HAL open science

Sobol-Hoeffding decomposition: bounds and extremes

Olivier Roustant

▶ To cite this version:

Olivier Roustant. Sobol-Hoeffding decomposition: bounds and extremes. 2018. hal-01953080

HAL Id: hal-01953080
https://hal.science/hal-01953080

Preprint submitted on 12 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sobol-Hoeffding decomposition: bounds and extremes

Olivier Roustant ⋆

Synthesis of works with

F. Barthe, J. Fruth, F. Gamboa, B. Iooss,
S. Kuhnt, C. Mercadier and T. Muehlenstaedt

⋆ Mines Saint-Étienne

Seminar of Statistics at IMT Toulouse, 2018 December 11

Outline

1. **Sobol-Hoeffding decomposition**
   - Definition and ANOVA
   - Supersets and application to screening
2. **Computational shortcuts based on derivatives**
   - Upper bounds with Poincaré inequalities
   - Lower bounds with geometry
3. **Connexion with extremes: the tail dependograph**

Part I

Sobol-Hoeffding decomposition

Sobol-Hoeffding decomposition

Framework. $X = (X_1, \ldots, X_d)$ is a vector of independent input variables with distribution $\mu_1 \otimes \cdots \otimes \mu_d$, and $g : \Delta \subseteq \mathbb{R}^d \to \mathbb{R}$ is such that $g(X) \in L^2(\mu)$.

Theorem [Hoeffding, 1948, Efron and Stein, 1981, Sobol, 1993]

There exists a unique expansion of g of the form

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

such that $E[g_I(X_I)|X_J] = 0$ for all $I \subseteq \{1, \ldots, d\}$ and all $J \subsetneq I$.

Sobol-Hoeffding decomposition

Framework. $X = (X_1, \dots, X_d)$ is a vector of independent input variables with distribution $\mu_1 \otimes \cdots \otimes \mu_d$, and $g : \Delta \subseteq \mathbb{R}^d \to \mathbb{R}$ is such that $g(X) \in L^2(\mu)$.

Theorem [Hoeffding, 1948, Efron and Stein, 1981, Sobol, 1993]

There exists a unique expansion of g of the form

$$g(X) = g_0 + \sum_{i=1}^{d} g_i(X_i) + \sum_{1 \leq i < j \leq d} g_{i,j}(X_i, X_j) + \cdots + g_{1,\dots,d}(X_1, \dots, X_d)$$

such that $E[g_I(X_I) | X_J] = 0$ for all $I \subseteq \{1, \dots, d\}$ and all $J \subsetneq I$. Moreover:

$$\begin{aligned}
g_0 &= \mathbb{E}[g(X)] \\
g_i(X_i) &= \mathbb{E}[g(X)|X_i] - g_0 \\
g_I(X_I) &= \mathbb{E}[g(X)|X_I] - \sum_{J \subsetneq I} g_J(X_J) \qquad \text{(recursion)} \\
&= \sum_{J \subseteq I} (-1)^{|I|-|J|} \mathbb{E}[g(X)|X_J] \qquad \text{(inclusion-exclusion)}
\end{aligned}$$

Variance decomposition

- The non-overlapping condition

$$\mathbb{E}[g_I(X_I)|X_J] = 0 \quad \text{for all} \quad J \subsetneq I$$

avoids one term to be considered as a more complex one.

Variance decomposition

- The non-overlapping condition

$$\mathbb{E}[g_I(X_I)|X_J] = 0 \quad \text{for all} \quad J \subsetneq I$$

avoids one term to be considered as a more complex one.

- It implies that *$g_I(X_I)$ is orthogonal to $L^2(X_J)$ such that $J \cap I \subsetneq I$*:

$$\begin{aligned} \mathbb{E}[g_I(X_I)h(X_J)] &= \mathbb{E}[\mathbb{E}[g_I(X_I)h_J(X_J)|X_J]] \\ &= \mathbb{E}[h(X_J)\mathbb{E}[g_I(X_I)|X_{J\cap I}]] = 0 \end{aligned}$$

Variance decomposition

- The non-overlapping condition

$$\mathbb{E}[g_I(X_I)|X_J] = 0 \quad \text{for all} \quad J \subsetneq I$$

avoids one term to be considered as a more complex one.

