Risk management with Local Least Squares Monte Carlo

D. Hainaut & A. Akbaraly & C. Tossou | DETRALYTICS

5th European
Congress of Actuaries
www.eca2024.org

About the speaker

- **AKBARALY Adnane –** *TAP2 Consultant, DETRALYTICS*
  Graduated with a Master’s degree in Actuarial science from ISFA in Lyon in 2022

- **TOSSOU Comlan –** *TAP2 Consultant, DETRALYTICS*
  Graduated with a Master’s degree in Actuarial science from ENSAE in Paris in 2022, and a statistical engineer degree from ENSAE Dakar in 2020

- Detralytics supports companies in the advancement of actuarial science and help them to solve their challenges. We guarantee to go beyond traditional consulting, by offering a unique combination of academic expertise, deep career and market knowledge

Introduction

The least squares Monte-Carlo method (LSMC) of Longstaff and Schwartz (2001) is a powerful and simple simulation method for pricing path dependent options.

The LSMC is not only useful for pricing but also for managing risk. E.g. Bauer et al. (2012) adapt the LSMC method for computing the required risk capital in the Solvency II framework.

Sensitive point : the LSMC requires a regression model predicting the responses (i.e. discounted CF's) as a function of risk factors (by a polynomial, by a combination of basis functions, by a neural network).

The main contribution of this article is to propose an alternative based on local regressions called the local least square Monte-Carlo, LLSMC.

Introduction

LLSMC in a nutshell:

1. Create clusters of responses the K-means algorithm and next to locally regress them on corresponding risk factors.
2. Fit a logistic regression model that a priori estimates the probability that a combination of risk factors belongs to each cluster.
3. A global regression function is obtained by weighting local models by these probabilities.

Main advantages of the LLSMC:

- Robustness,
- High level of interpretability.

LSMC and risk management

- We consider m risk factors, $\mathbf{X}_t = \left(X_t^{(1)}, ..., X_t^{(m)}\right)_{t\geq 0}$. They drives the value of assets managed by a financial institution.
- The total asset is a function of time and risk factors, $A(t, \mathbf{X}_t)$. The asset pays random cash-flows, C_k^A, at time $(t_k)_{k=0,...,d}$.
- $\mathbb{P}$ and $\mathbb{Q}$ are respectively the real and risk neutral measure.
- Using the cash account $(B_t)_{t\geq 0}$ as numeraire. $A(t, \mathbf{X}_t)$ is given by

$$A(t, \mathbf{X}_t) \quad = \quad a(\mathbf{X}_t) + \mathbb{E}^{\mathbb{Q}}\left(\sum_{k=0}^{d} \frac{B_t}{B_{t_k}} C_k^A \mathbf{1}_{\{t_k \geq t\}} | \mathbf{X}_t\right) \quad (1)$$

where $a(\mathbf{X}_t)$ is directly determined by the value of underlying risk factors.

