A variance reduction method for computing VaR

1. Computing Value at Risk by Monte Carlo simulations

2. Importance Sampling for variance reduction

3. Interacting Particle Systems for Importance Sampling (IPS-IS)

4. Simulation results

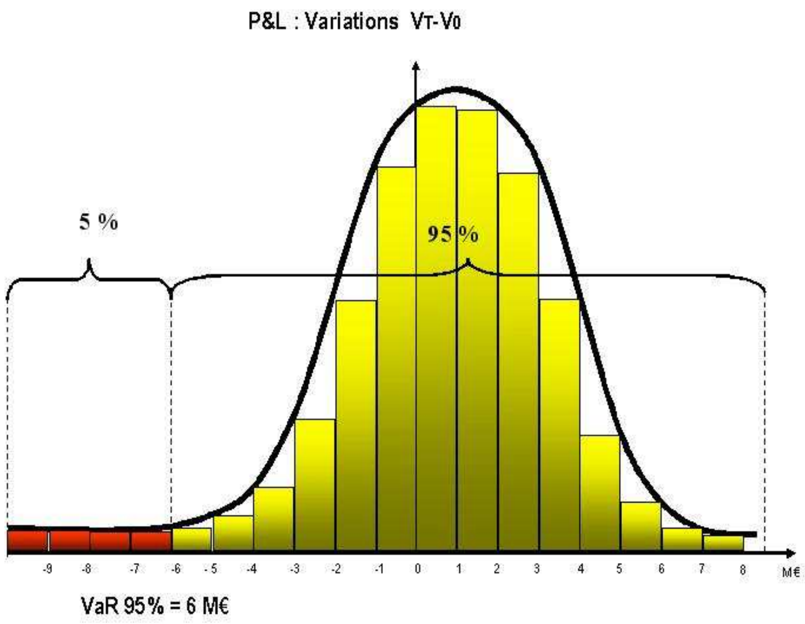
Nadia Oudjane - EDF R&D- Journées MAS 2008

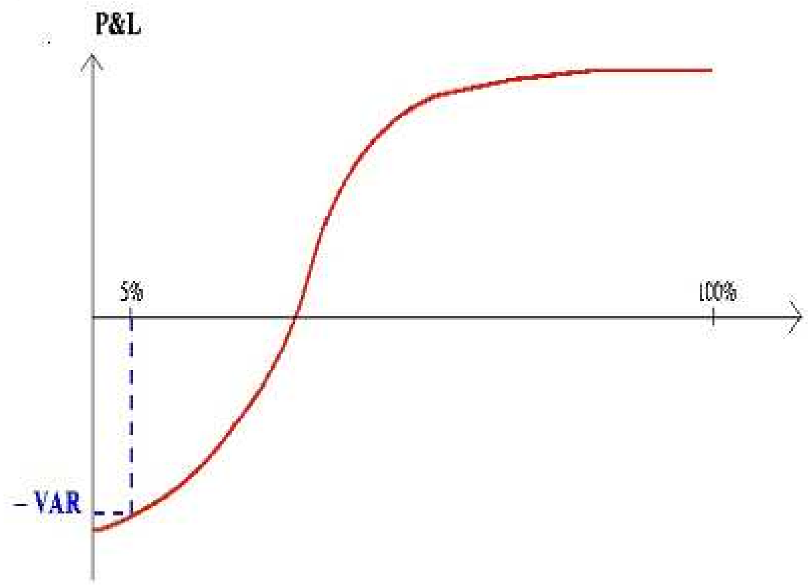
Value at Risk and quantile

► **P&L of a portfolio on $[0,T]$** $\qquad \Delta V(X) = V_T - V_0 + \int_0^T CF$

with $\quad X \in \mathbb{R}^d \quad$ **the risk factors impacting the portfolio value on** $[0,T]$

► **Value at Risk** $\quad VaR_\alpha = \left|\, \text{inf}\{s \in \mathbb{R} \,|\, \mathbb{P}(\Delta V \le s) \ge 1-\alpha\} \,\right|$





$$VaR_\alpha = |F^-(\alpha)| \,, \quad \textbf{where} \quad F(s) = \mathbb{P}(\Delta V \le s) \,, \quad \textbf{for all } s \in \mathbb{R}$$

