Stein’s method and zero bias transformation: Application to CDO pricing

Ying Jiao

ESILV and Ecole Polytechnique

Joint work with N. El Karoui
CDO — a portfolio credit derivative containing \( \approx 100 \) underlying names susceptible to default risk.

Key term to calculate: the cumulative loss
\[
L_T = \sum_{i=1}^{n} N_i (1 - R_i) \mathbb{1}_{\{\tau_i \leq T\}}
\]
where \( N_i \) is the nominal of each name and \( R_i \) is the recovery rate.

Pricing of one CDO tranche: call function \( \mathbb{E}[(L_T - K)^+] \) with attachment or detachment point \( K \).

Motivation: fast and precise numerical method for CDOs in the market-adopted framework.
Factor models

- Standard model in the market: factor model with conditionally independent defaults.

- Normal factor model: Gaussian copula approach

\[ \{ \tau_i \leq t \} = \{ \rho_i Y + \sqrt{1 - \rho^2_i} Y_i \leq \mathcal{N}^{-1}(\alpha_i(t)) \} \]

where \( Y_i, Y \sim N(0, 1) \) and independent, and \( \alpha_i(t) \) is the average probability of default before \( t \).

- Given \( Y \) (not necessarily normal), \( L_T \) is written as sum of independent random variables

- \( L_T \) follows binomial law for homogeneous portfolios

- Central limit theorems: Gaussian and Poisson approximations

Ying Jiao

Stein’s method and zero bias transformation for CDOs
Some approximation methods

- Gaussian approximation in finance:
  - Total loss of a homogeneous portfolio (Vasicek (1991)), normal approximation of binomial distribution
  - Option pricing: symmetric case (Diener-Diener (2004))
  - Difficulty: \( n \approx 100 \) and small default probability.

- Saddle point method applied to CDOs (Martin-Thompson-Browne, Antonov-Mechkov-Misirpashaev): calculation time for non-homogeneous portfolio

- Our solution: combine Gaussian and Poisson approximations and propose a corrector term in each case.

- Mathematical tools: Stein’s method and zero bias transformation
Stein method and Zero bias transformation
Goldstein-Reinert, 1997

- $X$ a r.v. of mean zero and of variance $\sigma^2$
- $X^*$ has the zero-bias distribution of $X$ if for all regular enough $f$,

$$\mathbb{E}[Xf(X)] = \sigma^2 \mathbb{E}[f'(X^*)] \quad (1)$$

-distribution of $X^*$ unique with density $p_{X^*}(x) = \frac{\mathbb{E}[X1_{\{X>x\}}]}{\sigma^2}$.

- Observation of Stein (1972): if $Z \sim N(0, \sigma^2)$, then

$$\mathbb{E}[Zf(Z)] = \sigma^2 \mathbb{E}[f'(Z)].$$
For the normal approximation of $\mathbb{E}[h(X)]$, Stein’s idea is to use $f_h$ defined as the solution of

$$h(x) - \Phi_\sigma(h) = xf(x) - \sigma^2 f'(x) \quad (2)$$

where $\Phi_\sigma(h) = \mathbb{E}[h(Z)]$ with $Z \sim N(0, \sigma^2)$.

**Proposition (Stein):** If $h$ is absolutely continuous, then

$$|\mathbb{E}[h(X)] - \Phi_\sigma(h)| = |\sigma^2 \mathbb{E}[f_h'(X^*) - f_h'(X)]|$$

$$\leq \sigma^2 \|f_h''\| \mathbb{E}[|X^* - X|].$$

To estimate the approximation error, it’s important to measure the “distance” between $X$ and $X^*$ and the sup-norm of $f_h$ and its derivatives.
Example (Asymmetric Bernoulli distribution):

\[ P(X = q) = p \] and \[ P(X = -p) = q \] with \[ q = 1 - p \]. So \[ E(X) = 0 \] and \[ Var(X) = pq \]. Furthermore, \( X^* \sim U[-p, q] \).

If \( X \) and \( X^* \) are independent, then, for any even function \( g \),

\[
E\left[g(X^* - X)\right] = \frac{1}{2\sigma^2} E\left[X^s G(X^s)\right]
\]

where \( G(x) = \int_0^x g(t) dt \), \( X^s = X - \tilde{X} \) and \( \tilde{X} \) is an independent copy of \( X \).

