

Using Least Squares Monte Carlo techniques in insurance with R

Sébastien de Valeriola

`sebastien.devaleriola@reacfin.com`



Amsterdam, June 29th 2015

The major difference between Solvency I and Solvency II directives is that the latter imposes a **fair value** valuation of assets and liabilities to insurance companies.

The new standards take an economic view of the balance sheet and require the computation of the **economic capital**, the minimum capital giving the insurance company a 99.5% survival probability over a one-year horizon.

Two approaches exist for the estimation of this economic capital:

- a standard approach (standard formula),
- an internal approach ([possibly partial] internal model).

The latter involves a finer analysis of the risks and requires the **distribution of capital consumption** to be defined over a one-year horizon.

Computing the SCR

In a flat-rate environment, the Solvency Capital Requirement at level $\alpha = 99.5\%$ can be computed as

$$SCR = \operatorname{argmin}_x \left\{ \mathbb{P} \left[AC_0 - \frac{AC_1}{1+r} > x \right] \leq 1 - \alpha \right\},$$

where AC_t is the available capital at time t :

$$AC_t = ANAV_t + X_t + \mathbb{E}_{\mathbb{Q}} \left[\sum_{i=t+1}^T \frac{X_i}{(1+r)^i} \mid Y_s, s \in [0, t] \right],$$

where $ANAV_t$ is the adjusted net asset value, X_t is the profit due to in-force business and Y_t is the market conditions at time t .

The computation of the SCR thus includes two projections: the first one from $t = 0$ to $t = 1$ (in the real world), and the second one from $t = 1$ to $t = T$ (in the risk-neutral world).

Monte Carlo nested simulations

The expectation appearing in the definition of Available Capital is generally **not computable analytically**, due to the very complex interactions between X and Y .

It can however be estimated using Monte Carlo simulations.

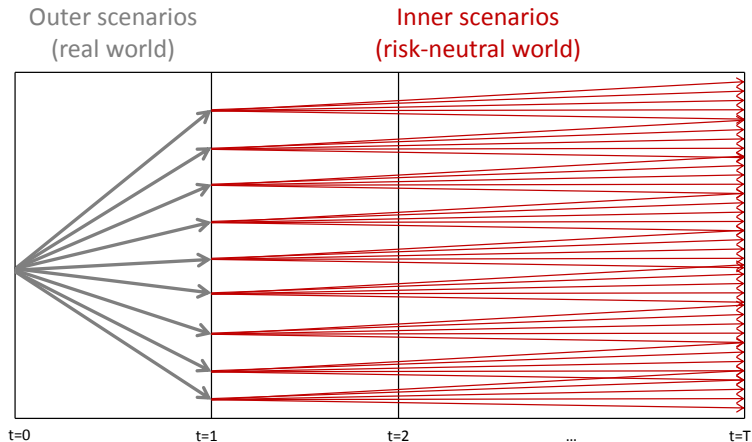
The expectation is then approximated by a sum over stochastic scenarios:

$$\mathbb{E}_{\mathbb{Q}} \left[\sum_{i=t+1}^T \frac{X_i}{(1+r)^t} \mid Y_s, s \in [0, t] \right] \approx \frac{1}{K} \sum_{k=1}^K \left(\sum_{i=t+1}^T \frac{X_i^k}{(1+r)^t} \right).$$

Two sets of stochastic scenarios have to be produced: **outer scenarios** for the evolution of all the variables during first year and **inner scenarios** for the expectation evaluation.

This approach is called the Nested Simulations.

Monte Carlo nested simulations



The Least Squares Monte Carlo approach

Nested Simulations are extremely time- and memory-consuming. Think about a mid-sized insurance portfolio with 100,000 policies, 10,000 outer scenarios and 1,000 inner scenarios... If we can project 1 policy (outer+inner) in 0.001 s., the total would take more than 30 years to process.

An alternative to the nested simulations approach is the Least Squares Monte Carlo approach.

