Structural Safety 32 (2010) 411–416

Contents lists available at ScienceDirect

Structural Safety

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

On new cautious structural reliability models in the framework of imprecise probabilities

L.V. Utkin [a], I. Kozine [b,*]

[a] Saint-Petersburg State Forest Technical Academy, Institutski per. 5, St. Petersburg 194021, Russia
[b] Technical University of Denmark, Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Keywords:
Structural reliability models
Imprecise probabilities
Bayesian inference

ABSTRACT

New imprecise structural reliability models are described in this paper. They are developed based on the imprecise Bayesian inference and are imprecise Dirichlet, imprecise negative binomial, gamma-exponential and normal models. The models are applied to computing cautious structural reliability measures when the number of events of interest or observations is very small. The main feature of the models is that prior ignorance is not modelled by a fixed single prior distribution, but by a class of priors which is defined by upper and lower probabilities that can converge as statistical data accumulate. Numerical examples illustrate some features of the proposed approach.

1. Introduction

A probabilistic model of *structural reliability* and safety has been introduced by Freudenthal [1]. Following his work, a number of studies have been carried out to compute the probability of failure under different assumptions about initial information. Briefly the problem of structural reliability can be stated as follows [2]. Let Y represent a random variable describing the *strength* of a system and let X represent a random variable describing the *stress* or load placed on the system. By assuming that X and Y are defined on $\Omega = [x_{\min}, x_{\max}]$ and $\Theta = [y_{\min}, y_{\max}]$, respectively, system failure occurs when the stress on the system exceeds the strength of the system: $\Psi = \{(x \in \Omega, y \in \Theta) : x \geqslant y\}$. Here Ψ is a region where the combination of system parameters leads to an unacceptable or unsafe system response. Then the *reliability* of the system is determined as $R = \Pr\{X \leqslant Y\}$, and the *unreliability* is determined as $U = \Pr\{X > Y\} = 1 - R$. Generally, we have a function $h(\mathbf{X})$ of n random variables $\mathbf{X} = (X_1, \ldots, X_n)$ characterizing the stress. In this case, $\Psi = \{(\mathbf{x} \in \Omega^n, y \in \Theta) : h(\mathbf{x}) \geqslant y\}$, where $\mathbf{x} = (x_1, \ldots, x_n)$ is a value of $\mathbf{X}$.

Uncertainty of parameters in engineering design has been successfully modelled by means of *interval analysis* [3,4]. Several papers, [5,6], describe the *fuzzy set* and *possibility* theories to cope with a lack of complete statistical information about the stress and strength. Several structural problems solved by means of *random set theory* have been scrutinized in [7–10]. The random set theory provides us with an appropriate mathematical model of uncertainty when the information about the stress and strength is not complete or when the result of each observation is not point-valued but set-valued, so that it is not possible to assume the existence of a unique probability measure.

A more general approach to the structural reliability analysis was proposed in [11,12]. This approach utilizes a wider class of partial information about structural parameters, which includes possible data about probabilities of arbitrary events, expectations of the random stress and strength and their functions. The main idea proposed in [11] is to use *imprecise probability theory* [13], whose general framework is provided by upper and lower previsions (expectations). They can model a variety of kinds of uncertainty, partial information, and ignorance. At the same time, this approach presupposes the existence of some probabilistic measures (precise or imprecise) of strength and stress. Often such characteristics do not exist and the analyst has only some judgments or measurements (observations) of values of stress and strength themselves. Therefore, the first question is how to utilize the available information and to compute the structural reliability. The second question is what to do if the number of judgments or measurements is very small.

In this paper we describe new models for computing structural reliability based on measurements of values of stress and strength and taking account of the fact that the number of observations may be rather small. The approach to developing the models is based on using the *imprecise Bayesian inference models* [14]. These models provide a rich supply of coherent imprecise inferences that are expressed in terms of posterior upper and lower probabilities. The probabilities are initially vacuous, reflecting prior ignorance, become more precise as the number of observations increase. All

* Corresponding author. Tel.: +45 4525 4548; fax: +45 4593 3435.
E-mail addresses: lev.utkin@mail.ru (L.V. Utkin), igko@man.dtu.dk (I. Kozine).