- It implies that *$g_I(X_I)$ is orthogonal to $L^2(X_J)$ such that $J \cap I \subsetneq I$*:

$$\begin{aligned}\mathbb{E}[g_I(X_I)h(X_J)] &= \mathbb{E}[\mathbb{E}[g_I(X_I)h_J(X_J)|X_J]] \\ &= \mathbb{E}[h(X_J)\mathbb{E}[g_I(X_I)|X_{J\cap I}]] = 0\end{aligned}$$

In particular *the decomposition is orthogonal (ANOVA)*:

$$D := \mathrm{Var}(g(X)) = \sum_{I \subseteq \{1,\dots,d\}} \mathrm{Var}(g_I(X_I))$$

Orthogonal projections

Property

For each $I \subseteq \{1, \dots, d\}$, the map $\Pi_I : g \mapsto g_I$ is an orthogonal projection

Orthogonal projections

Property

For each $I \subseteq \{1, \dots, d\}$, the map $\Pi_I : g \mapsto g_I$ is an orthogonal projection

Proof.
Using the non-overlapping condition:

- Projection: applying twice the decomposition leaves it unchanged.
- Orthogonality:

$$\begin{aligned} \langle \Pi_I g, h \rangle &= \mathbb{E}(g_I(X_I)h(X)) \\ &= \sum_{J \subseteq \{1,\dots,d\}} \mathbb{E}(g_I(X_I)h_J(X_J)) = \mathbb{E}(g_I(X_I)h_I(X_I)) = \langle g, \Pi_I h \rangle \end{aligned}$$

since if $J \neq I$, then $I \cap J \subsetneq I$ or $I \cap J \subsetneq J$, thus $\mathbb{E}(g_I(X_I)h_J(X_J)) = 0$.

Multivariate decompositions with commuting projections

S.-H. dec. is an example of multivariate decompositions obtained with *(a class of) commuting projections* $P_1, \dots, P_d$ ([Kuo et al., 2010]), here orthogonals:

$$P_j(g)(x) = \int g(x) d\mu_j(x_j) = \mathbb{E}[g(X) | X_{-j} = x_{-j}]$$

Multivariate decompositions with commuting projections

S.-H. dec. is an example of multivariate decompositions obtained with *(a class of) commuting projections $P_1, \dots, P_d$* ([Kuo et al., 2010]), here orthogonals:

$$P_j(g)(x) = \int g(x) d\mu_j(x_j) = \mathbb{E}[g(X) | X_{-j} = x_{-j}]$$

The form of the decomposition is simply obtained by expansion:

$$I_d \quad = \quad (P_1 + (I_d - P_1)) \dots (P_d + (I_d - P_d))$$

Multivariate decompositions with commuting projections

S.-H. dec. is an example of multivariate decompositions obtained with *(a class of) commuting projections $P_1, \dots, P_d$* ([Kuo et al., 2010]), here orthogonals:

$$P_j(g)(x) = \int g(x) d\mu_j(x_j) = \mathbb{E}[g(X)|X_{-j} = x_{-j}]$$

The form of the decomposition is simply obtained by expansion:

$$\begin{aligned} I_d &= (P_1 + (I_d - P_1)) \dots (P_d + (I_d - P_d)) \\ &= \sum_{I \subseteq \{1,\dots,d\}} \underbrace{\prod_{j \notin I} P_j \prod_{k \in I} (I - P_k)}_{\Pi_I} \end{aligned}$$

The non-overlapping condition is written here $P_i(g_I) = 0$ for all $i \in I$.
We find again that *Π_I is an orthogonal projection*.

