Sensitivity analysis based on Cramér von Mises distance

Fabrice Gamboa, Thierry Klein, Agnès Lagnoux

Sensitivity analysis based on Cramér von Mises distance

Fabrice Gamboa* Thierry Klein*† Agnès Lagnoux*

November 30, 2017

Abstract

In this paper, we first study a sensitivity index that is based on higher moments and generalizes the so-called Sobol one. Further, following an idea of Borgonovo (see [3]), we define and study a new sensitivity index based on the Cramér von Mises distance. This index appears to be more general than the Sobol one as it takes into account the whole distribution of the random variable and not only the variance. Furthermore, we study the statistical properties of its Pick and Freeze estimator.

Keywords: Sensitivity analysis, Cramér von Mises distance, Pick and Freeze method, functional delta-method.

1 Introduction

A very classical problem in the study of computer code experiments (see [26]) is the evaluation of the relative influence of the input variables on some numerical result obtained by a computer code. In this context, a sensitivity analysis is performed. Such a topic has been widely studied in the last decades and is still challenging nowadays (see for example [27, 25, 13] and references therein). More precisely, the result of the numerical code Y is seen as a function of the vector of the distributed input $(X_i)_{i=1,\cdots,d}$ ($d \in \mathbb{N}^*$). Statistically speaking, we are dealing here with the following unnoisy non parametric model

$$Y = f(X_1, \ldots, X_d),$$

where f is a regular unknown numerical function on the state space $E_1 \times E_2 \times \ldots \times E_d$ on which the distributed variables $(X_1, \ldots, X_d)$ are living. Generally, the random inputs are assumed to be independent and a sensitivity analysis is performed using the so-called Hoeffding decomposition (see [29, 1]). In this functional decomposition, f is expanded as an L^2-sum of uncorrelated functions involving only a part of the random inputs. This leads, for any subset v of $I_d := \{1, \ldots, d\}$, to an index called the Sobol index ([27]) that measures the amount of *randomness* (more precisely, the part of the variance) of Y due to the subset of input variables $(X_i)_{i\in v}$. Since nothing has been assumed on the nature of the inputs, one can consider the vector $(X_i)_{i\in v}$ as a single input. Without loss of generality, we then consider the case where v reduces to a singleton. More precisely, the numerator H_v^2 of the Sobol index related to the input X_v is

$$H_v^2 := \mathrm{Var}\left(\mathbb{E}\left[Y|X_v\right]\right)$$

while the denominator of the index is nothing more than the variance of Y. Notice that we also have:

$$H_v^2 = \mathbb{E}\left[\left(\mathbb{E}[Y|X_v] - \mathbb{E}[Y]\right)^2\right] = \mathrm{Var}(Y) - \mathbb{E}\left[\left(\mathbb{E}[Y] - \mathbb{E}\left[Y|X_v\right]\right)^2\right] \tag{1}$$

In order to estimate H_v^2, Sobol in [27] proposed to rewrite the variance of the conditional expectation as a covariance (see equation (3)). Further, a well tailored design of experiment called the Pick and Freeze scheme is considered [19]. More precisely, let X^v be the random vector such that $X_v^v = X_v$ and $X_i^v = X_i'$ if $i \neq v$ where X_i' is an independent copy of X_i. Then, setting

$$Y^v := f(X^v), \tag{2}$$

*Institut de Mathématiques de Toulouse, 118 Route de Narbonne 31062 Toulouse Cedex 9. France. firstname.lastname@math.univ-toulouse.fr

†ENAC - Ecole Nationale de l'Aviation Civile , Université de Toulouse, France

an obvious computation leads to the following relationship (see, e.g., [19])

$$\mathrm{Var}(\mathbb{E}[Y|X_v]) = \mathrm{Cov}\left(Y, Y^v\right). \tag{3}$$

The last equality leads to a natural Monte-Carlo estimator, the so-called Pick and Freeze estimator,

$$T^v_{N,\mathrm{Cl}} = \frac{1}{N}\sum_{j=1}^{N} Y_j Y_j^v - \left(\frac{1}{2N}\sum_{j=1}^{N}(Y_j + Y_j^v)\right)^2$$

where for $j = 1, \cdots, N$, Y_j (resp. Y_j^v) are independent copies of Y (resp. Y^v). The sharp statistical properties and some functional extensions of the Pick and Freeze method are considered in [19, 18, 12]. Notice that the Sobol indices and their Monte-Carlo estimation are order two methods since they derive from the L^2-Hoeffding functional decomposition. This is their main drawback. As an illustration consider the following example. Let X_1 and X_2 be two independent standardized random variables having the same third and fourth moments with $\mathbb{E}\left[X_1^5\right] \neq \mathbb{E}\left[X_2^5\right]$. Let us consider the following model

$$Y = X_1 + X_2 + X_1^2 X_2^2.$$

One gets

$$\mathrm{Var}\left(\mathbb{E}\left[Y|X_1\right]\right) = \mathrm{Var}(X_1 + X_1^2) = \mathrm{Var}(X_2 + X_2^2) = \mathrm{Var}\left(\mathbb{E}\left[Y|X_2\right]\right).$$

Y is an exchangeable function of the inputs but X_1 and X_2 do not share the same distribution. So that, X_1 and X_2 should not have the same importance. That shows the need to introduce a sensitivity index that takes into account all the distribution and not only the second order behavior. As pointed out before, Sobol indices are based on an L^2 decomposition. As a matter of fact, they are well adapted to measure the contribution of an input on the deviation around the mean of Y. Nevertheless, it seems very intuitive that the sensitivity of an extreme quantile of Y could depend on sets of variables that cannot be captured using only the variances. Thus the same index should not be used for any task and we need to define more general indices.

There are several ways to generalize the Sobol indices. For example, one can define new indices through contrast functions based on the quantity of interest (see [16]). Unfortunately the Monte-Carlo estimators of these indices are computationally very expensive. In [11], Da Veiga presents a way to define moment independent measures through dissimilarity distances. These measures define a unified framework that encompasses some known sensitivity indices. They are efficiently estimated in low dimensions but as claimed by the author "it is well known that density estimation suffers from the curse of dimensionality". Now, as pointed out in [3, 5, 6, 24, 23], there are situations where higher order methods give a sharper analysis on the relative influence of the input and allow finer screening procedures. Borgonovo *et al.* propose and study an index based on the total variation distance (see [3, 5, 6]); while Owen *et al.* suggest to use procedures based on higher moments (see [24, 23]).

Our paper follows these tracks. We will first revisit the work of Owen *et al.* by studying the asymptotic properties of the multiple Pick and Freeze estimation. Further, we propose a new natural index based on the Cramér von Mises distance between the distribution of the output Y and its conditional law when an input is fixed. We will show that this approach leads to natural self-normalized indices. Indeed, as for Sobol indices, the sum of all first order indices is uniformly bounded. Notice that these indices take into account the whole output distribution instead of only the order two moments and contrary to most of the other known indices, they are naturally defined for multivariate outputs. As a consequence, they are well-tailored to perform a sensitivity analysis for any vectorial output. Furthermore, we show that surprisingly a Pick and Freeze scheme is also available to estimate this new index. This scheme is not really expensive and easy to implement. The sample size required for the estimation is of the same order as the size needed for the classical Sobol index estimation allowing its use in concrete situations. As a consequence, considering a sample with the appropriate size, one can provide simultaneously the Cramér von Mises indices and the Sobol indices. Other advantage of the Cramér von Mises index with respect to the general ones presented in [11] is that the theoretical statistical properties of its estimation can be derived. Indeed, we prove a Central Limit Theorem for the estimator that allows one to exhibit confidence intervals.

The paper is organized as follows. In the next section, we will study the statistical properties of the multiple Pick and Freeze method proposed earlier by Owen et al ([24, 23]). Section 3 is devoted to the new index built on the Cramér von Mises distance. In the last section, we give some numerical simulations that illustrate the interest of the new index. Moreover, we revisit a real data example introduced in [10] and studied in [15, 7].

2 Higher-moment indices

In the sequel, for any integer k, the notation I_k stands for the set $\{1,\dots,k\}$. Following [24, 23], we generalize the numerator of the Sobol index defined in (1) by considering higher order moments: for any integer $q \geqslant 2$, and singleton $v \in I_d$, we set

$$H_v^q := \mathbb{E}\left[(\mathbb{E}[Y|X_v] - \mathbb{E}[Y])^q\right].$$

Properties Obviously, H_v^q is non negative only for even q and

$$|H_v^q| \leqslant \mathbb{E}\left[|Y - \mathbb{E}[Y]|^q\right].$$

Further, H_v^q is invariant by any translation of the output.

Estimation procedure In view of the estimation of H_v^q, we first notice that

$$H_v^q = \mathbb{E}\left[\prod_{i=1}^{q}\left(Y^{v,i} - \mathbb{E}[Y]\right)\right] = \sum_{l=0}^{q}\binom{q}{l}(-1)^{q-l}\mathbb{E}\left[Y\right]^{q-l}\mathbb{E}\left[\prod_{i=1}^{l}Y^{v,i}\right]$$

with the usual convention $\prod_{i=1}^{0} Y^{v,i} = 1$ and $\binom{q}{l} = q!/l!(q-l)!$. Here, $Y^{v,1} = Y$ and for $i = 2,\dots,q$, $Y^{v,i}$ is constructed independently as Y^v defined in Equation (2).
Second, we use a Monte-Carlo scheme and consider the following Pick and Freeze design constituted by a N-sample $\left(Y_j^{v,i}\right)_{(i,j)\in I_q\times I_N}$ of $\left(Y^{v,1},\dots,Y^{v,q}\right)$. The Monte-Carlo estimator is then

$$H_{q,N}^v = \sum_{l=0}^{q}\binom{q}{l}(-1)^{q-l}\left(\overline{P}_1^v\right)^{q-l}\overline{P}_l^v$$

where for any $N \in \mathbb{N}^*$, $j \in I_N$ and $l \in I_q$, we have defined

$$P_{l,j}^v = \binom{q}{l}^{-1}\sum_{k_1<\dots<k_l\in I_q}\left(\prod_{i=1}^{l}Y_j^{v,k_i}\right) \quad \text{and} \quad \overline{P}_l^v = \frac{1}{N}\sum_{j=1}^{N}P_{l,j}^v.$$

Notice that we generalize the estimation procedure of [18] and use all the available information by considering the means over the set of indices $k_1,\dots,k_l \in I_d$, $k_n \neq k_m$.