LSMC and risk management

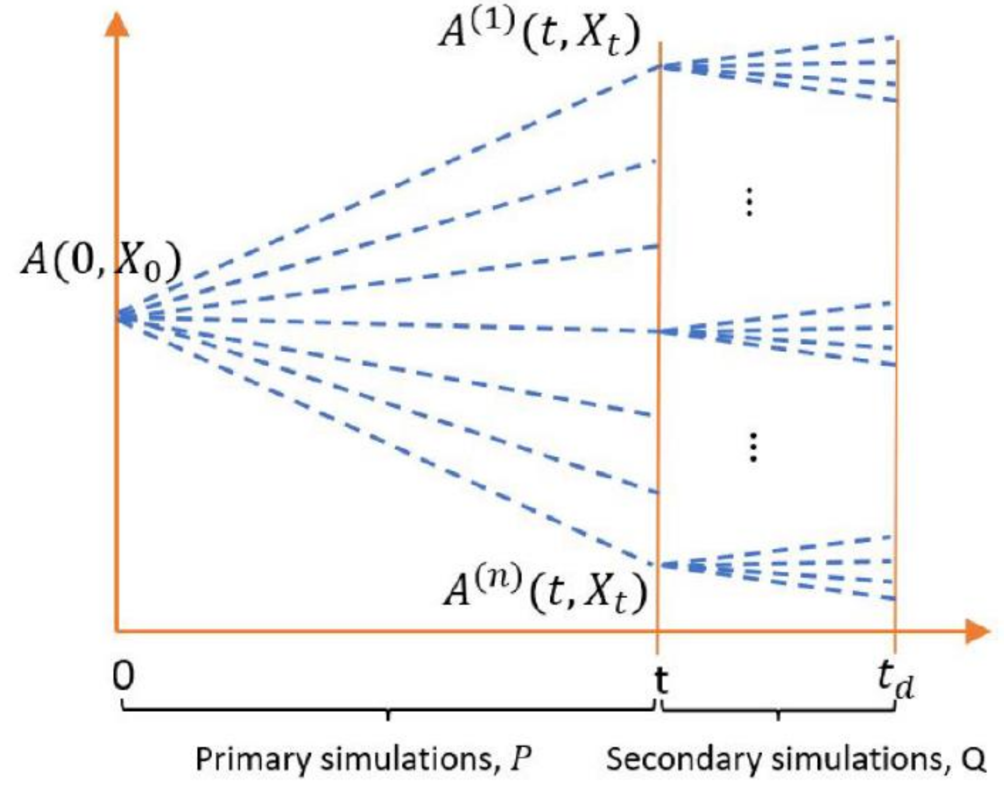
- Let us consider a risk measure denoted by $\rho(.)$. For risk management, we aim to calculate $\rho(A(t, \mathbf{X}_t))$.
- $\rho(.)$ is e.g. the value at risk (VaR) For $\alpha \in (0, 1)$, VaR is defined as

$$VaR_\alpha = \max \{x \in \mathbb{R} : \mathbb{P}(A(t, \mathbf{X}_t) \leq x) \leq \alpha\} ,$$

Problem: computing the risk-neutral expectation (1) is a challenging task because closed-form expressions are usually not available.

- **Solution 1 : simulations in simulations** (but computationally intensive)

LSMC and risk management



For each primary sample path of risk factors (under $\mathbb{P}$), we perform secondary simulations (under $\mathbb{Q}$). The value of $A(t, \mathbf{X}_t)$ is obtained by averaging the sums of discounted cash-flows of secondary scenarios.

LSMC and risk management

- **Solution 2: Least square Monte-Carlo.** Let us denote the discounted CF's

$$Y(t) = \sum_{k=0}^{d} \frac{B_t}{B_{t_k}} C_k^A \mathbf{1}_{\{t_k \geq t\}} ,$$

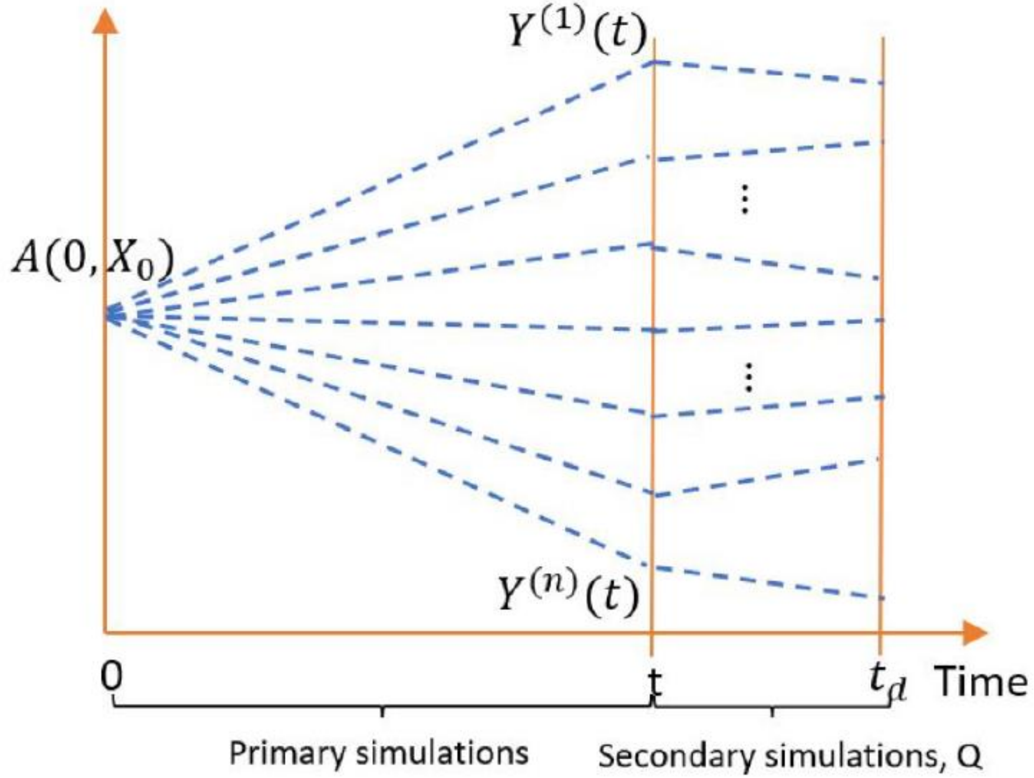
such that $A(t, \mathbf{X}_t) = a(\mathbf{X}_t) + \mathbb{E}^{\mathbb{Q}}(Y(t) \mid \mathbf{X}_t)$.

