Nested Simulation
in Portfolio Risk Measurement

Michael B. Gordy\(^1\) Sandeep Juneja\(^2\)

\(^1\)Federal Reserve Board  <michael.gordy@frb.gov>

\(^2\)Tata Institute of Fundamental Research  <juneja@tifr.res.in>

July 2008
On pricing derivatives

- Consider a very general derivatives portfolio: interest rate swaps, Treasury futures, equity options, default swaps, CDO tranches, etc.
- In many or even most cases, preferred pricing model requires simulation.
  - Models with analytical solution typically impose restrictive assumptions (Black-Scholes, most famously).
  - Simulation almost unavoidable for many path-dependent and basket derivatives.
Risk-management adds a new wrinkle

- Talking here about risk-measurement of portfolio at some chosen horizon.
  - Large loss exceedance probabilities.
  - Quantiles of the loss distribution (value-at-risk). Expected shortfall
- Simulation-based algorithm is nested:
  - **Outer step:** Draw paths for underlying prices to horizon and calculate implied cashflows during this period.
  - **Inner step:** Re-price each position at horizon conditional on drawn paths.
- Computational task perceived as burdensome because inner step simulation must be executed once for each outer step simulation.
- Practitioners invariably use rough pricing tools in the inner step in order to avoid nested simulation.
- We show the convention view is wrong – inner step simulation need not be burdensome.
Model framework

- The present time is normalized to 0 and the model horizon is $H$.
- Let $X_t$ be a vector of $m$ state variables that govern underlying prices referenced by derivatives.
  - interest rates, default intensities, commodity prices, equity prices, etc.
- Let $\xi$ be the information generated by $\{X_t\}$ on $t = (0, H]$.
- The portfolio consists of $K + 1$ positions.
- The price of position $k$ at horizon depends on $\xi$ and the contractual terms of the instrument.
- Position 0 represents the sub-portfolio of instruments for which there exist analytical pricing functions.
- Positions 1 through $K$ must be priced by simulation.
Portfolio loss

- “Loss” is defined on a mark-to-market basis
  - Current value less discounted horizon value, less PDV of interim cashflows.
- Let $W_k$ be the loss on position $k$; $Y = \sum_k W_k$ is the portfolio loss.
- Conditional on $\xi$, $W_k(\xi)$ is non-stochastic.
- Except for position 0, we do not observe $W_k(\xi)$, but rather obtain noisy simulation estimates $\tilde{W}_k(\xi)$ and $\tilde{Y}(\xi)$. 

Simulation framework

Let $L$ be number of outer step trials. For each trial $\ell = 1, \ldots, L$:

1. Draw a single path $X_t$ for $t = (0, H]$ under the **physical measure**.
   - Let $\xi$ represent the relevant information for this path.

2. Evaluate the value of each position at horizon.
   - Accrue interim cashflows to $H$.
   - Closed-form price at $H$ for instrument 0.
   - Simulation with $N$ “inner step” trials to price each remaining positions $k = 1, \ldots, K$. Here we use the **risk-neutral measure**.

3. Discount back to time 0, subtract from current value, get our position losses $W_0(\xi), \tilde{W}_1(\xi), \ldots, \tilde{W}_K(\xi)$.

4. Portfolio loss $\tilde{Y}(\xi) = W_0(\xi) + \tilde{W}_1(\xi) + \ldots + \tilde{W}_K(\xi)$.
Dependence in inner and outer steps

- Full dependence structure across the portfolio is captured in the period up to the model horizon.
- Inner step simulations are run independently across positions.
  - Value of position $k$ at time $H$ is simply a conditional expectation of its own subsequent cashflows.
  - Does not depend on future cashflows of other positions.
- Independent inner steps imply that pricing errors are independent across positions, and so tend to diversify away at portfolio level.
- Also reduces memory footprint of inner step: For position $k$, need only draw joint paths for the elements of $X_t$ upon which instrument $k$ depends.
Overview of our contribution

