Reliability Engineering and System Safety 155 (2016) 224–235

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

A surrogate method for density-based global sensitivity analysis ☆

Sharif Rahman

College of Engineering, The University of Iowa, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 6 December 2015
Received in revised form
27 June 2016
Accepted 2 July 2016
Available online 5 July 2016

Keywords:
f-Divergence
f-Sensitivity index
Hellinger distance
Kernel density estimation
Kullback–Leibler divergence
Polynomial dimensional decomposition
Total variational distance

ABSTRACT

This paper describes an accurate and computationally efficient surrogate method, known as the polynomial dimensional decomposition (PDD) method, for estimating a general class of density-based *f*-sensitivity indices. Unlike the variance-based Sobol index, the *f*-sensitivity index is applicable to random input following dependent as well as independent probability distributions. The proposed method involves PDD approximation of a high-dimensional stochastic response of interest, forming a surrogate input–output data set; kernel density estimations of output probability density functions from the surrogate data set; and subsequent Monte Carlo integration for estimating the *f*-sensitivity index. Developed for an arbitrary convex function *f* and an arbitrary probability distribution of input variables, the method is capable of calculating a wide variety of sensitivity or importance measures, including the mutual information, squared-loss mutual information, and $\mathcal{L}_1$-distance-based importance measure. Three numerical examples illustrate the accuracy, efficiency, and convergence properties of the proposed method in computing sensitivity indices derived from three prominent divergence or distance measures. A finite-element-based global sensitivity analysis of a leverarm was performed, demonstrating the ability of the method in solving industrial-scale engineering problems.

1. Introduction

Global sensitivity analysis rooted in the output variance has been a longstanding staple for modeling and simulation of complex systems [1–4]. In many applications, however, the variance provides a restricted summary of output uncertainty. Therefore, sensitivity measures derived solely from the variance should be guardedly interpreted. A more rational sensitivity analysis should account for the entire probability distribution of an output variable, meaning that the probabilistic characteristics above and beyond the variance should be considered [5–9]. Addressing this fundamental concern has triggered the development of a generalized density-based sensitivity index, referred to as the *f*-sensitivity index, which is founded on the well-known *f*-divergence between conditional and unconditional output probability measures [6]. Unlike the variance-based Sobol index [1], the *f*-sensitivity is applicable to random input following dependent as well as independent probability distributions. Since the class of *f*-divergences supports numerous divergence or distance measures, a bevy of density-based sensitivity measures are possible, including the mutual information [10], squared-loss mutual information [11], and $\mathcal{L}_1$-distance-based importance measure [5], to name a few, providing diverse choices to sensitivity analysis. A few researchers have applied existing divergence or distance measures for sensitivity analysis of engineering systems [12–15].

While the formulation of the *f*-sensitivity index is not overly complicated, its calculation in general is more intricate than the calculation of the variance-based sensitivity index. This is chiefly because the probability density functions required in defining the convex function *f* are harder to estimate than the variance. If the function *f* is already selected, resulting in a specific sensitivity index, then one can exploit the functional structure of *f* to devise accurate and efficient methods of calculation. This is exemplified in the works of Borgonovo [5], Liu and Homma [16], and Wei et al. [17], which present several estimation procedures for calculating the $\mathcal{L}_1$-distance-based importance measure. Here, the author delves into calculating the *f*-sensitivity index derived from a general convex function *f*, so that the method proposed is applicable to a host of density-based sensitivity indices. Nonetheless, if the sample size concomitant with a required accuracy in estimating the sensitivity index is very large, say, in the order of millions or more, then existing methods [5,16,17] will be limited to problems or functions that are inexpensive to evaluate. In a practical setting, however, the response function is often determined via time-consuming, costly finite-element analysis (FEA) or similar numerical calculations. In which case, an arbitrarily large sample size is no longer viable, and hence alternative routes to estimating the output probability densities, leading to the *f*-sensitivity index, should be explored.

☆Grant sponsor: U.S. National Science Foundation; Grant No. 1462385.
E-mail address: sharif-rahman@uiowa.edu

http://dx.doi.org/10.1016/j.ress.2016.07.002
0951-8320/

This paper presents an accurate and computationally efficient surrogate method, known as the polynomial dimensional decomposition (PDD) method, for estimating a general class of density-based *f*-sensitivity indices. The method is based on (1) PDD approximation of a high-dimensional stochastic response, forming a surrogate input-output data set; (2) kernel density estimations (KDEs) of output probability density functions from the surrogate data set; and (3) subsequent Monte Carlo integration for estimating the *f*-sensitivity index. Section 2 formally defines the *f*-sensitivity index, summarizes its fundamental properties, and cites a few special cases. Section 3 briefly reviews the PDD approximation and explains how the integration of the PDD approximation with KDE leads to calculating the *f*-sensitivity index. Numerical results from two mathematical functions, as well as from a computationally intensive solid-mechanics problem, are reported in Section 4. Finally, conclusions are drawn in Section 5.

