Reliability Engineering and System Safety 112 (2013) 82–93

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

Probabilistic sensitivity analysis of system availability using Gaussian processes

Alireza Daneshkhah, Tim Bedford *

Department of Management Science, University of Strathclyde, Glasgow G1 1QE, UK

A R T I C L E I N F O

Article history:
Received 7 January 2012
Received in revised form
30 October 2012
Accepted 1 November 2012
Available online 27 November 2012

Keywords:
Sensitivity analysis
Sobol indices
Dependency
Computer code model output
Emulator-based sensitivity
Gaussian process
System availability

A B S T R A C T

The availability of a system under a given failure/repair process is a function of time which can be determined through a set of integral equations and usually calculated numerically. We focus here on the issue of carrying out sensitivity analysis of availability to determine the influence of the input parameters. The main purpose is to study the sensitivity of the system availability with respect to the changes in the main parameters. In the simplest case that the failure repair process is (continuous time/discrete state) Markovian, explicit formulae are well known. Unfortunately, in more general cases availability is often a complicated function of the parameters without closed form solution. Thus, the computation of sensitivity measures would be time-consuming or even infeasible.

In this paper, we show how Sobol and other related sensitivity measures can be cheaply computed to measure how changes in the model inputs (failure/repair times) influence the outputs (availability measure). We use a Bayesian framework, called the Bayesian analysis of computer code output (BACCO) which is based on using the Gaussian process as an emulator (i.e., an approximation) of complex models/functions. This approach allows effective sensitivity analysis to be achieved by using far smaller numbers of model runs than other methods.

The emulator-based sensitivity measure is used to examine the influence of the failure and repair densities' parameters on the system availability. We discuss how to apply the methods practically in the reliability context, considering in particular the selection of parameters and prior distributions and how we can ensure these may be considered independent—one of the key assumptions of the Sobol approach. The method is illustrated on several examples, and we discuss the further implications of the technique for reliability and maintenance analysis.

1. Introduction

In this paper, we present a new approach to study the sensitivity analysis of availability. In general sensitivity analysis is concerned with understanding how changes in the model input (distribution parameters) would influence the output. Suppose that our deterministic model can be written as $y=f(\mathbf{x})$, where $\mathbf{x}$ is a vector of input variables (or parameters) and y is the model output. For example, the inputs could be considered as the parameters of the failure and repair densities, $\boldsymbol{\theta}$, and the output could be the availability $A(t,\boldsymbol{\theta})$ at time t.

The traditional method of examining sensitivity of a model with respect to the changes in its input variables is local sensitivity analysis which is based on derivatives of $f(\cdot)$ evaluated at some 'base-line' (or central estimate) $\mathbf{x}=\mathbf{x}_0$ and indicates how the output y will change if the base line input values are slightly perturbed (see [11] for the different local sensitivity measures commonly used in Bayesian analysis). This is clearly of limited value in understanding the consequences of real uncertainty about the inputs, which would in practice require more than infinitesimal changes in the inputs. Furthermore, these methods are computationally very expensive for complex models and usually require a considerable number of model runs if we use a Monte Carlo based method to compute these sensitivity measures.

For instance, Marseguerra et al. [21] used Monte Carlo simulation to calculate the first-order differential sensitivity indexes of the basic events characterising the reliability behaviour of a nuclear safety system. They reported that the computation of the sensitivity indexes for the system unavailability at the mission time by Monte Carlo simulation requires 10^7 iterations. In another study, Reedijk [28] reported that first order reliability methods and Monte Carlo simulation have certain disadvantages and some problems could not be solved with these methods.

This issue is particularly interesting in the case where the model is computationally expensive so that simply computing the output for any given set of input values is a non-trivial task.

* Corresponding author.
E-mail address: tim.bedford@strath.ac.uk (T. Bedford).

0951-8320/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2012.11.001

This is especially the case for large process models in engineering, environmental science, reliability analysis, etc. that may be implemented in complex computer codes requiring many minutes, hours or even days for a single run. However, in order to implement many of the standard sensitivity techniques discussed by [30] we require a very large number of model runs. In that case even for a model that takes just one second to run, many sensitivity analysis measures may take too long to compute.

The most frequently used sensitivity indices are due to Sobol [32]. However, these require an assumption of independence (as discussed by Bedford [1]). Hence in Section 3 we discuss here how one might go about choosing an appropriate parameterisation in which the sensitivity analysis can be carried out using independent uncertainty variables.

It should be noticed that the probabilistic sensitivity analyses are often effectively carried out with efficient sampling procedures (e.g., [14,15]), but these procedures are computationally very expensive. Therefore, we present an alternative computational tool to implement sensitivity analysis based on the work of [22]. This is a Bayesian approach of sensitivity analysis which unifies the various methods of probabilistic sensitivity analysis which will be briefly introduced in Section 2. This approach is computationally highly efficient and allows effective sensitivity analysis to be achieved by using very smaller numbers of model runs than Monte Carlo methods require. The range of tools used in this approach also enables us to do uncertainty analysis, prediction, optimisation and calibration. Section 4 presents this method.

