Structural Safety 51 (2014) 13–22

Contents lists available at ScienceDirect

Structural Safety

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

Probabilistic importance analysis of the input variables in structural systems

Cui Lijie, Lu Zhenzhou *, Zhou Changcong

School of Aeronautics, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, PR China

ARTICLE INFO

Article history:
Received 6 June 2012
Received in revised form 13 May 2013
Accepted 20 May 2014
Available online 24 June 2014

Keywords:
Importance analysis
Variance
Failure probability
Correlation
Copula transformation
Probability density evolution method

ABSTRACT

To measure the effects of input variables' realization on variance of the output performance function and on the failure probability of structural system, two new probabilistic importance measures (PIMs) are defined. As an input variable takes its realization according to its probability distribution, the two PIMs can quantify the possibility of reducing the variance of the output performance function and the possibility of improving the structural system reliability, respectively. After the properties of the PIMs are illuminated and proved in detail, a solution based on the probability density function evolution method (PDEM) is constructed to evaluate the PIMs. The solution is used to solve the PIMs with correlated input variables based on the Copula transformation. Examples demonstrate that the proposed solution on the PDEM can improve the computational efficiency greatly with acceptable precision, and the solution based on the conjunction of the PDEM and the Copula transformation can effectively solve the PIMs with correlated input variables.

1. Introduction

Due to the universal existence of uncertainty in engineering, the propagation of uncertainty is becoming a main concern in the literature. To properly address this concern, it is highly necessary to recognize the effects of the input uncertainty on the output uncertainty, and this is also named as the importance analysis of input variables [1], which is the focus of this contribution. At present, Helton et al. [2], Saltelli [3], Iman and Hora [4], Sobol [5], Miman and Pohl [6], Pörn [7], de Rocquigny et al. [8], Chun et al. [9] and Borgonovo [10] proposed their importance measures (IMs) to represent the effects of the input variables, and the corresponding importance analysis methods have been constructed. These IMs can be classified into three categories: the nonparametric technique, the variance-based IM and the moment-independent IM. By analyzing these existing IMs, it can be found that most of them measure and compare the average effects of the input variables' realizations on the different statistic characteristics of the output performance. The average values can measure effects of the input variables' realizations. However, the possibilities of reducing the variance of the output performance and improving the reliability cannot be given by the average effects. Therefore, in order to measure these possibilities as a result of realizations of the input variables, two probabilistic IMs (PIMs) are defined in the contribution, and their properties are discussed. After the definition and properties of the PIMs are introduced, the solutions of the PIMs are suggested. By transforming the solving of the PIMs into the traditional reliability analysis, a probability density evolution method (PDEM) [13] is employed to obtain the results of the PIMs. Furthermore, solutions of the PIMs for the structural system with correlated input variables are provided by use of the conditional sampling on the Copula transformation and its inverse transformation.

In this paper, before two PIMs are defined with their properties discussed, we first review the variance-based IM and the moment-independent IM on the failure probability in Section 2. Section 3 provides a highly efficient solution for the PIMs. In Section 4, two numerical examples and an engineering example are used to illustrate the effectiveness of the PIMs and the PDEM based solution. Section 5 is the closure with some conclusions.

2. Two PIMs

Suppose $Y = g(\boldsymbol{X}) = g(X_1, X_2, \ldots, X_n)$ is the output performance function of the structural system, where $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ is a n-dimension random input vector. Because the effects of the input variables on the variance of Y and on the failure probability are more remarkable in engineering, the variance-based IM [5] and

* Corresponding author. Tel./fax: +86 29 88460480.
E-mail addresses: cuilijie@mail.nwpu.edu.cn (L. Cui), zhenzhoulu@nwpu.edu.cn (Z. Lu), changcongzhou@mail.nwpu.edu.cn (C. Zhou).

http://dx.doi.org/10.1016/j.strusafe.2014.05.006
0167-4730/

the moment-independent IM on the failure probability [11] are presented respectively.