Monte Carlo method for VaR estimation

- **The distribution function F can be viewed as an expectation**

$$F(s) = \mathbb{E}[\mathbf{I}_{\Delta V(X) \leq s}] \,, \quad \text{for all} \quad s \in \mathbb{R}$$

- **Traditional Monte Carlo Method for computing VaR**

1. **Monte Carlo simulations give an approximation of $F(s)$:**

$$\hat{F}_N(s) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{I}(\Delta V(X_i) \leq s) \,, \quad \text{for all } s \in \mathbb{R}$$

$\Rightarrow$ **Too many evaluations of ΔV for a given accuracy**

2. **Inversion of $\hat{F}^N$ and interpolation for approximating VaR**

Importance Sampling for variance reduction

▶ **Change of measure** $p \longrightarrow q$ **where** q **dominates** Hp

$$m = \mathbb{E}_p[H(X)] = \mathbb{E}_q[H(Y)\frac{p}{q}(Y)]\,, \quad \text{where} \quad X \sim p \quad \text{and} \quad Y \sim q$$

▶ **Optimal change of measure** $p \longrightarrow q^*$ **achieves zero variance if** $H \geq 0$

$$q^* = \frac{Hp}{\int H(x)p(x)\,dx} = \frac{Hp}{\mathbb{E}_p[H(X)]} = H \cdot p$$

▶ **Monte Carlo approximation**

$$\mathbb{E}_p[H(X)] \approx \hat{m}_M^q = \frac{1}{M}\sum_{i=1}^{M} H(Y_i)\frac{p}{q}(Y_i)\,, \quad \text{where} \quad (Y_1, \cdots, Y_M) \quad \text{i.i.d.} \sim q$$

$\Rightarrow$ **How to simulate and evaluate approximately q^* ?**

Variance of the Importance Sampling estimate

► **Let q be a (possibly random) importance probability density dominating q^***

$$Var(\hat{m}_M^q) = \mathbb{E}\Big[Var[\hat{m}_M^q \,|\, \mathcal{F}_q]\Big] + \underbrace{Var\Big[\mathbb{E}[\hat{m}_M^q \,|\, \mathcal{F}_q]\Big]}_{=0}$$

$\mathcal{F}_q$ **denotes the sigma-algebra generated by the random variables involved in** q

► **The variance of the IS estimate depends on the "distance" between q and q^***

$$Var(\hat{m}_M^q) = \frac{m^2}{M}\mathbb{E}\left[\int [(q^* - q)\frac{q^*}{q}](x)dx\right]$$

► **Idea : use Interacting Particle Systems for Importance Sampling (IPS-IS) to approximate q^* by q^N based on an N-particle system to achieve**

$$Var(\hat{m}_M^{q^N}) \leq \frac{C}{MN^\alpha} \quad \text{with} \quad 0 < \alpha < 1/2$$

Some alternative approaches

- Large deviation approximation for rare events simulation

- Approximation of H to obtain a simple form fo q^*

ex : [Glasserman&al00] for computing VaR, Δ-Γ approximation of the portfolio

- Cross-entropy [Homem-de-Mello&Rubinstein02]

q^θ is chosen in a parametric family such as to minimize the entropy $K(q^\theta, q^*)$

- Interacting Particle Systems whithout Importance Sampling

[DelMoral&Garnier05], [Cerou&al06]

Interacting Particle Systems for Importance Sampling (IPS-IS) can be viewed as a non parametric version of cross entropy approach

Progressive correction [Musso&al01]

► We introduce a sequence of non negative functions $(G_k)_{0 \le k \le n}$ such that

for all $x \in \mathbb{R}^d$,
$$\begin{cases} G_0(x) = 1 \\ \text{The product} \quad G_0(x) \cdots G_n(x) = H(x) \\ \text{If} \quad G_k(x) = 0 \quad \text{then} \quad G_{k+1}(x) = 0 \end{cases}$$

► In our case $H(x) = \mathbf{I}(\Delta V(x) \le s)$ then we choose

$$G_k(x) = \mathbf{I}_{\Delta V(x) \le s_k}\,, \quad \text{with} \quad s = s_n \le \cdots \le s_0 = +\infty$$

