Structural Safety 64 (2017) 54–61

Contents lists available at ScienceDirect

Structural Safety

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

A general framework for the estimation of analytical fragility functions based on multivariate probability distributions

Irmela Zentner *,[1]

IMSIA, Institute for Mechanical Sciences and Industrial Applications, UMR 9219 CNRS-EDF-CEA-ENSTA ParisTech, France
EDF R&D Lab Saclay, France

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form 20 September 2016
Accepted 22 September 2016
Available online 1 October 2016

Keywords:
Fragility curve
Copula
Performance based earthquake engineering
Conditional probability
Time history analysis

ABSTRACT

A fragility curve is a function that expresses the probability of failure of a structure or component as a function of the intensity of external aggression. This paper proposes a general framework for the development of analytical fragility functions from data based on the copula approach. Such a model allows for any kinds of marginal distributions and dependence structures so that it can be applied to various types of fragility data, analytical or empirical. The fragility function is then derived from the joint distribution of intensity and damage measures. The Bayesian information criterion is used to select the most plausible model among the candidate joint distributions, given the data. The practical implementation of the methodology is illustrated by an analytical test case and by the evaluation of seismic fragility curves for a reinforced concrete building. Several candidate marginal distributions, in agreement with the nature and the physical properties of the variables (e.g. common intensity and damage measures take only positive values) are evaluated. In particular, seismic intensity measures are lognormal random variables according to seismological models. This paper is focused on bivariate distributions but the case of vector valued intensity measures can be treated accordingly.

1. Introduction

A fragility curve is a function that expresses the probability of failure of a structure or component as a function of the intensity of external aggression. The intensity of the aggression is characterized by a ground motion Intensity measure (IM). The most popular IM for the evaluation of fragility functions are the Peak Ground Acceleration (PGA) and the Pseudo-Spectral Acceleration (PSA) at a fundamental frequency of the structure. In the recent years, various methods to determine fragility curves have been proposed in the literature. The increased interest is, among others, motivated by the ongoing implementation of performance based earthquake engineering (PBEE) procedures in civil engineering design. In the US, the Applied Technology Council (ATC) [1] implemented the PBEE methodology to be used for civil structures, in agreement with the concepts developed at the Pacific Earthquake Engineering Research (PEER) Center [2]. A comprehensive set of procedures to estimate fragility functions from data is proposed in references [3,1,4]. In particular, the fragility analysis methods depend on the nature and number of available data. Epistemic uncertainties are paid particular attention to in Liel et al. [5] and Celik & Ellingwood [6]. In the nuclear sector, the seismic Probabilistic Risk Assessment (PRA) dates back to the 70s and is now widely used for the evaluation of plant safety [7].

Classical methods for the evaluation of fragility curves, when input–output samples are available, are [8]:

- Method of moments e.g. [9,3]. It can be used when a sample of capacities has been observed: this is the case in incremental dynamic analysis [10] or when qualification tests until failure are available.
- Maximum likelihood method, for example [11,12]. This method requires demand data before and beyond failure.
- Fitting of a linear seismic demand model in the log-scale by means of linear regression, e.g. [13,14]. This method is applicable for any kind of data but it means extrapolation of the behavior beyond yield if no failure is observed.

Both in PBEE and seismic PRA, the seismic fragility curves are modeled as lognormal cumulative distribution functions.

In a recent contribution [15], Gaussian kernel smoothing has been proposed to evaluate non parametric fragility functions. This allows to develop functional forms without any assumption on the

* Address: IMSIA, Institute for Mechanical Sciences and Industrial Applications, UMR 9219 CNRS-EDF-CEA-ENSTA ParisTech, France.

E-mail address: irmela.zentner@edf.fr

[1] 7, bvd Gaspard Monge, Palaiseau, France.

http://dx.doi.org/10.1016/j.strusafe.2016.09.003
0167-4730/

probability distribution. One botttleneck of this methodology is the dependence of the estimations on the choice of the bandwidth. The trade-off between robustness, over-fitting and accuracy is indeed a major issue in the framework of Gaussian kernel smoothing. Generally speaking, a non parametric approach allows to overcome model assumptions (such as lognormal probability distribution) but at the expense of robustness. Parametric models posses the advantage of reduced numerical cost and improved robustness. If the parametric model is adequate, then the precision is higher at the same cost (in terms of data). For example, it is well known that extreme value distributions are good candidates for modeling peak responses and common seismic motion intensity measures are lognormal according to Ground Motion Prediction Equations (GMPE) [16]. Any multivariate joint distribution can be written in terms of the marginal distribution functions and a copula which describes the dependence structure between the variables. This approach is pursued in this paper: the fragility function is derived according to its intrinsic definition as a conditional probability from the joint distribution of the IM and the damage measure (DM), where the latter is defined by the marginal distributions and a copula. Copula models have been applied for reliability analysis and hazard assessment by various authors in the recent years [17–24]. In a recent contribution, Goda [17] evaluates the joint distribution of peak and residual displacement seismic demands of a simple structure. Moreover, a comprehensive introduction to the copula technique is provided by that author. Tang et al. [19,20] investigate the impact of different copula models when constructing bivariate distributions for parallel systems and slope reliability analysis.