0167-4730/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.strusafe.2010.08.004

the models described in this paper are based on the assumption that we find ourselves in the state of complete ignorance prior to observations, which means that prior probabilities are vacuous, i.e., the lower is equal to 0 while the upper is equal to 1.

In Sections 2–4 some basic and relevant concepts for the developed imprecise structural reliability models are briefly described. Sections 5–8 describe the four new models. Some concluding notes are given in Section 9.

2. Failure probability under complete information about the stress and strength

Complete information about the stress and strength means that there exist precisely known probability distribution functions or densities of all random variables specifying the structural reliability. Moreover, the second condition is that all the random variables are statistically independent. The latter assumption is rather restrictive and the results derived based on it should be used with a great care if there are grounds to assume that the variables are dependent. There are two pathways to overcome this assumption. One is to assume complete ignorance of whether the variables are dependent or not. This is how it was done in [11] for the case of some known interval-valued probability measures of stress and strength. The other one is to assess the correlation between the variables that will enter the expressions explicitly. The either way is feasible but difficult to realise. Numerical computations are probably the only possibility to obtain the results. In the current study we did not attempt to explore this problem deeper.

Let ρ_i be the probability density function of a random variable X_i, $\rho_{\mathbf{x}}$ be the joint density of a random vector $\mathbf{X}$. Due to independence of random variables, we can write $\rho_{\mathbf{x}}(\mathbf{x}) = \rho_1(x_1)\cdots\rho_n(x_n)$, $\forall \mathbf{x} \in \Omega^n$.

Let ρ_Y be the density function of a random variable Y. The cumulative distribution functions of the considered random variables are denoted as F_i, $F_{\mathbf{x}}(\mathbf{x})$, F_Y. Then the structural unreliability is defined as the $(n+1)$-multiple integral on the set Ψ, i.e.

$$U = \int \cdots \int_{\Psi} \rho_{\mathbf{x}}(\mathbf{x})\rho_Y(y)\,\mathbf{dx}\,\mathbf{d}y.$$

For every vector $\mathbf{x}$, we can determine a set of values y which lead to system failure, i.e. $(\mathbf{x}, y) \in \Psi$. If we denote this set $\Psi(\mathbf{x})$, then it is obvious that $\Psi(\mathbf{x}) = [0, h(\mathbf{x})]$. Now the integral can be rewritten as follows:

$$\begin{aligned} U &= \int \cdots \int_{\Omega^n} \int_{\Psi(\mathbf{x})} \rho_{\mathbf{x}}(\mathbf{x})\rho_Y(y)\,\mathbf{dx}\,\mathbf{d}y \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \rho_{\mathbf{x}}(\mathbf{x}) \left(\int_0^{h(\mathbf{x})} \rho_Y(y)\,\mathbf{d}y \right) \mathbf{dx} \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \rho_{\mathbf{x}}(\mathbf{x}) F_Y(h(\mathbf{x}))\,\mathbf{dx}. \end{aligned}$$

The failure probability U can be regarded as the expectation of function $F_Y(h(\mathbf{X}))$ of a random vector $\mathbf{X}$ having the density function $\rho_{\mathbf{x}}$, i.e. $U = \mathbb{E}_{\rho_X} F_Y(h(\mathbf{X}))$.

Stress $\mathbf{X}$and strength Y can be both negative and positive depending on application. For the purpose of presentation we consider nonnegative variables without loss in generality.