► Dynamical system on the space of probability measures $(\nu_k)_{0 \le k \le n}$

$$\begin{cases} \nu_0 = p\,dx \\ \nu_k = \dfrac{G_k \nu_{k-1}}{\int_{\mathbb{R}^d} G_k(x)\nu_{k-1}(x)\,dx} = G_k \cdot \nu_{k-1}\,, \quad \text{for all} \quad 1 \le k \le n \end{cases}$$

$\Rightarrow \nu_n = q^*\,dx$

Space exploration

▶ **We introduce a sequence of Markov kernels** $(Q_k)_{0\leq k\leq n}$ **such that**

$$\nu_k \approx \nu_k Q_k \quad \text{i.e.} \quad \nu_k(dx) \approx \int_{\mathbb{R}^d} \nu_k(du) Q_k(u,dx) \; , \quad \text{for all} \quad x \in \mathbb{R}^d$$

▶ **In our case where** $G_k(x) = \mathbf{I}_{\Delta V(x)\leq s_k}$ **, if** p **is Gaussian then** Q_k **is easily obtained from a Gaussian kernel** Q **reversible for** p**,**

$$Q_k(x,dx') = Q(x,dx')\mathbf{I}_{\Delta V(x)\leq s_k} + \left[1 - Q(x, \Delta V^{-}((-\infty, s_k]))\right]\delta_x(dx')$$

▶ **Dynamical system on the space of probability measures** $(\nu_k)_{0\leq k\leq n}$

$$\begin{cases} \nu_0 = p\,dx \\ \nu_k = G_k \cdot (\nu_{k-1}Q_{k-1}) \; , \quad \text{for all} \quad 1 \leq k \leq n \end{cases}$$

$\Rightarrow \nu_n = q^*\,dx$

Approximation of the dynamical system

► **The idea is to replace at each iteration k, $\nu_{k-1} Q_{k-1}$ by its N-empirical measure $S^N(\nu_{k-1} Q_{k-1})$ such that**

$$S^N(\nu_{k-1} Q_{k-1}) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^i} \quad \text{where} \quad (X_k^1, \cdots, X_k^N) \text{ are i.i.d.} \sim \nu_{k-1} Q_{k-1}$$

► **Dynamical system on the space of dicrete probability measures $(\nu_k^N)_{0 \leq k \leq n}$**

$$\begin{cases} \nu_0^N = S^N(\nu_0) \\ \nu_k^N = G_k \cdot S^N(\nu_{k-1}^N Q_{k-1}), \quad \text{for all} \quad 1 \leq k \leq n \end{cases}$$

$\Rightarrow$ **One can show that $\nu_n^N \approx q^* \, dx$ [DelMoral]**

Algorithm

- **Initialization : Generate independently**

$$(X_0^1, \cdots, X_0^N) \quad \textbf{i.i.d.} \quad \sim \quad p \quad \textbf{then set} \quad \nu_0^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_0^i}$$

- **Selection : Generate independently**

$$(\tilde{X}_k^1, \cdots, \tilde{X}_k^N) \quad \textbf{i.i.d.} \quad \sim \quad \nu_k^N = \sum_{i=1}^N \omega_k^i \, \delta_{X_k^i}$$

- **Mutation : Generate independently for each** $i \in \{1, \cdots, N\}$**,**

$$X_{k+1}^i \quad \sim \quad Q_k(\tilde{X}_k^i, \cdot)$$

- **Weighting : For each particle** $i \in \{1, \cdots, N\}$**, compute**

$$\omega_{k+1}^i = \frac{G_{k+1}(X_{k+1}^i)}{\sum_{j=1}^N G_{k+1}(X_{k+1}^j)} \quad \textbf{then set} \quad \nu_{k+1}^N = \sum_{i=1}^N \omega_{k+1}^i \, \delta_{X_{k+1}^i}$$

Adaptive choice of the sequence $(G_k)_{0\leq k\leq n}$

[Musso&al01], [Hommem-de-Mello&Rubinstein02], [Cérou&al06]

► The performance of Interacting particle systems is known to deteriorate when the quantities $\frac{\max G_k}{S^N(\nu_{k-1}^N Q_{k-1})(G_k)}$ are big

The idea is then to chose G_k such that $\frac{1}{N}\sum_{i=1}^{N} G_k(X_k^i)$ is not to small

