Reliability Engineering and System Safety 169 (2018) 437–450

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Copula-based decomposition approach for the derivative-based sensitivity of variance contributions with dependent variables

Pan Wang [a,*], Zhenzhou Lu [b], Kaichao Zhang [b], Sinan Xiao [b], Zhufeng Yue [a]

[a] School of Mechanics and Civil & Architecture, Northwestern Polytechnical University, China
[b] School of Aeronautics, Northwestern Polytechnical University, China

ARTICLE INFO

Keywords:
Variance contribution
Dependence
Copula
Kernel function
SDP

ABSTRACT

Variance-based sensitivity analysis with dependent variables represents how the uncertainties and dependence of variables influence the output uncertainty. Since the distribution parameters of variables are difficult to be given precisely, this work defines the derivative-based sensitivity of variance contribution with respect to the distribution parameters, which reflects how small variation of distribution parameters influences the variance contributions. By introducing the copula functions to describe the dependence of variables, the derivative of variance contributions can be transformed into those of marginal PDF and copula function, which can be defined by kernel function and copula kernel function. Then the derivative-based sensitivity of variance contributions can be decomposed into the independent part and dependent part. Since the derivatives of marginal PDF and copula function can be given analytically, the proposed derivative-based sensitivity can be computed with no additional computational cost, which is seen as the 'by-product' of variance-based sensitivity analysis. To calculate the proposed sensitivity, two computational methods, numerical method and SDP (state dependent parameter) method are presented for comparison. Several examples are used to demonstrate the reasonability of the proposed sensitivity and the accuracy of the applied method.

1. Introduction

Sensitivity analysis (SA) provides an important tool to explore the complex model behaviors, and it is often critical towards understanding the physical mechanisms and modifying the design to mitigate and manage risk [1]. Generally, sensitivity analysis can be classified into two groups: local SA and global SA [2]. Local SA usually investigates how small variations of one or multiple distribution parameters around their reference point changes the value of the output. One classical local SA is the derivative-based SA by defining the derivative of probabilistic statistics of output with respect to the distribution parameters of variables. The main drawbacks of the derivative based SA are that it depends on the choice of the reference point [3] and it can be quite expensive if large number of inputs are considered. Global SA studies how the uncertainty in the output of a computational model can be decomposed according to the input sources of uncertainty [4–7]. Contrary to the local SA, global SA explores the whole range of uncertainty of the model variables by letting them vary simultaneously. At present, the variance-based SA is the most popular global SA technique which has been applied to design under uncertainty problems and are capable of identifying the contributions of any variable [8,9].

Traditionally, the variance-based SA are investigated in the case of independent variables and the well-rounded methodology has been established by many scholars [4–7]. However, in many cases, the marginal variables are dependent with each other, and the dependence of variables may lead to a different ranking of variance-based sensitivity indices. Thus, for the case of dependent variables, Saltelli et al. [10,11] proposed approaches to perform the SA with dependent variables, but these researches only provide an overall sensitivity of one variable, which cannot distinguish the independent or dependent influence of one variable. To make a better understanding of the dependence in variance-based SA, Xu and Gertner [12] and Li et al. [13,14] divided the variance contribution of individual variable into the independent part and the dependent part, but this approach is constructed based on the approximation of linear model. Li et al. [15] decomposed the total variance contribution of one variable or a set of variables into the structural contribution and correlative contribution based on the covariance decomposition, which can deal with both the linear and nonlinear models. Mara and Tarantola [16] proposed a set of variance-based sensitivity indices based on a specific orthogonalization of the variables and ANOVA-representation of model output. Kucherenko et al. [17] proposed the generalized sensitivity indices which can preserve the advantages of the original variance-based sensitivities without necessity of determining functional decomposition or orthogonalization of the variable space. Therefore, for the models with or without the dependence in the variables, a comprehensive research of the variance-based SA has been

* Corresponding author.
E-mail address: panwang@nwpu.edu.cn (P. Wang).

https://doi.org/10.1016/j.ress.2017.09.012
Received 13 January 2017; Received in revised form 20 August 2017; Accepted 26 September 2017
Available online 28 September 2017
0951-8320/

investigated and different sensitivity indices are proposed to measure the contributions of interests.

Nevertheless, it is noticed that in the variance-based SA with or without dependence, the distribution parameters of variables are assumed to be available precisely. But the distribution parameters are always estimated, measured with uncertainty, or guesses [18,19], if the variation of one distribution parameter leads to a considerable change to the variance contributions, the computational results of the variance based SA will be vulnerable and less reliable. Thus, it is significant to identify how the distribution parameters influence the variance contributions. For the models with independent variables, the derivative-based sensitivity are proposed in references [20,21] by defining the derivatives of variance contributions with respect to the distribution parameters, which can provide the information on how small variation of the distribution parameter around the reference point changes the results of the variance contributions. By introducing the kernel function (or score function) of marginal distributions, the computation of the derivatives of variance contributions can be transformed into that of the derivatives of marginal distributions which needs no additionally computational cost. While for the models with dependent variables, the existing variance-based SA are mainly discussed based on the Pearson correlation coefficient [12–17], thus the definition of derivative-based sensitivity of variance contributions brings a challenge that it needs to compute the derivatives of joint distributions and conditional distributions, which are difficult to be achieved.