Asymptotic properties of $H_{q,N}^v$

Theorem 2.1. *$H_{q,N}^v$ is strongly consistent and asymptotically Gaussian:*

$$\sqrt{N}\left(H_{q,N}^v - H_q^v\right) \underset{N\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}\left(0,\sigma^2\right)$$

where

$$\sigma^2 = q\left[\mathrm{Var}(Y) + (q-1)\mathrm{Cov}(Y,Y^{v,2})\right]\left(\sum_{l=1}^{q}a_l b_l\right)^2,$$

$$a_l = \frac{l}{q}\mathbb{E}[Y]^{l-1}, \qquad l = 1,\dots,q$$

$$b_1 = (-1)^{q-1}q(q-1)\mathbb{E}[Y]^{q-1} + \sum_{l=2}^{q-1}\binom{q}{l}(-1)^{q-l}(q-l)\mathbb{E}[Y]^{q-l-1}\mathbb{E}\left[\prod_{i=1}^{l}Y^{v,i}\right]$$

and

$$b_l = \binom{q}{l}(-1)^{q-l}\mathbb{E}[Y]^{q-l}, \qquad l = 1,\dots,q.$$

Interpretation and comments The collection of all indices $(H_v^q)_q$ is much more informative than the classical Sobol index with respect to v. Nevertheless, it has several drawbacks. First, these indices are moment-based and it is well known that they are not stable when the moment order increases. Second, they may be negative when q is odd. To overcome this fact, one may introduce $\mathbb{E}\left[|\mathbb{E}[Y|X_v] - \mathbb{E}[Y]|^q\right]$ but the Pick and Freeze estimation procedure is then lost. Third, the Pick and Freeze estimation procedure is computationally expensive and may be unstable: it requires a $q \times N$-sample of the output Y. In order to have a good idea of the influence of an input on the law of the output, we need to estimate the first $K-1$ indices H_v^q: $H_v^2, \ldots, H_v^K$. Hence, we need to run the code $K \times N$ times.
In a nutshell, these indices are not attractive in a practical point of view. In the next section, we then introduce a new sensitivity index that is based on the conditional distribution of the output and requires only $3 \times N$ runs. Concretely, it compares the output distribution to the conditional one whereas the q higher-order moment indices only compare the q-th output moment to the conditional one.

3 The Cramér von Mises index

In this section, the code will be denoted by $Z = f(X_1, \ldots, X_d) \in \mathbb{R}^k$. It is worth noticing that here we can consider multivariate outputs unlike in Section 2 and [7], e.g., Let F be the distribution function of Z:

$$F(t) = \mathbb{P}\left(Z \leqslant t\right) = \mathbb{E}\left[\mathbb{1}_{\{Z \leqslant t\}}\right], \text{ for } t = (t_1, \ldots, t_k) \in \mathbb{R}^k$$

and F^v be the conditional distribution function of Z conditionally on X_v:

$$F^v(t) = \mathbb{P}\left(Z \leqslant t|X_v\right) = \mathbb{E}\left[\mathbb{1}_{\{Z \leqslant t\}}|X_v\right], \text{ for } t = (t_1, \ldots, t_k) \in \mathbb{R}^k.$$

Notice that $\{Z \leqslant t\}$ means that $\{Z_1 \leqslant t_1, \ldots, Z_k \leqslant t_k\}$. Obviously, $\mathbb{E}\left[F^v(t)\right] = F(t)$. Now, we define $Y(t) = \mathbb{1}_{\{Z \leqslant t\}}$ and take $p = 2$. Since for any fixed $t \in \mathbb{R}^k$, $Y(t)$ is a real-valued random variable, we apply the framework presented in Section 2. More precisely, for any $v \in I_p$ let $\sim v$ be $I_p \setminus \{v\}$ and we first perform the Hoeffding decomposition of $Y(t)$:

$$Y(t) = \mathbb{1}_{\{Z \leqslant t\}} = \mathbb{E}[Y(t)] + (\mathbb{E}[Y(t)|X_v] - \mathbb{E}[Y(t)]) + (\mathbb{E}[Y(t)|X_{\sim v}] - \mathbb{E}[Y(t)]) + R(t, v) \tag{4}$$

where

$$R(t, v) = Y(t) - \mathbb{E}[Y(t)] - (\mathbb{E}[Y(t)|X_v] - \mathbb{E}[Y(t)]) - (\mathbb{E}[Y(t)|X_{\sim v}] - \mathbb{E}[Y(t)]).$$

As done usually, we compute the variance of both sides of (4) that leads to

$$\begin{aligned}
\mathrm{Var}(Y(t)) &= F(t)(1 - F(t)) \\
&= \mathrm{Var}\left(\mathbb{E}[Y(t)|X_v] - \mathbb{E}[Y(t)]\right) + \mathrm{Var}\left(\mathbb{E}[Y(t)|X_{\sim v}] - \mathbb{E}[Y(t)]\right) + \mathrm{Var}(R(t, v)) \\
&= \mathrm{Var}\left(F^v(t)\right) + \mathrm{Var}\left(F^{\sim v}(t)\right) + \mathrm{Var}(R(t, v)) \\
&= \mathbb{E}\left[\left(F^v(t) - F(t)\right)^2\right] + \mathbb{E}\left[\left(F^{\sim v}(t) - F(t)\right)^2\right] + \mathrm{Var}(R(t, v))
\end{aligned} \tag{5}$$

by the decorrelation of the different terms involved in the Hoeffding decomposition.

Remark 3.1. A straightforward application of the results of Section 2 provides for any fixed $t \in \mathbb{R}^k$ a consistent and asymptotically normal procedure for the estimation of

$$\mathbb{E}\left[\left(F^v(t) - F(t)\right)^2\right] = \mathrm{Var}\left(F^v(t)\right) \quad \text{and} \quad \mathbb{E}\left[\left(F^{\sim v}(t) - F(t)\right)^2\right] = \mathrm{Var}\left(F^{\sim v}(t)\right).$$

Now we integrate the terms in (5) in $t \in \mathbb{R}^k$ with respect to the distribution of Z:

$$\begin{aligned}
&\int_{\mathbb{R}^k} F(t)(1 - F(t))dF(t) \\
&= \int_{\mathbb{R}^k} \mathbb{E}\left[\left(F^v(t) - F(t)\right)^2\right] dF(t) + \int_{\mathbb{R}^k} \mathbb{E}\left[\left(F^{\sim v}(t) - F(t)\right)^2\right] dF(t) + \int_{\mathbb{R}^k} \mathrm{Var}(R(t, v))dF(t)
\end{aligned} \tag{6}$$

This integration has to be understood in the Riemmann-Stieltjes sense (see, e.g., [28]). Notice that the first term in the right hand side of (6) represents a Cramér Von Mises type distance of order 2 between

the distribution $\mathcal{L}(Z)$ of Z and the distribution $\mathcal{L}(Z|X_v)$ of Z given X_v.

Following the classical way of defining Sobol indices, we normalize the previous equation by

$$\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)$$

leading to

$$1 = \frac{\int_{\mathbb{R}^k} \mathbb{E}\left[(F^v(t)-F(t))^2\right] dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)} + \frac{\int_{\mathbb{R}^k} \mathbb{E}\left[(F^{\sim v}(t)-F(t))^2\right] dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)} + \frac{\int_{\mathbb{R}^k} \mathrm{Var}(R(t,v))dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)} \quad (7)$$

Then we define the Cramér Von Mises indices with respect to v and $\sim v$ by

$$S^v_{2,CVM} := \frac{\int_{\mathbb{R}^k} \mathbb{E}\left[(F(t)-F^v(t))^2\right] dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)} \quad \text{and} \quad S^{\sim v}_{2,CVM} := \frac{\int_{\mathbb{R}^k} \mathbb{E}\left[(F(t)-F^{\sim v}(t))^2\right] dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)}.$$

Properties 3.2. *These new indices are naturally adapted to multivariate outputs and they share the same properties as the classical Sobol index. Namely,*

1. *as seen in (7), the different contributions sum to 1.*
2. *they are invariant by translation, by any isometry and by any non degenerated scaling of the components of Y.*

Remark 3.3.

1. We could have defined the following indices instead

$$\int_{\mathbb{R}^k} \frac{\mathbb{E}\left[(F(t)-F^v(t))^2\right]}{F(t)(1-F(t))} dF(t) \quad \text{and} \quad \int_{\mathbb{R}^k} \frac{\mathbb{E}\left[(F(t)-F^{\sim v}(t))^2\right]}{F(t)(1-F(t))} dF(t).$$

   normalizing by $F(t)(1-F(t))$ (like in the Anderson-Darling statistic) before the integration phase. Nevertheless, the previous integrals might be not defined. Moreover, even if the integrals are well defined, one may encounter numerical explosion during the estimation procedure that might be produced for small and large values of t since the normalizing factor then cancels.

2. In this paper, we only consider first-order sensitivity indices as well for the classical Sobol indices and for the Cramér von Mises indices. Anyway, as well as for the Sobol indices, one may define higher-order and total Cramér von Mises indices. The construction of the former is straightforward taking v no longer a singleton. For example, if one is interested in the second-order Cramer von Mises index with respect to the first and second inputs, it suffices to take $v = \{1,2\}$. Concerning the latter, the total Cramér von Mises index $S^{Tot,v}_{2,CVM}$ with repect to v is defined by

$$S^{Tot,v}_{2,CVM} := 1 - S^{\sim v}_{2,CVM} = 1 - \frac{\int_{\mathbb{R}^k} \mathbb{E}\left[(F(t)-F^{\sim v}(t))^2\right] dF(t)}{\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t)}.$$

3. To use the Hoeffding decomposition, the inputs are required to be independent. Anyway, one can compute the Cramér von Mises index when the inputs are dependent. Nevertheless, there are then difficult to interpret.

3.1 General comments on the Cramér von Mises indices

Cramér von Mises indices versus Sobol indices

Cramér von Mises and Sobol indices are both based on the Hoeffding decomposition and sum to 1. Nevertheless, the former are based on the whole distribution of the output, in contrast with the latter that are only based on the order-two moments. Notice that two variables that have a different influence on the output may have the same Sobol indices (just as two random variables with different distribution can have the same variance). This point represents one limitation of Sobol indices and does not occur with

the Cramér von Mises indices as one can see in Section 4.1.

In addition, remark that a null value for a Sobol index does not imply that the input is unimportant whereas a null value for a Cramér von Mises index means that the input is unimportant. Moreover, by definition, a large Cramér von Mises index means that the input variable under concern has a great influence on the output in regions where the output has a large distribution mass. That is why we advice the practitioner to use them in a general context. Nevertheless, when one is interested in the mean output behavior, the Sobol indices are more adapted. Indeed, as noted in [16], the Sobol indices minimize the contrast associated to the mean. In the same spirit, if one is interested in specific feature of the output (for example an α-quantile), one should use the index based on the associated contrast. See [16] for more details on the notion of contrast and the results therein.

In contrast, the indices based on the whole distribution partially get rid of such limitations and pathological patterns. However, one can build an example based, e.g., on two input variables that leads to the same indices $S^1_{2,CVM}$ and $S^2_{2,CVM}$ once the integration with respect to t has been done.

Cramér von Mises indices versus moment independent indices

There already exists several moment-independent indices: some of them have been introduced by Borgonovo et *al.* (density-based indices [5], cumulative distribution function based indices [9]). See also [4] for other indices and references therein. More recently, Da Veiga [11] shows that those indices are special cases of a class of sensitivity indices based on the Csizár f-divergence. A lot of classical "distances" between probability measures as, e.g., the Kullback-Leibler divergence, the Hellinger distance and the total variation distance belong to this family of divergences. Other dissimilarity measures exist to compare probability distributions: in particular, integral probability metrics [20].

In comparison with the indices defined in Equation (17) in [8], we can notice that the integration is done with respect to the distribution of the output in the former indices while the integration is done with respect to the Lebesgue measure in the latter indices. Our method represents at least two advantages: (i) the index always exist whatever the output distribution (ii) such an integration weights the support of the output distribution.