An example: separable functions

Consider $g(x) = f_1(x_1)\dots f_d(x_d)$, and denote $m_j = \mathbb{E}(X_j)$. Then:

$$\begin{aligned} g_I(x_I) &= \prod_{i \in I}(f_i(x_i) - m_i)\prod_{j \notin I} m_j \\ g_I^{\text{tot}}(x) &= \prod_{i \in I}(f_i(x_i) - m_i)\prod_{j \notin I} f_j(x_j) \end{aligned}$$

Proof. The Sobol-Hoedding decomposition is obtained by expanding:

$$g(x) = ((f_1(x_1) - m_1) + m_1)\dots((f_d(x_d) - m_d) + m_d)$$

For each bracket,

- for g_I, choose $(f_i(x_i) - m_i)$ if $i \in I$, and m_j otherwise
- for g_I^{tot}, choose $(f_i(x_i) - m_i)$ if $i \in I$

Sensitivity indices

Sobol indices

- Partial variances: $D_I = \mathrm{Var}(g_I(X_I))$, and *Sobol indices* $S_I = D_I / D$

$$D = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- $D_i^{\text{tot}} = \sum_{J \supseteq \{i\}} D_J, \qquad S_i^{\text{tot}} = \frac{D_i^{\text{tot}}}{D}$ *Total index*
- $D_I^{\text{tot}} = \sum_{J \supseteq \{I\}} D_J, \qquad S_I^{\text{tot}} = \frac{D_I^{\text{tot}}}{D}$ *Total interaction, superset importance*

Sensitivity indices

Sobol indices

- Partial variances: $D_I = \mathrm{Var}(g_I(X_I))$, and *Sobol indices* $S_I = D_I / D$

$$D = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- $D_i^{\text{tot}} = \sum_{J \supseteq \{i\}} D_J,\qquad S_i^{\text{tot}} = \frac{D_i^{\text{tot}}}{D}\qquad$ *Total index*
- $D_I^{\text{tot}} = \sum_{J \supseteq \{I\}} D_J,\qquad S_I^{\text{tot}} = \frac{D_I^{\text{tot}}}{D}\qquad$ *Total interaction, superset importance*

Derivative Global Sensitivity Measure (DGSM)

$$\nu_i = \int \left(\frac{\partial g(x)}{\partial x_i} \right)^2 d\mu(x), \qquad \nu_I = \int \left(\frac{\partial^{|I|} g(x)}{\partial x_I} \right)^2 d\mu(x)$$

Usage for screening

Assume that:

- g is continuous on $\Delta = [0,1]^d$
- for all i, the support of μ_i contains $[0,1]$

- **Variable screening**
  If either $D_i^{tot} = 0$ or $\nu_i = 0$, then X_i is non influential

Usage for screening

Assume that:

- g is continuous on $\Delta = [0,1]^d$
- for all i, the support of μ_i contains $[0,1]$

- **Variable screening**
  If either $D_i^{tot} = 0$ or $\nu_i = 0$, then X_i is non influential

- **Interaction screening**
  If either $D_{i,j}^{tot} = 0$ or $\nu_{i,j} = 0$, then $(x_i, x_j) \mapsto g(x)$ is additive

  Total interactions can be visualized on the *FANOVA graph*, where the edge size is proportionnal to the index value.

Illustration on a toy example

8D g-Sobol function, with uniform inputs on $[0, 1]$:

$$g(x) = \prod_{j=1}^{8} \frac{|4x_j - 2| + a_j}{1 + a_j}$$

with $a = c(0, 1, 4.5, 9, 99, 99, 99, 99)$.

Illustration on a toy example

8D g-Sobol function, with uniform inputs on $[0, 1]$:

$$g(x) = \prod_{j=1}^{8} \frac{|4x_j - 2| + a_j}{1 + a_j}$$

with $a = c(0, 1, 4.5, 9, 99, 99, 99, 99)$.

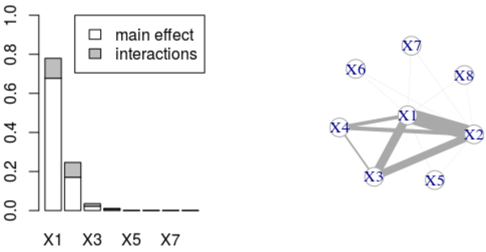


Figure: 1st order analysis (left) and 2nd order analysis (right) with 10^5 simulated data

Illustration on a toy example

A 6D block-additive function, with uniform inputs on $[-1, 1]$:

$$g(x) = \cos([1, x_1, x_2, x_3]^\top \beta) + \sin([1, x_4, x_5, x_6]^\top \gamma))$$

with $\beta = (-0.8, -1.1, 1.1, 1)^\top$ and $\gamma = (-0.5, 0.9, 1, -1.1)^\top$.