2. A general sensitivity measure

Let $\mathbb{N}:=\{1, 2, \ldots\}$, $\mathbb{N}_0:=\mathbb{N}\cup\{0\}$, $\mathbb{R}:=(-\infty, \infty)$, and $\mathbb{R}_0^+:=[0, \infty)$ represent the sets of positive integer (natural), non-negative integer, real, and non-negative real numbers, respectively. For $k\in\mathbb{N}$, denote by $\mathbb{R}^k$ the *k*-dimensional Euclidean space and by $\mathbb{N}_0^k$ the *k*-dimensional multi-index space. These standard notations will be used throughout the paper.

2.1. *f*-Divergence

Let $(\Psi, \mathcal{G})$ be a measurable space, where Ψ is a sample space and $\mathcal{G}$ is a σ-algebra of the subsets of Ψ, and μ be a σ-finite measure on $(\Psi, \mathcal{G})$. Let $\mathcal{P}$ be a set of all probability measures on $(\Psi, \mathcal{G})$, which are absolutely continuous with respect to μ. For two such probability measures $P_1, P_2\in\mathcal{P}$, let $dP_1/d\mu$ and $dP_2/d\mu$ denote the Radon–Nikodym derivatives of P_1 and P_2 with respect to the dominating measure μ, that is, $P_1<<\mu$ and $P_2<<\mu$. For absolutely continuous measures in terms of probability theory, take Ψ to be the real line and μ to be the Lebesgue measure, that is, $d\mu=d\xi$, $\xi\in\mathbb{R}$, so that $dP_1/d\mu$ and $dP_2/d\mu$ are simply probability density functions, denoted by $f_1(\xi)$ and $f_2(\xi)$, respectively.

Let f: $[0, \infty)\to(-\infty, \infty]$ be an extended real-valued function, which is (1) continuous on $[0, \infty)$ and finite-valued on $(0, \infty)$; (2) convex on $[0, \infty)$; (3) strictly convex at $t=1$; and (4) equal to *zero* at $t=1$, that is, $f(1)=0$. The *f*-divergence, describing the difference or discrimination between two probability measures P_1 and P_2, is defined by the integral [18–21]

$$D_f(P_1\|P_2):=\int_{\mathbb{R}} f\left(\frac{f_1(\xi)}{f_2(\xi)}\right)f_2(\xi)d\xi. \tag{1}$$

The divergence measure in Eq. (1) was introduced in the 1960s by Csiszár [18,19], Ali and Silvey [20], and Morimoto [21].

2.2. *f*-sensitivity index

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is a sample space, $\mathcal{F}$ is a σ-field on Ω, and P: $\mathcal{F}\to[0, 1]$ is a probability measure. With $\mathcal{B}^N$ representing the Borel σ-field on $\mathbb{R}^N$, $N\in\mathbb{N}$, consider an $\mathbb{R}^N$-valued absolutely continuous random vector $\mathbf{X}:=(X_1, \ldots, X_N)$: $(\Omega, \mathcal{F})\to(\mathbb{R}^N, \mathcal{B}^N)$, describing the statistical uncertainties in all system and input parameters of a general stochastic problem. The probability law of $\mathbf{X}$, which may comprise independent or dependent random variables, is completely defined by its joint probability density function $f_{\mathbf{X}}: \mathbb{R}^N\to\mathbb{R}_0^+$. Let u be a non-empty subset of $\{1, \ldots, N\}$ with the complementary set $-u:=\{1, \ldots, N\}\backslash u$ and cardinality $1\le|u|\le N$, and let $\mathbf{X}_u=(X_{i_1}, \ldots, X_{i_{|u|}})$, $1\le i_1<\cdots<i_{|u|}\le N$, be a subvector of $\mathbf{X}$ with $\mathbf{X}_{-u}:=\mathbf{X}_{\{1,\ldots,N\}\backslash u}$ defining its complementary subvector. Then, for a given $\emptyset\ne u\subseteq\{1, \ldots, N\}$, the marginal density function of $\mathbf{X}_u$ is $f_{\mathbf{X}_u}(\mathbf{x}_u):=\int_{\mathbb{R}^{N-|u|}}f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}_{-u}$.

