Estimating a Probability of Failure with Adaptive Multilevel Splitting

Lucie Bernard
LPSM - Univ. Pierre et Marie Curie, Paris, France ; LMPT - Univ. François Rabelais, Tours, France ; STMicroelectronics, Tours, France

Supervisor(s): A. Guyader (UPMC), P. Leduc (STMicroelectronics), F. Malrieu (Univ. Tours)

Ph.D. expected duration: October 2015 - September 2018

Adress: UPMC, 4 Place Jussieu, 75005 Paris

Email: lucie.bernard@upmc.fr

Abstract:

Given a hyperrectangle $A = [a, b] = \prod_{i=1}^{d} [a_i, b_i]$, we consider the stochastic model $Y = g(\mathbf{X})$, where $\mathbf{X}$ is a random vector with support in A and $g : A \to \mathbb{R}$ is a black-box function that can be computed at any point $\mathbf{x} \in A$ but which is costly to evaluate. We typically assume that $\mathbf{X}$ has a density which is known up to a normalizing constant, so that Metropolis-Hastings (MH) techniques can be applied to simulate from the law of $\mathbf{X}$ or from its restriction to a subset of A.

Next, given a threshold $T \in \mathbb{R}$, we want to estimate the small probability of failure p defined as:

$$p = \mathbb{P}(Y > T) = \mathbb{P}\left(g(\mathbf{X}) > T\right) = \mathbf{P}_{\mathbf{X}}(\Gamma),$$

where $\Gamma = \{\mathbf{x} \in A,\ g(\mathbf{x}) > T\}$. A naive Monte Carlo method consists in simulating n i.i.d. realizations $\mathbf{X}_1, \ldots, \mathbf{X}_n$ from $\mathbf{P}_{\mathbf{X}}$ and set $\widehat{p} = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{g(\mathbf{X}_i) > T}$. Since p is small and any evaluation of g is very expensive, this method is clearly intractable in the present context.

The iterative algorithm we propose assumes that g is Lipschitz with a known constant $L > 0$, which means that for all $(\mathbf{x}, \mathbf{x}') \in A \times A$, we have $|g(\mathbf{x}) - g(\mathbf{x}')| \leq L \left\|\mathbf{x} - \mathbf{x}'\right\|_\infty$. This assumption is in fact the same as in [3] where the objective is to find the minimum of a function g.

Indeed, our method combines two main ingredients: the fact that g is Lipschitz and the ability to simulate according to the restriction of the law of $\mathbf{X}$ to any subset of A via an Adaptive Multilevel Splitting (AMS) method, also called Subset Simulation (see e.g. [1, 2]).

Let n be the maximal number of calls to g. We propose an iterative procedure that determines at each step $k = 1, \ldots, n$ the point $\mathbf{x}_k$ to query g and returns an estimation $\widehat{p}_k$ of p. Let us stress that n is the critical parameter of the problem since any computation of g is very costly.

For $k = 1, 2, \ldots n$, we denote by $\tilde{A}_{k-1}$ the union of all subdomains (hyperrectangles) of A where g is likely to exceed T. We start with $\tilde{A}_0 = A$. Let A_k be a subdomain of $\tilde{A}_{k-1}$, we set $A_k = \left[a^k, b^k\right] = \prod_{i=1}^{d} \left[a_i^k, b_i^k\right]$. We initialize $A_1 = A$. The probability that $\mathbf{X}$ belongs to any subdomain A_k of $\tilde{A}_{k-1}$ is estimated on the fly, hence in particular $\mathbb{P}(\mathbf{X} \in A_1) = 1$. We also defined $\tilde{\Gamma}_k$ an increasing sequence of subdomains of A contained in Γ and initialize $\tilde{\Gamma}_0 = \emptyset$.

For $k = 1, 2, \ldots n$, (see also Figure 1)

- Find the hyperrectangle A_k of $\tilde{A}_{k-1}$ with the highest probability estimate.
- Evaluate g at the center $\mathbf{x}_k = \left((a_i^k + b_i^k)/2\right)_{1 \leq i \leq d}$ of A_k. Under the Lipschitz condition:

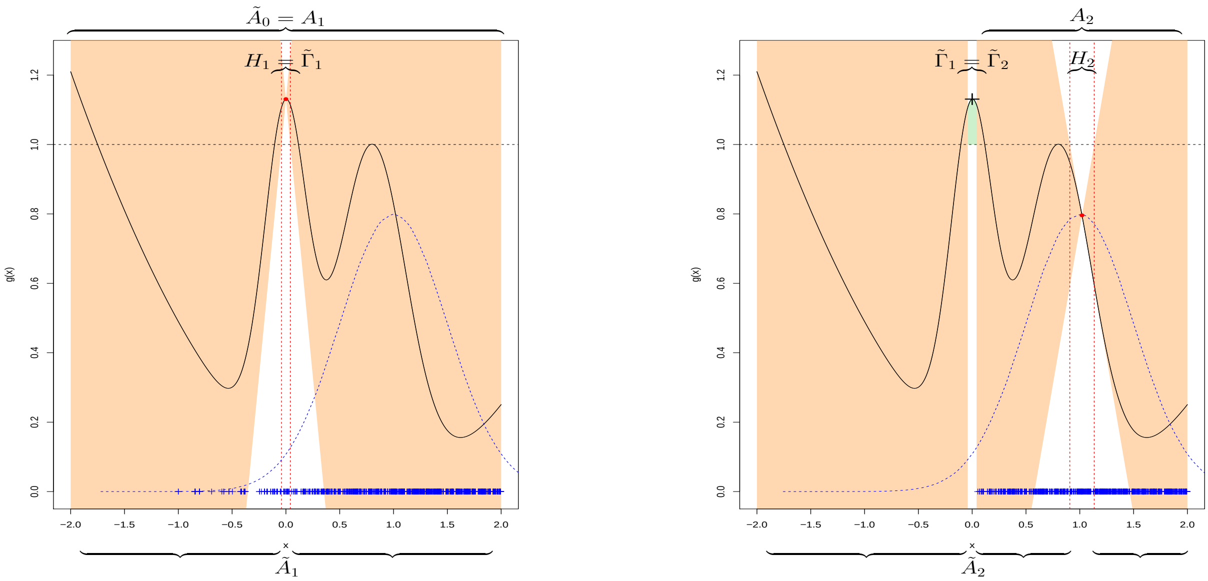

Figure 1: Left: step $k = 1$; $A_1 = [-2, 2]$; $x_1 = 0$. Right: step $k = 2$; $A_2 = [0.04, 2]$; $x_2 = 1.02$. The line - - - is the density of $\mathbf{X}$; $+ + +$ is the sample $\mathbf{X}_1, \ldots \mathbf{X}_N \backsim \mathcal{L}(\mathbf{X}|\mathbf{X} \in A_k)$.

  - If $g(\mathbf{x}_k) + L \left\| b^k - \mathbf{x}_k \right\|_\infty \leq T$, it follows that for all $\mathbf{x} \in A_k$, $g(\mathbf{x}) \leq T$, hence the update $\tilde{A}_k = \tilde{A}_{k-1} \setminus A_k$ and $\tilde{\Gamma}_k = \tilde{\Gamma}_{k-1}$. No other evaluation of g in A_k is necessary.
  - Else, if $g(\mathbf{x}_k) - L \left\| b^k - \mathbf{x}_k \right\|_\infty > T$, it follows that for all $\mathbf{x} \in A_k$, $g(\mathbf{x}) > T$, hence the update $\tilde{A}_k = \tilde{A}_{k-1} \setminus A_k$ and $\tilde{\Gamma}_k = \tilde{\Gamma}_{k-1} \cup A_k$. No other evaluation of g in A_k is necessary.
  - Else there exists a hyperrectangle H_k in A_k, with center $\mathbf{x}_k$, where g is always above or below the threshold, depending on $g(\mathbf{x}_k)$. Indeed, if $g(\mathbf{x}_k) > T$ (resp. $g(\mathbf{x}_k) \leq T$), then for all $\mathbf{x} \in H_k$, $g(\mathbf{x}) > T$ and $\tilde{\Gamma}_k = \tilde{\Gamma}_{k-1} \cup H_k$ (resp. $g(\mathbf{x}) < T$ and $\tilde{\Gamma}_k = \tilde{\Gamma}_{k-1}$). In other words, there exist hyperrectangles of A_k where g is still likely to exceed T, so that $\tilde{A}_k = \tilde{A}_{k-1} \setminus H_k$.

- For H_k and for all hyperrectangle D of A_k where g may exceed T, use an AMS method to estimate $\mathbb{P}(\mathbf{X} \in D)$. First, draw a Monte Carlo sample $\mathbf{X}_1, \ldots, \mathbf{X}_N \backsim \mathcal{L}(\mathbf{X}|\mathbf{X} \in A_k)$ for example by applying MH algorithm. Then, estimate $\mathbb{P}(\mathbf{X} \in D) = \mathbb{P}(\mathbf{X} \in D|\mathbf{X} \in A_k)\mathbb{P}(\mathbf{X} \in A_k)$, where $\mathbb{P}(\mathbf{X} \in A_k)$ was estimated before. Now, any subset of $\tilde{A}_k$ and $\tilde{\Gamma}_k$ is associated with a probability.
- Update the estimation $\widehat{p}_k^N$ of p as follows: $\widehat{p}_k^N = \widehat{p}_k^N(\tilde{A}_k) + \widehat{p}_k^N(\tilde{\Gamma}_k)$, where $\widehat{p}_k^N(\tilde{A}_k)$ and $\widehat{p}_k^N(\tilde{\Gamma}_k)$ are the estimations of $\mathbb{P}(\mathbf{X} \in \tilde{A}_k)$ and $\mathbb{P}(\mathbf{X} \in \tilde{\Gamma}_k)$, that are computed on the fly. Note that at the end of step n, one can refine $\widehat{p}_n^N$ by running again the algorithm with a larger N but with the same sets $\tilde{A}_k$ and $\tilde{\Gamma}_k$, $1 \leq k \leq n$.

As a conclusion, our algorithm provides a closer and closer upper-bound of p. The precision of this estimation only depends on the budget n and the size N of the Monte Carlo sample.

References

[1] S.K. Au and J.L. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Journal of Probabilistic Engineering Mechanics*, Vol. 16:263–277, 2001.

[2] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808, 2012.

[3] A. Cohen, R. Devore, G. Petrova, and P. Wojtaszczyk. Finding the minimum of a function. *Methods Appl. Anal.*, 20(4):365–381, 2013.

Short biography – I graduated from University of Rennes 1 in 2014 with a Master 2 in Statistical research. I started my PhD at STMicroelectronics in October 2015 which is about developing tools to estimate failure probabilities.