Illustration on a toy example

A 6D block-additive function, with uniform inputs on $[-1, 1]$:

$$g(x) = \cos([1, x_1, x_2, x_3]^\top \beta) + \sin([1, x_4, x_5, x_6]^\top \gamma))$$

with $\beta = (-0.8, -1.1, 1.1, 1)^\top$ and $\gamma = (-0.5, 0.9, 1, -1.1)^\top$.

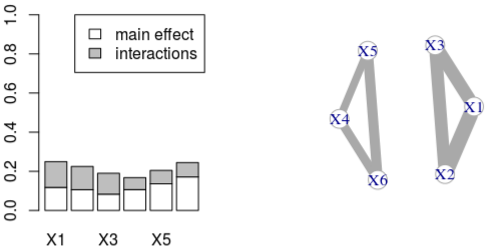


Figure: 1st order analysis (left) and 2nd order analysis (right) with 10^5 simulated data

Part II

Upper bounds for Sobol indices

Variance-based and derivative-based measures

- **Usage for screening.**
  If either $D_i^{tot} = 0$ or $\nu_i = 0$, then X_i is non influential

- **Advantages / Drawbacks**

	Computational cost	Interpretability
Sobol indices	-	+
DGSM	+	-

Variance-based and derivative-based measures

- **Usage for screening.**
  If either $D_i^{tot} = 0$ or $\nu_i = 0$, then X_i is non influential

- **Advantages / Drawbacks**

	Computational cost	Interpretability
Sobol indices	-	+
DGSM	+	-

$$\downarrow$$

Can we use DGSM to do screening based on Sobol indices?

Poincaré inequality

Poincaré inequality (1-dimensional case)

A distribution μ satisfies a Poincaré inequality if the energy in $L^2(\mu)$ sense of any centered function is controlled by the energy of its derivative:

For all h in $L^2(\mu)$ such that $\int h(x)d\mu(x) = 0$, and $h'(x) \in L^2(\mu)$:

$$\int h(x)^2 d\mu(x) \leq C(\mu) \int h'(x)^2 d\mu(x)$$

The best constant is denoted $C_{\mathrm{P}}(\mu)$.

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Proof 1. Denote $g_i^{\text{tot}}(x) := \sum_{J \supseteq \{i\}} g_J(x_J)$. Then:

$$\frac{\partial g(x)}{\partial x_i} = \frac{\partial g_i^{\text{tot}}(x)}{\partial x_i}$$

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Proof 1. Denote $g_i^{\text{tot}}(x) := \sum_{J \supseteq \{i\}} g_J(x_J)$. Then:

$$\frac{\partial g(x)}{\partial x_i} = \frac{\partial g_i^{\text{tot}}(x)}{\partial x_i}$$

$$\begin{aligned} D_i^{\text{tot}} = \mathrm{Var}(g_i^{\text{tot}}(x)) &= \int \left(g_i^{\text{tot}}(x)\right)^2 d\mu(x) \\ &\leq C(\mu_i) \int \left(\frac{\partial g_i^{\text{tot}}(x)}{\partial x_i}\right)^2 d\mu(x) = C(\mu_i)\nu_i \end{aligned}$$

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Proof 2. Denote $g_{i,j}^{\text{tot}}(x) := \sum_{J \supseteq \{i,j\}} g_J(x_J)$. Then:

$$\frac{\partial^2 g(x)}{\partial x_i \partial x_j} = \frac{\partial^2 g_{i,j}^{\text{tot}}(x)}{\partial x_i \partial x_j}$$

Link between total Sobol indices and DGSM

Theorem [Lamboni et al., 2013], [Roustant et al., 2014]

If μ_i and μ_j admit a Poincaré inequality, then:

$$D_i \leq D_i^{\text{tot}} \leq C(\mu_i)\nu_i, \qquad D_{i,j} \leq D_{i,j}^{\text{tot}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}$$