Since the space of the probability measures on $\mathbb{R}^k$ is of infinite dimension, the different distances on this space are not equivalent; hence they are very difficult to compare. Each index is constructed on a specific distance and has its own interest. Despite the fact that the Cramér von Mises indices have no clear dual formulation, they present the following remarkable advantages. As we will see in the next sections, one can easily estimate them with a low simulation cost that does not depend on the dimension of the output. The sample required for their estimation also provide Sobol indicies estimation. In addition, these estimators are asymptotically normal and converge at the rate $\sqrt{N}$ which allows the practitioner to build confidence intervals.

The rest of the section is dedicated to the estimation of $S^v_{2,CVM}$ (and $S^{\sim v}_{2,CVM}$). One has to estimate both the numerator and the denominator of the indices. Nevertheless, when the output Z has independent coordinates that are absolutely continuous with respect to the Lebesgue measure, we have

$$\int_{\mathbb{R}^k} F(t)(1-F(t))dF(t) = \mathbb{E}[F(Z)(1-F(Z)] = \frac{1}{2^k} - \frac{1}{3^k}.$$

Thus the normalizing factor reduces to $\frac{1}{2^k} - \frac{1}{3^k}$. As a consequence, we propose two versions of Central Limit Theorems: the first one deals with the numerator's estimator and can be applied when the output Z has independent coordinates that are absolutely continuous with respect to the Lebesgue measure whereas the second one concerns the general estimator and may apply in any other cases.

3.2 Numerator estimation and its asymptotic properties

We denote the numerator of $S^v_{2,CVM}$ by $N^v_{2,CVM}$. Notice that it can be rewritten as

$$N^v_{2,CVM} = \mathbb{E}_{\tilde{Z}}\left[\mathbb{E}_{X_v}\left[\left(F(\tilde{Z}) - F^v(\tilde{Z})\right)^2\right]\right]$$

where $\tilde{Z}$ is an independent copy of Z.
Then we proceed to a double Monte-Carlo scheme for the estimation of $N^v_{2,CVM}$ and consider the following design of experiment consisting in:

1. The classical Pick and Freeze sample, that is two N-samples of Z: $(Z_j^{v,1}, Z_j^{v,2})$, $1 \leqslant j \leqslant N$;
2. A third N-sample of Z independent of $(Z_j^{v,1}, Z_j^{v,2})_{1\leqslant j\leqslant N}$: W_k, $1 \leqslant k \leqslant N$.

The empirical estimator of $N^v_{2,CVM}$ is then given by

$$\widehat{N}^v_{2,CVM} = \frac{1}{N}\sum_{k=1}^{N}\left\{\frac{1}{N}\sum_{j=1}^{N}\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}}\mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}} - \left[\frac{1}{2N}\sum_{j=1}^{N}\left(\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}}\right)\right]^2\right\}. \quad (8)$$

Now we established the consistency of $\widehat{N}^v_{2,CVM}$ that follows directly from an auxiliary lemma (see Section 6).

Corollary 3.4. *$\widehat{N}^v_{2,CVM}$ is strongly consistent as N goes to infinity.*

Now we turn to the asymptotic normality of $\widehat{N}^v_{2,CVM}$. We follow van der Vaart [29] to establish the following proposition (more precisely Theorems 20.8 and 20.9, Lemma 20.10 and Example 20.11).

Theorem 3.5. *The sequence of estimators $\widehat{N}^v_{2,CVM}$ is asymptotically Gaussian in estimating $N^v_{2,CVM}$. That is, $\sqrt{N}\left(\widehat{N}^v_{2,CVM} - N^v_{2,CVM}\right)$ converges in distribution towards the centered Gaussian law with a limiting variance ξ^2 whose explicit expression can be found in the proof.*

Remark 3.6. Thanks to Theorem 3.5, we are now able to provide asymptotic confidence intervals for the estimation of $N^v_{2,CVM}$. They are of the form $(\widehat{N}^v_{2,CVM} \pm z_\alpha \xi/\sqrt{N})$, where z_α is the $1-\alpha/2$ quantile of a standard normal distribution. Unfortunately, the variance ξ^2 is unknown but thanks to its explicit form it is easy to replace it by a consistent estimator $\widehat{\xi}$ and use Slutsky's Lemma to have an asymptotic confidence interval.

3.3 Estimation of the general index and its asymptotic properties

In order to estimate the general index $S^v_{2,CVM}$, we first estimate its numerator as in Subsection 3.2 and then its denominator that we denote $D^v_{2,CVM}$. Notice that it can be rewritten as

$$D^v_{2,CVM} = \mathbb{E}\left[F(Z)(1-F(Z))\right]$$

and estimated using the design of experiment already introduced for the estimation of the numerator by

$$\widehat{D}^v_{2,CVM} = \frac{1}{N}\sum_{k=1}^{N}\left\{\frac{1}{2N}\sum_{j=1}^{N}\left(\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}}\right) - \left[\frac{1}{2N}\sum_{j=1}^{N}\left(\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}}\right)\right]^2\right\}. \quad (9)$$

Proceeding as in Subsection 3.2, we have

Corollary 3.7. *$\widehat{S}^v_{2,CVM}$ is strongly consistent as N goes to infinity.*

The following Central Limit Theorem comes from the functional Delta method.

Theorem 3.8. *The sequence of estimators $\widehat{S}^v_{2,CVM}$ is asymptotically Gaussian in estimating $S^v_{2,CVM}$. That is, $\sqrt{N}\left(\widehat{S}^v_{2,CVM} - S^v_{2,CVM}\right)$ converges in distribution towards the centered Gaussian law with a limiting variance that can be computed.*

3.4 Practical advices

In a general setting, for all the nice properties of the Cramér von Mises indices and their efficient estimation easy to implement, we recommend to use the Cramér von Mises indices. As a consequence, considering a sample with the appropriate size, one can estimate once at a time the Cramér von Mises indices and the Sobol indices. More precisely, if one wants to estimate p Sobol indices a sample size of $(p+1)N$ is required. With only N more output evaluations, we get both the p Sobol indices and the Cramér von Mises ones. Furthermore, the theoretical theorems provides confidence intervals that controlled the accuracy of the estimations. Anyway, when the practitioner is interested in a specific feature (e.g., mean behavior or quantile) of the output, he should use more suited indices (e.g., the classical Sobol indices for the mean or the indices introduced in [16] for the quantile).

4 Numerical applications

4.1 A flavor of the method applied on a toy model

Let us consider the quite simple linear model

$$Y = \alpha X_1 + X_2, \quad \alpha > 0,$$

where X_1 has a Bernoulli distribution with success probability $0 < p < 1$ and X_1, X_2 are independent. Assume further that X_2 has a continuous distribution F_2 on $\mathbb{R}$ such that $\mathbb{E}[X_2] = \alpha p$ and with finite variance $\mathrm{Var}(X_2) = \alpha^2 p(1-p)$. With these choices, the random variables αX_1 and X_2 share the same expectation and the same variance. Thus X_1 and X_2 have the same first order Sobol indices all equal to $1/2$.
We present a general closed formula to compute our new indices and show that in some particular cases an exact formula is available. Then we perform a simulation study in order to illustrate the Central Limit Theorem and analyse the practical behaviour of our estimators.

4.1.1 General closed formula

On one hand, the distribution of Y given $X_1 = 0$ and the distribution of Y given $X_1 = 1$ are given by

$$\begin{cases} \mathcal{L}(Y|X_1 = 0) = \mathcal{L}(X_2) \\ \mathcal{L}(Y|X_1 = 1) = \mathcal{L}(X_2 + \alpha). \end{cases}$$

On the other hand, the conditional distribution of Y given X_2 is

$$\mathbb{P}\left(Y = \alpha + X_2 \mid X_2\right) = 1 - \mathbb{P}\left(Y = X_2 | X_2\right) = p.$$

Hence, the distribution function of Y is the mixture $pF_2(\cdot - \alpha) + (1-p)F_2(\cdot)$. Tedious computations lead to

$$S^1_{2,CVM} = 6p(1-p) \int_{\mathbb{R}} \left(F_2(t) - F_2(t-\alpha)\right)^2 \left[(1-p)dF_2(t) + pdF_2(t-\alpha)\right]$$

and

$$S^2_{2,CVM} = 1 - 6p(1-p)\left[\frac{1}{2} - \int_{\mathbb{R}} F_2(t-\alpha)dF_2(t)\right]$$

(the normalizing factor being $1/6$ as explained before).
As p goes to 0 (and α goes to infinity), $(S^1_{2,CVM}, S^2_{2,CVM})$ goes to $(0,1)$ while the two classical Sobol indices remain equal to $1/2$. Our new indices shed lights on the fact that, for small p, X_2 has much more influence on Y than X_1 which follows the intuition. This fact is not detected by the classical Sobol indices.

Similarly we can compute the indices of order q ($q \geqslant 2$):

$$H^q_1 = \alpha^q \left[p(1-p)^q + (-p)^q(1-p)\right] \quad \text{and} \quad H^q_2 = \mathbb{E}[(X_2 - \mu)^q].$$

Particular cases (i) if X_2 is a centered Gaussian random variable with variance $\mathrm{Var}(X_2) = \alpha^2 p(1-p)$, one can easily derive an explicit formula for H_2^q:

$$H_2^q = \mathbb{E}[(X_2 - m)^q] = \frac{q!}{2^{q/2} \cdot (q/2)!} \mathbb{1}_{q \in 2\mathbb{N}^*}.$$

(ii) if X_2 is uniformly distributed on $[0, b]$ with $b = 2\alpha\sqrt{3p(1-p)}$, one can easily derive an explicit formula for the indices introduced before:

$$S^1_{2,CVM} = 6p(1-p) \times \left(\left(\frac{\alpha}{b}\right)^2 \left(1 - \frac{2}{3}\frac{\alpha}{b}\right) \mathbb{1}_{\alpha \leqslant b} + \frac{1}{3}\mathbb{1}_{\alpha > b} \right)$$
$$S^2_{2,CVM} = 1 - 3p(1-p)\left(1 - \left(\frac{b-\alpha}{b}\right)^2 \mathbb{1}_{\alpha \leqslant b}\right).$$

Moreover, $H_2^q = \mathbb{E}[(X_2 - \mu)^q] = (b/2)^q/(q+1)\mathbb{1}_{q \in 2\mathbb{N}^*}$.

(iii) if X_2 is exponentially distributed with mean $1/\lambda = \alpha\sqrt{p(1-p)}$, one can easily derive an explicit formula for the indices introduced before:

$$S^1_{2,CVM} = 2p(1-p)(1 - e^{-\lambda\alpha})^2 \quad \text{and} \quad S^2_{2,CVM} = 1 - 3p(1-p)(1 - e^{-\lambda\alpha}).$$

Moreover, $H_2^q = \mathbb{E}[(X_2 - \mu)^q] = q!\lambda^{-q}/2$.

4.1.2 Simulation study

A numerical illustration with sample sizes N=100 and 500 is presented in Figures 1 and 2 (remind that in order to estimate both indices we compute $4N$ values of the output function). We consider the case where the random variable X_2 is uniformly distributed (for the other cases the simulations provide the same kind of results). We estimate the Cramér von Mises indices thanks to Equation (8) and renormalize it by the factor $1/6$ since the output has a continuous distribution. Then we estimate the limiting variance in (12) in order to provide asymptotic confidence intervals. In Figures 1 and 2, the blue line represents the true value of index $D^1_{2,CVM}$ (first row) or $D^2_{2,CVM}$ (second row). The red dashed line (resp. the red dashed line with $+$) represents the index estimation based on (8) (resp. the confidence interval). In the left column, α is fixed to $1/2$ and p varies while in the right one, p is fixed to $1/4$ and α varies.

