Computers and Structures 142 (2014) 54–63

Contents lists available at ScienceDirect

Computers and Structures

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

An efficient method for the estimation of structural reliability intervals with random sets, dependence modeling and uncertain inputs

Diego A. Alvarez *, Jorge E. Hurtado

Universidad Nacional de Colombia, Apartado 127, Manizales, Colombia

ARTICLE INFO

Article history:
Received 12 November 2013
Accepted 9 July 2014
Available online 30 July 2014

Keywords:
Reliability interval
Reliability bounds
Random sets
Probability boxes
Possibility distributions
Probability distributions

ABSTRACT

A general method for estimating the bounds of the reliability of a system in which the input variables are described by random sets (probability distributions, probability boxes, or possibility distributions), with dependence modeling is proposed. The method is based on an analytical property of the so-called design point vector; this property is exploited by constructing a nonlinear projection of Monte Carlo samples of the input variables in a two-dimensional diagram from which the analyst can easily extract the relevant samples for computing both the lower and upper bounds of the failure probability using random set theory. The method, which is illustrated with some examples, represents a dramatic reduction in the number of focal element evaluations performed when applying the Monte Carlo method to random set theory.

1. Introduction

Uncertainty analysis in engineering should ideally be a part of routine design because the variables and supposedly constant parameters are either random or known with imprecision. In some cases the uncertainty can be very large, such as the case of natural actions provoking disasters or modeling errors leading to technological catastrophes. In approaching the estimation of the risk of a given engineering problem, use is traditionally made of cumulative distributions functions (CDFs) defining the input variables and then, by means of analytic or synthetic methods (i.e. Monte Carlo) the probability of not exceeding undesirable thresholds, is computed [1,2].

One of the main problems in applying the probabilistic approach is that the CDFs of the input variables are usually known with imprecision. This is normally due to the lack of sufficient data for fitting the model to each input random variable. For this reason, the parameters of the input distributions are commonly known up to confidence intervals, and even these latter are not wholly certain. This hinders the application of the probability-based approach in actual design practice [3]. Even if the information is abundant, there remains the problem of the high sensitivity of the usually small probabilities of failure to the parameters of the distribution functions [4–6]. Such a sensitivity is due to the fact that the estimation of a probability density function from empirical data is an ill-posed problem [7,8]. This means that small changes in the empirical sample affects the parameters defining the model being fitted, with serious consequences in the tails, which are just the most important zones of the distribution functions for probabilistic reliability methods [9–11].

These and other considerations have fostered the research on alternative methods for incorporating uncertainty in the structural analysis, such as fuzzy sets and related theories [12–16], anti-optimization or convex-set modeling [5,10,17], interval analysis [10,18–27], random sets [28–30], ellipsoid modeling [31,32] and worst-case scenarios [33]. Also comparisons have been made between probabilistic and the alternative methods [34–36] or their combination has been explored [37–40].

Taking into account that the first- and second-order reliability methods (FORM and SORM) can be very inaccurate in many cases e.g. [2,41–46], the focus of present paper is the determination of the reliability intervals under uncertain input variables by means of Monte Carlo simulation. In this regard, attention is called to [23] where an interval finite-element approach for linear structural analysis and a Monte Carlo method for calculating intervals of the failure probability is proposed and to [28,29] who developed an even more general method of computing the bounds of the probability of failure under the general framework of random set theory and that comprised uncertainty modeled in the form of probability boxes, possibility distributions, CDFs, Dempster-Shafer structures or intervals; in addition the method allows to model dependence between the input variables.

* Corresponding author. Tel.: +57 68879300; fax: +57 68879334.
E-mail addresses: daalvarez@unal.edu.co (D.A. Alvarez), jehurtadog@unal.edu.co (J.E. Hurtado).

http://dx.doi.org/10.1016/j.compstruc.2014.07.006
0045-7949/

Present paper is aimed to the goal of facilitating the Monte Carlo solution of the interval reliability computation, which is much more computationally demanding than the conventional computation of a single reliability value [23,28,29,47]. In particular, a method based on random set theory is proposed that allows selecting all relevant samples for a Monte Carlo estimation of the bounds of the failure probability from a large mass of input variable realizations generated from the uncertain distributions. Hence, the method avoids the large number of sample evaluations with null contribution to the failure probability estimation, which is the typical case in using plain Monte Carlo simulation.