- Key insight of paper is that mean-zero pricing errors have minimal effect on estimation. Can set $N$ small!
- For finite $N$, estimators of exceedance probabilities, VaR and ES are biased (typically upwards).
- We obtain bias and variance of these estimators.
- Can allocate fixed computational budget between $L$, $N$ to minimize mean square error of estimator.
- Large portfolio asymptotics ($K \to \infty$).
- Jackknife method for bias reduction.
- Dynamic allocation scheme for greater efficiency.
Estimating probability of large losses

- Goal is efficient estimation of $\alpha = P(Y(\xi) > u)$ via simulation for a given $u$ (typically large).
- If analytical pricing formulae were available, then for each generated $\xi$, $Y(\xi)$ would be observable.
- In this case, outer step simulation would generate iid samples $Y_1(\xi_1), Y_2(\xi_2), \ldots, Y_L(\xi_L)$, and we would take average

$$\frac{1}{L} \sum_{i=1}^{L} 1[Y_i(\xi_i) > u]$$

as an estimator of $\alpha$. 
Pricing errors in inner step

- When analytical pricing formulae unavailable, we estimate $Y(\xi)$ via inner step simulation.
- Let $\zeta_{ki}(\xi)$ be zero-mean pricing error associated with $i^{th}$ “inner step” trial for position $k$.
- Let $Z_i(\xi)$ be the zero-mean portfolio pricing error associated with this inner step trial, i.e., $Z_i(\xi) = \sum_{k=1}^{K} \zeta_{ki}(\xi)$.
- Average portfolio error across trials is $\bar{Z}^N(\xi) = \frac{1}{N} \sum_{i=1}^{N} Z_i(\xi)$.
- Instead of $Y(\xi)$, we take as surrogate $\tilde{Y}(\xi) \equiv Y(\xi) + \bar{Z}^N(\xi)$.
- By the law of large numbers,
  \[ \bar{Z}^N(\xi) \rightarrow 0 \quad \text{a.s.} \quad \text{as} \ N \rightarrow \infty \]
  i.e., pricing error vanishes as $N$ grows large.
We generate iid samples \( (\tilde{Y}_1(\xi_1), \ldots, \tilde{Y}_L(\xi_L)) \) via outer and inner step simulation, and take average

\[
\hat{\alpha}_{LN} = \frac{1}{L} \sum_{\ell=1}^{L} 1[\tilde{Y}_\ell(\xi_\ell) > u].
\]

Let \( \alpha_N \equiv P(\tilde{Y}(\xi) > u) = E[\hat{\alpha}_{LN}] \).

Mean square error decomposes as

\[
E[\hat{\alpha}_{LN} - \alpha]^2 = E[\hat{\alpha}_{LN} - \alpha_N + \alpha_N - \alpha]^2 = E[\hat{\alpha}_{LN} - \alpha_N]^2 + (\alpha_N - \alpha)^2.
\]

\( \hat{\alpha}_{LN} \) has binomial distribution, so variance term is

\[
E[\hat{\alpha}_{LN} - \alpha_N]^2 = \frac{\alpha_N(1 - \alpha_N)}{L}.
\]
Approximation for bias

Proposition:

$$\alpha_N = \alpha + \frac{\theta}{N} + O\left(\frac{1}{N^{3/2}}\right)$$

where

$$\theta = \frac{-1}{2} \frac{d}{du} f(u) E[\sigma^2_\xi | Y = u],$$

and where $\sigma^2_\xi = V[Z_1 | \xi]$ is the conditional variance of the portfolio pricing error, and $f(u)$ is density of $Y$.

- Our approach follows Gouriéroux, Laurent and Scaillet (JEF, 2000) and Martin and Wilde (Risk, 2002) on sensitivity of VaR to portfolio allocation.
- Similar approximations for bias in VaR and ES.
Highly stylized example for which RMSE has analytical expression.

Homogeneous portfolio of $K$ positions.

Let $X \sim \mathcal{N}(0, 1)$ be a market risk factor.

