

A technique for multivariate stratified sampling

Natalie Packham

Frankfurt School of Finance & Management

Frankfurt Mathfinance Workshop, 26-27 March 2007

27 March 2007

Motivation

- Valuation of financial claim with Monte Carlo simulation:
 - ▶ generate sample paths of the evolution of underlying financial securities (under the risk-neutral measure)
 - ▶ expectation/ sample mean of claim's payoff is estimator of claim's fair value
- Require estimator to be unbiased and consistent
- Variance of estimator is key figure for rate of convergence
~> Variance reduction techniques
- Here: Claims
 - ▶ that depend on several underlying financial securities
 - ▶ that are sensitive to dependence structure of underlyings
- Examples: First-to-default credit baskets, CDO tranches

Outline

- Univariate stratified sampling
- Bivariate sampling strategy
- Asymptotic behaviour of estimator
- Example (FTD basket)
- Further research

Univariate stratified sampling

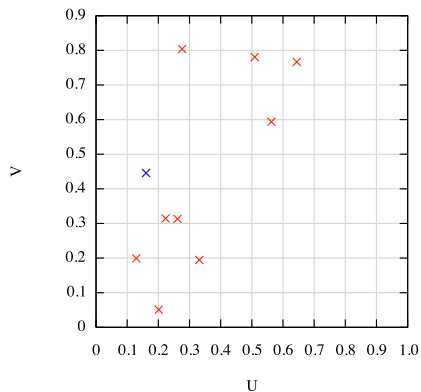
- Principal idea of stratified sampling:
 - ▶ Uniform random variable $U \sim U(0, 1)$
 - ▶ Partition A_1, \dots, A_K of $[0, 1]$
 - ▶ Then, for (payoff) function g

$$\mathbb{E}g(U) = \sum_{i=1}^K \mathbb{E}(g(U)|U \in A_i) \mathbf{P}(U \in A_i)$$

- ▶ Stratified estimator given by sampling U given $U \in A_i$, $i = 1, \dots, K$
 - ▶ For example:
 - ★ Choose A_i such that $\mathbf{P}(U \in A_i) = 1/K$, for $i = 1, \dots, K$
 - ★ Select one sample from each partition.
- See ([Glasserman, 2004](#)), Chapter 4.3.

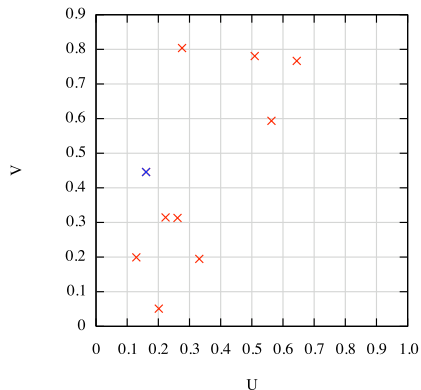
Bivariate sampling strategy I

Uniform bivariate samples, correlated with $\rho=0.5$

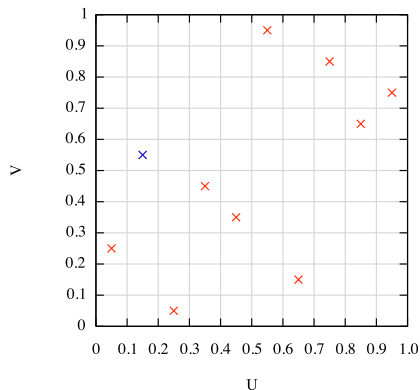


Bivariate sampling strategy I

Uniform bivariate samples, correlated with $\rho=0.5$



Stratified samples



Bivariate sampling strategy II

Definition 1 (Rank statistic)

Let X^1, \dots, X^n be iid random variables. Reorder them such that $X^{(1)} < \dots < X^{(n)}$. The index of X^k within $X^{(1)}, \dots, X^{(n)}$ is the *k-th rank statistic*, given by

$$h_{k,n}(X^1, \dots, X^n) := \sum_{i=1}^n \mathbf{1}_{\{X^i \leq X^k\}}.$$

Bivariate sampling strategy III

Bivariate stratified sampling strategy

Let (U, V) be a random vector (possibly with some dependence). For n independent samples of the bivariate random vector, $(U^1, V^1), \dots, (U^n, V^n)$, set

$$\tilde{U}_n^k := \frac{h_{k,n}(U^1, \dots, U^n) - 1/2}{n}$$
$$\tilde{V}_n^k := \frac{h_{k,n}(V^1, \dots, V^n) - 1/2}{n}.$$