As an example, when $h(\mathbf{x}) = x$ and $n = 1$, we get the well-known expressions for the failure probability

$$\begin{aligned} U &= \int_0^{\infty} \rho_X(y)F_Y(y)\,\mathbf{d}y = \int_0^{\infty} \rho_Y(y)(1 - F_X(y))\,\mathbf{d}y \\ &= 1 - \int_0^{\infty} \rho_Y(y)F_X(y)\,\mathbf{d}y. \end{aligned} \tag{1}$$

3. Bayesian inference

If we assume that a random variable has a probability distribution with the vector of unknown parameters $\mathbf{b}$, then these parameters could be regarded as random variables with some probability density $\pi(\mathbf{b})$. In this case, the Bayesian approach could be applied for computing distribution function of the random variable $F(x) = \int_{\Theta} F(x|\mathbf{b}) \cdot \pi(\mathbf{b})\mathbf{db}$. Here Θ is the set of values of $\mathbf{b}$.

More generally, the Bayesian approach is concerned with generating the posterior distribution of the parameters of interest given both the data and some prior density for these parameters. Suppose that the prior distribution $\pi(\mathbf{b})$ represents our uncertainty about the possible values of $\mathbf{b}$ prior to collecting any information about the values of $\mathbf{x} = (x_1, \ldots, x_n)$ that could, for example, be interpreted as observed successive intervals between failures. Let $p(\mathbf{x}|\mathbf{b})$ be the probability density function for the observed data x given $\mathbf{b}$. Then the posterior distribution $\pi(\mathbf{x}|\mathbf{b})$ as the conditional distribution of $\mathbf{b}$ given the observed data x is computed as $\pi(\mathbf{b}|\mathbf{x}) \propto p(x_1|\mathbf{b}) \cdots p(x_n|\mathbf{b}) \cdot \pi(\mathbf{b})$.

The prior distribution is often chosen in such a way to facilitate the calculation of the posterior, especially through the use of *conjugate priors*. When the posterior distribution $\pi(\mathbf{b}|x)$ and the prior distribution $\pi(\mathbf{b})$ both belong to the same distribution family, π is the conjugate prior for p.

A critical feature of any Bayesian analysis is the choice of a prior or the choice of the parameters of the prior probability distribution, especially when one is completely ignorant about the parameters. In this case, a non-informative prior has to be constructed. There is a set of methods for determining the non-informative priors in the literature. But every method has some arguable shortcomings and can be used only in special cases.

Another approach to defining non-informative prior models, which challenges conventional Bayesian analysis and which we advocate, is based on defining a class $\mathcal{M}$ of prior distributions π that is manifested through the lower $\underline{P}$ and upper $\overline{P}$ probabilities of an event A as

$$\underline{P}(A) = \sup\{P_{\pi}(A) : \pi \in \mathcal{M}\},$$
$$\overline{P}(A) = \inf\{P_{\pi}(A) : \pi \in \mathcal{M}\}.$$

As pointed out by Walley [14], the class $\mathcal{M}$ under some conditions is "Not a class of reasonable priors, but a reasonable class of priors". This means that each single member of the class is not a reasonable model for prior ignorance, because no single distribution can model ignorance satisfactory. But the whole class is a reasonable model for prior ignorance. When we have little prior information, the upper probability of a non-trivial event should be close to one and the lower probability should be close to zero. This means that the prior probability of the event may be arbitrary from 0 to 1. The implication of using vacuous priors is twofold. On the one hand, quite a large number of observations will be needed to reach a high enough precision in the posterior probabilities to make them useful in practice. On the other hand, it is compelling to avoid introducing any unjustifiable assumptions prior to observations and specifying probability bounds on priors.

Following this setting the models described in this paper have been developed.

4. Reliability under imprecise information in the form of p-boxes

The outcome of the imprecise Bayesian inference is the lower $\underline{F}$ and upper $\overline{F}$ distribution function of a random variable X such that $\underline{F}(x) \leqslant F(x) \leqslant \overline{F}(x),\ \forall x \in \mathbb{R}$. These bounding functions define a set $\mathcal{M}$ of distributions called also *p-box*.