Mechanical Systems and Signal Processing 100 (2018) 782–801

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Estimation of the lower and upper bounds on the probability of failure using subset simulation and random set theory

Diego A. Alvarez [a,*], Felipe Uribe [b], Jorge E. Hurtado [a]

[a] Departamento de Ingeniería Civil, Universidad Nacional de Colombia, Sede Manizales, Carrera 27 # 64-60, 170004 Manizales, Colombia
[b] Engineering Risk Analysis Group, Technische Universität München, Arcisstraße 21, 80333 München, Germany

ARTICLE INFO

Article history:
Received 30 March 2017
Received in revised form 7 July 2017
Accepted 19 July 2017

Keywords:
Lower and upper probabilities of failure
Imprecise probabilities
Random set theory
Copulas
Interval Monte Carlo
Subset simulation
Isoprobabilistic transformation

ABSTRACT

Random set theory is a general framework which comprises uncertainty in the form of probability boxes, possibility distributions, cumulative distribution functions, Dempster-Shafer structures or intervals; in addition, the dependence between the input variables can be expressed using copulas. In this paper, the lower and upper bounds on the probability of failure are calculated by means of random set theory. In order to accelerate the calculation, a well-known and efficient probability-based reliability method known as subset simulation is employed. This method is especially useful for finding small failure probabilities in both low- and high-dimensional spaces, disjoint failure domains and nonlinear limit state functions. The proposed methodology represents a drastic reduction of the computational labor implied by plain Monte Carlo simulation for problems defined with a mixture of representations for the input variables, while delivering similar results. Numerical examples illustrate the efficiency of the proposed approach.

1. Introduction

The treatment of uncertainties and the calculation of the probability of limit state violations are of primary concern in modern engineering systems. Different modes of failure can be grouped in the so-called *limit state function* (LSF) $g : \mathscr{X} \to \mathbb{R}$, which depends on a set of uncertain system parameters $\boldsymbol{x} \in \mathscr{X} \subseteq \mathbb{R}^d$. In the framework of reliability assessment, the *failure surface* $g(\boldsymbol{x}) = 0$ splits the $\mathscr{X}-$ space in two domains, namely the *safe set* $S = \{\boldsymbol{x} \in \mathscr{X} : g(\boldsymbol{x}) > 0\}$ and the *failure set* $F = \{\boldsymbol{x} \in \mathscr{X} : g(\boldsymbol{x}) \leqslant 0\}$. The probability measure of $F \subseteq \mathscr{X}$, also known as the *probability of failure*, is defined as

$$P_{\mathrm{f}} = \int_{\mathscr{X}} \mathbb{I}_F[\boldsymbol{x}] \mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{\mathscr{X}} \mathbb{I}[g(\boldsymbol{x}) \leqslant 0] \mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{x}) \qquad (1)$$

where, $F_{\boldsymbol{X}}(\boldsymbol{x})$ is the joint cumulative distribution function (CDF) of the input variables and $\mathbb{I}[\cdot]$ stands for the indicator function, which takes the values $\mathbb{I}_F[\boldsymbol{x}] = 1$ when either $\boldsymbol{x} \in F$ or the condition in square brackets is true, and $\mathbb{I}_F[\boldsymbol{x}] = 0$ otherwise. When $F_{\boldsymbol{X}}(\boldsymbol{x})$ is sufficiently differentiable, the associated joint probability density function (PDF) $f_{\boldsymbol{X}}(\boldsymbol{x})$ exists, and in this case Eq. (1) can be expressed also as

$$P_{\mathrm{f}} = \int_{\mathscr{X}} \mathbb{I}_F[\boldsymbol{x}] f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_{\mathscr{X}} \mathbb{I}[g(\boldsymbol{x}) \leqslant 0] f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

* Corresponding author.
E-mail addresses: daalvarez@unal.edu.co (D.A. Alvarez), felipe.uribe@tum.de (F. Uribe), jehurtadog@unal.edu.co (J.E. Hurtado).

http://dx.doi.org/10.1016/j.ymssp.2017.07.040
0888-3270/

One of the main drawbacks of applying the probabilistic approach to reliability analysis of structures is that the CDFs of the input variables $\boldsymbol{x} \in \mathscr{X}$ are usually known with imprecision. This is normally due to the lack of sufficient data for fitting the model to each input random variable. Even if the information is plenty, there remains the problem of the high sensitivity of the usually small probabilities of failure to the parameters of the CDFs (see e.g. [23,35]). In addition, the information about the input variables, is not always given as joint CDFs, but they can be expressed as well, in terms of probability boxes, possibility distributions, Dempster-Shafer structures, intervals, among other representations of uncertainty. All those reasons complicate the application of the efficient methods that exist in the realm of probability theory for the estimation of the reliability of structural and mechanical systems.