Loss on position $k$ is $W_k = (X + \epsilon_k)/K$ per unit exposure where the $\epsilon_k$ are iid $\mathcal{N}(0, \nu^2)$.

- Scale exposures by $1/K$ to ensure that portfolio loss distribution converges to $\mathcal{N}(0, 1)$ as $K \to \infty$.

Implies portfolio loss $Y \sim \mathcal{N}(0, 1 + \nu^2/K)$.

Assume pricing errors $\zeta_k$ iid $\mathcal{N}(0, \eta^2)$, so portfolio pricing error has variance $\sigma^2 = \eta^2/K$ for each inner step trial.

Implies $\tilde{Y} = Y + \tilde{Z}^N \sim \mathcal{N}(0, 1 + \nu^2/K + \sigma^2/N)$.
Density of the loss distribution

Parameters: $\nu = 3$, $\eta = 10$, $K = 100$. 
Variance of $Y$ is $s^2 = 1 + \nu^2/K$, variance of $\tilde{Y}$ is $\tilde{s}^2 = s^2 + \sigma^2/N$.

Exact bias is
$$\alpha_N - \alpha = \Phi(-u/\tilde{s}) - \Phi(-u/s)$$

where $\Phi$ is the standard normal cdf.

Apply Proposition to approximate $\alpha_N - \alpha \approx \theta/N$ where
$$\theta = \phi(-u/s)\frac{u\sigma^2}{2s^3}$$

where $\phi$ is the standard normal density.
Parameters: $\nu = 3$, $\eta = 10$, $K = 100$, $u = F^{-1}(0.99)$. 
Optimal allocation of workload

- Total computational effort is $L(N\gamma_1 + \gamma_0)$ where
  - $\gamma_0$ is average cost to sample $\xi$ (outer step).
  - $\gamma_1$ is average cost per inner step sample.
- Fix overall computational budget $\Gamma$.
- Minimize mean square error subject to $\Gamma = L(N\gamma_1 + \gamma_0)$.
- For $\Gamma$ large, get

$$N^* \approx \left(\frac{2\theta^2}{\alpha(1-\alpha)\gamma_1}\right)^{1/3} \Gamma^{1/3}$$

$$L^* \approx \left(\frac{\alpha(1-\alpha)}{2\gamma_1^2\theta^2}\right)^{1/3} \Gamma^{2/3}$$

- Similar results in Lee (1998).
- Analysis for VaR and ES proceeds similarly, also find $N^* \propto \Gamma^{1/3}$. 
Approximate $\Gamma \propto N \cdot L$. Parameters: $\nu = 3$, $\eta = 10$, $K = 100$, $u = F^{-1}(0.99)$. 

RMSE in Gaussian example
Approximate $\Gamma \propto N \cdot L$. Parameters: $\nu = 3$, $\eta = 10$, $K = 100$, $u = F^{-1}(0.99)$. 
Optimal $N$ depends on exceedance threshold

Quantiles of the distribution of $Y$ marked in basis points. Budget is $\Gamma = N \cdot L$.
Parameters: $\nu = 3$, $\eta = 10$ and $K = 100$. 
Consider an infinite sequence of exchangeable positions. Let $\bar{Y}^K$ be average loss per position on a portfolio consisting of the first $K$ positions, i.e.,

$$\bar{Y}^K = \frac{1}{K} \sum_{k=1}^{K} W_k$$

Assume budget is $\chi K^\beta$ for $\chi > 0$ and $\beta \geq 1$.

Assume fixed cost per outer step is $\psi(m, K)$, so budget constraint is

$$L(KN\gamma_1 + \psi(m, K)) = \chi K^\beta$$

Proposition: For $\beta \leq 3$, $N^* \rightarrow 1$ as $K \rightarrow \infty$, specifically,

$$N^* = \max \left(1, \left( \frac{2\ddot{\theta}^2 \chi}{\alpha_u(1 - \alpha_u)\gamma_1} \right)^{1/3} K^{\beta/3-1} \right)$$
Jackknife estimators for bias correction