- $Y(t)$ is called the "response" at time t.
- The LSMC method is based on property that the conditional expectation $\mathbb{E}^{\mathbb{Q}}(Y(t) \mid \mathbf{X}_t)$ is a function $h(\mathbf{X}_t)$ such that

$$h(\mathbf{X}_t) = \arg \min_{h \in \mathcal{B}(,\mathbb{R},\mathbb{R}^m)} \mathbb{E}^{\mathbb{Q}}\left(\left(h(\mathbf{X}_t) - Y(t) \right)^2 \right) .$$

LSMC and risk management

In practice, it means that we only need a single (or a few) secondary simulations under $\mathbb{Q}$ and to approximate $h(\mathbf{X}_t)$.



LSMC and risk management

- The LSMC algorithm consists in simulating a sample denoted by

$$\mathcal{S} = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \ , \qquad (2)$$

of n realizations of $(\mathbf{X}_t, Y(t))$ and in regressing responses on risk factors.

- We recall that $\mathbf{X}_t$ is simulated up to time t under the real measure $\mathbb{P}$ while the response $Y(t)$ is obtained by simulations from t up to t_d, under the risk neutral measure $\mathbb{Q}$.

- Let us denote by $\mathcal{P}_h$ the set of polynomials $\widehat{h}(\mathbf{x})$ of degree d_h approximating $h(\mathbf{x})$. It is estimated by least squares minimization:

$$\widehat{h}(.) \ = \ \arg\min_{\widehat{h}\in\mathcal{P}_h} \left(\sum_{(\mathbf{x}_i, y_i)\in\mathcal{S}} \left(y_i - \widehat{h}(\mathbf{x}_i)\right)^2 \right) . \qquad (3)$$

LSMC and risk management

- After calibration of $\widehat{h}(\mathbf{x})$ $(\approx \mathbb{E}^{\mathbb{Q}}(Y(t) \,|\, \mathbf{X}_t = \mathbf{x}))$, we calculate

$$\widehat{a}_i = a(\mathbf{x}_i) + \widehat{h}(\mathbf{x}_i) \approx A(t, \mathbf{x}_i)\,,$$

  the approximated value of total assets for a given vector of risk factors $\mathbf{X}_t = \mathbf{x}_i$.

- The VaR is then the α-quantile of $\hat{a}_i$'s.
- Main challenge: find a global polynomial approximation which captures non-linearities but does not have an unrealistic behaviour for extreme values of risk factors.