4.2 A non linear model

Now, let us consider the following non linear model

$$Y = \exp\{X_1 + 2X_2\}, \tag{10}$$

where X_1 and X_2 are independent standard Gaussian random variables. The distribution of Y is log-normal and we can derive both its density and its distribution functions:

$$f_Y(y) = \frac{1}{\sqrt{10\pi}y} e^{-(\ln y)^2/10} \mathbb{1}_{\mathbb{R}^+}(y) \quad \text{and} \quad F_Y(y) = \Phi\left(\frac{\ln y}{\sqrt{5}}\right)$$

where Φ stands for the distribution function of the standard Gaussian random variable. Its density function will be denoted by f in the sequel. Then tedious computations lead to the Cramér von Mises indices $S^1_{2,CVM}$ and $S^2_{2,CVM}$.

Proposition 4.1. *Assume that Y is defined by Equation (10) then*

$$S^1_{2,CVM} = \frac{6}{\pi}\arctan 2 - 2 \approx 0.1145$$

and

$$S^2_{2,CVM} = \frac{6}{\pi}\arctan\sqrt{19} - 2 \approx 0.5693.$$

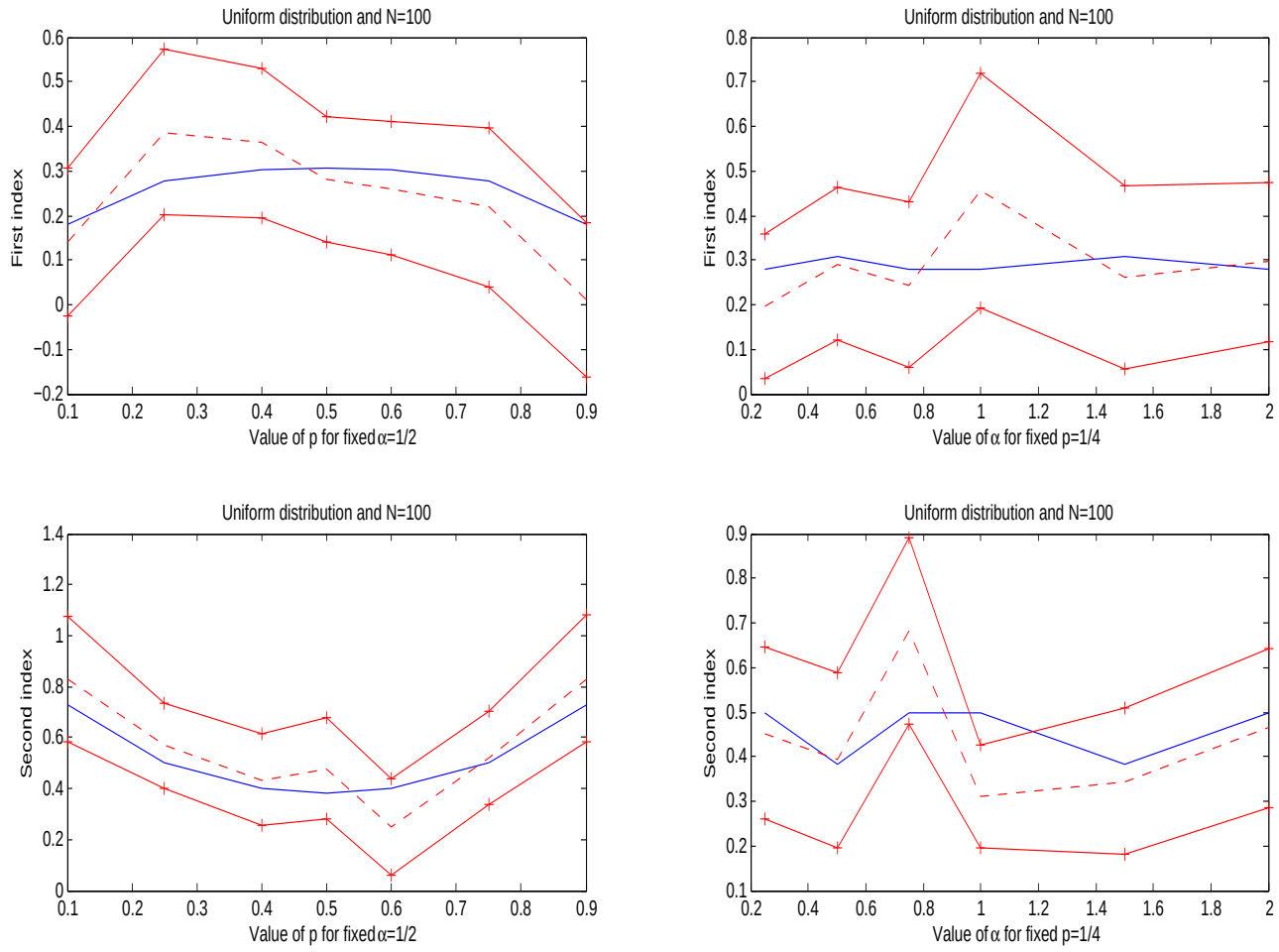

Figure 1: Example 1 - X_2 uniformly distributed and N=100.

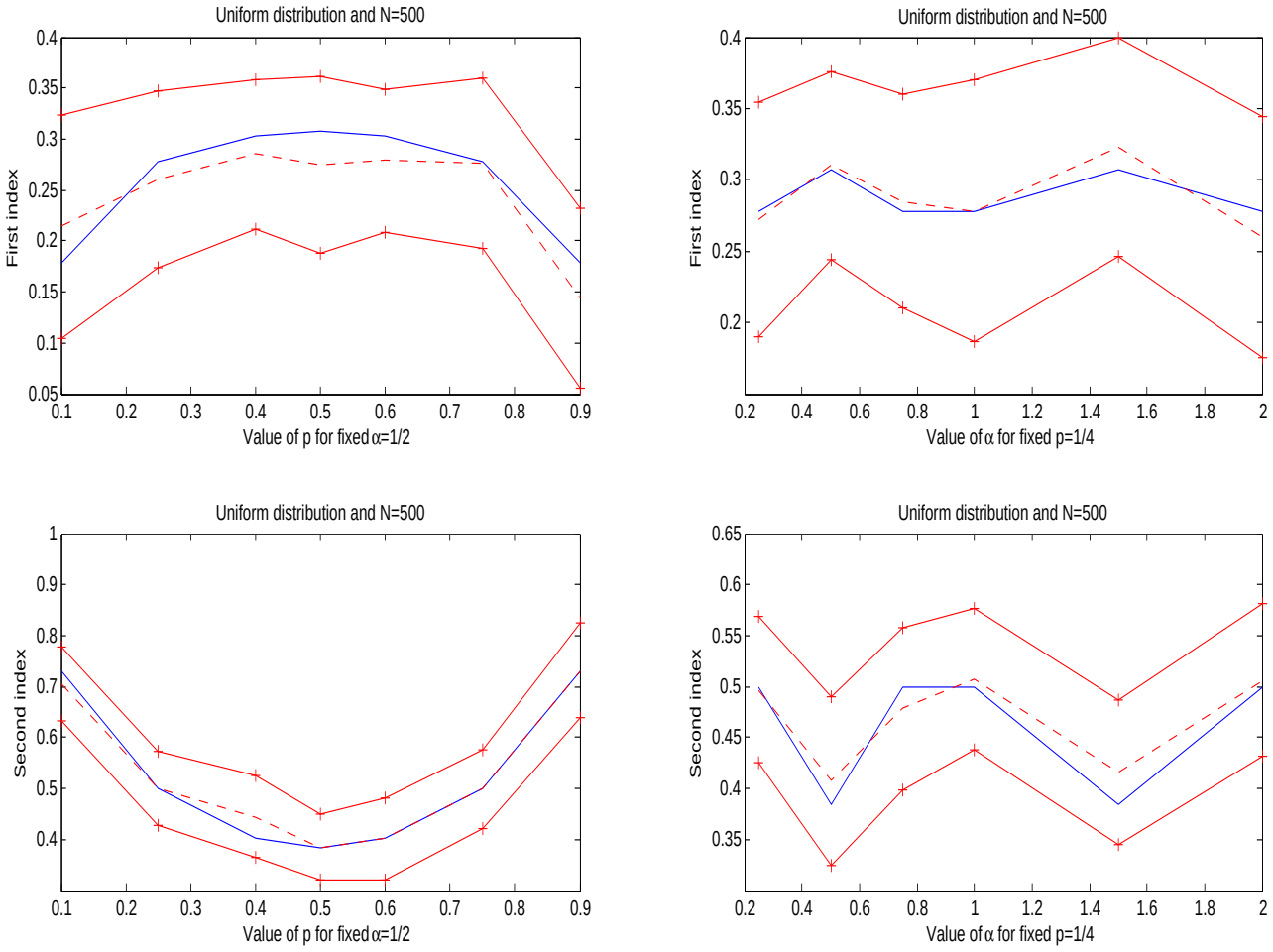

Figure 2: Example 1 - X_2 uniformly distributed and N=500.

Remark 4.2. In this simple example, one can compute the indices of order q ($q \geqslant 2$):

$$H_1^q = \mathbb{E}\left[(e^{X_1+2} - e^{5/2})^q\right] \quad \text{and} \quad H_2^q = \mathbb{E}\left[(e^{2X_1+1/2} - e^{5/2})^q\right].$$

The Sobol indices and their estimation based on the Pick-Freeze scheme with a sample of size N are computed using equation (6) in [19]. We also compute the Cramér von Mises indices and their estimation based on (8). Moreover, we estimate the limiting variances in both cases (see equation (12) for the Cramér von Mises indices and equation (12) in [19] for the Sobol indices) in order to provide confidence intervals. The results are presented in Table 1.

		Cramér von Mises		Sobol indices	
		$D^1_{2,CVM}$	$D^2_{2,CVM}$	S^1	S^2
	True values	0.1145	0.5693	0.0118	0.3738
$N = 10^2$	Est. values	0.1287	0.6097	0.0425	0.1954
	CI 5%	[-0.0601,0.3175]	[0.4692,0.7503]	[0.0265,0.0585]	[0.0430,0.3477]
$N = 10^3$	Est. values	0.1358	0.6007	0.1198	0.2345
	CI 5%	[0.07861,0.19297]	[0.54897,0.65242]	[-0.5633,0.8030]	[0.1343,0.3347]
$N = 10^4$	Est. values	0.1166	0.5585	0.01685	0.26252
	CI 5%	[0.09930,0.13382]	[0.54150,0.57540]	[0.0010,0.0327]	[-1.2744, 1.7994]

Table 1: Model (10). The Cramér von Mises and Sobol indices, their estimations based on (8) and (6) in [19] and the associated 5%-confidence intervals.

As a conclusion, with only $N = 10^3$, the statistical method provides a precise estimation of the different indices. Moreover, in this example, the Sobol and Cramér von Mises indices give the same influence ranking of the two random inputs. Nevertheless, the estimation of the Cramér von Mises indices seems to be more efficient to give the true ranking.