Let $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ be bounded, measurable and a.s. continuous. An estimator of $\mathbb{E}g(U, V)$ is then given by

$$\frac{1}{n} \sum_{k=1}^n g(\tilde{U}_n^k, \tilde{V}_n^k).$$

Remark: The function g can be viewed as a payoff function.

Bivariate sampling strategy IV

- Variant I: Define

$$\tilde{U}_n^k := \frac{h_{k,n}(U^1, \dots, U^n) - 1 + B_{h_{k,n}(\mathbf{u})}^n(U^k)}{n}$$
$$\tilde{V}_n^k := \frac{h_{k,n}(V^1, \dots, V^n) - 1 + B_{h_{k,n}(\mathbf{v})}^n(V^k)}{n},$$

where $B_i^n(x)$ denotes the beta distribution functions with parameters i and n . We can show that $\tilde{U}_n^k \sim U(0, 1)$ and $\tilde{V}_n^k \sim U(0, 1)$.

Bivariate sampling strategy IV

- Variant I: Define

$$\tilde{U}_n^k := \frac{h_{k,n}(U^1, \dots, U^n) - 1 + B_{h_{k,n}(\mathbf{u})}^n(U^k)}{n}$$
$$\tilde{V}_n^k := \frac{h_{k,n}(V^1, \dots, V^n) - 1 + B_{h_{k,n}(\mathbf{v})}^n(V^k)}{n},$$

where $B_i^n(x)$ denotes the beta distribution functions with parameters i and n . We can show that $\tilde{U}_n^k \sim U(0, 1)$ and $\tilde{V}_n^k \sim U(0, 1)$.

- Variant II: Let U^1, \dots, U^n and V^1, \dots, V^n be independent uniform random variables. Define

$$\tilde{U}_n^k := \frac{h_{k,n}(U^1, \dots, U^n) - 1 + U^1}{n}$$
$$\tilde{V}_n^k := \frac{h_{k,n}(V^1, \dots, V^n) - 1 + V^1}{n}$$

Then $\tilde{U}_n^k \sim U(0, 1)$ and $\tilde{V}_n^k \sim U(0, 1)$.

Asymptotic behaviour of estimator I

From Weak Law of Large Numbers we know:

For bounded and a.s. continuous $g : [0, 1]^2 \rightarrow \mathbb{R}$:

$$\frac{1}{n} \sum_{k=1}^n g(U^k, V^k) \xrightarrow{\mathbf{P}} \mathbb{E}g(U, V).$$

Asymptotic behaviour of estimator I

From Weak Law of Large Numbers we know:

For bounded and a.s. continuous $g : [0, 1]^2 \rightarrow \mathbb{R}$:

$$\frac{1}{n} \sum_{k=1}^n g(U^k, V^k) \xrightarrow{\mathbf{P}} \mathbb{E}g(U, V).$$

Does this still hold for $g(\tilde{U}_n^k, \tilde{V}_n^k)$?

Asymptotic behaviour of estimator I

From Weak Law of Large Numbers we know:
For bounded and a.s. continuous $g : [0, 1]^2 \rightarrow \mathbb{R}$:

$$\frac{1}{n} \sum_{k=1}^n g(U^k, V^k) \xrightarrow{\mathbf{P}} \mathbb{E}g(U, V).$$

Does this still hold for $g(\tilde{U}_n^k, \tilde{V}_n^k)$?

Proposition 2

Let $g : [0, 1]^2 \rightarrow \mathbb{R}$ be bounded, measurable and a.s. continuous. Then

$$\frac{1}{n} \sum_{k=1}^n g(\tilde{U}_n^k, \tilde{V}_n^k) \xrightarrow{\mathbf{P}} \mathbb{E}g(U, V).$$

Asymptotic behaviour of estimator II

Remark: It suffices to show that

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{1}{n} \sum_{k=1}^n g(\tilde{U}_n^k, \tilde{V}_n^k) - \frac{1}{n} \sum_{k=1}^n g(U^k, V^k) \right| > \varepsilon \right) = 0,$$

because the second sum converges \mathbf{P} -a.s. to $\mathbb{E}g(U, V)$ by the Strong Law of Large Numbers.