To overcome this deficiency, this work sets out from the copula theory, which is used as an important approach to describe the dependence of variables for decades [22–27]. Copula functions are used to combine the joint distributions with marginal distributions, thus the derivatives of joint distributions can be transformed into those of copula functions and marginal distributions. Since the common copula function and marginal distribution are always given analytically, their derivatives can be achieved conveniently. Therefore, for the models with dependent variables, the derivative-based sensitivity of variance contributions can be also computed with no additional computational cost. More importantly, the derivative-based sensitivity of variance contributions can be divided into two parts: one depends on the copula function and the other depends on the marginal distributions, which can make a clear insight for the effect of the dependence of variables. While for the computation of the proposed sensitivity, since a great number of approaches have been presented to perform the variance-based SA [5–9], this work mainly use the SDP method, which has been already used to compute the variance-based sensitivity indices with or without the dependence of variables [28,29].

The remainder of this work is organized as follows: Section 2 gives a brief review of variance-based SA and proposes the definition of the derivative-based sensitivity of variance contributions. Section 3 introduces the copula theory first, then makes an analytical derivations for the proposed sensitivity of variance contributions with copula functions. Since the copula function and marginal distributions can be achieved analytically, Section 4 gives the analytical expressions of the derivatives of frequently-used copula functions and marginal distributions. In Section 5, the Monte Carlo method and SDP method are used to compute the proposed sensitivity of variance contributions. In Section 6, numerical examples are first employed to validate the proposed sensitivity and computational methods, then one engineering practice of roof truss is analyzed. Finally, some conclusions are given in Section 7.

2. Sensitivity of variance contributions

2.1. Brief review of variance-based sensitivity

Consider a square integrable function $Y = g(\boldsymbol{X})$ defined in the hypercube H^d, where $\boldsymbol{X} = (X_1, X_2, ..., X_d)$ are d-dimensional independent variables. According to the idea of HDMR (high-dimensional model representation), the function $g(\boldsymbol{X})$ can be expanded into terms of increasing dimensions:

$$g(\boldsymbol{X}) = g_0 + \sum_{i=1}^{d} g_i(X_i) + \sum_{1\le i<j\le d}^{d} g_{ij}(X_i, X_j) + \cdots + g_{1,2,\ldots,d}(X_1, X_2, ..., X_d) \tag{1}$$

This decomposition is unique when all the terms are orthogonal and with an expectation of zero (except for g_0) [4]. As a consequence, these terms can be calculated by the conditional expectations of the model output, which can be given as:

$$\begin{aligned} g_0 &= E(y) \\ g_i &= E(y|x_i) - g_0 \\ g_{ij} &= E(y|x_i, x_j) - g_i - g_j - g_0 \end{aligned} \tag{2}$$

where the higher-order items can be obtained analogously.

In order to qualify the contribution of variables to the uncertainty of output, the variance-based sensitivity indices proposed by Sobol [7] are defined as

$$S_{i_1\cdots i_k} = \frac{V(g_{i_1\cdots i_k})}{V(Y)} \tag{3}$$

where k represents the order of the sensitivity index. Specially, the most popular first-order sensitivity index is given as

$$S_i = \frac{V_i}{V} = \frac{V(g_i)}{V(Y)} = \frac{V(E(Y|X_i))}{V(Y)} \tag{4}$$

where $V_i = V(E(Y|X_i))$ is the first-order variance contribution.

In the case of independent variables, the first-order sensitivity index S_i indicates how much one could reduce, on average, the output variance if X_i could be fixed. However, when the dependence is present among variables, the variance contribution V_i of an individual variable X_i consists of not only the contribution resulting from the variable itself, but also contains the dependent contribution resulting from the dependence between variable X_i and other variables. For this case, a large amount of literature [12–17] is presented to discuss the variance contributions V_i with dependent variables, and make a clear distinction between the independent contribution and dependent contribution of variable X_i. Therefore, this work focuses on investigating the sensitivity of the variance contribution with dependent variables.

2.2. Definition of the sensitivity of variance contribution

As discussed in Section 1, sensitivity of variance contribution can be defined as the derivative of the variance contribution with respect to the distribution parameters. Without any lack of generality, it is supposed that each variable possesses only one distribution parameter μ_i (actually, the common distributions possess more than one parameter, i.e. normal distribution possesses the mean value and standard deviation) in order to simplify the notation in the following.

Before measuring the influences of distribution parameters on the variance contributions, those on the total variance should be investigated at first. This is because the total variance consists of various variance contributions, if the distribution parameters have no significant effect on the total variance, whatever contributions will appear would not matter. Thus, the derivative-based sensitivity of total variance can be defined as

$$SV_{\mu_i} = \frac{\partial V}{\partial \mu_i} \tag{5}$$

The derivative-based sensitivity SV_{μ_i} can identify the influential distribution parameters which are worthy to be investigated further for the variance contributions.

Then, in order to measure the influences of distribution parameters on the variance contributions with dependent variables, the derivative-based sensitivity of variance contribution V_i is defined as

$$SVC^i_{\mu_j} = \frac{\partial V_i}{\partial \mu_j} \tag{6}$$