Using Poisson processes for rare event simulation

C. Walter
Université Paris Diderot

Supervisor(s): Prof. Josselin Garnier (Université Paris Diderot) and Gilles Defaux (CEA)
Ph.D. expected duration: 2013-2016
Address: CEA, DAM, DIF, F-91297 Arpajon Cedex, France
Email: clement.walter@cea.fr

Abstract:

In the context of reliability analysis, complex models are used to determine failure and safety mode of a given installation. This is determined by a threshold not to be overpassed. These models are often very time consuming and no analytical expression is available. Thus, the evaluation of a failure probability given a threshold or the estimation of a quantile given a targeted probability cannot be done by usual computation tools. Among other approaches we focus here on sequential Monte-Carlo methods [3, 6], also called Subset Simulation, which allow to decrease the number of necessary calls to the model; in this context [6] have shown that the optimal algorithm works with limit subsets, i.e. fixing the current threshold at the minimum of the working population. Another approach is to use the given computational budget to fit a surrogate model and then to use it instead of the real model to evaluate quantities of interest [5, 4, 1]. However optimal sequential Monte-Carlo suffers from the fact that it disables parallel computation of the algorithm while still requiring a lot of calls to the model. On the other hand even though the surrogate models are cheap compared to the original one, approaches using a naive Monte Carlo method or an iid sampling [4, 5] may fail because of the huge size of the required population when \( p \ll 1 \) (say \( p < 10^{-7} \)). To circumvent this limitation some researchers have tried to combine both strategies [2, 7] but optimal settings are still under investigation.

We present here a new approach to these methods in terms of a Poisson process related to any real-valued random variable with continuous cdf [9]. Indeed, let us consider a real-valued random variable \( X \), (for instance \( X = g(U) \) can be the output of a complex deterministic code \( g \) with random input parameters \( U \)). Then the Markov chain such that each state is greater than the previous one and simulated according to the truncated distribution: \( X_{n+1} \sim \mu_X(\cdot|X > X_n) \) is a Poisson process with intensity measure \( \nu((-\infty, x]) = -\log(P[X > x]) \). Practically speaking, this means that one requires on average \( \log 1/p \) simulations to get a failing sample for the computer code. This is to be compared with \( 1/p \) in the case of an iid sampling. Then from the simulation of several Markov chains (possibly in parallel) we are able to find back the optimal Multilevel Splitting estimator which appears as an application of the Lehmann-Scheffé theorem [8, 10]. If the deterministic code \( g \) is replaced by a kriging-based metamodel \( \xi \), then the random quantity of interest becomes \( P[\xi(U) > q] \) and one seeks for estimating its conditional expectation given the data. We also apply the Poisson process framework to this case: more precisely we manage to estimate the conditional expectation as well as the SUR criteria [1] with the simulation of several Poisson processes, thus limiting drastically the size of the working population and the computational cost of SUR criteria. These results are illustrated on usual test-cases as well as on a numerical code from CEA simulating a spherical tank under pressure.
References


Short biography – Clément Walter, 26. My PhD thesis is hosted by the Laboratoire de Probabilités et Modèles Aléatoires at Université Paris Diderot and funded by the CEA. Before starting my PhD in November 2013 I studied at Mines ParisTech where I got specialized in Geostatistics.