The proposed approach is based on a property of FORM [48], which consists in that the design point vector points to a direction of steep evolution of the limit state function [49–51]. This property also holds for functions arising from the perturbation represented by the interval uncertainty in the distribution parameters. Therefore, in spite of FORM's inaccuracy in many reliability problems [2,41,44,45], its design point vector emerges as a powerful clustering device, because of the way that the performance function evolves in this direction. Then, such a property is exploited by constructing a nonlinear transformation of the reliability problem from d dimensions to a bi-dimensional space of two independent variables whose marginal and joint density functions are explicitly derived. The main characteristic of this transformation is that it makes evident the organizing property mentioned above in a bi-dimensional representation of the entire set of random numbers, allowing the selection of the relevant samples for interval or single reliability computations on almost a blind basis. The proposed approach is illustrated with detailed structural examples. The paper ends with some conclusions and suggestions for future work.

2. A brief introduction to random sets

Random set theory is a mathematical tool, which can effectively unify a wide range of theories for coping with aleatory and epistemic uncertainty. It is an extension of probability theory to set-valued rather than point-valued maps. In the following paragraphs a brief summary of some of the most important concepts on random sets required in the ensuing discussion is presented.

2.1. Copulas

A *copula* is a d-dimensional CDF $C : [0,1]^d \to [0,1]$ such that each of its marginal CDFs is uniform on the interval $[0,1]$.

According to Sklar's theorem (see Refs. [52,53]), copulas are functions that relate a joint CDF with its marginals, carrying in this way the dependence information in the joint CDF; Sklar's theorem states that a multivariate CDF $F_{X_1,X_2,\ldots,X_d}(x_1,\ldots,x_d) = P[X_1 \leqslant x_1,\ldots,X_d \leqslant x_d]$ of a random vector $(X_1,X_2,\ldots,X_d)$ with marginals $F_{X_i}(x_i) = P[X_i \leqslant x_i]$ can be written as $F_{X_1,X_2,\ldots,X_d}(x_1,\ldots,x_d) = C(F_{X_1}(x_1),\ldots,F_{X_d}(x_d))$, where C is a copula. The copula C contains all information on the dependence structure between the components of $(X_1,X_2,\ldots,X_d)$ whereas the marginal cumulative distribution functions F_{X_i} contain all information on the marginal distributions.

In the following we will denote as μ_C, the Lebesgue-Stieltjes measure corresponding to the copula C (see [54] for details).

The reader is referred to [55] for the standard introduction to copulas.

2.2. Definition of a random set

Let us consider a universal set $\mathcal{X} \neq \emptyset$ and its power set $\mathcal{P}(\mathcal{X})$, a probability space $(\Omega,\sigma_\Omega,P_\Omega)$ and a measurable space $(\mathcal{F},\sigma_\mathcal{F})$ where $\mathcal{F} \subseteq \mathcal{P}(\mathcal{X})$. In the same spirit as the definition of a random variable, a *random set* (RS) Γ is a $(\sigma_\Omega - \sigma_\mathcal{F})$-measurable mapping $\Gamma : \Omega \to \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$. In other words, a random set is like a random variable whose realization is a set in $\mathcal{F}$, not a number; let us call each of those sets $\gamma := \Gamma(\alpha) \in \mathcal{F}$ a *focal element* while $\mathcal{F}$ is a *focal set*.

Similarly to the definition of a random variable, the random set can be used to define a probability measure on $(\mathcal{F},\sigma_\mathcal{F})$ given by $P_\Gamma := P_\Omega \circ \Gamma^{-1}$. In other words, an event $\mathcal{R} \in \sigma_\mathcal{F}$ has the probability

$$P_\Gamma(\mathcal{R}) = P_\Omega\{\alpha \in \Omega : \Gamma(\alpha) \in \mathcal{R}\}. \tag{1}$$

The random set Γ will be called henceforth also as $(\mathcal{F},P_\Gamma)$.

