Available online at www.sciencedirect.com
SCIENCE DIRECT®
Reliability Engineering and System Safety 91 (2006) 1210–1218
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

An approximate sensitivity analysis of results from complex computer models in the presence of epistemic and aleatory uncertainties

Bernard Krzykacz-Hausmann*

Gesellschaft für Anlagen-und Reaktorsicherheit (GRS), 85748 Garching, Germany

Available online 10 January 2006

Abstract

This paper focuses on sensitivity analysis of results from computer models in which both epistemic and aleatory uncertainties are present. Sensitivity is defined in the sense of "uncertainty importance" in order to identify and to rank the principal sources of epistemic uncertainty. A natural and consistent way to arrive at sensitivity results in such cases would be a two-dimensional or double-loop nested Monte Carlo sampling strategy in which the epistemic parameters are sampled in the outer loop and the aleatory variables are sampled in the nested inner loop. However, the computational effort of this procedure may be prohibitive for complex and time-demanding codes. This paper therefore suggests an approximate method for sensitivity analysis based on particular one-dimensional or single-loop sampling procedures, which require substantially less computational effort. From the results of such sampling one can obtain approximate estimates of several standard uncertainty importance measures for the aleatory probability distributions and related probabilistic quantities of the model outcomes of interest. The reliability of the approximate sensitivity results depends on the effect of all epistemic uncertainties on the total joint epistemic and aleatory uncertainty of the outcome. The magnitude of this effect can be expressed quantitatively and estimated from the same single-loop samples. The higher it is the more accurate the approximate sensitivity results will be. A case study, which shows that the results from the proposed approximate method are comparable to those obtained with the full two-dimensional approach, is provided.

Keywords: Sensitivity analysis; Aleatory and epistemic uncertainty; Uncertainty importance; Conditional expectation

1. Introduction

It is a common practice in many scientific fields to analyze the impact of aleatory uncertainties on model results by Monte Carlo sampling procedures. From the sample values of the model outcome generated in this way one may determine statistical estimates of the probabilities of the process states of interest and other useful probabilistic quantities expressing aleatory uncertainty. In this context the frequentistic concept of probability interpretation is applied.

Often, however, the exact types of the random laws involved, their distributional parameters, the values of many other model parameters and input data for the model application etc. are not known precisely, i.e. are subject to epistemic uncertainty. These uncertainties may also be quantified by probability distributions representing the respective subjective state of knowledge and being interpreted according to the subjectivistic or degree-of-belief concept of probability.

It is widely acknowledged that these two types of uncertainty must very carefully be distinguished and that this distinction must be maintained throughout the analysis and displayed in the final results.

It is also intuitively clear and has often been pointed out [1] that the most natural Monte-Carlo-based method consistent with the principle of separating the two types of uncertainty is the "double-loop" nested sampling procedure, also called "two-stage" or "two-dimensional" sampling. It consists of

(1) an outer loop where the values of the epistemic parameters are sampled according to their epistemic marginal probability distributions and

*Fax: + +49 89 32004 301.
E-mail address: krb@grs.de.

0951-8320/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2005.11.019

(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 generated in the outer loop.

Each inner loop provides a conditional empirical aleatory distribution of the model outcome of interest. From this distribution the probabilities of the process states of interest and other useful probabilistic quantities can be statistically estimated. Finally, both loops together provide a sample of such empirical aleatory distributions and, correspondingly, a sample of probabilities of the process states or other probabilistic quantities of interest expressing aleatory uncertainty.

This sample, along with the underlying parameter sample generated in the outer loop, can be used to perform a standard sensitivity analysis, i.e. to compute appropriate sensitivity indices for the probabilistic quantities of interest with respect to all the epistemic uncertainties.

It is clear that the computational effort for such two-dimensional sampling may not be feasible, particularly if the model is computationally expensive and small probabilities are to be estimated from the inner loop samples. For computationally not too expensive models the sampling within the inner loop may be modified using variance-reducing sampling methods [2] or may even replaced by analytical methods like FORM/SORM [3] or Fault-/Event-Tree analysis [4]. For further references on the subject of separating uncertainties at reduced computational costs, cf. [5,6].