This paper extends work carried out by Daneshkhah and Bedford [10] where emulators were used to examine the influence of failure and repair densities' parameters on the system availability of a simple one component repairable system where Exponential and Weibull were considered as distributions for the failure and repair rates. Here, we consider the sensitivity analysis of repairable systems with more than one component, chosen so that we can use numerical integration to compute availability. The systems are: a parallel system with two components where the failure and repair distributions are exponentials; a well known standby-redundancy system where there are three parameters to be examined and the failure and repair distributions are also exponentials; and move-drive system with eight components and 17 parameters, where the repair rate is constant but the failure distributions are Weibulls. There are closed forms for the availability functions for the first two systems and we use them to validate the method. But there is no closed form for the third system, and to evaluate the system availability at the selected parameter values, an expensive numerical method required. We present some conclusions and possible future developments in Section 6.

2. Probabilistic sensitivity analysis

Local sensitivity analysis evaluates the influence of uncertain inputs around a point in the input space and generally relies on the estimation, at this point, of the partial derivatives of the output with respect to the inputs. This is known as a one-at-a-time measure of sensitivity because it measures the effect on the response of varying one input alone by a fixed fraction of its value (assumed known). As a consequence, if the model is nonlinear, the relative importance of the model inputs depends on the chosen point. Several investigators tried to get around this limitation by evaluating averages of the partial derivatives at different points in the input space.

Conversely, global sensitivity analysis of model output evaluates the relative importance of inputs when they are varied generously, i.e., when their uncertainty is acknowledged over a wide range. One approach to global sensitivity analysis is the analysis of variance of the model response originally proposed by [32]. In this setting nonlinearity in the model is not an issue. The approach can capture the fraction of the model response variance explained by a model input on its own or by a group of model inputs. In addition, it can also provide the total contribution to the output variance of a given input, that is its marginal contribution and its cooperative contribution. There are different computational techniques to perform this sensitivity analysis, e.g., [32,30,33]. We will focus on the emulator-based method to calculate this sensitivity measure presented by [22].

Borgonovo et al. [4] introduced the moment-independent sensitivity methods which are recently attracted by practitioners (particulary, in Environmental sciences). Similar to the emulator-based sensitivity method, these methods provide a thorough way of investigating the sensitivity of model output under uncertainty. However, their estimation is challenging, especially in the presence of computationally intensive models. Borgonovo et al. [4] suggest to replace the original model by a metamodel to lower the computation burden. They utilise the emulator proposed in [27] as an efficient metamodel. Their results show that the emulator allows an accurate estimation of density-based sensitivity measures, when the main structural features of the original model are captured (see also [3]).

Ratto et al. [26] also reported that emulators (also denoted as metamodelling in the literature) provide an efficient means for doing a sensitivity analysis for large and expensive models. They provide some tools and applications of sensitivity analysis in the context of environmental modelling.

Caniou and Sudret [6] propose an alternative method against the variance-based sensitivity methods mentioned above (e.g., Sobol indices) which are relatively expensive because of the conditional moments estimation for quantifying the uncertainty of the model output due to the uncertain input variable. They substitute the initial model by an alternative metamodel called polynomial chaos expansion. The process is applied to numerical test cases. Their obtained results are discussed and compared with the reference obtained with variance-based methods. In 2011 [7], these authors developed this work to the situation where the input parameters are correlated.

In this paper, we focus on the method suggested by [22]. We shall consider how a function $f(\mathbf{x})$ depends on its input variables, and in our case f will typically be the function that computes system availability as a function of a vector of quantities such as failure and repair distribution parameters. Note that we would consider availability at different time points as distinct functions, so time is not considered as an input variable. We also assume the existence of a distribution G representing the uncertainty of the parameters. We discuss below appropriate choices for the vector of inputs and the prior distribution G. For the purpose of computing Sobol indices it is often assumed that G has to be an independent distribution, and we discuss this further below.

Some notation is introduced first. Write a d-dimensional random vector as $\mathbf{X}=(X_1,\ldots,X_d)$, where X_i is the ith element of $\mathbf{X}$, the subvector (X_i,X_j) is shown by $\mathbf{X}_{i,j}$, and in general if p is a set of indices then write $\mathbf{X}_p$ for the subvector of $\mathbf{X}$ whose elements have those indices. We also denote $\mathbf{X}_{-i}$ as the subvector of $\mathbf{x}$ containing all elements except x_i. Similarly, $\mathbf{x}=(x_1,\ldots,x_d)$ denote the corresponding observed random vector $\mathbf{X}$.

2.1. Main effects and interactions

Since f is a function of uncertain quantities $\mathbf{X}$ we can consider its expected value (when f is availability at a given time, then the expected value is the predictive availability at that time).