Sensitivity Analysis of Monte Carlo Estimates from Computer Models in the Presence of Epistemic and Aleatory Uncertainties

Bernard Krzykacz-Hausmann

Gesellschaft für Anlagen- und Reaktorsicherheit (GRS)
85748 Garching
GERMANY
E-mail: krb@grs.de
fax : ++49-89-32004 301
phone: ++49-89-32004 394

Abstract

The effect of model input variables subject to aleatory uncertainty ("random behavior") is very frequently analyzed by Monte Carlo simulation using a computer model. To this end the aleatory variables are sampled according to their random laws and the corresponding model runs are performed. The results are then summarized in form of empirical distribution functions representing the aleatory uncertainty of the process outcomes of interest. From these empirical distributions statistical estimates of the probabilities of the process states of interest and other useful probabilistic quantities (expectations etc.) are obtained.

Often, however, the exact types of the random laws and their distributional parameters as well as the relevance of phenomena, the model formulations, the values of model parameters and input data of the model application are not known precisely, i.e. they are subject to epistemic ("lack-of-knowledge") uncertainty. These uncertainties are represented by subjective probability distributions which quantify the respective states of knowledge. The process variables from the model are therefore subject to both epistemic and aleatory uncertainties.

An epistemic sensitivity analysis in this case should quantify the effect of the epistemic uncertainty of an input parameter on the epistemic uncertainty of the probability of the process states of interest or other probabilistic quantities representing aleatory uncertainty (expectations etc).

It is intuitively clear and has often been pointed out by many authors that the natural method to appropriately account for both types of uncertainty by Monte Carlo simulation is a double-loop nested sampling procedure. It consists of (1) an "outer loop" where the values of the epistemic parameters are sampled according to their epistemic marginal probability distributions and of (2) a nested "inner loop" where the values of the aleatory variables are sampled according to their aleatory conditional probability distributions given the values of

the epistemic variables chosen in the outer loop. Since each "inner loop" provides an empirical conditional aleatory distribution function of the process outcome of interest, a sample of such empirical distribution functions is finally obtained from this nested double-loop procedure. This sample can be used to perform a standard epistemic sensitivity analysis for the process state probabilities and other probabilistic quantities of interest with respect to the underlying uncertain parameters.

However, if the code representing the model is very complex and time-consuming to run, and this is often the case in many important applications, e.g. in probabilistic safety analysis (PSA) of nuclear power plants, the computational effort to perform all the model runs necessary for the nested double-loop Monte Carlo simulation will be prohibitive.

Therefore, an approach of an approximate epistemic sensitivity analysis is suggested in this contribution. Instead of the nested double-loop sampling procedure this approach employs a simple single-loop sampling procedure where both epistemic parameters and aleatory variables are sampled "simultaneously" according to their joint probability distribution. It turns out that many of the commonly used sensitivity measures, e.g. the standard correlation-regression related measures as well as the variance-based measures etc., can be estimated from this single-loop sample, as well. In particular, using the helpful notion of conditional expectation, it will be shown that many of the population sensitivity measures to be estimated from a full double-loop nested sample are uniformly proportional to the corresponding sensitivity measures estimated from this single-loop sample. The respective population parameter importance ranking will therefore be the same in both cases. Consequently, the sample based parameter importance ranking obtained from the single-loop approach may approximately be used as the importance ranking asked for. Given the sample size N the accuracy of this approximation depends on the value of the proportionality constant C which relates the corresponding sensitivity measures and which can be interpreted as an indicator of the impact of the epistemic uncertainties on the overall joint outcome uncertainty. In many cases this constant can be approximated from the underlying single-loop sample. In more difficult situations where this approximation is unsatisfactory or impracticable an additional appropriate single-loop sample may be generated to obtain a better statistical estimate of the constant C.

Analytical/numerical examples and simple applications will be presented that demonstrate the practicability and usefulness as well as the limitations of this approach.