Note that when every element of $\mathcal{F}$ is a singleton, then Γ becomes a random variable X, and the focal set $\mathcal{F}$ is said to be *specific*; in other words, if $\mathcal{F}$ is a specific set then $\Gamma(\alpha) = X(\alpha)$ and the probability of occurrence of the event F, is $P_X(F) := (P_\Omega \circ X^{-1})(F) = P_\Omega\{\alpha : X(\alpha) \in F\}$ for every $F \in \sigma_X$. In the case of random sets, it is not possible to compute exactly $P_X(F)$ but its upper and lower probability bounds. [56] defined those upper and lower probabilities by,

$$\begin{aligned} \mathrm{LP}_{(\mathcal{F},P_\Gamma)}(F) &:= P_\Omega\{\alpha : \Gamma(\alpha) \subseteq F, \Gamma(\alpha) \neq \emptyset\} \\ &= P_\Gamma\{\gamma : \gamma \subseteq F, \gamma \neq \emptyset\}, \end{aligned} \tag{2a}$$

$$\mathrm{UP}_{(\mathcal{F},P_\Gamma)}(F) := P_\Omega\{\alpha : \Gamma(\alpha) \cap F \neq \emptyset\} = P_\Gamma\{\gamma : \gamma \cap F \neq \emptyset\}, \tag{2b}$$

where

$$\mathrm{LP}_{(\mathcal{F},P_\Gamma)}(F) \leqslant P_X(F) \leqslant \mathrm{UP}_{(\mathcal{F},P_\Gamma)}(F). \tag{3}$$

Note that the equality in (3) holds when $\mathcal{F}$ is specific. The reader is referred to Refs. [57,58] a complete survey on random sets.

2.3. Relationship between random sets and probability boxes, CDFs and possibility distributions

Definition in Section 2.2 is very general; [28,59] showed that making the particularizations $\Omega := (0,1]^d, \sigma_\Omega := (0,1]^d \cap \mathcal{B}^d$ and $P_\Gamma \equiv \mu_C$ for some copula that contains the dependence information within the joint random set, and using intervals and d-dimensional boxes as elements of $\mathcal{F}$, it is enough to model possibility distributions, probability boxes, intervals, CDFs and Dempster-Shafer structures or their joint combinations; these are some of the most popular engineering representations of uncertainty. Let us denote by $P_\Gamma \equiv \mu_C$ the fact that P_Γ is the probability measure generated by P_Ω which is defined by the Lebesgue-Stieltjes measure corresponding to the copula C, i.e. μ_C. In other words, $P_\Gamma(\Gamma(G)) = \mu_C(G)$ for $G \in \sigma_\Omega$; also $\mathcal{B}$ will stand for the Borel σ-algebra on $\mathbb{R}$.

In the rest of this section, $(\Omega,\sigma_\Omega,P_\Omega)$ will stand for a probability space with $\Omega := (0,1], \sigma_\Omega := (0,1] \cap \mathcal{B} := \cup_{\theta\in\mathcal{B}}\{(0,1] \cap \theta\}$ and P_Ω will be a probability measure corresponding to the CDF of a random variable $\tilde{\alpha}$ uniformly distributed on $(0,1]$, i.e. $F_{\tilde{\alpha}}(\alpha) := P_\Omega[\tilde{\alpha} \leqslant \alpha] = \alpha$ for $\alpha \in (0,1]$; that is, P_Ω is a Lebesgue measure on $(0,1]$.

Probability boxes, CDFs and possibility distributions can be interpreted as random sets, as will be explained in the following:

2.3.1. Probability boxes

A *probability box* or *p-box* (see e.g. [60]) $\langle \underline{F}, \overline{F} \rangle$ is a set of CDFs $\{F : \underline{F}(x) \leqslant F(x) \leqslant \overline{F}(x), F \text{ is a CDF}, x \in \mathbb{R}\}$ delimited by lower and upper CDF bounds $\underline{F}$ and $\overline{F} : \mathbb{R} \to [0,1]$. It can be represented as the random set $\Gamma : \Omega \to \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$ (i.e. $(\mathcal{F},P_\Gamma)$) defined on $\mathbb{R}$ where $\mathcal{F}$ is the class of focal elements $\Gamma(\alpha) := \langle \underline{F}, \overline{F} \rangle^{(-1)}(\alpha) := [\overline{F}^{(-1)}(\alpha), \underline{F}^{(-1)}(\alpha)]$ for $\alpha \in \Omega$ with $\underline{F}^{(-1)}(\alpha)$ and $\overline{F}^{(-1)}(\alpha)$ denoting the quasi-inverses of $\underline{F}$ and $\overline{F}$ (the *quasi-inverse* of the CDF F is defined by $F^{(-1)}(\alpha) := \inf\{x : F(x) \geqslant \alpha\}$) and P_Γ is specified by (1). This is a good point to mention that