In particular, \( E[|X^* - X|] = \frac{1}{4\sigma^2} E[|X^s|^3] \) \( \sim O\left(\frac{1}{\sqrt{n}}\right) \).

\[
E[|X^* - X|^k] = \frac{1}{2(k+1)\sigma^2} E[|X^s|^{k+2}] \sim O\left(\frac{1}{\sqrt{n^k}}\right).
\]
Let \( W = X_1 + \cdots + X_n \) where \( X_i \) are independent r.v. of mean zero and \( \text{Var}(W) = \sigma_W^2 < \infty \). Then

\[
W^* = W^{(I)} + X_i^*,
\]

where \( \mathbb{P}(I = i) = \sigma_i^2 / \sigma_W^2 \). \( W^{(i)} = W - X_i \). Furthermore, \( W^{(i)} \), \( X_i \), \( X_i^* \) and \( I \) are mutually independent.

Here \( W \) and \( W^* \) are not independent.

\[
\mathbb{E}[|W^* - W|^k] = \frac{1}{2(k+1)\sigma_W^2} \sum_{i=1}^{n} \mathbb{E}[|X_i^s|^k] \sim O\left(\frac{1}{\sqrt{n^k}}\right).
\]
To obtain an addition order in the estimations, for example,
\[ \mathbb{E}[X_i|X_1, \ldots, X_n] = \sum_{i=1}^{n} \frac{\sigma_i^2}{\sigma_w^2} X_i, \text{ then,} \]
\[ \mathbb{E}[^{\mathbb{E}}[X_i|X_1, \ldots, X_n]^2] = \sum_{i=1}^{n} \frac{\sigma_i^6}{\sigma_w^4} \text{ is of order } O\left(\frac{1}{n^2}\right). \]
However, \( \mathbb{E}[X_i^2] \) is of order \( O\left(\frac{1}{n}\right) \).

This technic is crucial for estimating \( \mathbb{E}[g(X_i, X^*_i)] \) and \( \mathbb{E}[f(W)g(X_i, X^*_i)] \).
First order Gaussian correction

Theorem

Normal approximation $\Phi_{\sigma_W}(h)$ of $\mathbb{E}[h(W)]$ has corrector:

$$C_h = \frac{1}{2\sigma_W^4} \sum_{i=1}^{n} \mathbb{E}[X_i^3] \Phi_{\sigma_W} \left( \left( \frac{x^2}{3\sigma_W^2} - 1 \right) xh(x) \right)$$

(3)

where $h$ has bounded second order derivative. The corrected approximation error is bounded by

$$\left| \mathbb{E}[h(W)] - \Phi_{\sigma_W}(h) - C_h \right| \leq \alpha(h, X_1, \cdots, X_n)$$

where $\alpha(h, X_1, \cdots, X_n)$ depends on $\|h''\|$ and moments of $X_i$ up to fourth order.
Remarks

- For i.i.d. asymmetric Bernoulli: corrector of order $O\left(\frac{1}{\sqrt{n}}\right)$; corrected error of order $O\left(\frac{1}{n}\right)$.
- In the symmetric case, $\mathbb{E}[X^*_j] = 0$, so $C_h = 0$ for any $h$. $C_h$ can be viewed as an asymmetric corrector.
- Skew play an important role.
Ingredients of proof

- development around the total loss $W$

\[
\mathbb{E}[h(W)] - \Phi_{\sigma_W}(h) = \sigma_W^2 \mathbb{E}[f''_h(W)(X_i^* - X_i)] \\
+ \sigma_W^2 \mathbb{E}\left[f_h^{(3)}(\xi W + (1 - \xi)W^*)(\xi(W^* - W)^2)\right].
\]

- by independence, $\mathbb{E}[f''_h(W)X_i^*] = \mathbb{E}[X_i^*] \mathbb{E}[f''_h(W)] = \mathbb{E}[X_i^*] \Phi_{\sigma_W}(f''_h) + \mathbb{E}[X_i^*] (\mathbb{E}[f''_h(W)] - \Phi_{\sigma_W}(f''_h))$.