4.3 Application: The Giant Cell Arthritis Problem

Context and goal

In this subsection, we consider the realistic problem of management of suspected giant cell arthritis posed by Bunchbinder and Detsky in [10]. More recently, this problem was also studied by Felli and Hazen [15] and Borgonovo *et al.* [7]. As explained in [10], "giant cell arthritis (GCA) is a vasculitis of unknown etiology that affects large and medium sized vessels and occurs almost exclusively in patients 50 years or older". This disease may lead to severe side effects (loss of visual accuity, fever, headache,...) whereas the risks of not treating it include the threat of blindness and major vessels occlusion. A patient with suspected GCA can receive a therapy based on Prednisone. Unfortunately, a treatment with high Prednisone doses may cause severe complications. Thus when confronted to a patient with suspected GCA, the clinician must adopt a strategy. There is a considerable literature on sensitivity analysis for these sorts of models, based on the utility of learning a model input before choosing a treatment strategy (see, e.g., [14] and [22]). In [10], the authors considered four different strategies:

A : Treat none of the patients;

B : Proceed to the biopsy and treat all the positive patients;

C : Proceed to the biopsy and treat all the patients whatever their result;

D : Treat all the patients.

The clinician wants to adopt the strategy optimizing the patient outcomes measured in terms of utility. The reader is referred to [21] for more details on the concept of utility. The basic idea is that a patient with perfect health is assigned a utility of 1 and the expected utility of the other patients (not perfectly healthy) is calculated subtracting some "disutilities" from this perfect score of 1. These strategies are represented in Figures 3 to 6 with the different inputs involved in the computation of the utilities.
For example in strategy A (see Figure 3), the utility of a patient having GCA and developing severe GCA complications is given by $1 - d_s - du_{gc} - du_{dx}$. His entire sub-path is then

$$g \times gc \times (1 - d_s - du_{gc} - du_{dx}).$$

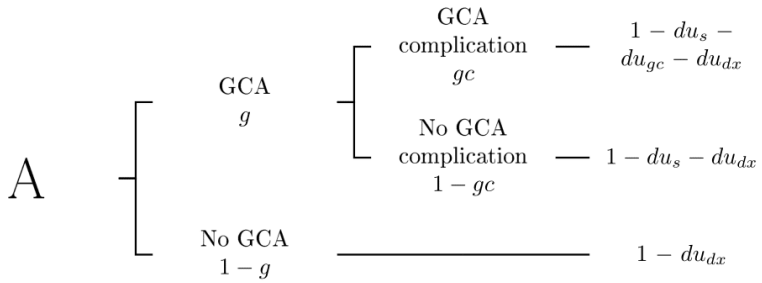

Figure 3: The decision tree for the treat none alternative

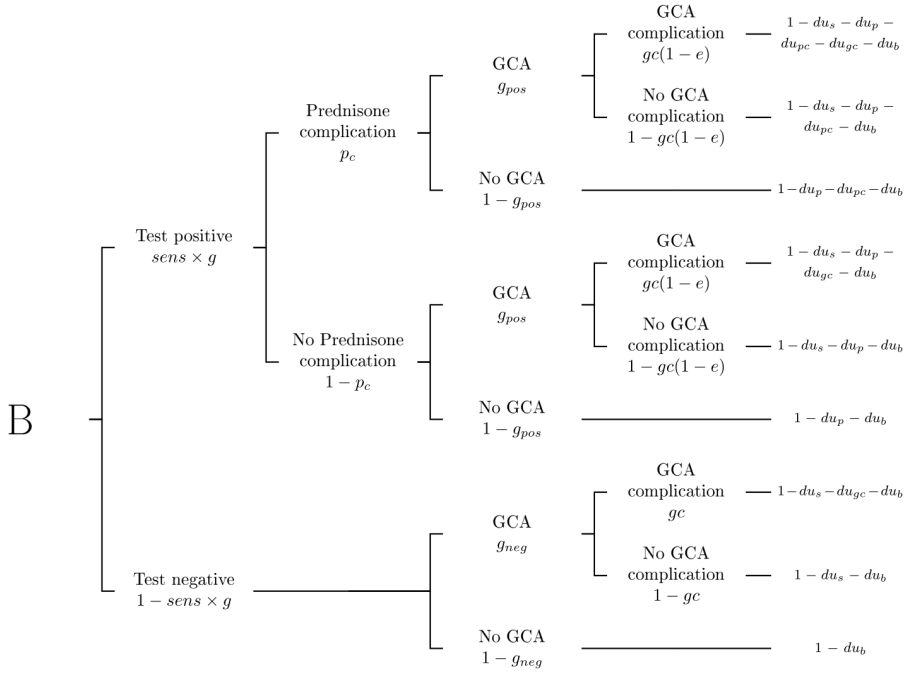

Figure 4: The decision tree for the biopsy and the treat positive alternative

The input parameters and the modelisation of the random ones

As seen in Figures 3 to 6, the different strategies involve input parameters like, e.g., the proportion g of patients having GCA or the probability gc for a patient to develop severe GCA complications (fixed at 0.8 as done in [10]) or even the disutility associated to having GCA symptoms. Table 2 summarizes the input parameters involved.

The values $\mathbb{P}[\cdot]$ and $D(\cdot)$ refer respectively to the probability of an event and to the disutility associated with an event. The minimum and maximum values m and M depict each parameter's range for the sensitivity analysis. The base values are provided by a clinician expertise. The utilities of the different strategies when all the input parameters are set to their base value are summarized in Table 3.

The base value of some input parameters are reliable while the others are really uncertain that leads us to consider them as random. As a consequence, if Y_A, Y_B, Y_C and Y_D represent the outcomes corresponding to the four different strategies A to D, the clinician aims to determine

$$\max\{\mathbb{E}[Y_A], \mathbb{E}[Y_B], \mathbb{E}[Y_C], \mathbb{E}[Y_D]\} \tag{11}$$

with the uncertain model input presented in Table 2. A sensitivity analysis is then performed to determine the most influential input variables on the outcome.

As done in [15, 7], all the random inputs will be independently based on Beta distributions. The Beta density parameters corresponding to each random input are determined by fitting the base value as their mean and capturing 95% of the probability mass in the range defined by the minimum and maximum.

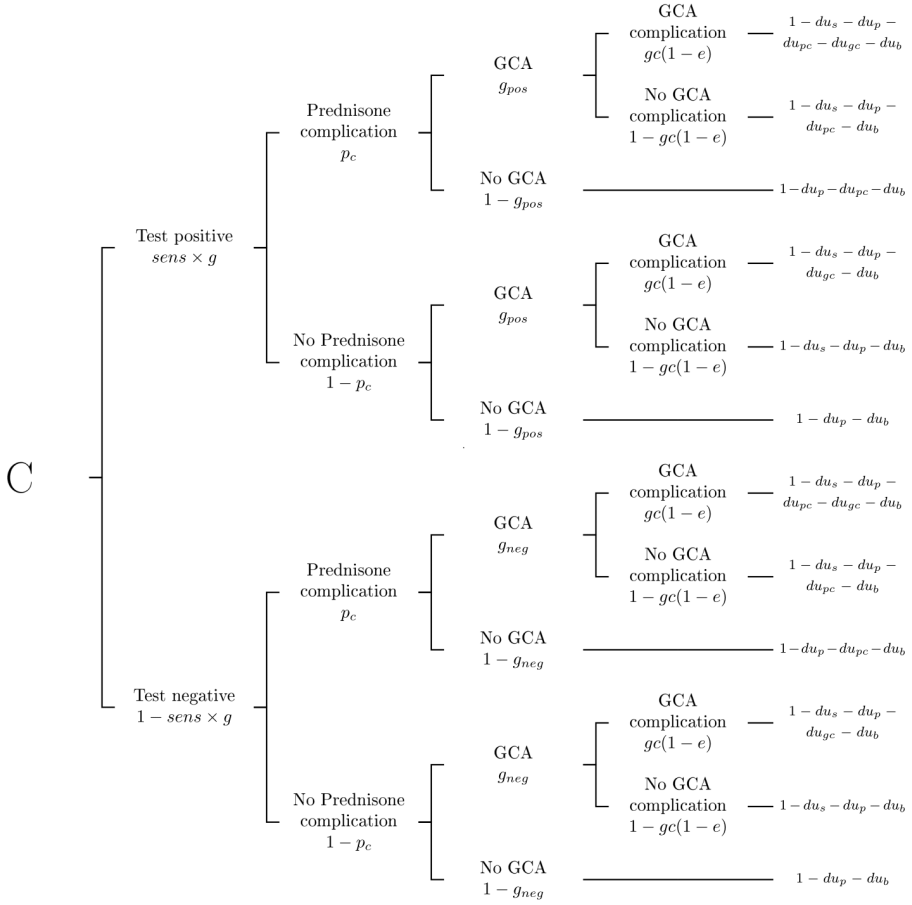

Figure 5: The decision tree for the biopsy and the treat all alternative

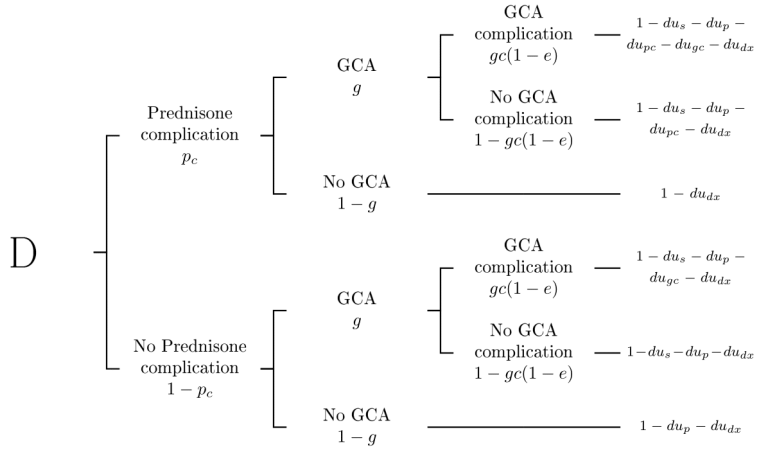

Figure 6: The decision tree for the treat all alternative

The remaining 5% will be equally distributed to either side of this range if possible. Concretely, each random input will be distributed as

$$Z\mathbb{1}_{m\leqslant Z<M} + U\mathbb{1}_{m>Z} + V\mathbb{1}_{Z\geqslant M}$$

where Z, U and V are independent random variables. Z is Beta distributed with parameters (α, β). U and V are uniform random variables on $[0, m]$ and $[M, 1]$ respectively.

Fixed parameters	Symbols	Fixed value				
$\mathbb{P}$[having GCA]	g	0.8	–	–	–	–
D(having symptoms of GCA)	du_s	0.12	–	–	–	–
D(having a temporal artery biopsy)	du_b	0.005	–	–	–	–
D(not knowing the true diagnosis)	du_{dx}	0.025	–	–	–	–
					Beta(α,β)	
Uncertain parameters	Symbols	Base	Min. m	Max. M	α	β
$\mathbb{P}$[developing severe complications of GCA]	gc	0.3	0.05	0.5	4.179	11.011
$\mathbb{P}$[developing severe iatrogenic side effects]	pc	0.2	0.05	0.5	2.647	10.589
Efficacy of high dose Prednisone	e	0.9	0.8	1	27.787	3.087
Sensitivity of temporal artery biopsy	$sens$	0.83	0.6	1	7.554	1.547
D(major complication from GCA)	du_{gc}	0.8	0.3	0.9	27.454	6.864
D(Prednisone therapy)	du_p	0.08	0.03	0.2	4.555	52.380
D(major iatrogenic side effect)	du_{pc}	0.3	0.2	0.9	15.291	35.680

Table 2: The data used by Buchbinder and Detsky [10] in their analysis

Treatment alternative	Utilility	Expectation
A Treat none	0.6870	0.6991
B Biopsy and treat positive	0.7575	0.7570
C Biopsy and treat all	0.7398	0.7371
D Treat all	0.7198	0.7171

Table 3: The utilities of the different strategies when all the input parameters are set to their base value (second column) and their expectation when they are random (third column).