Proof 2. Denote $g_{i,j}^{\text{tot}}(x) := \sum_{J \supseteq \{i,j\}} g_J(x_J)$. Then:

$$\frac{\partial^2 g(x)}{\partial x_i \partial x_j} = \frac{\partial^2 g_{i,j}^{\text{tot}}(x)}{\partial x_i \partial x_j}$$

$$\begin{aligned} D_{i,j}^{\text{tot}} = \mathrm{Var}(g_{i,j}^{\text{tot}}(x)) &= \int \left(g_{i,j}^{\text{tot}}(x)\right)^2 d\mu(x) \\ &\leq C(\mu_i) \int \left(\frac{\partial g_{i,j}^{\text{tot}}(x)}{\partial x_i}\right)^2 d\mu(x) \\ &\leq C(\mu_i)C(\mu_j) \int \left(\frac{\partial}{\partial x_j} \frac{\partial g_{i,j}^{\text{tot}}(x)}{\partial x_i}\right)^2 d\mu(x) = C(\mu_i)C(\mu_j)\nu_{i,j} \end{aligned}$$

Getting optimal Poincaré constants on intervals

Assume that $d\mu_1(t)/dt = e^{-V(t)} > 0$ on a bounded interval $[a, b]$. Then, the smallest Poincaré constant $C(\mu_1)$ is obtained by solving a *spectral problem*:

$$Lf := f'' - V'f' = -\lambda f \qquad \text{with} \quad f'(a) = f'(b) = 0$$

Comments.

- For some (rare) pdf, $C(\mu_1)$ can be computed semi-analytically.
- For many other ones, a finite element method can be used.
- Adaptations are possible for unbounded intervals and pdf vanishing at the boundaries.

See technical details in [Roustant et al., 2017].

Optimal Poincaré constants: Examples

pdf	Support	C_{opt}	Form of $f_{\mathrm{opt}}(x)$
Uniform	$[a, b]$	$(b-a)^2/\pi^2$	$\cos\left(\frac{\pi(x-a)}{b-a}\right)$
$\mathcal{N}(\mu, \sigma^2)$	$\mathbb{R}$	σ^2	$x-\mu$
	$[r_{n,i}, r_{n,i+1}]$	$1/(n+1)$	$H_{n+1}(x)$
	$[a, b]$	*related to Kummer*	*hypergeom. func.*
Db. exp. $e^{-\lvert x\rvert}dx/2$	$\mathbb{R}$	4	$\times$
(*)	$[a, b], ab > 0$	$\left(\frac{1}{4}+\omega^2\right)^{-1}$	$e^{x/2}\cos(\omega x+\phi)$
(*, **)	$[a, b], ab \leq 0$	$> \left(\frac{1}{4}+\omega^2\right)^{-1}$	$e^{\lvert x\rvert/2} \times$ *trig. spline*
Logistic $\frac{e^x}{(1+e^x)^2}dx$	$\mathbb{R}$	4	$\times$
Triangular	$[-1, 1]$	≈ 0.1729	*linked to Bessel* J_0

(*) For the truncated Exponential on $[a, b] \subseteq \mathbb{R}^+$, we use $\omega = \pi/(b-a)$
(**) If $a < 0 < b$, the spectral gap is the zero in $]0, \min(\pi/|a|, \pi/|b|)[$ of $x \mapsto \mathrm{cotan}(|a|x) + \mathrm{cotan}(|b|x) + 1/x$

Optimal Poincaré constants: Examples

Truncated normal distribution – Symmetric case: I = [-b,b]

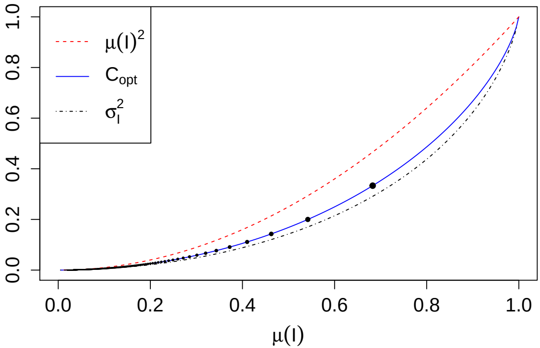
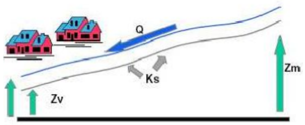


Figure: Poincaré constant of $\mu = \mathcal{N}(0,1)$ truncated on $I = [-b, b]$, vs $\mu(I)$

σ_I^2 : variance of the truncated normal on I – Black points: Hermite polynomials of even degree.