- use the conditional expectation technic

\[
\mathbb{E}[f''_h(W)X_i] = \text{cov}(f''_h(W), \mathbb{E}[X_i|X_1, \cdots, X_n]).
\]

- write $\Phi_{\sigma_W}(f''_h)$ in term of $h$ and obtain

\[
C_h = \sigma_W^2 \mathbb{E}[X_i^*] \Phi_{\sigma_W}(f''_h) = \frac{1}{\sigma_W^2} \mathbb{E}[X_i^*] \Phi_{\sigma_W}\left(\left(\frac{x^2}{3\sigma_W^2} - 1\right)xh(x)\right).
\]
Call function $C_k(x) = (x - k)^+$: absolutely continuous with $C'_k(x) = \mathbf{1}_{\{x>k\}}$, but no second order derivative. Hypothesis not satisfied.

**Same corrector**

$$\left| \mathbb{E}[(W - k)^+] - \Phi_{\sigma_W}((x - k)^+) - \frac{1}{3} \mathbb{E}[X_i^*] k \phi_{\sigma_W}(k) \right| \sim O\left(\frac{1}{n}\right).$$

Error bounds depend on $|f'_{C_k}|, |xf''_{C_k}|$.

**Ingredients of proof:**

- write $f''_{C_k}$ as more regular function by Stein’s equation
- Concentration inequality (Chen-Shao, 2001)

**Corrector** $C_h = 0$ when $k = 0$, extremal values when $k = \pm \sigma^2$
The methodology works for other distributions.

Stein’s method in Poisson case (Chen 1975): a r.v. $\Lambda$ taking positive integer values follows Poisson distribution $\mathcal{P}(\lambda)$ if $\mathbb{E}[\Lambda g(\Lambda)] = \lambda \mathbb{E}[g(\Lambda + 1)] := A_P g(\Lambda)$.

Poisson zero bias transformation $X^*$ for $X$ with $\mathbb{E}[X] = \lambda$: $\mathbb{E}[X g(X)] = \mathbb{E}[A_P g(X^*)]$

Stein’s equation:

$$h(x) - \mathcal{P}_\lambda(h) = x g(x) - A_P g(x)$$ (4)

where $\mathcal{P}_\lambda(h) = \mathbb{E}[h(\Lambda)]$ with $\Lambda \sim \mathcal{P}(\lambda)$. 
Poisson correction

- Poisson corrector of $P_{\lambda W}(h)$ for $\mathbb{E}[h(W)]$:

$$C^P_h = \frac{\lambda W}{2} \mathbb{P}_{\lambda W}(\Delta^2 h) \mathbb{E}[X^*_I - X_I]$$

where $\Delta h(x) = h(x + 1) - h(x)$ and $\mathbb{P}(l = i) = \lambda_i/\lambda_W$.

- Proof: combinatorial techniques for integer valued r.v.
- For $h = (x - k)^+$, $\Delta^2 h(x) = 1_{\{x=k-1\}}$, then

$$C^P_h = \frac{\sigma^2_W - \lambda_W}{2([k] - 1)!} e^{-\lambda_W} \lambda_{W}^[k]^{-1}.$$
Numerical tests: $\mathbb{E}[(W_n - k)^+]$

- $\mathbb{E}[(W_n - k)^+]$ in function of $n$
- $\sigma_W = 1$ and $k = 1$ (maximal Gaussian correction) for homogeneous portfolio.
- Two graphs: $p = 0.1$ and $p = 0.01$ respectively.
- Oscillation of the binomial curve and comparison of different approximations.
Numerical tests: $\mathbb{E}[(W_n - k)^+]$

Legend: binomial (black), normal (magenta), corrected normal (green), Poisson (blue) and corrected Poisson (red). $p = 0.1, n = 100$.  

Ying Jiao
Stein’s method and zero bias transformation for CDOs
Legend: binomial (black), normal (magenta), corrected normal (green), Poisson (blue) et corrected Poisson (red).

$p = 0.01, n = 100.$

Corrected Poisson approximation remains precise even for $p = 0.1$. [Diagram showing numerical tests: $E[(W_n - k)^+]$.]