Sensitivity analysis

As already mentioned, the clinician wants to determine the highest utility. In [4], the authors then consider the highest utility as output and lead a sensitivity analysis to determine the input having the largest influence on this output. Since we are able to treat multivariate outputs, we consider a more general framework in this paper: the output is the four-dimensional random variable $Y = (Y_A, Y_B, Y_C, Y_D)$ where Y_S represents the outcome corresponding to strategy S.

We compare three different methodologies and indices. First, we consider the Sobol indices introduced in [17] (Multivariate). Second, we consider the indices constructed in this paper, based on the Cramér von Mises distance and estimated by the ratios of the numerator estimator (8) and the denominator estimator (9). Third, we consider the index presented in [4] and named β defined by

$$\beta_i = \mathbb{E}[\sup_{y \in \mathcal{Y}} \{|F_Y(y) - F_{Y|X_i}(y)|\}].$$

Then we use the estimator given in [7, Table 1] adapted to the multivariate case that is based on the tedious and costly estimation of conditional expectations.

Results

Table 4 summarizes the sensitivity measures of the seven random inputs on the multivariate output with the three different methodologies while Table 5 presents the associated ranks. It is worth mentioning that the same total sample size has been used to compare properly the three methodologies.

As a conclusion, in this example, unlike the indices defined by Borgonovo *et al.*, the multivariate sensitivity indices and the Cramér von Mises indices provide the same ranking. The main advantage of the Cramér von Mises sensitivity methodology with respect to the one of Borgonovo *et al.* is that one can use the Pick and Freeze estimation scheme which provides an accurate estimation (see (8)) simple to implement. Notice that in [7], the authors study a slightly different model that explains the numerical differences between their results and the ones of the present paper. Furthermore, they perform a sensitivity analysis on the best alternative with the greater mean instead of considering the multivariate output.

	Sensitivity meas.	1	2	3	4	5	6	7	Cputime
$N = 10^2$	Multivariate	0.3690	0.0193	0.0105	-0.0821	-0.0617	0.1150	-0.0751	0.0624
	Borgonovo *et al.*	0.1195	0.1047	0.1064	0.1022	0.1046	0.1063	0.1027	1.5132
	Cramér von Mises	0.3496	0.0745	0.0206	-0.0010	0.0084	0.1042	0.0105	0.9048
$N = 10^3$	Multivariate	0.4024	0.1201	0.0516	-0.0190	-0.0043	0.2403	0.0093	0.0156
	Borgonovo *et al.*	0.1788	0.1192	0.1009	0.1007	0.1044	0.1195	0.1028	57.8452
	Cramér von Mises	0.3494	0.0750	0.0209	-0.0008	0.0086	0.1045	0.0109	10.1089
$N = 10^4$	Multivariate	0.3828	0.1333	0.0618	-0.0016	0.0100	0.3182	0.0217	0.0312
	Borgonovo *et al.*	0.3842	0.1572	0.1033	0.0930	0.0986	0.1775	0.1061	$5.1988\ 10^3$
	Cramér von Mises	0.3494	0.0775	0.0232	0.0011	0.0108	0.1056	0.0124	436.8028

Table 4: Sensitivity measures. The estimation of the Cramér von Mises indices is the ratio of (8) and (9).

	Sensitivity meas.	Ranking
$N = 10^2$	Multivariate	1 6 2 3 5 7 4
	Borgonovo *et al.*	1 3 6 2 5 7 4
	Cramér von Mises	1 6 2 3 7 5 4
$N = 10^3$	Multivariate	1 6 2 3 7 5 4
	Borgonovo *et al.*	1 6 2 5 7 3 4
	Cramér von Mises	1 6 2 3 7 5 4
$N = 10^4$	Multivariate	1 6 2 3 7 5 4
	Borgonovo *et al.*	1 6 2 7 3 5 4
	Cramér von Mises	1 6 2 3 7 5 4

Table 5: Ranks. The estimation of the Cramér von Mises indices is the ratio of (8) and (9).

5 Conclusion

In this paper, we first study the asymptotic properties of the multiple Pick and Freeze scheme proposed by Owen *et al.* for the estimation of higher order Sobol indices. This index has several drawbacks that lead us to propose a new natural index based on the Cramér von Mises distance between the distribution of the output Y and the conditional law when an input is fixed. This new index contains all the distributional information, is naturally defined for multivariate outputs and provides a rigorous sharper way for a fast screening of complex computer codes. Furthermore, our approach is generic and may be extended and implemented for general outputs (vectorial, valued on a manifold, functional, ...). Concerning its estimation, we show that surprisingly a Pick and Freeze scheme is also available for the estimation procedure and prove that it is efficient in a theoretical point of view as well as in a practical one. More precisely, we establish a Central Limit Theorem that confirms the good statistical properties of our estimator and allows us to build confidence intervals. Furthermore, the estimation is well working with moderate sample sizes as shown in toy examples. Finally, the performance of the method is proven on a real data example.

Acknowledgement

The authors are greatly indebted to the referees for their fruitful and detailed suggestions or comments which permit us to greatly improve our paper.
Part of this research was conducted within the frame of the Chair in Applied Mathematics OQUAIDO and the ANR project PEPITO (ANR-14-CE23-0011).

6 Proofs

6.1 Proof of Theorem 2.1

Proof of Theorem 2.1. The consistency follows from a straightforward application of the strong law of large numbers. The asymptotic normality is derived by two successive applications of the delta method [29] .

(1) Let $W_j^1 := (Y_j^{v,1}, \dots, Y_j^{v,p})^T$ $(j = 1, \dots, N)$ and g^1 be the mapping from $\mathbb{R}^p$ to $\mathbb{R}^p$ whose l-th coordinate is given by

$$g_l^1(x_1, \dots, x_p) = \binom{p}{l}^{-1} \sum_{\substack{k_1 < \dots < k_l \\ k_i \in I_p, i = 1, \dots, l}} \left(\prod_{i=1}^{l} x_{k_i} \right).$$

Then $(W_j^1)_{j=1,\dots,N}$ is an i.i.d. sample distributed as $W^1 := (Y^{v,1}, \dots, Y^{v,p})^T$.
Let Σ^1 be the covariance matrix of W_j^1. Clearly, one has $\Sigma_{ii}^1 = \mathrm{Var}(Y)$ for $i \in I_p$ while for $i \neq j$, $\Sigma_{ij}^1 = \mathrm{Cov}(Y^{v,i}, Y^{v,j}) = \mathrm{Cov}(Y, Y^{v,2})$. The multidimensional Central Limit Theorem gives that

$$\sqrt{N} \left(\frac{1}{N} \sum_{j=1}^{N} W_j^1 - m \right) \underset{N \to \infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}_p \left(0, \Sigma^1 \right),$$

where $m := (\mathbb{E}[Y], \dots, \mathbb{E}[Y])^T$. We then apply the so-called delta method to W^1 and g^1 so that

$$\sqrt{N} \left(g^1 \left(\overline{W}_N^1 \right) - g^1 \left(\mathbb{E}\left[W^1 \right] \right) \right) \underset{N \to \infty}{\xrightarrow{\mathcal{L}}} \mathcal{N} \left(0, J_{g^1} \left(\mathbb{E}\left[W^1 \right] \right) \Sigma^1 J_{g^1} \left(\mathbb{E}\left[W^1 \right] \right)^T \right)$$

where $J_{g^1} \left(\mathbb{E}\left[W^1 \right] \right)$ is the Jacobian of g^1 at point $\mathbb{E}\left[W^1 \right]$. Notice that for $i \in I_p$ and $k \in I_p$,

$$\frac{\partial g_l^1}{\partial x_k} \left(\mathbb{E}\left[W^1 \right] \right) = \frac{\binom{p-1}{l-1}}{\binom{p}{l}} m^{l-1} = \frac{l}{p} \mathbb{E}[Y]^{l-1} =: a_l.$$

Thus $\Sigma^2 := J_{g^1} \left(\mathbb{E}\left[W^1 \right] \right) \Sigma^1 J_{g^1} \left(\mathbb{E}\left[W^1 \right] \right)^T$ is given by

$$\Sigma_{ij}^2 = p a_i a_j \left(\Sigma_{11}^1 + (p-1) \Sigma_{12}^1 \right).$$

(2) Now consider $W_j^2 := (P_j^{v,1}, \dots P_j^{v,p})^T$ $(j = 1, \dots, N)$ and g^2 the mapping from $\mathbb{R}^p$ to $\mathbb{R}$ defined by

$$g^2(y_1, \dots, y_p) = \sum_{l=0}^{p} \binom{p}{l} (-1)^{p-l} y_1^{p-l} y_l.$$

Then $(W_j^2)_{j=1,\dots,N}$ is an i.i.d. sample distributed as $W^2 := (P^{v,1}, \dots P^{v,p})^T$.
We apply once again the delta method to W^2 so that

$$\sqrt{N} \left(g^2 \left(\overline{W}_N^2 \right) - g^2 \left(\mathbb{E}\left[W^2 \right] \right) \right) \underset{N \to \infty}{\xrightarrow{\mathcal{L}}} \mathcal{N} \left(0, J_{g^2} \left(\mathbb{E}\left[W^2 \right] \right) \Sigma^2 J_{g^2} \left(\mathbb{E}\left[W^2 \right] \right)^T \right)$$

where $J_{g^2} \left(\mathbb{E}\left[W^2 \right] \right)$ is the Jacobian of g^2 at point $\mathbb{E}\left[W^2 \right]$. Notice that for $k \in I_p$,

$$\begin{aligned} \frac{\partial g^2}{\partial y_1} \left(\mathbb{E}\left[W^2 \right] \right) &= (-1)^{p-1} p(p-1) \mathbb{E}[Y]^{p-1} \\ &\quad + \sum_{l=2}^{p-1} \binom{p}{l} (-1)^{p-l} (p-l) \mathbb{E}[Y]^{p-l-1} \mathbb{E}\left[\prod_{i=1}^{l} Y^{v,i} \right] \end{aligned}$$

and

$$\frac{\partial g^2}{\partial y_l} \left(\mathbb{E}\left[W^2 \right] \right) = \binom{p}{l} (-1)^{p-l} \mathbb{E}[Y]^{p-l}.$$

Thus the limiting variance is

$$\sigma^2 := J_{g^2}\left(\mathbb{E}\left[W^2\right]\right)\Sigma^2 J_{g^2}\left(\mathbb{E}\left[W^2\right]\right)^T = p\left(\Sigma_{11}^1 + (p-1)\Sigma_{12}^1\right)\left(\sum_{i=1}^p a_i b_i\right)^2,$$

where b_i is the i-th coordinate of $\nabla g^2\left(\mathbb{E}\left[W^2\right]\right)$. □

6.2 An auxiliary result and the proofs of the results of Section 3

Lemma 6.1. *Let G and H be two measurable functions. Let $(U_j)_{j\in I_N}$ and $(V_k)_{k\in I_N}$ be two independent samples of i.i.d. random variables. Assume that $G(U_1, V_1)$ and $H(U_1, U_2, V_1)$ are both integrable and centered. We define S_N and T_N by*

$$S_N = \frac{1}{N^2}\sum_{j,k=1}^N G(U_j, V_k) \quad \text{and} \quad T_N = \frac{1}{N^3}\sum_{i,j,k=1}^N H(U_i, U_j, V_k).$$

Then S_N and T_N converge a.s. to 0 as N goes to infinity.