Local Least Square MC

- Instead of fitting a global polynomial predicting responses $(y_i)_{i=1,\ldots,n}$, we partition the domain of $Y(t)$ into K clusters $(\mathcal{Y}_k)_{k=1,\ldots,K}$ (K-means algorithm).
- We next define local regression function

$$h_k(\mathbf{x}) \quad = \quad \mathbb{E}^{\mathbb{Q}}\left(Y(t)\,|\,\mathbf{X}_t = \mathbf{x}\,,\; Y(t) \in \mathcal{Y}_k\right), k = 1, \ldots, K,$$

  equal to the conditional expectation of responses, knowing that $\mathbf{X}_t = \mathbf{x}$ and $Y(t) \in \mathcal{Y}_k$.
- Using standard properties of the conditional expectation, we can rewrite the function $h(\mathbf{x})$ as a weighted sum of $h_k(.)$:

$$\begin{aligned} h(\mathbf{x}) &= \mathbb{E}^{\mathbb{Q}}\left(Y(t)\,|\,\mathbf{X}_t = \mathbf{x}\right) \\ &= \sum_{k=1}^{K} \mathbb{Q}\left(Y(t) \in \mathcal{Y}_k\,|\,\mathbf{X}_t = \mathbf{x}\right) h_k(\mathbf{x})\,. \end{aligned}$$

Local Least Square MC

- Based on this decomposition, we approximate the K unknown functions $h_k(.)$ by polynomial regression $\widehat{h}_k$ of $Y(t) \in \mathcal{Y}_k$ on risk factors.

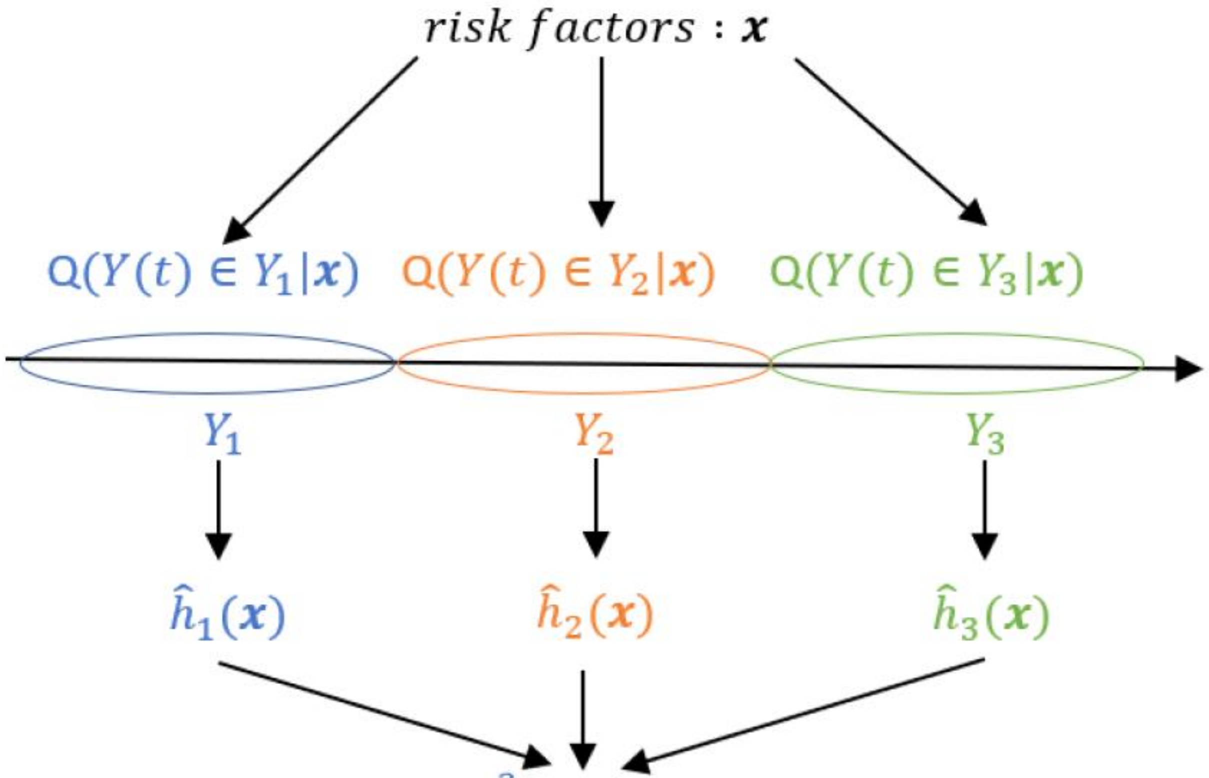
$$\widehat{h}_k = \min_{\widehat{h}_k \in \mathcal{P}_h} \left(\sum_{i \,|\, y_i \in \mathcal{Y}_k} \left(y_i - \widehat{h}_k(\mathbf{x}_i) \right)^2 \right) .$$