These difficulties have fostered the research on alternative methods (coined under the term of *imprecise probabilities*) for incorporating uncertainty in the analysis of engineering systems, such as possibility theory, info-gap theory, convex models, interval analysis, ellipsoid modeling, credal sets, Dempster-Shafer evidence theory, random sets, random fuzzy sets, fuzzy random sets, probability boxes, sets with parametrized probability measures, among other methods. The reader is referred to the state-of-the-art review [12] and references within for further discussion and additional information.

In this context, random set theory appears as a unifying framework which comprises several types of uncertainty representations, either aleatory or epistemic. This approach allows us to estimate the lower and upper bounds on the probability of events, and thus, it can be used to bound the probability of failure. Within the random set approach to structural reliability, research has been done for introducing the epistemic uncertainty in the probabilistic models. In this direction, Tonon [44] used random set theory to calculate the reliability bounds for the challenge problem proposed by Oberkampf et al. [36]. Du [19] developed a methodology, termed the unified uncertainty analysis method for reliability assessment of structural and mechanical systems. The approach uses a double loop optimization process which contains probabilistic and interval analysis, and employs the first-order reliability method (FORM) for the solution of the reliability problem. Similarly to the sampling methods developed in Alvarez [1,3], Zhang and co-workers proposed the so-called interval Monte Carlo simulation method [52,51]; also they developed an interval importance sampling method [49] and an interval quasi-Monte Carlo method [50]. Xiao et al. [47] proposed an efficient saddle-point approximation to speed up the results of the interval Monte Carlo simulation method. Alvarez and Hurtado [5] proposed a method based on the reliability plot to estimate in a parsimonious way the lower and upper probabilities of failure. Recently, Yang et al. [48] used a surrogate kriging model to accelerate the computations of the failure probability bounds.

As it will be seen in the paper, the estimation of the lower and upper probabilities of failure can be postulated as two standard reliability assessment problems, and consequently, any method for the estimation of the probability of failure that only uses probabilistic information can be applied. In particular, we will illustrate this methodology using one of the most popular and efficient methods, namely, subset simulation [8,9]. During the last decade, subset simulation has established itself as one of the leading algorithms for the estimation of failure probabilities. Therefore, the engineering research community has focused on the enhancement and generalization of the method; some of the most recent contributions include: Bayesian post-processor for subset simulation [53], combination of subset simulation with machine learning-based surrogate models [13,37], and the subset simulation enhancements proposed by Papaioannou et al. [38] and Au and Patelli [11]. Perhaps one of the most significant developments was proposed by Walter [46], who applied the concept of moving particles; the approach presents two main results in the context of subset simulation (referred to as multilevel splitting): first, the number of samples required to populate F is considerably reduced, and second, the adaptive selection of the intermediate levels is no longer required since by construction the nested subsets do not exist anymore.

In this contribution, we will use subset simulation in conjunction with random set theory in order to estimate the upper and lower bounds on the probability of failure when the input variables are defined in terms of probability boxes, possibility distributions, CDFs, Dempster-Shafer structures, or intervals. In fact, the proposed approach is so general that any method for assessing the probability of failure can be used as well.

The plan of this work is as follows. The document begins with a succinct introduction to copulas and random sets in Sections 2 and 3, respectively. Then in Section 4, we introduce the mathematical formulation for the estimation of the probability of failure and its relationship with random set theory. Specifically, we will see that the calculation of the lower and upper probabilities of failure will correspond to the evaluation of two integrals that compute the probability of failure for two different LSFs. Section 5 will introduce the Monte Carlo simulation method for the estimation of the probability of failure, and Section 6 will introduce the subset simulation algorithm. The proposed methodology will be illustrated in Section 7; there, we explain how to use subset simulation to estimate the lower and upper probabilities of failure provided by random set theory after applying a suitable isoprobabilistic transformation. Section 8 demonstrates the advantages of the proposed approach with three numerical examples. The paper ends with the discussion of results, conclusions, some open problems and the corresponding acknowledgements.

2. An introduction to copulas

2.1. Overview

This concise review of some important concepts about copulas follows the exposition presented in Nelsen [33]. A *copula* C is a d-dimensional CDF, $C : [0,1]^d \to [0,1]$, whose univariate marginal CDFs are uniform on the interval $[0,1]$. The main