Ying Jiao

Stein’s method and zero bias transformation for CDOs
Higher order corrections

- The regularity of $h$ plays an important role. Second order approximation is given by

\[
C(2, h) = \Phi_{\sigma_W}(h) + C_h
\]

\[
+ \Phi_{\sigma_W} \left( \left( \frac{x^6}{18\sigma_W^8} - \frac{5x^4}{6\sigma_W^6} + \frac{5x^2}{2\sigma_W^4} \right) (h(x) - \Phi_{\sigma_W}(h)) \right) \mathbb{E}[X_i^*]^2
\]

\[
+ \frac{1}{2} \Phi_{\sigma_W} \left( \left( \frac{x^4}{4\sigma_W^6} - \frac{3x^2}{2\sigma_W^4} \right) h(x) \right) \left( \mathbb{E}[(X_i^*)^2] - \mathbb{E}[X_i^2] \right).
\]

(5)

- Since the call function is not regular enough, it’s difficult to control errors.
Applications to CDOs

In collaboration with David KURTZ (Calyon, Londres).
Normalizing for loss

- **Percentage loss** $l_T = \frac{1}{n} \sum_{i=1}^{n} (1 - R_i) \mathbb{1}_{\{\tau_i \leq T\}}$
  
  $K = 3\%, 6\%, 9\%, 12\%$.

- **Normalization for** $\xi_i = \mathbb{1}_{\{\tau_i \leq T\}}$ : let $X_i = \frac{\xi_i - p_i}{\sqrt{np_i(1-p_i)}}$ where $p_i = \mathbb{P}(\xi_i = 1)$. Then $\mathbb{E}[X_i] = 0$ and $\text{Var}[X_i] = \frac{1}{n}$.

- Suppose $R_i = 0$ and $p_i = p$ for simplicity. Let $p(Y) = p_i(Y) = \mathbb{P}(\tau_i \leq T | Y)$. Then
  
  $$
  \mathbb{E}[(l_T - K)^+ | Y] = \sqrt{\frac{p(Y)(1 - p(Y))}{n}} \mathbb{E}[(W - k(n, p(Y)))^+ | Y].
  $$

  where $W = \sum_{i=1}^{n} X_i$ and $k(n, p) = \frac{(K-p)\sqrt{n}}{\sqrt{p(1-p)}}$.

- **Constraint in Poisson case:** $R_i$ deterministic and proportional.

---

Ying Jiao  
Stein’s method and zero bias transformation for CDOs
Conditional error of $\mathbb{E}[(l_T - K)^+]$, in function of $K/p$, $n = 100$.

Inhomogeneous portfolio with log-normal $p_i$ of mean $p$. 
Numerical tests: Domain of validity

- **Legend**: corrected Gaussian error (blue), corrected Poisson error (red). np=5 and np=20.

- **Domain of validity**: $np \approx 15$

- **Maximal error**: when $k$ near the average loss; overlapping area of the two approximations.
Comparison with saddle-point method

- **Legend**: Gaussian error (red), saddle-point first order error (blue dot), saddle-point second order error (blue). np=5 and np=20.

![Graph for np=5](image1.png)

![Graph for np=20](image2.png)
Numerical tests: stochastic $R_i$

- Conditional error of $\mathbb{E}[(l_T - K)^+]$, in function of $K/p$ for Gaussian approximation error with stochastic $R_i$.
- $R_i$: Beta distribution with expectation 50% and variance 26%, independent (Moody’s).
- Corrector depends on the third order moment of $R_i$.
- Confidence interval 95% by $10^6$ Monte Carlo simulation.
Gaussian approximation better when $p$ is large as in the standard case. $np=5$ and $np=20$. 

**Numerical tests : stochastic $R_i$**
Real-life CDOs pricing

- Parametric correlation model of $\rho(Y)$ to match the smile.
- Method adapted to **all conditional independent models**, not only in the normal factor model case.
- Numerical results compared to the recursive method (Hull-White).
- Calculation time largely reduced: $1 : 200$ for one price.
Real-life CDOs pricing

- Error of prices in bp ($10^{-4}$) for corrected Gauss-Poisson approximation with realistic local correlation.
- Expected loss 4% – 5%

![Break Even Error Chart]

Stein’s method and zero bias transformation for CDOs
Conclusion remark and perspective

- Tests for VaR and sensitivity remain satisfactory
- Perspective: remove the deterministic recovery rate condition in the Poisson case.