This paper proposes a new, general framework for the development of analytical fragility functions from data (continuous or discrete, simulated or observed) based on the copula approach. In what follows, we first give a short overview over common lognormal fragility function fitting with a focus on analytical approaches that do not require the scaling of accelerograms. Secondly, the copula based fragility function estimation is presented. The approach is illustrated by an analytical test case and the application to seismic analysis of a reinforced concrete building. For this purpose, a Gaussian copula model is chosen.

2. Computation of analytical lognormal fragility curves

A fragility curve is a function that expresses the probability that a structure or component submitted to an excitation of intensity α, fails or reaches a previously defined damage state. As it is common use, in what follows, random variables will be denoted by capital letters while their realization are lower case. More precisely, if damage or failure is assumed to occur when the design variable or damage measure D exceeds a threshold d_0 (which can be itself a random variable), then the fragility curve is given by the following conditional probability:

$$P_f(\alpha) = P(\mathcal{D} > d_0 \mid IM = \alpha). \tag{1}$$

When discrete damage states are considered, then the damage measure boils down to a Bernoulli random variable X, taking the value $x = 1$ when the damage state is reached and otherwise $x = 0$. This is the case, for example, when failure is defined as the buckling of a structural element or for damage data from post-earthquake field observations, where only the membership to a certain damage state (e.g. "no damage", "slight damage", ..., "failure") can be assigned. In the latter framework, the fragility function formally reads:

$$P_f(\alpha) = P(x = 1 \mid IM = \alpha). \tag{2}$$

When a lognormal model is adopted, then the fragility curve can be expressed as:

$$P_f(\alpha) = \Phi\left(\frac{\ln(\alpha/A_m)}{\beta}\right), \tag{3}$$

where Φ denotes standard normal (Gaussian) cumulative distribution function (cdf) with standard deviation β and mean value ln A_m. The variable A_m is the median of the lognormal distribution and is also called median capacity.

2.1. Maximum likelihood estimator

The maximum likelihood method is of particular interest when the damage (or failure) is not defined by a continuous but a discrete state variable. In this framework, the dataset consists of N pairs $\{\alpha_i, x_i\}$. If the failure or damage criterion is reached, then $x_i = 1$ and otherwise (no failure) it equals zero: $x_i = 0$. These events arrive with probability $P_f(\alpha)$ and respectively $1 - P_f(\alpha)$. For the lognormal model, the latter can be calculated by expression (3). The likelihood function for this problem reads:

$$\mathcal{L} = \prod_{i=1}^{N} [P_f(\alpha_i)]^{x_i} [1 - P_f(\alpha_i)]^{1-x_i} \tag{4}$$

The estimators of parameters A_m and β are solution of the following optimization problem:

$$(\hat{A}_m, \hat{\beta}) = \arg\min_{A_m,\beta}(-\ln\mathcal{L}). \tag{5}$$

As pointed out in [4], the maximum likelihood estimator is equivalent to generalized linear regression with a Probit link function. Prior information on median structural capacity A_m or standard deviation β can be introduced by Bayesian updating, e.g. [25,26]. This is the case, if either an expert opinion or previous analysis are available. The profiled likelihood allows to evaluate confidence intervals [27,28]. However, when the damage measure is a continuous design variable, then information is lost by the transformation to the discrete damage indicator. Indeed, only the information on damage state membership is used while the distance to the threshold is disregarded. Moreover, this methodology can not be applied when only few or no failures (or occurrences of the damage state) are observed.

2.2. Linear regression

The linear regression approach [13,14] assumes a continuous damage measure of the form

$$D = b\alpha^c\eta \tag{6}$$

where η is a lognormal random variable with unit median and log-standard deviation equal to β. The linear regression is performed in log space:

$$\ln D = \ln b + c\ln\alpha + \epsilon, \tag{7}$$

where ϵ is a centered normal random variable with standard deviation σ_ϵ. The regression parameters $\ln b$, c and σ_ϵ read: $c = r\sigma_{\ln D}/\sigma_{\ln IM}$ and $\ln b = \mu_{\ln D} - r\,\sigma_{\ln D}/\sigma_{\ln IM}\mu_{\ln IM}$, where $\mu_{\ln D}$, $\sigma_{\ln D}$ and $\mu_{\ln IM}$, $\sigma_{\ln IM}$ are the mean and standard deviations of variables $\ln D$ and $\ln IM$ and r is the linear correlation coefficient. They have to be approximated by their statistical estimator. Denoting by e_i the regression residual, an unbiased estimation of σ_ϵ^2 is given by the formula [29]:

$$s_\epsilon^2 = \frac{1}{n-2}\sum_n e_i^2.$$