Asymptotic behaviour of estimator III

Idea of proof:

Asymptotic behaviour of estimator III

Idea of proof:

- We show $\tilde{U}_n^k \xrightarrow{\text{P-a.s.}} U^k$ and $\tilde{V}_n^k \xrightarrow{\text{P-a.s.}} V^k$, $1 \leq k \leq n$.

Asymptotic behaviour of estimator III

Idea of proof:

- We show $\tilde{U}_n^k \xrightarrow{\mathbf{P}\text{-a.s.}} U^k$ and $\tilde{V}_n^k \xrightarrow{\mathbf{P}\text{-a.s.}} V^k$, $1 \leq k \leq n$.
- For g bounded, measurable and a.s. continuous, we obtain $g(\tilde{U}_n^k, \tilde{V}_n^k) \xrightarrow{\mathbf{P}\text{-a.s.}} g(U^k, V^k)$.

Asymptotic behaviour of estimator III

Idea of proof:

- We show $\tilde{U}_n^k \xrightarrow{\text{P-a.s.}} U^k$ and $\tilde{V}_n^k \xrightarrow{\text{P-a.s.}} V^k$, $1 \leq k \leq n$.
- For g bounded, measurable and a.s. continuous, we obtain $g(\tilde{U}_n^k, \tilde{V}_n^k) \xrightarrow{\text{P-a.s.}} g(U^k, V^k)$.
- We then show that

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{1}{n} \sum_{k=1}^n g(\tilde{U}_n^k, \tilde{V}_n^k) - \frac{1}{n} \sum_{k=1}^n g(U^k, V^k) \right| > \varepsilon \right) = 0,$$

which holds iff

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\left| \frac{1}{n} \sum_{k=1}^n g(\tilde{U}_n^k, \tilde{V}_n^k) - g(U^k, V^k) \right| \wedge \max_{x,y} (g(x,y)) \right) = 0$$

and we can show that the latter holds by employing the Dominated Convergence Theorem.

Asymptotic behaviour of estimator XIII

Variance of stratified estimator compared to simple estimator?

Asymptotic behaviour of estimator XIII

Variance of stratified estimator compared to simple estimator???

Asymptotic behaviour of estimator XIII

Variance of stratified estimator compared to simple estimator???

- Stratified samples are no longer independent
- Calculation of total variance requires that covariance be identified
- Sampling into strata reduces variance in each dimension
- Bias to be identified as estimator is unbiased only in the limit

Asymptotic behaviour of estimator XIV

An example: Let U, V be independent, fix $n \in \mathbb{N}$, choose u, v such that $\lfloor nu + 1/2 \rfloor / n = u$ and $\lfloor nv + 1/2 \rfloor / n = v$ and let $g(x, y) = \mathbf{1}_{\{x \leq u, y \leq v\}}$. Then

$$\text{Var} \left(\frac{1}{n} \sum_k \mathbf{1}_{\{U^k \leq u, V^k \leq v\}} \right) = \frac{uv - (uv)^2}{n}$$

and

$$\begin{aligned} \text{Var} \left(\frac{1}{n} \sum_k \mathbf{1}_{\{\tilde{U}_n^k \leq u, \tilde{V}_n^k \leq v\}} \right) &= \frac{\mathbf{P}(\tilde{U}_n^1 \leq u, \tilde{V}_n^1 \leq v)}{n} \\ &+ \frac{n-1}{n} \mathbf{P}(\tilde{U}_n^1 \leq u, \tilde{V}_n^1 \leq v, \tilde{U}_n^2 \leq u, \tilde{V}_n^2 \leq v) - \left(\mathbf{P}(\tilde{U}_n^1 \leq u, \tilde{V}_n^1 \leq v) \right)^2 \\ &= \frac{uv}{n} + \frac{u(nu-1)v(nv-1)}{n(n-1)} - (uv)^2 \\ &= \frac{uv - (uv)^2}{n} - \frac{uv(u+v-2uv)}{n-1} + \frac{uv - (uv)^2}{n(n-1)}. \end{aligned}$$