Proof of Lemma 6.1. Notice that if $\mathbb{E}[S_N^4] = O\left(\frac{1}{N^2}\right)$ then by Borel-Cantelli lemma, S_N converges a.s. to 0. Now,

$$\mathbb{E}[S_N^4] = \frac{1}{N^8}\sum \mathbb{E}[G(U_{i_1}, V_{j_1})G(U_{i_2}, V_{j_2})G(U_{i_3}, V_{j_3})G(U_{i_4}, V_{j_4})]$$

where the sum is taken over all the indices i_1, i_2, i_3, i_4, j_1, j_2, j_3, j_4 from 1 to N. The only cases leading to terms in $O\left(\frac{1}{N}\right)$ or even in $O(1)$ appear when we sum over indices that are all different except two i's or two j's or over indices that are all different. Nevertheless, in those cases, at least one term of the form $\mathbb{E}[G(U_i, V_j)]$ appears. Since the function G is centered, those cases are then discarded.

The proof of the result concerning T_N follows the same tracks. □

Proof of Corollary 3.4. The proof is based on Lemma 6.1. First, we define $Z_j = \left(Z_j^{v,1}, Z_j^{v,2}\right)$,

$$\begin{aligned}
G(Z_j, W_k) &= \mathbb{1}_{\{Z_j^{v,1} \leqslant W_k\}}\mathbb{1}_{\{Z_j^{v,2} \leqslant W_k\}},\\
F(Z_j, W_k) &= \frac{1}{2}\left(\mathbb{1}_{\{Z_j^{v,1} \leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2} \leqslant W_k\}}\right),\\
H(Z_i, Z_j, W_k) &= F(Z_i, W_k)F(Z_j, W_k).
\end{aligned}$$

Second, we proceed to the following decomposition

$$\widehat{N}^v_{2,CVM} = \frac{1}{N}\sum_{k=1}^{N}\left\{\frac{1}{N}\sum_{j=1}^{N}\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}}\mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}} - \left[\frac{1}{2N}\sum_{j=1}^{N}\left(\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}}\}\right)\right]^2\right\}$$

$$= \frac{1}{N^2}\sum_{j,k=1}^{N}\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}}\mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}} - \frac{1}{4N^3}\sum_{i,j,k=1}^{N}\left(\mathbb{1}_{\{Z_i^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_i^{v,2}\leqslant W_k\}}\right)\left(\mathbb{1}_{\{Z_j^{v,1}\leqslant W_k\}} + \mathbb{1}_{\{Z_j^{v,2}\leqslant W_k\}}\right)$$

$$= \frac{1}{N^2}\sum_{j,k=1}^{N}G(Z_j,W_k) - \frac{1}{N^3}\sum_{i,j,k=1}^{N}H(Z_i,Z_j,W_k)$$

$$= \frac{1}{N^2}\sum_{j,k=1}^{N}\{G(Z_j,W_k) - \mathbb{E}[G(Z_j,W_k)]\} - \frac{1}{N^3}\sum_{i,j,k=1}^{N}\{H(Z_i,Z_j,W_k) - \mathbb{E}[H(Z_i,Z_j,W_k)]\}$$
$$+ \frac{1}{N^2}\sum_{j,k=1}^{N}\mathbb{E}[G(Z_j,W_k)] - \frac{1}{N^3}\sum_{i,j,k=1}^{N}\mathbb{E}[H(Z_i,Z_j,W_k)]$$

$$= \frac{1}{N^2}\sum_{j,k=1}^{N}\{G(Z_j,W_k) - \mathbb{E}[G(Z_j,W_k)]\} - \frac{1}{N^3}\sum_{i,j,k=1}^{N}\{H(Z_i,Z_j,W_k) - \mathbb{E}[H(Z_i,Z_j,W_k)]\}$$
$$+ \mathbb{E}[G(Z_1,W_1)] - \left(1-\frac{1}{N}\right)\mathbb{E}[H(Z_1,Z_2,W_1)] - \frac{1}{N}\mathbb{E}[H(Z_1,Z_1,W_1)].$$

The two first sums converge almost surely to 0 by Lemma 6.1. The remaining term goes to $\mathbb{E}[G(Z_1,W_1)] - \mathbb{E}[H(Z_1,Z_2,W_1)]$ as N goes to infinity.

It remains to show that $N^v_{2,CVM} = \mathbb{E}[G(Z_1,W_1)] - \mathbb{E}[H(Z_1,Z_2,W_1)]$. On the one hand,

$$\begin{aligned}N^v_{2,CVM} &= \int_{\mathbb{R}}\mathbb{E}[(F(t)-F^v(t))^2]dF(t) = \mathbb{E}[H_v^2(W_1)]\\ &= \mathbb{E}[\mathrm{Cov}(\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}},\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}})]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}] - \mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}]^2].\end{aligned}$$

On the other hand,

$$\begin{aligned}&\mathbb{E}[G(Z_1,W_1)] - \mathbb{E}[H(Z_1,Z_2,W_1)]\\ &= \mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}] - \frac{1}{4}\mathbb{E}[(\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}} + \mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}})(\mathbb{1}_{\{Z_2^{v,1}\leqslant W_1\}} + \mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}})]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}}]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}}|W_1]]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}|W_1]\mathbb{E}[\mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}}|W_1]]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}|W_1]]\mathbb{E}[\mathbb{E}[\mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}}|W_1]]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}]\mathbb{E}[\mathbb{1}_{\{Z_2^{v,2}\leqslant W_1\}}]\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}]] - \mathbb{E}[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}]^2\\ &= \mathbb{E}_W[\mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}\mathbb{1}_{\{Z_1^{v,2}\leqslant W_1\}}] - \mathbb{E}_Z[\mathbb{1}_{\{Z_1^{v,1}\leqslant W_1\}}]^2]\end{aligned}$$

that completes the proof. □

Proof of Theorem 3.5. We define for $t \in \mathbb{R}$,

$$\mathbb{G}_N^{1,2}(t,t) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{\{Z_j^{v,1} \leqslant t\}} \mathbb{1}_{\{Z_j^{v,2} \leqslant t\}},$$
$$\mathbb{G}_N^{i}(t) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{\{Z_j^{v,i} \leqslant t\}},\ i = 1,2,$$
$$\mathbb{F}_N(t) = \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\{W_k \leqslant t\}}$$

and we rewrite $\widehat{N}^v_{2,CVM}$ as a regular function depending on the four empirical processes defined above:

$$\widehat{N}^v_{2,CVM} = \int \left[\mathbb{G}_N^{1,2} - \left(\frac{\mathbb{G}_N^1 + \mathbb{G}_N^2}{2}\right)^2\right] d\mathbb{F}_N.$$

By Donsker's theorem,

$$\sqrt{N}\left(\mathbb{G}_N^{1,2} - \widetilde{G}, \mathbb{G}_N^1 - F, \mathbb{G}_N^2 - F, \mathbb{F}_N - F\right) \underset{N\to\infty}{\overset{\mathcal{L}}{\to}} \mathbb{G} = (\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3, \mathbb{G}_4)$$

where $G(t,s) = \mathbb{P}\left(Z^{v,1} \leqslant t,\ Z^{v,2} \leqslant s\right)$, $\widetilde{G}(t) = G(t,t)$ and $\mathbb{G}$ is a centered Gaussian process of dimension 4 with covariance function defined by

$$\Pi(t,s) = \mathbb{E}\left(A_t A_s^T\right) - \mathbb{E}\left(A_t\right)\mathbb{E}\left(A_s\right)^T, \text{ for } \quad (t,s) \in \mathbb{R}^2$$

and $A_t := \left(\mathbb{1}_{\{Z^{v,1}\leqslant t\}}\mathbb{1}_{\{Z^{v,2}\leqslant t\}}, \mathbb{1}_{\{Z^{v,1}\leqslant t\}}, \mathbb{1}_{\{Z^{v,2}\leqslant t\}}, \mathbb{1}_{\{W\leqslant t\}}\right)^T$.

Since these processes are càd-làg functions of bounded variation, we introduce the maps $\psi_1,\ \psi_2 : BV_1[-\infty,+\infty]^2 \mapsto \mathbb{R}$ and $\Psi : BV_1[-\infty,+\infty]^4 \mapsto \mathbb{R}$ defined by

$$\psi_i(F_1,F_2) = \int (F_1)^i dF_2,\ i = 1,2 \quad \text{and} \quad \Psi(F_1,F_2,F_3,F_4) = \psi_1(F_1,F_4) - \psi_2\left(\frac{F_2+F_3}{2}, F_4\right),$$

where $BV_M[a,b]$ is the set of càd-làg functions of variation bounded by M. Hence,

$$\widehat{N}^v_{2,CVM} = \Psi\left(\mathbb{G}_N^{1,2}, \mathbb{G}_N^1, \mathbb{G}_N^2, \mathbb{F}_N\right),$$

Now using the chain rule 20.9 and Lemma 20.10 in [29], the map Ψ is Hadamard-differentiable from the domain $BV_1[-\infty,+\infty]^4$ into $\mathbb{R}$ whose derivative is given by

$$(h_1,h_2,h_3,h_4) \mapsto D\psi_1(F_1,F_4)(h_1,h_4) - D\psi_2\left(\frac{F_2+F_3}{2}, F_4\right)\left(\frac{h_2+h_3}{2}, h_4\right)$$

where the derivative of ψ_i are given by Lemma 20.10:

$$(h_1,h_2) \mapsto h_2\varphi_i \circ F_1|_{-\infty}^{+\infty} - \int h_{2-} d\varphi_i \circ F_1 + \int \varphi_i'(F_1) h_1 dF_2$$

with $\varphi_i(x) = x^i$ and h_- is the left-continuous version of a càd-làg function h.