2.1. The variance-based IM and the moment-independent IM on the failure probability

(1) The variance-based IM [5]

The variance-based IM of an input variable or a group of input variables is defined as the ratio of the variance of the output expectation given $\boldsymbol{X_I}$ and the unconditional output variance, i.e.

$$S_{\boldsymbol{I}} = \frac{Var(E_{\boldsymbol{X_I}}(Y|\boldsymbol{X_I}))}{Var(Y)} = \frac{Var(Y) - E_{\boldsymbol{X_I}}[Var(Y|\boldsymbol{X_I})]}{Var(Y)} = \frac{E_{\boldsymbol{X_I}}[Var(Y) - Var(Y|\boldsymbol{X_I})]}{Var(Y)} \quad (1)$$

where $S_{\boldsymbol{I}}$ is the variance-based IM. $\boldsymbol{X_I}$ denotes a variable X_i or a group of variables $(X_{i_1}, \ldots, X_{i_g})(1 \leqslant i_1 \leqslant \cdots \leqslant i_g \leqslant n)$. $Var(Y)$ is the unconditional variance of Y, and $Var(Y|\boldsymbol{X_I})$ is the conditional variance of Y given $\boldsymbol{X_I}$.

(1) The moment-independent IM on the failure probability [11]

Noting that the failure probability is an important index of the structural system, the moment-independent IM $\eta_{\boldsymbol{I}}$ defined in Ref. [11] represents the effect of $\boldsymbol{X_I}$ on the failure probability,

$$\eta_{\boldsymbol{I}} = \frac{1}{2}E_{\boldsymbol{X_I}}[|P_f - P_{f|\boldsymbol{X_I}}|] = \frac{1}{2}\int_{-\infty}^{+\infty}\left|\int_F f_Y(y)dy - \int_F f_{Y|\boldsymbol{X_I}}(y)dy\right| f_{\boldsymbol{X_I}}(\boldsymbol{x_I})d\boldsymbol{x_I} = \frac{1}{2}\int_{-\infty}^{+\infty}|P_f - P_{f|\boldsymbol{X_I}}| f_{\boldsymbol{X_I}}(\boldsymbol{x_I})d\boldsymbol{x_I} \quad (2)$$

where F denotes the failure domain defined by the output performance function $F = \{\boldsymbol{X} : Y = g(\boldsymbol{X}) \leqslant 0\}$. P_f is the unconditional failure probability, and $P_{f|\boldsymbol{X_I}}$ is the conditional failure probability given $\boldsymbol{X_I}$. $f_Y(y)$ is the unconditional probability density function (PDF) of Y, $f_{Y|\boldsymbol{X_I}}(y)$ is the conditional PDF of Y given $\boldsymbol{X_I}$, and $f_{\boldsymbol{X_I}}(\boldsymbol{x_I})$ is the joint PDF of $\boldsymbol{X_I}$, and $d\boldsymbol{x}_i = dx_{i_1}dx_{i_2}...dx_{i_g}$.

2.2. Discussion about the above IMs

In Eq. (1), it can be found that $S_{\boldsymbol{I}}$ reflects the average effect of the realization $\boldsymbol{X_I}$ on the variance of the output performance function Y. Because $S_{\boldsymbol{I}}$ is defined as the average difference between $Var(Y)$ and $Var(Y|\boldsymbol{X_I})$, the partial effect of $\boldsymbol{X_I}$ on $Var(Y)$ is possibly concealed since the values of $Var(Y) - Var(Y|\boldsymbol{X_I})$ are possible positive or negative in the random domain of $\boldsymbol{X_I}$. As is shown in Fig. 1, when the input variables X_i and X_j uniformly distribute in the interval $[a, b]$, the difference between $Var(Y|X_i)$ and $Var(Y)$ is completely counteracted as $X_i \sim U(a, b)$. For the case shown in Fig. 1, it is obviously that $S_i = 0$ but $S_j > 0$, which is to say the effect of X_j on the variance of Y is more significant than that of X_i, and X_i has no effect on the variance of Y. Obviously, this conclusion generated from the definition shown in Eq. (1) is insufficient to evaluate the contributions of input variables X_i and X_j to the variance of the performance function. Because $Var(Y) - Var(Y|\boldsymbol{X_I})$ is random and it varies with value of X_i, the average only reflects the partial information of the random characteristics of $Var(Y) - Var(Y|\boldsymbol{X_I})$.

Similar problem also exists in the definition shown in Eq. (2). There is not counteraction effect of the variables' realizations on the failure probability in the definition of $\eta_{\boldsymbol{I}}$, however, since $|P_f - P_{f|\boldsymbol{X_I}}|$ is a random variable of which the values vary with $\boldsymbol{X_I}$, it is also inappropriate to evaluate the effect only from the average by Eq. (2), and it is shown in Fig. 2.