- In a second step, we use a multinomial logistic regression to estimate the probabilities $\mathbb{Q}\left(Y(t) \in \mathcal{Y}_k \,|\, \mathbf{X}_t\right)$ for $k = 1, \ldots, K$.

$$\mathbb{Q}\left(Y(t) \in \mathcal{Y}_k \,|\, \mathbf{X}_t = \mathbf{x}\right) = \begin{cases} \frac{e^{-\widehat{\gamma}_k(\mathbf{x})}}{1+\sum_{j=2}^{K} e^{-\widehat{\gamma}_j(\mathbf{x})}} & k = 2, \ldots, K\,, \\ \frac{1}{1+\sum_{j=2}^{K} e^{-\widehat{\gamma}_j(\mathbf{x})}} & k = 1\,, \end{cases}$$

where $\widehat{\gamma}_k(\mathbf{x})$ is a polynomial of risk factors.

Local Least Square MC



$$\hat{a} = a(\boldsymbol{x}) + \sum_{j=1}^{3} Q(Y(t) \in Y_j|\boldsymbol{x})\, \hat{h}_j(\boldsymbol{x})$$

Local Least Square MC

- It may appear counterintuitive to partition the dataset using responses instead of risk factors. Two reasons motivate this:

1. local regressions based on hard clusters of risk factors generate discontinuities in predicted $\mathbb{E}^{\mathbb{Q}}(Y(t) \mid \mathbf{X}_t)$ on borders of clusters.

2. This prevents to observe the Simpson's paradox. This is when a trend appears in several groups of data but disappears or reverses when the groups are combined.

Case study 1: Heston model

- Calculation of the VaR in 1 year of a butterfly option (maturity 2 years) in the Heston model with the LSMC and LLSMC.
- The stock price, noted $(S_t)_{t\geq 0}$, is ruled by a Brownian diffusion with a stochastic variance, $(V_t)_{t\geq 0}$:

$$\begin{cases} dS_t &= \mu\, S_t\, dt + S_t\sqrt{V_t}\left(\rho dW_t^v + \sqrt{1-\rho^2}dW_t^s\right), \\ dV_t &= \kappa\,(\gamma - V_t)\, dt + \sigma\sqrt{V_t}dW_t^v\,. \end{cases}$$

- Risk factors: the normed stock price and volatility.

This case study is available on our Detranote and will not be presented here because it's a financial case study. The case study 2 is an actuarial case study

Case study 2: Life insurance

- We compare the performance of LSMC and LLSMC for assessing the risk of a participating pure endowment
- The stock price indice, the interest rate and the force of mortality are respectively denoted by $(S_t)_{t\geq 0}$, $(r_t)_{t\geq 0}$ and $(\mu_{x+t})_{t\geq 0}$:

$$\begin{pmatrix} dS_t \\ dr_t \\ d\mu_{x+t} \end{pmatrix} = \begin{pmatrix} \mu\, S_t \\ \kappa_r\,(\gamma_r(t) - r_t) \\ \kappa_\mu\,(\gamma_x(t) - \mu_{x+t}) \end{pmatrix} dt + \begin{pmatrix} S_t\sigma_S & 0 & 0 \\ 0 & \sigma_r & 0 \\ 0 & 0 & \sigma_x(t) \end{pmatrix} \Sigma \begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \\ dW_t^{(3)} \end{pmatrix}$$

- The matrix Σ is the (upper) Choleski decomposition of the correlation matrix.

Case study 2: Life insurance

- We consider a contract subscribed by a x- years old individual that promises at expiry (date T) the maximum between a capital C_T and the value of the stock indice S_T, in case of survival.
- The benefit is nevertheless upper bounded by C_M.
- If we denote by $\tau \in \mathbb{R}^+$, the random time of insured's death, the value of such a policy is

$$V_t = \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} \mathbf{1}_{\{\tau \geq T\}} \left(C_T + (S_T - C_T)_+ - (S_T - C_M)_+ \right) | \mathcal{F}_t \right)$$

- This contract admits a closed form-expression for its price (but long...see Detranote for details)

Case study 2: Life insurance

- We fit a Nelson-Siegel model to the Belgian state yield curve on the 23/11/22.
- Initial survival probabilities are described by a Makeham's model adjusted to male Belgian mortality rates.
- Other market parameters are estimated from time series of the Belgian stock index BEL 20 and of the 1 year Belgian state yield from the 26/11/10 to the 23/11/22.
- As we do not have enough data, the correlations $\rho_{S\,\mu}$ and $\rho_{r\,\mu}$ are set to -5% and 0%.

