Computational Statistics and Data Analysis 80 (2014) 153–166

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Bounding rare event probabilities in computer experiments

Yves Auffray[a], Pierre Barbillon[b,c,*], Jean-Michel Marin[d]

[a] Dassault Aviation, France
[b] AgroParisTech / UMR INRA MIA, France
[c] INRA UMR 518, France
[d] Institut de Mathématiques et Modélisation de Montpellier, Université Montpellier 2, France

ARTICLE INFO

Article history:
Received 21 March 2013
Received in revised form 25 June 2014
Accepted 28 June 2014
Available online 5 July 2014

Keywords:
Computer experiments
Rare events
Kriging
Importance sampling
Bayesian estimates
Risk assessment with fighter aircraft

ABSTRACT

Bounding probabilities of rare events in the context of computer experiments is an important concern in reliability studies. These rare events depend on the output of a physical model with random input variables. Since the model is only known through an expensive black box function, standard efficient Monte Carlo methods designed for rare events cannot be used. That is why a strategy based on importance sampling methods is proposed. This strategy relies on Kriging meta-modeling and manages to achieve sharp upper confidence bounds on the rare events probabilities. The variability due to the Kriging meta-modeling step is properly taken into account. The proposed methodology is applied to an artificial example and compared with more standard Bayesian bounds. Eventually, a challenging real case is analyzed. It consists in finding an upper bound for the probability that the trajectory of an airborne load will collide with the aircraft that released it.

1. Introduction

When studying a complex system, computer experiments that describe its behavior are more and more a surrogate to physical experiments since those physical experiments are often too costly or even impracticable. Computer experiments are used in many fields: chemical kinetics (Sacks et al., 1989b), radiological protection (Kennedy and O'Hagan, 2000), particle physics (Higdon et al., 2004), thermodynamics (Higdon et al., 2008; Liu et al., 2008), hydrogeologic (Marrel et al., 2008; Reichert et al., 2011), climate science (Rougier, 2007), food safety (Johnson et al., 2011).... A computer experiment (Welch et al., 1992; Koehler and Owen, 1996) consists in an evaluation of a black box function which describes a physical model,

$$y = f(\mathbf{x}), \tag{1.1}$$

where $y \in \mathbb{R}$ and $\mathbf{x} \in E$, E being a compact subset of $\mathbb{R}^d$. In this paper, we shall assume that the function f is deterministic, each of its evaluation is time-consuming and that only N evaluations of f can be performed.

Rare events are a major concern in the reliability of complex systems (Heidelberg, 1995; Shahabuddin, 1995). For instance, a company has to warrant that the failures of its products are very rare and often that the probability of failure is smaller than a probability level fixed by the law or by customers. The risk assessment can be performed thanks to the computer code where the only source of uncertainty is due to its inputs. Two main types of uncertainties affect the inputs: environmental uncertainty and lack of precision. Some inputs deal with the environmental context such as temperature or humidity. Since these inputs cannot be fixed in the real world, a probability distribution models their variations. When

* Correspondence to: 16 rue Claude Bernard, 75005 Paris, France.
E-mail address: pierre.barbillon@agroparistech.fr (P. Barbillon).

http://dx.doi.org/10.1016/j.csda.2014.06.023

inputs are controllable (set by the user), they can still suffer from an uncertainty on their measure and that is why they are also modeled by a random variable. We denote by $\mathbf{X}$ the random vector of inputs. Its distribution is assumed to be known or at least such that we can simulate realizations of this vector. Hence, the uncertainties on inputs propagate to the output of the computer code: $Y = f(\mathbf{X})$ is also a random variable. In this paper, the failure occurs when this output is smaller than a limit threshold. Our goal is to propose an upper bound for this probability:

$$\pi_\rho = \mathbb{P}(f(\mathbf{X}) < \rho) = \mathbb{P}(\mathbf{X} \in R_\rho) = \mathbb{P}_{\mathbf{X}}(R_\rho),$$

where R_ρ is defined as the following subset of E: $R_\rho = \{\mathbf{x} : f(\mathbf{x}) < \rho\}$ where $\rho \in \mathbb{R}$ is a given threshold.