A case study for global sensitivity analysis



A simplified flood model [Iooss, 2011], [Iooss and Lemaitre, 2015].

- 1 output: maximal annual overflow (in meters), denoted by S:

$$S = Z_v + H - H_d - C_b \quad \text{with} \quad H = \left(\frac{Q}{BK_s \sqrt{\frac{Z_m - Z_v}{L}}} \right)^{0.6}$$

where H is the maximal annual height of the river (in meters).

A case study for global sensitivity analysis

- 8 inputs variables assumed to be independent r.v., with distributions:

Input	Description	Unit	Probability distribution
$X_1 = Q$	Maximal annual flowrate	m^3/s	Gumbel $\mathcal{G}(1013, 558)$, truncated on $[500, 3000]$
$X_2 = K_s$	Strickler coefficient	-	Normal $\mathcal{N}(30, 8^2)$, truncated on $[15, +\infty[$
$X_3 = Z_v$	River downstream level	m	Triangular $\mathcal{T}(49, 50, 51)$
$X_4 = Z_m$	River upstream level	m	Triangular $\mathcal{T}(54, 55, 56)$
$X_5 = H_d$	Dyke height	m	Uniform $\mathcal{U}[7, 9]$
$X_6 = C_b$	Bank level	m	Triangular $\mathcal{T}(55, 55.5, 56)$
$X_7 = L$	River stretch	m	Triangular $\mathcal{T}(4990, 5000, 5010)$
$X_8 = B$	River width	m	Triangular $\mathcal{T}(295, 300, 305)$

- *Aim: To detect unessential X_i's, to quantify the influence of X_i's on S, ...*

A case study for global sensitivity analysis

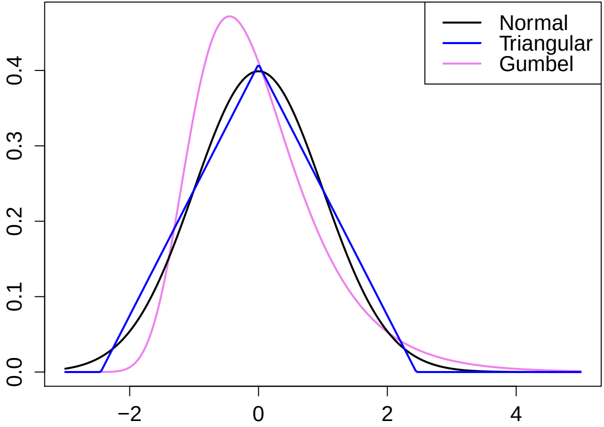
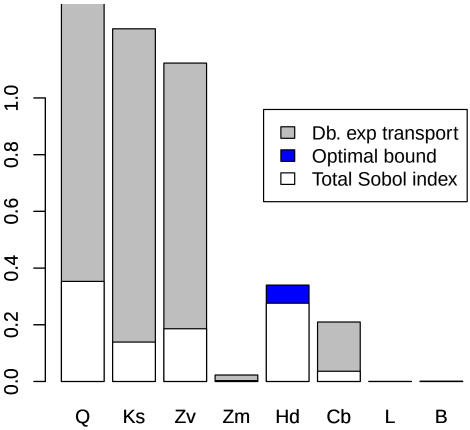
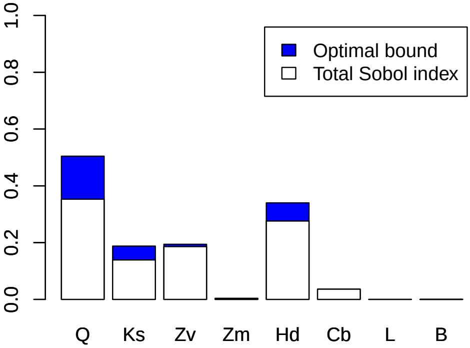