Nevertheless, there are still many computationally demanding applications, e.g. in nuclear safety, where all these methods are not feasible. For an uncertainty and sensitivity analysis of results from such models appropriate methods are needed, which keep the overall number of model runs as small as possible.

The sampling method for an approximate sensitivity analysis presented in this paper consists of solely two single-loop Monte Carlo samples. Compared with the full two-stage nested sampling this is a substantial reduction of computational effort.

A slightly different sampling method was proposed in [5,6] for the purposes of an approximate uncertainty analysis, only. However, it will turn out that with the present approach both an approximate uncertainty and an approximate sensitivity analysis can be conducted at the same low computational costs and even under less restrictive conditions.

A similar sampling strategy was also proposed in [7,10] in the context of estimating the "first-order effect" and the "total effect" variance-based sensitivity indices. It will turn out to be a special case of the proposed sampling method under the assumption of independence between the variables involved.

2. Fundamentals of the approach

The terminology notation and presentation in this section is adopted from [5].

Any scalar process variable or model outcome Y, which is subject to epistemic and aleatory uncertainties, may be represented as

$$Y = h(\mathbf{U}, \mathbf{V}),$$

with $\mathbf{U} = (U_1, \ldots, U_n)$ is the set of all epistemic uncertainties (**u**ncertain parameters), $\mathbf{V} = (V_1, \ldots, V_\mathrm{m})$ the set of all aleatory uncertainties (random **v**ariables), $h(\mathbf{U},\mathbf{V})$ the computational model considered as a function of both aleatory and epistemic uncertainties $\mathbf{U}$ and $\mathbf{V}$.

This situation can probabilistically be characterized by two probability distributions (PDF):

(1) $f_\mathbf{U}(\mathbf{u})$: marginal epistemic PDF of $\mathbf{U}$ which quantifies the epistemic uncertainty of the parameters $\mathbf{U}$. It is very often obtained from expert judgment.
(2) $f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{U} = \mathbf{u})$: conditional aleatory PDF of $\mathbf{V}$ given $\mathbf{U} = \mathbf{u}$ (= family of distributions indexed by $\mathbf{u}$). It describes the random laws of the aleatory variables $\mathbf{V}$ for a given value $\mathbf{u}$ of the uncertain parameters $\mathbf{U}$. This distribution is usually part of the computational model and may depend on $\mathbf{u}$ also via intermediate model results.

If $\mathbf{U} = \mathbf{u}$, i.e. if the epistemic variables $\mathbf{U}$ are held fixed at $\mathbf{u}$, the resulting outcome Y is a function of the aleatory uncertainties $\mathbf{V}$ alone. Its distribution is the conditional distribution of Y given $\mathbf{U} = \mathbf{u}$ and quantifies the corresponding conditional aleatory uncertainty in Y. The expected value of this distribution is

$$\begin{aligned} E[Y|\mathbf{U} = \mathbf{u}] &= E[\text{conditional distribution of } Y \text{ given } \mathbf{U} = \mathbf{u}] \\ &= \int h(\mathbf{u}, \mathbf{v}) f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{U} = \mathbf{u})\, \mathrm{d}\mathbf{v}, \end{aligned}$$

i.e. $\mathbf{v}$ is "integrated out" from $h(\mathbf{u},\mathbf{v})$.

$E[Y|\mathbf{U} = \mathbf{u}]$ can be viewed as a quantity depending on $\mathbf{u}$ and representing the conditional aleatory uncertainty of the outcome Y given $\mathbf{U} = \mathbf{u}$. In the sequel the standard concise notation

$$E[Y|\mathbf{U}]$$

will be used to denote the above conditional expectation $E[Y|\mathbf{U} = \mathbf{u}]$ considered as function of the epistemic uncertainties $\mathbf{U}$ alone.

Using expectation to represent conditional aleatory uncertainty must not be regarded as very restrictive. Many of the standard distributional parameters characterizing aleatory uncertainty can be viewed as expectations of appropriate outcome functions. E.g. the value $F_Y(y|\mathbf{U} = \mathbf{u})$ of the conditional distribution function of an outcome variable Y at any given point y as well as the probability of any other "system state of interest" may be represented as conditional expectation $E[I_{\{\}}|\mathbf{U} = \mathbf{u}]$ of an appropriately