Asymptotic behaviour of estimator XV

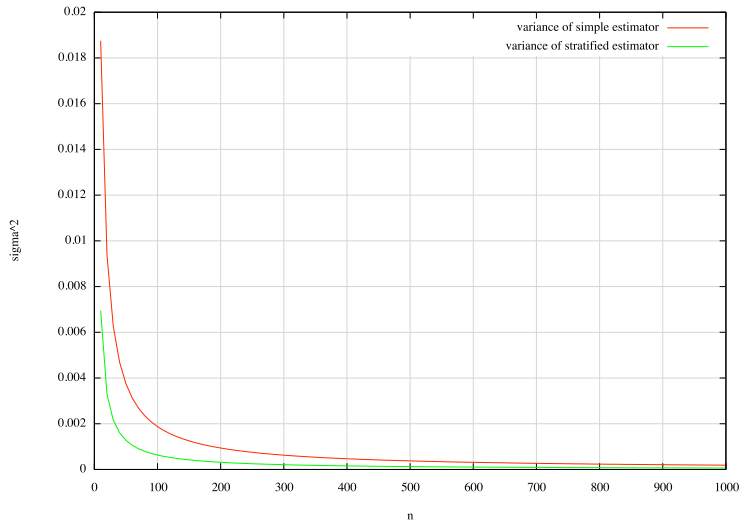


Figure: Variance of simple estimator over stratified estimator for $u = v = 0.5$.

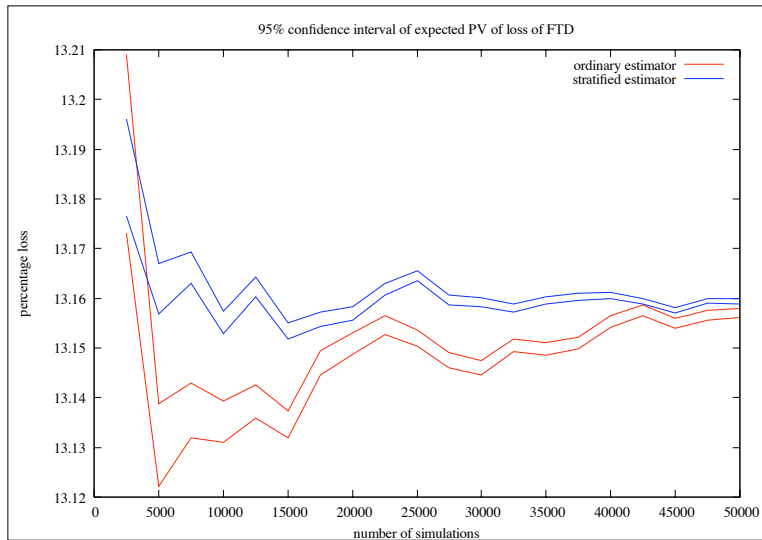
FTD Example I

- First-to-default (FTD) basket insures against loss of the first default in underlying credit portfolio.
- Premium leg provides regular payments (“the spread”) until first default event or maturity (whichever takes place first).
- Default leg pays loss-given-default in case of first default.
- FTD is sensitive to dependence structure among underlying credits.

FTD Example II

- Example credit portfolio contains 5 homogeneous credits, each of which has
 - ▶ a CDS spread of 1%,
 - ▶ a recovery rate of 50%.
 - ▶ Dependence of default times among credits modelled by Normal copula with correlation of 0.5.
- FTD has a maturity of 5 years.
- Expected PV of loss is 13.159%, fair spread is 3.39%.

FTD Example III



Further research

- Quantify variance reduction, possibly restricted to some copulas and some payoffs.
- Investigate overall efficiency by analyzing trade-off between variance reduction and running time of simulation.
- Compare to other variance reduction techniques (Latin hypercube sampling, importance sampling, low discrepancy methods, ...).

Bibliography

Glasserman, P. (2004).

Monte Carlo Methods in Financial Engineering.

Springer.