Let $y(\mathbf{X}):=y(X_1, \ldots, X_N)$, a real-valued, continuous, measurable transformation on $(\Omega, \mathcal{F})$, define a general, square-integrable stochastic response function of interest. Define $Y:=y(\mathbf{X})$ to be the associated output random variable. Denote by P_Y and $P_{Y|\mathbf{X}_u}$ the output probability measures and by $f_Y(\xi)$ and $f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)$ the probability density functions of random variables Y and $Y|\mathbf{X}_u$, respectively, where $Y|\mathbf{X}_u$ stands for Y conditional on $\mathbf{X}_u$, which is itself random. Setting $P_1=P_Y$, $f_1=f_Y$, $P_2=P_{Y|\mathbf{X}_u}$, and $f_2=f_{Y|\mathbf{X}_u}$ in Eq. (1), the *f*-divergence between the unconditional and conditional probability measures of Y becomes

$$D_f(P_Y\|P_{Y|\mathbf{X}_u}):=\int_{\mathbb{R}} f\left(\frac{f_Y(\xi)}{f_{Y|\mathbf{X}_u}(\xi|\mathbf{X}_u)}\right)f_{Y|\mathbf{X}_u}(\xi|\mathbf{X}_u)d\xi. \tag{2}$$

For density-based importance analysis, suppose that the sensitivity of Y with respect to a subset $\mathbf{X}_u$, $\emptyset\ne u\subseteq\{1, \ldots, N\}$, of input variables $\mathbf{X}$ is desired. As unveiled in a prequel [6], such a sensitivity measure can be linked to the *f*-divergence in Eq. (2). However, the *f*-divergence is random because $\mathbf{X}_u$ is random. Therefore, apply the expectation operator with respect to the probability measure of $\mathbf{X}_u$ on Eq. (2), thereby defining the *f*-divergence-rooted *f*-sensitivity index [6]

$$H_{u,f}:=\mathbb{E}_{\mathbf{X}_u}\left[D_f(P_Y\|P_{Y|\mathbf{X}_u})\right] \tag{3}$$

of Y for $\mathbf{X}_u$. Applying the definition of the expectation operator in Eq. (3) and then substituting the expression of $D_f(P_Y\|P_{Y|\mathbf{X}_u})$ from Eq. (2) yields the *f*-sensitivity index

$$\begin{aligned} H_{u,f} &= \int_{\mathbb{R}^{|u|}} D_f(P_Y\|P_{Y|\mathbf{X}_u=\mathbf{x}_u})f_{\mathbf{X}_u}(\mathbf{x}_u)d\mathbf{x}_u \\ &= \int_{\mathbb{R}^{|u|}\times\mathbb{R}} f\left(\frac{f_Y(\xi)}{f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)}\right)f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)f_{\mathbf{X}_u}(\mathbf{x}_u)d\mathbf{x}_u d\xi \\ &= \int_{\mathbb{R}^{|u|}\times\mathbb{R}} f\left(\frac{f_Y(\xi)f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)}\right)f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)d\mathbf{x}_u d\xi, \end{aligned} \tag{4}$$

where $P_{Y|\mathbf{X}_u=\mathbf{x}_u}$ and $f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)$ are the probability measure and the probability density function, respectively, of Y conditional on $\mathbf{X}_u=\mathbf{x}_u$, and $f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)$ is the joint probability density function of $(\mathbf{X}_u, Y)$. The last equality in Eq. (4) is formed by recognizing $f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)=f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)f_{\mathbf{X}_u}(\mathbf{x}_u)$ and is useful for calculating the sensitivity index, to be discussed in Section 3.

2.3. General properties

It is important to emphasize a few general properties of the *f*-sensitivity index $H_{u,f}$ inherited from the *f*-divergence measure. The properties, originally derived in a prior work [6], are described in conjunction with six propositions as follows.

Proposition 1 (*Non-negativity*). *The f-sensitivity index $H_{u,f}$ of Y for $\mathbf{X}_u$, $\emptyset\ne u\subseteq\{1, \ldots, N\}$, is non-negative and vanishes when Y and $\mathbf{X}_u$ are statistically independent.*

Proposition 2 (*Range of values*). *The range of values of $H_{u,f}$ is*

$$0\le H_{u,f}\le f(0)+f^*(0), \tag{5}$$

where $f(0)=\lim_{t\to0_+}f(t)$ and $f^*(0)=\lim_{t\to0_+}tf(1/t)=\lim_{t\to\infty}f(t)/t$.

Proposition 3 (*Importance of all input variables*). *The f-sensitivity index $H_{\{1,\ldots,N\},f}$ of Y for all input variables $\mathbf{X}=(X_1, \ldots, X_N)$ is*

$$H_{\{1,\ldots,N\},f}=f(0)+f^*(0), \tag{6}$$