Case study 2: Life insurance

Parameters			
μ	0.04642	σ_S	0.18470
κ_r	0.20482	σ_r	0.00774
ρ_{Sr}	-0.03957	r_0	0.0235
α	8.5277e-7	β	0.11094
κ_μ	0.83925	μ_0	3.325e-03
$\rho_{S\mu}$	-0.05000	$\rho_{r\mu}$	0.00000
t	5 years	T	10 years
S_0	100	C_T	100
x	50	C_M	$100(1+3\%)^{10}$

We perform 10000 primary simulations.

Case study 2: Life insurance

- R^2, MSE, MSE($\mathcal{V}$) of regressions of Y_t on $\mathbf{X}_t$ in the LSMC model. $\sqrt{\text{EMSE}}$ is the MSE valued with analytical prices. d.f. is the number of parameters.

d_h	R^2	$\sqrt{\text{MSE}(\mathcal{V})}$	$\sqrt{\text{MSE}}$	$\sqrt{\text{EMSE}}$	d.f.
2	0.3767	3.02	11.34	1.59	10
3	0.3855	2.47	11.27	1.04	20
4	0.3868	3.94	11.28	1.12	35
5	0.3912	4.85	11.26	1.02	56
6	0.3947	6.47	11.26	1.10	84

- The validation set counts 1000 triplets of risk factors. We consider combinations of 10 empirical quantiles of risk factors for probabilities from 1% to 5% and from 95% to 99% by step of 1%.

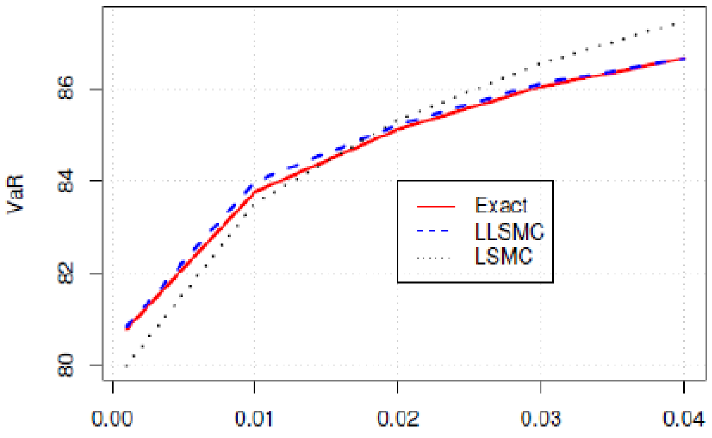
Case study 2: Life insurance

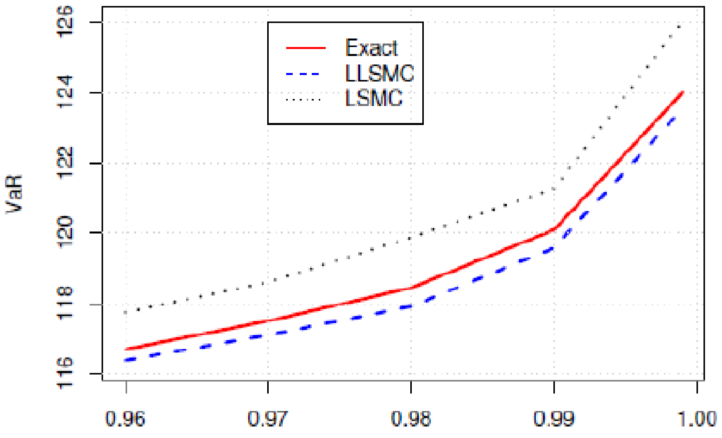
K	d_γ	d_h	R^2	$\sqrt{\text{MSE}(\mathcal{V})}$	$\sqrt{\text{MSE}}$	$\sqrt{\text{EMSE}}$	d.f.	R^2_{loc}
2	3	2	0.39	0.69	11.24	0.67	40	0.87
3	3	2	0.39	0.79	11.27	0.55	70	0.93
4	2	2	0.38	0.87	11.32	0.77	70	0.95
5	2	2	0.38	0.91	11.34	0.75	90	0.96
3	2	2	0.38	0.93	11.29	0.76	50	0.93