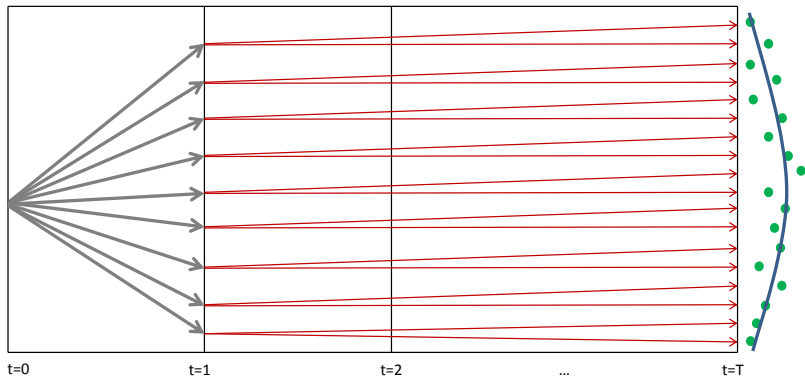
It consists in producing only a small number of inner scenarios and thereafter performing a least squares regression of the obtained values on the market conditions at time $t = 1$.

The expectation is then computed as the average of the predictions given by the fit.

The Least Squares Monte Carlo approach

Outer scenarios
(real world)

LSMC scenarios
(risk-neutral world)



First uses of the LSMC techniques

The Least Squares Monte Carlo techniques have first been used in the valuation of American options by Longstaff and Schwarz in 2001.

The first authors to present an application of the LSMC to the calculation of an insurance economic capital were Bauer, Bermann and Reuss in 2010.

The application of LSMC techniques to financial option pricing has been implemented in R.

→ package `LSMonteCarlo`.

However, such an implementation does not exist for the insurance framework.

The regression of the average LSMC scenarios on the stock values is the core of the technique.

Theoretical justification: the random variable $\mathbb{E}[f(X)|Y] \in L^2$ can be approximated using a truncation of its Hilbert decomposition

$$\mathbb{E}[f(X)|Y] \approx \sum_{i=0}^D c_i g^i \quad c_i \in \mathbb{R}, g^i \in L^2.$$

The orthonormal basis $\{g_0, g_1, \dots\}$ is often chosen among weighted orthonormal polynomials, such as Laguerre, Hermite, Legendre or Chebyshev polynomials.

→ package `orthopolynom`.

We consider N outer scenarios and K LSMC scenarios (K is intended to be rather small).

In order to obtain the available capital at time $t = 1$:

1. Produce N outer scenarios in real world : $s^i (i = 1, \dots, N)$,
2. For each i , produce K LSMC scenarios in risk-neutral world : $s^{i,j} (i = 1, \dots, N ; j = 1, \dots, K)$,
3. Perform a least squares regression of these $s^{i,j}$ on some functions of the market values at $t = 1$,
4. Compute the final expectation as the mean of the predictions given by the model.

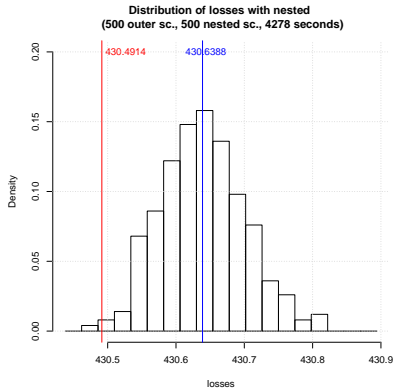
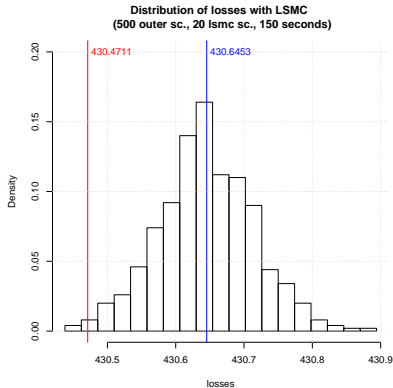
Example of results: toy model

Let us consider a very simple model: a life insurance company sells 10 years-pure endowment contracts (guaranteed rate of 3.5%), with profit share (90% of profits are redistributed).

Its entire assets portfolio is invested in stocks (modelled with a GBM).

We assume that the only source of randomness is the evolution of the company's assets: the interest rate is not stochastic and mortality is not taken into account.

