Reliability-oriented sensitivity analysis with multiple importance sampling

Mboko, Jonathan
Morio, Jérôme
Prosvirnova, Tatiana
Seguin, Christel
Fédération ENAC ISAE-SUPAERO ONERA, ONERA/DTIS, Université de Toulouse, F-31000 Toulouse, France

Chaudemar, Jean-Charles
Fédération ENAC ISAE-SUPAERO ONERA, ISAE-SUPAERO/DISC, Université de Toulouse, F-31000 Toulouse, France

Critical systems demand strong guaranties of safety throughout their mission. According to the type of critical system considered, these guaranties are expressed as the fulfillment of quantitative and/or qualitative requirements. One of the common key quantitative requirement is to ensure that the probability of failure P_f of the system before the end of its mission is inferior to a target threshold (typically $P_f \leq 10^{-4}$ for the aerospace applications we consider). The critical system is here represented by a numerical input-output model with random inputs. The system failure is associated to an output threshold exceedance. Reliability-oriented sensitivity analysis [5] aims at evaluating the sensitivity of the inputs on the output failure. In this work, we are more precisely interested in estimating the influence of the input distribution parameters on the failure probability with a variance-based approach through the estimation of Sobol indices.

The input distribution parameters are usually fixed in numerical models, with the source of uncertainties in the models limited to the known input distributions. In reality the parameters values that best represent the system behavior may be unknown. This lack of knowledge about parameters values constitutes another level of uncertainty in the modelling of the system. Quantifying how much these uncertainties affect the probability of failure P_f can help us identify which input distribution parameters should be precisely estimated for a better estimation of P_f. With the introduction of these parameteric uncertainties, naive estimation of the Sobol indices becomes very expensive, needing many calls to the model to obtain input-output samples. We propose a method to estimate the Sobol indices with adaptive enrichment of the samples.

We consider a numerical model $\mathcal{M}$ as a deterministic black-box, function of a random vector $\mathbf{X}$ of d independent random inputs with a real output $\mathcal{M}(\mathbf{X})$. $\mathbf{X}$ is characterized by a probability distribution function $f_{\mathbf{X}|\boldsymbol{\theta}}$ where $\boldsymbol{\theta}$ is a distribution parameter vector. The failure event is represented by the variable $\mathbb{1}_{\mathcal{M}(\mathbf{X})\leq T}$ with T the threshold characterizing the failure event. The quantity of interest is $P_f(\boldsymbol{\theta}) = P(\mathcal{M}(\mathbf{X}) \leq T)$. When $\boldsymbol{\theta}$ is fixed, $P_f(\boldsymbol{\theta})$ is an unknown deterministic quantity. To represent epistemic uncertainty, the variability of $\boldsymbol{\theta}$ is modeled with a continuous random variable $\boldsymbol{\Theta}$. The failure probability $P_f(\boldsymbol{\Theta})$ becomes a random variable. The Sobol indices on $P_f(\boldsymbol{\Theta})$ associated to the independant components of $\boldsymbol{\Theta}$ can be estimated with the pick-freeze estimator [2], based on an iid N-sample $(\boldsymbol{\Theta}_i)_{i=1,\ldots,N}$ from $\boldsymbol{\Theta}$. However two difficulties arise for the computation of this estimator. First, a high number of accurate probability estimations $P_f(\boldsymbol{\Theta}_i)_{i=1,\ldots,N}$ have to be performed. Second, we are in the case of rare event estimations, meaning that classical Monte Carlo are not efficient to provide accurate probability estimation.

A possible solution is proposed in [1] for an estimation of all the Sobol indices with reverse importance sampling (RIS). For a given $\boldsymbol{\theta}_0$, the failure probability $P_f(\boldsymbol{\theta}_0)$ is estimated with Monte Carlo method or importance sampling with sampling density g. RIS enables then to estimate $P_f(\boldsymbol{\Theta}_i)_{i=1,\ldots,N}$ without any calls to $\mathcal{M}$. The corresponding Sobol indices can then be derived at a limited cost. Nevertheless the accuracy of $P_f(\boldsymbol{\Theta}_i)$ with RIS estimates depends mainly on the Kullback–Leibler divergence between the optimal sampling density $\propto \mathbb{1}_{\mathcal{M}(\mathbf{X})\leq T} f_{\mathbf{X}|\boldsymbol{\theta}_i}$ and g. When the variability of $\boldsymbol{\Theta}$ around $\boldsymbol{\theta}_0$ is too high, the RIS approximation is not sufficiently accurate and can lead to a misestimation of the Sobol indices.