While the MSE's of LLSMC and LSMC are comparable, the $\sqrt{\text{MSE}(\mathcal{V})}$ and $\sqrt{\text{EMSE}}$ are clearly reduced by the LLSMC.

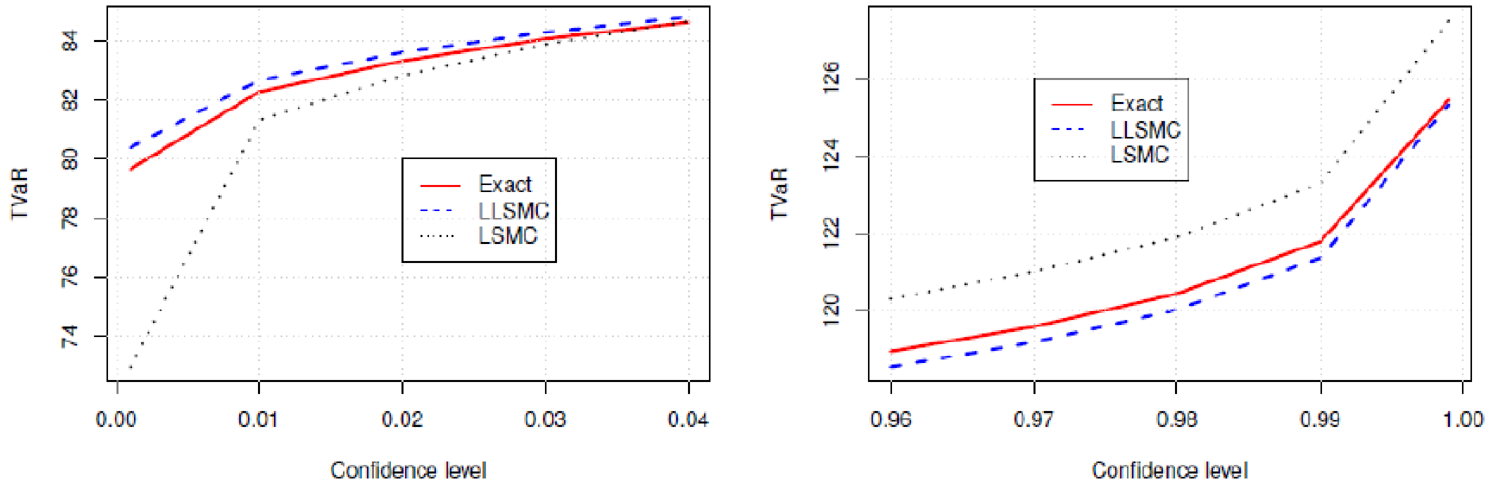
Case study 2: Life insurance

As the contract can be valued analytically, we compare exact VaR and TVaR to these computed with LSMC (order 3) and LLSMC ($K = 3$, $d_h = 3$ and $d_\gamma = 2$).





Case study 2: Life insurance



The LLSMC clearly yields VaR and TVaR estimates closer to the the exact ones.

Conclusions

- The LLSMC combines local and logistic regressions. It presents several interesting features:
  - Computational cost is not prohibitive compared to LSMC
  - Robustness and easily interpretable
  - Better predictions of prices for extreme values of risk factors
- Warning: be cautious when you design a LSMC or a LSMC model: R^2 (or Mallow's Cp), MSE are poor metrics. Working with a validation sample is more than recommended.
- Full research report : https://detralytics.com/wp-content/uploads/2023/02/DetraNote-2023-1_Risk-management-with-local-least-squares-Monte-Carlo.pdf

Thank you

LinkedIn :

Adnane AKBARALY | Comlan Rodrigue TOSSOU

Detralytics :

https://detralytics.com/