- In simplest version, divide inner step sample into two subsamples of \( N/2 \) each.
- Let \( \hat{\alpha}_j \) be the estimator of \( \alpha \) based on subsample \( j \).
- Observe that the bias in \( \hat{\alpha}_j \) is \( \theta/(N/2) \) plus terms of order \( O(1/N^{3/2}) \).
- We define the jackknife estimator \( a_{LN} \) as
  \[
  a_{LN} = 2\hat{\alpha}_{LN} - \frac{1}{2}(\hat{\alpha}_1 + \hat{\alpha}_2)
  \]
- Jackknife estimator requires no additional simulation work.
- Can generalize by dividing the inner step sample into \( I \) overlapping subsamples of \( N - N/I \) trials each.
The bias in $a_{LN}$ is

$$E[a_{LN}] - \alpha = 2\alpha_N - \alpha_{N/2} - \alpha$$

$$= 2(\alpha + \theta/N + O(1/N^{3/2})) - (\alpha + \theta/(N/2) + O(1/N^{3/2})) - \alpha$$

$$= \theta \left( \frac{2}{N} - \frac{1}{N/2} \right) + O(1/N^{3/2}) = O(1/N^{3/2}).$$

- First-order term in the bias is eliminated.
- Variance of $a_{LN}$ depends on covariances among $\hat{\alpha}_{LN}, \hat{\alpha}_1, \hat{\alpha}_2$.
  Tedious but tractable. Find $\text{Var}[a_{LN}] > \text{Var}[\hat{\alpha}_{LN}]$.
- Optimal choice of $N^*$ and $L^*$ changes because bias is a lesser consideration and variance a greater consideration.
  - Find $N^* \propto \Gamma^{1/4}$ (versus 1/3 for uncorrected estimator) and $L^* \propto \Gamma^{3/4}$ (versus 2/3).
Dynamic allocation

- Through dynamic allocation of workload in the inner step we can further reduce the computational effort in the inner step while increasing the bias by a negligible controlled amount.
- Consider the estimation of large loss probabilities $P(Y > u)$.
- We form a preliminary estimate $\tilde{Y}^N$ based on the average of a small number $N$ of inner step trials.
- If this estimate is much smaller or much larger than $u$, it may be a waste of effort to generate many more samples in the inner simulation step.
- If this average is close to $u$, it makes sense to generate many more inner step samples in order to increase the probability that the estimated $1[\tilde{Y}(\xi) > u]$ is equal to the true value $1[Y(\xi) > u]$. 
Proposed allocation scheme

- For each trial \( \ell \) of the outer step, we generate \( \delta N \) inner step trials.
- If the resultant loss estimate \( Y + \tilde{Z}_{\delta N} < u - \epsilon \) for some well chosen \( \epsilon > 0 \) then we terminate the inner step and our sample output is zero.
- Otherwise, we generate additional \( (1 - \delta)N \) samples and our sample output is \( 1[Y + \tilde{Z}_N > u] \).
The additional bias can be bounded

- The additional bias is bounded above by

\[ P(\bar{Z}_{\delta N} \leq -\epsilon) + P(\bar{Z}_{\delta N, N} > \frac{\delta}{1 - \delta \epsilon}). \]

- Hoeffding’s inequality can be used to develop exact bounds if increments are bounded.

- Alternatively, by assuming that each \( Z_i \) is approximately Normally distributed (as it is a sum of zero mean noises from \( K \) positions), by batching the sum of a few \( Z_i \)'s if necessary, one can develop upper bounds.
Conclusion

- Large errors in pricing individual position can be tolerated so long as they can be diversified away.
  - Inner step gives errors that are zero mean and independent. Ideal for diversification!
  - In practice, large banks have many thousands of positions, so might have $N^* \approx 1$.
- Results suggest current practice is misguided.
  - Use of short-cut pricing methods introduces model misspecification.
  - Errors hard to bound and do not diversify away at portfolio level.
  - Practitioners should retain best pricing models that are available, run inner step with few trials.
- Dynamic allocation is robust and easily implemented in a setting with many state prices and both long and short exposures.