► In our case where $G_k(x) = \mathbf{I}_{\Delta V(x)\leq s_k}$, the threshold s_k is chosen as a r.v. depending on the current particle system and on a parameter $\rho \in (0,1)$:

$$s_k = \inf\left\{ s \quad \textbf{such that} \quad \sum_{i=1}^{N} \mathbf{I}_{\Delta V(X^i)\leq s} \geq \rho N \right\}$$

► This choice of s_k is not prooved to guarantee that the algorithms ends in a finite number of iterations but this point does not seem to be a problem in our simulations

Density estimation

- **At the end of the algorithm, we get** $\nu_n^N \approx q^* \, dx$

But Importance Sampling requires a smooth approximation with density q^N

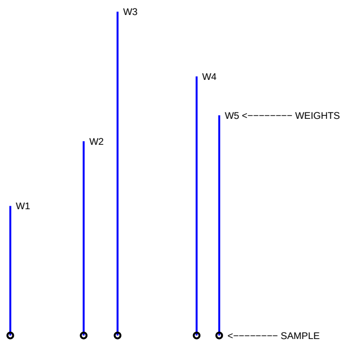
- **Kernel of order 2** $\quad K$

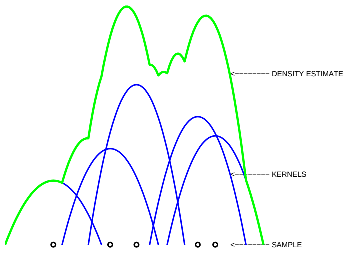
$$K \geq 0 \qquad \int K = 1 \qquad \int x_i \, K = 0 \qquad \int |x_i \, x_j| \, K < \infty$$

- **Rescaled kernel** $\quad K_h \qquad K_h(x) = \frac{1}{h^d} \, K(\frac{x}{h})$

- $$\nu^N = \sum \omega^i \, \delta_{X^i} \xrightarrow[K_h * \cdot]{\text{Density estimation}} q^{N,h} = \sum \omega^i \, K_h(\cdot - X^i)$$

- **Optimal choice of** h $\; => \quad \mathbb{E}\|q^N - q^*\|_1 \leq \frac{C}{N^{\frac{4}{2(d+4)}}}$





Some simulation results : Variance ratio

- Several test cases depending on the form of function $x \mapsto \Delta V(x)$ have been studied : results are all comparable
- X is a d dimensional Gaussian variable and $m = \mathbb{E}_p[\mathbf{I}_{\Delta V(X) \leq s}]$
- Particles $N = 500$ Iterations $n \approx 10$ to 60 Simulations $M = 10\,000$

	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
$m = 10^{-2}$	150 10^{-1}	50	50	30	25
$m = 10^{-3}$	1000 2	300	300	200	140
$m = 10^{-6}$	2.10^5 200	10^5 400	10^5 300	5.10^4 460	2.10^4 480

	$d=6$	$d=7$	$d=8$	$d=9$	$\cdots$	$d=30$
$m = 10^{-2}$	22	14	11	8	$\cdots$	5.10^{-3}
$m = 10^{-3}$	100	70	55	40	$\cdots$	10^{-3}
$m = 10^{-6}$	10^4 250	2.10^3 480	2.10^3 300	4.10^3 300	$\cdots$	1 360

References

► [Glasserman&al00] Glasserman, P. and Heidelberger, P. and Shahabuddin, P. *Reduction Variance techniques for estimating Value at Risk*, *Management Science, Vol. 46, No 10, 2000*.

► [DelMoral&Garnier05] Del Moral, P. and Garnier, J. *Genealogical particle analysis of rare events*, *Annals of Applied Probability, 2005*.

► [Cerou&al06] Cerou, F. and Del Moral, P. and Le Gland, F. and Guyader, P. and Lezaud, H. Topart, *Some recent improvements to importance splitting*, *Proceedings of the 6th International Workshop on Rare Event Simulation, Bamberg, October 9-10, 2006*.

► [Homem-de-Mello&Rubinstein02] Homem-de-Mello, T. and Rubinstein, R.Y. *Estimation of rare event probabilities using cross-entropy*, *Proceedings of the Winter Simulation Conference, 2002*. ► [Musso&al01] Musso, C. and Oudjane, N. and Le Gland, F. , *Improving regularized particle filters*, in Sequential Monte Carlo Methods in Practice, Doucet A& al. editors, *Statistics for Engineering and Information Science, 2001*.