Applying the functional delta method 20.8 in [29] we get the weak convergence of $\sqrt{N}\left(\widehat{N}^v_{2,CVM} - N^v_{2,CVM}\right)$ to the following limit distribution

$$\int \mathbb{G}_{4-} d(F^2 - \widetilde{G}) + \int \mathbb{G}_1 dF - \int F(\mathbb{G}_2 + \mathbb{G}_3) dF.$$

Since the map Ψ is continuous on the whole space $BV_1[-\infty,+\infty]^4$, the delta method in its stronger form 20.8 in [29] implies that the limit variable is the limit in distribution of the sequence

$$D\Psi(\widetilde{G},F,F,F)\left(\sqrt{N}\left(\mathbb{G}_N^{1,2}-\widetilde{G},\mathbb{G}_N^1-F,\mathbb{G}_N^2-F,\mathbb{F}_N-F\right)\right)$$
$$=\sqrt{N}\left[\int(\mathbb{F}_N-F)_-\,d\left(F^2-\widetilde{G}\right)+\int\left(\mathbb{G}_N^{1,2}-\widetilde{G}-F\left(\mathbb{G}_N^1+\mathbb{G}_N^2-2F\right)\right)dF\right].$$

We define

$$U:=\int\mathbb{1}_{\{W<t\}}d(F^2(t)-\widetilde{G}(t)=\widetilde{G}(W)-F(W)^2,$$
$$V:=\int\left[\mathbb{1}_{\{Z^{v,1}\leqslant t\}}\mathbb{1}_{\{Z^{v,2}\leqslant t\}}-\left(\mathbb{1}_{\{Z^{v,1}\leqslant t\}}+\mathbb{1}_{\{Z^{v,2}\leqslant t\}}\right)F(t)\right]dF(t)$$
$$=\frac{1}{2}\left(F(Z^{v,1})^2+F(Z^{v,2})^2\right)-F(Z^{v,1}\vee Z^{v,2}).$$

By independence, the limiting variance ξ^2 is

$$\xi^2=\mathrm{Var}U+\mathrm{Var}V. \tag{12}$$

□

6.3 Proof of Proposition 4.1

Proof of Proposition 4.1. First of all, the distribution function of Y conditioned on X_1 is given by

$$F^{(1)}(t)=\mathbb{P}(Y\leqslant t|X_1)=\Phi\left(\frac{\ln t-X_1}{2}\right).$$

Then

$$N_{2,CVM}^1=\int_{\mathbb{R}}\mathbb{E}\left[(F^{(1)}(t)-F_Y(t))^2\right]f_Y(t)dt$$
$$=\int_{\mathbb{R}^+}\mathbb{E}\left[\left(\Phi\left(\frac{\ln t-X_1}{2}\right)-\Phi\left(\frac{\ln y}{\sqrt{5}}\right)\right)^2\right]\frac{1}{\sqrt{10\pi t}}e^{-(\ln t)^2/10}dt$$
$$=\int_{\mathbb{R}}\mathbb{E}\left[\left(\Phi\left(\frac{\sqrt{5}z-X_1}{2}\right)-\Phi\left(z\right)\right)^2\right]e^{-z^2/10}\frac{dz}{\sqrt{2\pi}}$$
$$=\mathbb{E}\left[\left(\Phi(X_2)-\Phi\left(\frac{\sqrt{5}X_2-X_1}{2}\right)\right)^2\right]$$

where X_1 and X_2 are independent standard Gaussian random variables. In the same way,

$$N_{2,CVM}^2=\mathbb{E}\left[(\Phi(X_2)-\Phi\left(\sqrt{5}X_2-2X_1\right))^2\right].$$

Thus we are lead to compute the bivariate function:

$$\varphi(\alpha,\beta):=\mathbb{E}\left[(\Phi(X_2)-\Phi\left(\alpha X_2-\beta X_1\right))^2\right]$$

for $(\alpha,\beta)=(\sqrt{5}/2,1/2)$ and $(\alpha,\beta)=(\sqrt{5},2)$. The term $\mathbb{E}\left[\Phi(X_2)^2\right]$ is

$$\mathbb{E}\left[\Phi(X_2)^2\right]=\int\Phi(z)^2f(z)dz=\left[\frac{1}{3}\Phi(z)^3\right]_{-\infty}^{+\infty}=\frac{1}{3}.$$

We introduce three independent random variables U, U' and V distributed as standard Gaussian random variables. Then the term $\mathbb{E}\left[\Phi\left(\alpha X_2 - \beta X_1\right)^2\right]$ can be rewritten as

$$\begin{aligned}
\mathbb{E}\left[\Phi\left(\alpha X_2 - \beta X_1\right)^2\right] &= \mathbb{E}\left[\Phi\left(\sqrt{\alpha^2+\beta^2}V\right)^2\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{U\leqslant\sqrt{\alpha^2+\beta^2}V}|V\right]^2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{U\leqslant\sqrt{\alpha^2+\beta^2}V}|V\right]\mathbb{E}\left[\mathbb{1}_{U'\leqslant\sqrt{\alpha^2+\beta^2}V}|V\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{U\leqslant\sqrt{\alpha^2+\beta^2}V}\mathbb{1}_{U'\leqslant\sqrt{\alpha^2+\beta^2}V}|V\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_{U\leqslant\sqrt{\alpha^2+\beta^2}V}\mathbb{1}_{U'\leqslant\sqrt{\alpha^2+\beta^2}V}\right] = \mathbb{P}\left(U\leqslant\sqrt{\alpha^2+\beta^2}V,\, U'\leqslant\sqrt{\alpha^2+\beta^2}V\right) \\
&=: G(\sqrt{\alpha^2+\beta^2}).
\end{aligned}$$

Integrating by parts, we have

$$\begin{aligned}
G'(a) &= 2\int_{\mathbb{R}} z\Phi(az)e^{-(a^2+1)z^2/2}\frac{dz}{2\pi} \\
&= -\frac{1}{\pi(a^2+1)}\left(\left[\Phi(az)e^{-(a^2+1)z^2/2}\right]_{-\infty}^{+\infty} - a\int_{\mathbb{R}} f(az)e^{-(a^2+1)z^2/2}dz\right) \\
&= \frac{a}{\pi(a^2+1)}\frac{1}{\sqrt{2a^2+1}}.
\end{aligned}$$

Since $G(1) = 1/3$, we get

$$G(a) = \frac{1}{3} + \int_1^a \frac{x}{\pi(x^2+1)}\frac{1}{\sqrt{2x^2+1}}dx = \frac{1}{3} + \frac{1}{\pi}(\arctan\sqrt{1+2a^2} - \arctan\sqrt{3}) = \frac{1}{\pi}\arctan\sqrt{1+2a^2}$$

and

$$\mathbb{E}\left[\Phi\left(\alpha X_2 - \beta X_1\right)^2\right] = \frac{1}{3} + \frac{1}{\pi}(\arctan\sqrt{1+2(\alpha^2+\beta^2)} - \arctan\sqrt{3}) = \frac{1}{\pi}\arctan\sqrt{1+2(\alpha^2+\beta^2)}.$$

In the same way, the last term $\mathbb{E}\left[\Phi(X_2)\Phi\left(\alpha X_2 - \beta X_1\right)\right]$ is given by

$$\mathbb{E}\left[\Phi(X_2)\Phi\left(\alpha X_2 - \beta X_1\right)\right] = \mathbb{P}\left(U\leqslant V,\, \sqrt{\frac{1+\beta^2}{\alpha^2}}U'\leqslant V\right)$$

where U, U' and V are independent standard Gaussian random variables. Remind that we only need to consider $(\alpha,\beta) = (\sqrt{5}/2, 1/2)$ and $(\alpha,\beta) = (\sqrt{5}, 2)$ in which cases $\sqrt{\frac{1+\beta^2}{\alpha^2}} = 1$. Thus the last term equals $1/3$ in both cases. It remains to divide by the normalizing factor $1/6$ to get the result. □

Remark 6.2. In the previous proof, we establish that

$$G(a) = \mathbb{P}\left(U\leqslant aV,\, U'\leqslant aV\right)$$

is equal to $\frac{1}{\pi}\arctan\sqrt{1+2a^2}$ where U, U' and V are independent standard Gaussian random variables. Actually, this result is also a straightforward consequence of Lemma 4.3 in [2] with $X = (aV-U)/\sqrt{a^2+1}$ and $Y = (aV-U')/\sqrt{a^2+1}$. Nevertheless, since our proof is different and elegant, we decide not to skip it.

References

[1] A. Antoniadis. Analysis of variance on function spaces. *Statistics: A Journal of Theoretical and Applied Statistics*, 15(1):59–71, 1984.

[2] J.M. Azaïs and M. Wschebor. *Level sets and extrema of random processes and fields.* John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[3] E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, 2007.

[4] E. Borgonovo and M. Baucells. Invariant probabilistic sensitivity analysis. *Management Science*, 59(11):2536–2549, 2013.

[5] E. Borgonovo, W. Castaings, and S. Tarantola. Moment independent importance measures: New results and analytical test cases. *Risk Analysis*, 31(3):404–428, 2011.

[6] E. Borgonovo, W. Castaings, and S. Tarantola. Model emulation and moment-independent sensitivity analysis: An application to environmental modelling. *Environmental Modelling & Software*, 34:105–115, 2012.

[7] E. Borgonovo, G. Hazen, and E. Plischke. Probabilistic sensitivity measures: Foundations and estimation. *Submitted*, pages 1–24, 2014.

[8] E. Borgonovo and B. Iooss. *Moment Independent Importance Measures and a Common Rationale*. Preprint, 2015.

[9] E. Borgonovo and B. Iooss. *Moment-Independent and Reliability-Based Importance Measures*, pages 1–23. Springer International Publishing, Cham, 2016.

[10] R. Buchbinder and A. S. Detsky. Management of suspected giant cell arteritis: A decision analysis. *J. Rheumatology*, 19(9):1220–1228, 1992.

[11] S. Da Veiga. Global sensitivity analysis with dependence measures. *J. Stat. Comput. Simul.*, 85(7):1283–1305, 2015.

[12] Y. De Castro and A. Janon. Randomized pick-freeze for sparse Sobol indices estimation in high dimension. *ArXiv e-prints*, March 2014.

[13] E. De Rocquigny, N. Devictor, and S. Tarantola. *Uncertainty in industrial practice*. Wiley Online Library, 2008.

[14] J.C. Felli and G. Hazen. Sensitivity analysis and the expected value of perfect information. *Med. Decis. Making*, 18(1):95–109, 1998.

[15] J.C. Felli and G. Hazen. Javelin diagrams: A graphical tool for probabilistic sensitivity analysis. *Decision Analysis*, 1(2):93–107, 2004.

[16] J.-C. Fort, T. Klein, and N. Rachdi. New sensitivity analysis subordinated to a contrast. *Communications in Statistics - Theory and Methods*, 2015 to appear.

[17] F. Gamboa, A. Janon, T. Klein, and A. Lagnoux. Sensitivity analysis for multidimensional and functional outputs. *Electronic Journal of Statistics*, 8:575–603, 2014.

[18] F. Gamboa, A. Janon, T. Klein, A. Lagnoux-Renaudie, and C. Prieur. Statistical inference for Sobol pick freeze Monte Carlo method, March 2013.

[19] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364, 1 2014.

[20] A. Müller. Integral probability metrics and their generating classes of functions. *Adv. in Appl. Probab.*, 29(2):429–443, 1997.

[21] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton, NJ. Princeton University Press, 1953.

[22] J. E. Oakley. Decision-theoretic sensitivity analysis for complex computer models. *Technometrics*, 51(2):121–129, 2009.

[23] A.B. Owen. Variance components and generalized sobol' indices. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):19–41, 2013.

[24] A.B. Owen, J. Dick, and S. Chen. Higher order sobol' indices. *Information and Inference*, 3(1):59–81, 2014.

[25] A. Saltelli, K. Chan, and E.M. Scott. *Sensitivity analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000.

[26] T. J. Santner, B. Williams, and W. Notz. *The Design and Analysis of Computer Experiments*. Springer-Verlag, 2003.

[27] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, 1(4):407–414 (1995), 1993.

[28] H. J. ter Horst. On Stieltjes integration in Euclidean space. *J. Math. Anal. Appl.*, 114(1):57–74, 1986.

[29] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.