We propose an improvement of this solution with adaptive enrichment of the sample to improve the estimation of the Sobol indices while limiting the additional calls to $\mathcal{M}$. We aim to improve the Sobol estimation by improving the estimation of the probabilities $P_f(\mathbf{\Theta}_i)_{i=1,\dots,N}$. We chose a criterion to estimate the accuracy of the estimations of these different probabilities (e.g. the estimated coefficient of variation of the estimator or the effective sample size [6]). After a first estimation of the Sobol with the method of [1], we perform an IS for the least well estimated $P_f(\mathbf{\Theta}_i)$, obtaining a new sample density, new input samples and associated model outputs. These new data fit for this particular $\mathbf{\Theta}_i$ allow us to improve the estimation of $P_f(\mathbf{\Theta}_i)$, but we also use these new data in combination with the previous available data to improve as much as possible the estimations of all the $P_f(\mathbf{\Theta}_i)_{i=1,\dots,N}$. Using multiple importance sampling (mIS) [3], we search for each $P_f(\mathbf{\Theta}_i)$ the combination of available data that yields the best probability estimation. mIS does not require additional calls to M and allow us to potentially improve on all the probability estimations. We repeat this process of selection of $P_f(\mathbf{\Theta}_i)$, adapted resampling and mIS until all the $P_f(\mathbf{\Theta}_i)_{i=1,\dots,N}$ are considered sufficiently well estimated.

As case study, we will consider a drone operation safety evaluation model test case. The numerical model considered is a timed automaton based on a functional modelling of a drone. This model is created using the safety modelling language AltaRica 3.0 [4], we consider the model as a black box for our study. The model takes as input vector of dimension 40 with one distribution parameter for each dimension, corresponding to the failure time of the various components, and outputs the time of failure of the drone after the start of operations. We are interested in studying the probability that the drone fails within one hour of operation, and how the uncertainties about the input distribution affect this probability.

References

[1] Jérôme Morio. "Influence of Input PDF Parameters of a Model on a Failure Probability Estimation". In: *Simulation Modelling Practice and Theory* 19.10 (Nov. 2011), pp. 2244–2255. ISSN: 1569190X. DOI: `10.1016/j.simpat.2011.08.003`.

[2] Fabrice Gamboa et al. *Statistical Inference for Sobol Pick Freeze Monte Carlo Method*. Mar. 2013. DOI: `10.48550/arXiv.1303.6447`. arXiv: `1303.6447 [math, stat]`.

[3] Art B. Owen. *Monte Carlo theory, methods and examples*. `https://artowen.su.domains/mc/`, 2013. Chap. 9.

[4] Tatiana Prosvirnova. "AltaRica 3.0: a Model-Based approach for Safety Analyses". Theses. Ecole Polytechnique, Nov. 2014. URL: `https://pastel.hal.science/tel-01119730`.

[5] Vincent Chabridon. "Reliability-oriented sensitivity analysis under probabilistic model uncertainty – Application to aerospace systems". Theses. Université Clermont Auvergne [2017-2020], Nov. 2018. URL: `https://theses.hal.science/tel-02087860`.

[6] Víctor Elvira, Luca Martino, and Christian P. Robert. "Rethinking the Effective Sample Size". In: *International Statistical Review* 90.3 (Apr. 2022), pp. 525–550. ISSN: 1751-5823. DOI: `10.1111/insr.12500`. URL: `http://dx.doi.org/10.1111/insr.12500`.

[Jonathan Mboko; Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, F-31000 Toulouse, France]
[`jonathan.mboko@onera.fr` –]