Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 105 (2014) 132–143
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original Article

The ordering importance measure of random variable and its estimation

Lijie Cui [a,b,*], Zhenzhou Lu [b], Pan Wang [b], Weihu Wang [b]

[a] *School of Equipment Management and Safety Engineering, Air Force Engineering University, Xi'an, Shaanxi 710038, PR China*
[b] *School of Aeronautics, Northwestern Ploytechnical University, Xi'an, Shaanxi 710072, PR China*

Received 3 August 2010; received in revised form 18 June 2014; accepted 19 June 2014
Available online 27 June 2014

Abstract

Based on the importance analysis with independent random variables, an ordering importance measure is proposed to evaluate the effects of variables on the uncertainty of output response first, in which not only independent random variables but also correlated ones are included, and it provides a theoretical basis to improve a system or a model. Secondly, the sampling strategy of the conditional probability density function is provided by the Copula transformation, which could solve the vital problem of the importance analysis effectively with correlated random variables. What's more, Due to the low efficiency and tremendous computational cost of the Monte Carlo method, the probability density function evolution method (PDEM) is utilized to solve the ordering importance measure. Finally, some examples in cases of independent random variables and correlated random variables are employed to demonstrate the feasibility and reasonability of the proposed measure, test the precision of the probability density function evolution method, even.

Keywords: Importance analysis; Correlated variables; Copula transformation; Moment-independent importance measure; Probability density function evolution

1. Introduction

In several disciplines, engineers benefit from the utilization of quantitative models. These models are complex machines due to the intricacy of the phenomena under investigation. In utilizing these models, one needs to account for uncertainties generated by lack of knowledge in work conditions, environments, material properties (etc).In reliability engineering, the Bayesian approach is often assumed to uncertainty analysis. Thus, one can model uncertainty by assigning distributions to the uncertain variables. In reliability analysis, the uncertainty which influences the system behaviors or the model response can be assumed as random variable [29]. Generally, the random variables consist of physical dimensions, material characters, and loads et al. Identifying and representing the effects of random variables

* Corresponding author at: School of Equipment Management and Safety Engineering, Air Force Engineering University, Xi'an, Shaanxi 710038, PR China. Tel.: +86 13488192369

E-mail addresses: cuilijie@mail.nwpu.edu.cn (L. Cui), zhenzhoulu@nwpu.edu.cn (Z. Lu), panwang@mail.nwpu.edu.cn (P. Wang), wwh2286@163.com (W. Wang).

http://dx.doi.org/10.1016/j.matcom.2014.06.003

on the uncertainty of the output response are defined as uncertainty importance analysis. According to the Pareto rules [12], less than 20% of the elements in the risk model contributes to more than 80% of the total risk, thus it is a vital task to identify the key variable that influences the most to the output response. A number of importance measures and techniques have been suggested in the literature, e.g. Helton et al. [13], Saltelli [26], Iman and Hora [16], Sobol [27], Miman and Pohl [23], Porn [24], de Rocquigny et al. [25], Chun et al. [9] and Borgonovo [3,4]. Totally, the presented importance measures can be separated into three categories:

i. Nonparametric technique (input-output correlation)
ii. Variance-based importance measure
iii. Moment-independent importance measure

Concretely, the importance measure based on the nonparametric technique is built on evaluating the correlation of the inputs and the output, which uses the correlation to evaluate the effects of input variables on the uncertainty of output. The importance measure based on variance reflects the effect reducing of the variance of output by eliminating the uncertainty of one or more variables. Though the Variance-based importance measure has been widely applied in engineering, it inevitably loses some statistical information due to that it uses the variance to represent the whole uncertainty of the output. Cox [10] and next Huber [15] illustrate the pitfalls that "*mean-variance decision-making violates the principle that a rational decisionmaker should prefer higher to lower probabilities of receiving a fixed gain, all else being equal*". Recently, Borgonovo et al. [3,4] proposed the moment-independent importance measure, which reflects the influences of input variables on the uncertainty of the output at the perspective of the probabilistic distribution of the output, and it is superior to the above two measures in the respect of reflecting the global statistics property of the output.

Definitions and estimations presented in [3,4] assume that random variables are independent from each other, and moment independent importance measures with correlated inputs are defined in [5,2]. In fact, both independent and correlated input variables are concerned in engineering. Thus, it is necessary to propose a universal and direct importance measure and a common algorithm to deal with problems with independent and correlated random variables.

To overcome drawbacks of the current importance measures, the moment-independent importance measure is extended to the importance analysis with correlated random variables without additional computational cost, and an ordering importance measure is proposed on the probabilistic distribution of the ranking of the moment-independent measure. Furthermore, the Monte Carlo method (MCM) utilized in obtaining the current importance measures is simple to program, but its computational cost is so tremendous that it is unbearable in complex systems and tests. The probability density evolution method (PDEM) [7,8] recently applied in analyzing the stochastic dynamic response is introduced into solving the ordering importance measure, which can improve the computational efficiency largely without any loss of precision.

2. Definition of the ordering importance measure

2.1. The moment-independent importance measure

The moment-independent importance measure satisfies four requirements that an importance measure should meet, i.e. "*global, quantitative, model free and moment independent*", so it has been the focus of the importance analysis.

Given a performance response function $Y=g(X_1, X_2, \cdots X_n)$, where Y is the model output, and $X_1, X_2, \ldots, X_n$ are the input variables with uncertainty. Denote $f_Y(y)$ and $f_{Y|X_i}(y)$ as the unconditional probability density function (PDF) and conditional PDF respectively, while $f_{Y|X_i}(y)$ can be obtained by fixing the input variable X_i at a realization value. The absolute value of the difference between $f_Y(y)$ and $f_{Y|X_i}(y)$, which is denoted as $s(x_i)$, can represent the effect of X_i on the distribution of Y when X_i takes values x_i, and the geometrical sketch map of $s(x_i)$ is shown in Fig. 1. When y varies from $-\infty$ to $+\infty$, the cumulative effect of X_i on the distribution of Y can be measured by the shadow area $s(x_i)$ shown in Fig. 1.

It is given by the following integral

$$s(x_i) = \int_{-\infty}^{+\infty} \left| f_Y(y) - f_{Y|X_i}(y) \right| dy \tag{1}$$