2.3. Two PIMs

The above IMs quantify the average contributions of input variable to different statistical characteristics of the output performance function. They cannot reflect the possibilities of reducing the variance and improving the reliability if the input variable takes its realized value. To analyze those possibilities, two probabilistic IMs, of which one is on the variance and the other is on the failure probability, are proposed as follows.

$$\varepsilon_{\boldsymbol{I}}^{Var} = P\{Var(Y) - Var(Y|\boldsymbol{X_I}) > 0\} \quad (3)$$

$$\varepsilon_{\boldsymbol{I}}^{P_f} = P\{P_f - P_{f|\boldsymbol{XI}} > 0\} \quad (4)$$

where the PIM $\varepsilon_{\boldsymbol{I}}^{Var}$ reflects the possibility that the conditional variance is smaller than the unconditional variance as $\boldsymbol{X_I}$ taking its realized values according to the PDF $f_{\boldsymbol{X_I}}(\boldsymbol{x_I})$, and the PIM $\varepsilon_{\boldsymbol{I}}^{P_f}$ reflects the possibility that the conditional failure probability is smaller than the unconditional failure probability as $\boldsymbol{X_I}$ taking its realized values according to the PDF $f_{\boldsymbol{X_I}}(\boldsymbol{x_I})$. The most notable difference between the PIMs and the two IMs in Section 2.1 lies in that the former aims at the possibility of reducing the output variance or failure probability by fixing an input at its random values, while the IMs aims at the expectation of such reduction.

Similarly, the PIM $\varepsilon_{\boldsymbol{J}|\boldsymbol{I}}^{Var}$ can be defined to reflect the possibility that the conditional variance $Var(Y|\boldsymbol{X_I}\boldsymbol{X_J})$ given $\boldsymbol{X_I}$ and $\boldsymbol{X_J}$ is smaller than the conditional one $Var(Y|\boldsymbol{X_I})$ given $\boldsymbol{X_I}$, and the PIM $\varepsilon_{\boldsymbol{J}|\boldsymbol{I}}^{P_f}$ can be defined to reflect the possibility that the conditional failure probability $P_{f|\boldsymbol{X_I}\boldsymbol{X_J}}$ given $\boldsymbol{X_I}$ and $\boldsymbol{X_J}$ is smaller than the conditional one $P_{f|\boldsymbol{X_I}}$ given $\boldsymbol{X_I}$, i.e.

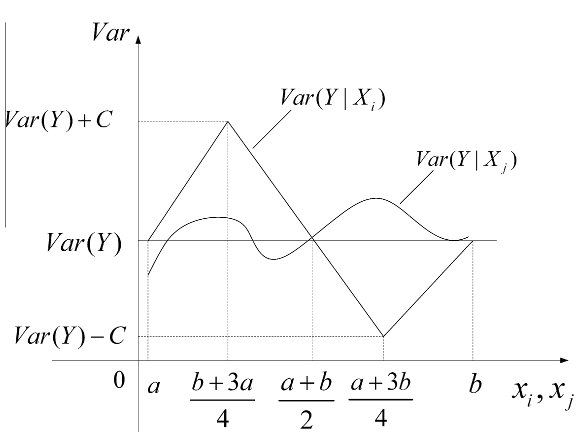

Fig. 1. Sketch map of the conditional variance varying with the input variables (where *a*, *b*, *c* are positive real).

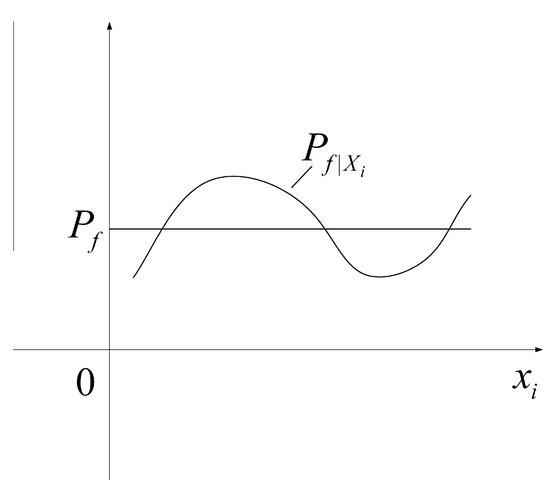

Fig. 2. Sketch map of the conditional failure probability varying with the input variable.