II

$\alpha = \frac{1}{\varepsilon}$	BMC	Importance Sampling
0.1	0(0)	$9.1 * 10^{-7}(7 * 10^{-8})$
1	0(0)	$7.5 * 10^{-6}(6 * 10^{-7})$
10	0(0)	$2.4 * 10^{-5}(2 * 10^{-6})$
50	$1 * 10^{-4}(1 * 10^{-4})$	$2.9 * 10^{-5}(3 * 10^{-6})$
100	$1 * 10^{-4}(1 * 10^{-4})$	$2.7 * 10^{-5}(2 * 10^{-6})$

Parameters are $S_{10} = S_{20} = 100, B_1 = 30, B_2 = 20, m = \pi/4, \nu = 0.5$.

Note the effect of correlation.

Conclusions

- Simple yet efficient importance sampling methods are proposed, justified by large deviation theory.
- Full expansion of joint default probability under stochastic correlation by singular perturbation.
- Of course all these ideas can be applied to option pricing, complex model calibration, etc.

Acknowledgments

- S.-J. Sheu(Academia Sinica), N.-R. Shieh(NTU), Y.L. Chen (NTHU), Doug Vestal(UBS, US).

Thank You