We are more interested in a sharp upper bound than in an estimate of the probability. Indeed, if the bound is found to be small enough, the complex system can then be declared to be safe and to meet legal or industrial requirements.

1.1. Crude Monte Carlo approach

This approach is presented to illustrate the need for a more efficient methodology to fit the constraint of a limited number of possible runs of the code. The Monte Carlo estimator of π_ρ is obtained as:

$$\hat{\pi}_{\rho,N} = \frac{\Gamma(f, \mathbf{X}_{1:N}, \rho)}{N}, \tag{1.2}$$

where $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ is defined by

$$\Gamma(f, \mathbf{X}_{1:N}, \rho) = \sum_{i=1}^{N} \mathbb{I}_{(-\infty,\rho)}\big(f(\mathbf{X}_i)\big), \tag{1.3}$$

where $\mathbb{I}_{(-\infty,\rho)} : E \to \mathbb{R}$ is the indicator function of the set $(-\infty, \rho)$ defined as

$$\mathbb{I}_{(-\infty,\rho)}(z) = \begin{cases} 1 & \text{if } z \in (-\infty, \rho) \\ 0 & \text{otherwise,} \end{cases}$$

and $\mathbf{X}_{1:N} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ is an N-sample of random variables with the same distribution as $\mathbf{X}$.

The expectation and variance of $\hat{\pi}_{\rho,N}$ are:

$$\mathbb{E}(\hat{\pi}_{\rho,N}) = \mathbb{P}(\mathbf{X} \in R_\rho) = \pi_\rho, \qquad \mathbb{V}(\hat{\pi}_{\rho,N}) = \frac{1}{N}\pi_\rho(1 - \pi_\rho).$$

Since $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ has a Binomial distribution with parameters N and π_ρ, an exact confidence upper bound on π_ρ:

$$\mathbb{P}\Big(\pi_\rho \le b\left(\Gamma(f, \mathbf{X}_{1:N}, \rho), N, \alpha\right)\Big) \ge 1 - \alpha,$$

is available.

Indeed, let T be a random variable having a Binomial distribution with parameters N and p. Then, for any real number $\alpha \in [0, 1]$, we can show that the upper confidence bound b on p:

$$\mathbb{P}_T\Big(p \le b(T, N, \alpha)\Big) \ge 1 - \alpha,$$

is such that:

$$\begin{cases} b = 1 & \text{if } T = N \\ b \text{ is the solution of equation } \displaystyle\sum_{k=0}^{T} \binom{N}{k} b^k (1-b)^{N-k} = \alpha & \text{otherwise.} \end{cases} \tag{1.4}$$

A closed form expression for this upper bound is not available but it is tractable.

In the case where $\Gamma(f, \mathbf{X}_{1:N}, \rho) = 0$, which happens with probability $(1 - \pi_\rho)^N$, the $(1 - \alpha)$-confidence interval is $[0, 1 - (\alpha)^{1/N}]$. As an example, if the realization of $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ is equal to 0, an upper confidence bound at level 0.9, $\pi_\rho \le 10^{-5}$ can be warranted only if more than 230,000 calls to f are performed. When the purpose is to assess the reliability of a system under the constraint of a limited number of calls to f, there is a need for a sharper upper bound on π_ρ. Since Monte Carlo estimation works better for frequent events, the first idea is to change the crude scheme in such a manner that the event becomes less rare. It is what importance sampling and splitting methods schemes try to achieve. For example, L'Ecuyer et al. (2007) showed that randomized quasi-Monte Carlo can be used jointly with splitting and/or importance sampling. By analyzing a rare event as a cascade of intermediate less rare events, Del Moral and Garnier (2005) developed a genealogical particle system approach to explore the space of inputs E. Cérou and Guyader (2007a,b) proposed an adaptive multilevel splitting also based on particle systems. An adaptive directional sampling method is presented by Munoz Zuniga et al. (2011) to accelerate the Monte Carlo simulation method. However, these methods may still need too many calls to f and the importance distribution is hard to set for an importance sampling method.