Results



The relative difference between the two quantiles is 0.0047 %.

Of course, real insurers are exposed to more than one risk factor. Several variables are then used as regressors for the conditional expectation.

Such risk factors include:

- Market prices of financial securities,
- Short, medium and long term rates,
- Exchange rates,
- Prices inflation,
- Health costs inflation,
- Lapses,
- Mortality,
- ...

The regression with multiple risk factors

When using multiple risk factors, multidimensional polynomials are often used for the regression.

In this framework, the choice of the variables to include and the degree of the polynomials to use is a major issue.

In real life situations, the number of regressors can be huge:

D x R	1	2	3	4	5	6	7	8
1	1	2	3	4	5	6	7	8
2	2	5	9	14	20	27	35	44
3	3	9	19	34	55	83	119	164
4	4	14	34	69	125	209	329	494
5	5	20	55	125	251	461	791	1286
6	6	27	83	209	461	923	1715	3002
7	7	35	119	329	791	1715	3431	6434
8	8	44	164	494	1286	3002	6434	12869

A lot of econometric techniques exist for the selection of the best model from a specific set of models. Among them,

- Selection (forward),
- Elimination (backward),
- Bidirectional (mixed).

Among the various criteria for significance (R^2 , AIC , BIC , C_p , ...), Bauer et al show that the Mallow criterion C_p is well suited for LSMC, as it is good at handling heteroskedasticity of residuals.

→ package **leaps**.

Choice of the number of simulations and the maximal degree of the polynomials

If all the relevant risk factors are considered,

$$\hat{E}^{LSMC}(N, D) \longrightarrow E = \mathbb{E}_{\mathbb{Q}} \left[\sum_{i=t+1}^T \frac{X_i}{(1+r)^t} \mid Y_s, s \in [0, t] \right]$$

as $N, D \rightarrow \infty$.

We need to quantify the speed of the convergence in order to select values for N and D .

The first idea would be to fix a threshold $\tau > 0$ and keep increasing N and D as long as

$$\left| \hat{E}^{LSMC}(N, D) - E \right| > \tau.$$

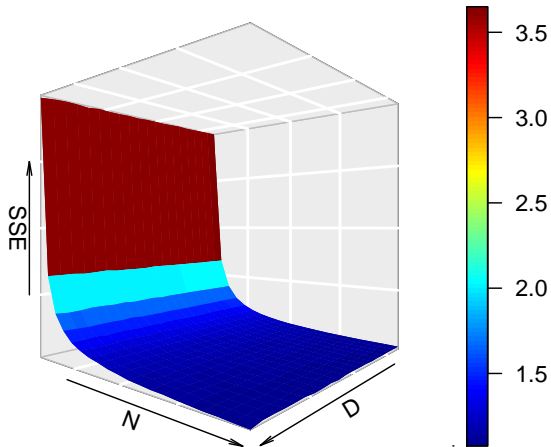
However the limit E is of course unknown.

To overcome this difficulty, the method of the *validation scenarios* is widely used.

It consists in selecting some scenarios (generally as few as 20) by expert judgement based on information obtained on the distribution so as to cover its main segments.

A common technique is to define shocks for every risk factor and cross them to build validation scenarios.

Sum of squared error



As the LSMC techniques aim at reducing the time necessary to compute simulations, other tools with the same purpose are often simultaneously used.

Among them, let us mention

- Variance reduction techniques, such as antithetic variates, control variates, stratified sampling, ...
- Parallel computing (package `parallel` among others),
- Just-in-time compiler (package `compiler`),
- Implementation of some parts of the code in a compiled language (package `Rcpp`),
- ...

Some references:

- F. A. Longstaff and E. S. Schwartz, Valuing American Options by Simulation : A Simple Least Squares Approach, *The Review of Financial Studies*, **14**, 113 (2001).
- E. Clément, D. Lamberton, and P. Protter, An analysis of a least squares regression method for American option pricing, *Finance and Stochastics*, **6**, 4 (2002).
- D. Bauer, D. Bergmann, and A. Reuss, Solvency II and nested simulations. A least squares Monte Carlo approach, In *Proceedings of the 2010 ICA congress* (2010).

Thank you for your attention!