Figure: The 3 distributions types of the case study, here with mean 0 and variance 1

Results with optimal Poincaré constants



Results with optimal Poincaré constants



Part III

Lower bounds for Sobol indices

Ongoing work with F. Gamboa and B. Iooss

Principle

Without loss of generality, *assume* $g_0 = 0$. Define:

$$\begin{array}{lcll} F_1 & = & \{g \in L^2(\mu) \text{ s.t. } g = g_1\} & \text{functions } \textit{depending exactly on } x_1 \\ F_1^{\text{tot}} & = & \{g \in L^2(\mu) \text{ s.t. } g = g_1^{\text{tot}}\} & \text{functions } \textit{depending at least on } x_1 \end{array}$$

Notice that g_1 and g_1^{tot} are obtained from g by orthogonal projection

$$\begin{array}{lcl} g_1 & = & \Pi_{F_1}(g) = \mathbb{E}[g(X)|X_1 = .] \\ g_1^{\text{tot}} & = & \Pi_{F_1^{\text{tot}}}(g) = g - \mathbb{E}[g(X)|X_2 = ., \dots, X_d = .] \end{array}$$

Hence, $D_1 = \|\Pi_{F_1}(g)\|^2$ and $D_1^{\text{tot}} = \|\Pi_{F_1^{\text{tot}}}(g)\|^2$.

Lower bounds of D_1, D_1^{tot} are obtained by *projecting onto subspaces* of F_1, F_1^{tot}

Main result

Let $\phi_1, \dots, \phi_m$ be orthonormal functions in F_1^{tot}. Then:

$$D_1^{\text{tot}} \geq \sum_{j=1}^{m} \left(\int g(x)\phi_j(x) d\mu(x) \right)^2$$

with equality iff g has the form $g(x) = \sum_{j=1}^{m} \alpha_m \phi_m(x) + h(x_2, \dots, x_m)$.

If all the ϕ_j's belong to F_1 then the lower bound is for D_1.

Proof.

- $D_1^{\text{tot}} = \|g_1^{\text{tot}}\|^2 = \|\Pi_{F_1^{\text{tot}}}(g)\|^2 \geq \|\Pi_{\phi_1,\dots,\phi_m}(g)\|^2 = \sum_{j=1}^{m} \left(\langle g, \phi_j \rangle\right)^2$
- Equality is when $g_1^{\text{tot}} = \Pi_{\phi_1,\dots,\phi_m}(g)$, leading to the condition above.
- Same arguments when all the ϕ_j's are in F_1

Tensor-based lower bounds

For all j, let $\psi_{j,0} = 1, \psi_{j,1}, \dots, \psi_{j,n_j-1}$ *be orthonormal functions in* $L^2(\mu_j)$.
Consider tensors, i.e. separable functions:

$$\phi_{\underline{\ell}}(x) = \prod_{j=1}^{d} \psi_{j,\ell_j}(x_j)$$

where $\underline{\ell} = (\ell_1, \dots, \ell_d)$ is a multi-index.

Let $\mathcal{T}_1 = \{\underline{\ell} \text{ s.t. } \ell_1 \geq 1\}$*, the set of tensors* $\phi_{\underline{\ell}}$ *involving* x_1. Then:

$$D_1^{\text{tot}}(f) \geq \sum_{\underline{\ell} \in \mathcal{T}_1} \left(\int f(x) \phi_{\underline{\ell}}(x) \nu(dx) \right)^2$$

with equality iff f has the form $f(x) = \sum_{\underline{\ell} \in \mathcal{T}_1} \alpha_{\ell} \phi_{\underline{\ell}}(x) + g(x_2, \dots, x_d)$.

Tensor-based lower bounds

As an illustration, if μ_i admit the first two moments, denote:

$$\psi_i(x) = (x_i - m_i)/s_i$$

where m_i is the mean and s_i the s.d.
Then $\psi_1, \psi_1\psi_2, \dots, \psi_1\psi_j$ *are orthonormal functions of* F_1^{tot}.
Hence:

$$D_1^{\text{tot}} \geq \underbrace{\left(\int g(x)\psi_1(x)d\mu(x)\right)^2}_{\text{lower bound for } D_1} + \sum_{j=2}^{m}\left(\int g(x)\psi_1(x)\psi_j(x)d\mu(x)\right)^2$$

Derivative-based lower bounds

All the integrals above can involve derivatives by integrating by part.
But this often induce weights; Here is a partial solution to avoid weights.

Derivative-based lower bounds

All the integrals above can involve derivatives by integrating by part.
But this often induce weights; Here is a partial solution to avoid weights.

Assume that μ_j *is continuous with pdf* $p_j \in \mathcal{H}^1(\mu_j)$ vanishing at the boundaries but not inside, and such that $p_j' \not\equiv 0$ and $p_j'/p_j \in L^2(\mu_j)$. Denote:

$$Z_j(X_j) = (\ln p_j)'(X_j), \qquad I_j = \mathrm{Var}(Z_j(X_j))$$

Then:

$$D_1^{\text{tot}} \geq \underbrace{I_1^{-1} c_1^2}_{\text{lower bound for } D_1} + I_1^{-1} \sum_{j=2}^{d} I_j^{-1} c_{1,j}^2$$

with

$$c_1 = \int g(x) Z_1(x_1) d\mu(x) = -\int \frac{\partial g(x)}{\partial x_1} d\mu(x)$$

$$c_{1,j} = \int g(x) Z_1(x_1) Z_j(x_j) d\mu(x) = -\int \frac{\partial g(x)}{\partial x_1} Z_j(x_j) d\mu(x) = \int \frac{\partial^2 g(x)}{\partial x_1 \partial x_j} d\mu(x)$$

Derivative-based lower bounds: examples

For normal variables $N(m_j, s_j^2)$:

$$D_1^{\text{tot}} \geq \underbrace{s_1^2 \left(\int \frac{\partial g(x)}{\partial x_1} d\mu(x) \right)^2}_{\text{lower bound for } D_1} + s_1^2 \sum_{j=2}^{d} s_j^2 \left(\int \frac{\partial^2 g(x)}{\partial x_1 \partial x_j} d\mu(x) \right)^2$$

Dist. name	Support	p	Z	I
Normal	$\mathbb{R}$	$\frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-m)^2}{s^2}\right)$	$-(X-m)/s^2$	$1/s^2$
Laplace	$\mathbb{R}$	$\frac{1}{2s} \exp\left(\frac{\lvert x-m\rvert}{s}\right)$	$-\mathrm{sgn}(X-m)/s$	$1/s^2$
Cauchy	$\mathbb{R}$	$\frac{1}{\pi}\frac{s}{(x-x_0)^2+s^2}$	$\frac{-2(x-x_0)}{(x-x_0)^2+s^2}$	$1/(2s^2)$

Improvements on existing works

According to results given in the review [Kucherenko and Iooss, 2017],

- For normal distributions, we improve on:

$$D_1^{\text{tot}} \geq D_1 \geq s_1^2 \left(\int \frac{\partial g(x)}{\partial x_1} d\mu(x) \right)^2.$$

Improvements on existing works

According to results given in the review [Kucherenko and Iooss, 2017],

- For normal distributions, we improve on:

$$D_1^{\text{tot}} \geq D_1 \geq s_1^2 \left(\int \frac{\partial g(x)}{\partial x_1} d\mu(x) \right)^2 .$$

- For uniforms on $[0,1]$ using the orthonormal function obtained from x_1^m, and an integration by part, we obtain:

$$D_1^{\text{tot}} \geq D_1 \geq \frac{2m+1}{m^2} \left(\int (g(1, x_{-1}) - g(x))dx - w_1^{(m+1)} \right)^2$$

where $w_1^{(m+1)} = \int \frac{\partial g(x)}{\partial x_1} x_1^{m+1} dx$. This improves on the known lower bound which has the same form, with the smaller multiplicative constant $\frac{2m+1}{(m+1)^2}$.

Improvements on existing works

According to results given in the review [Kucherenko and Iooss, 2017],

- For normal distributions, we improve on:

$$D_1^{\text{tot}} \geq D_1 \geq s_1^2 \left(\int \frac{\partial g(x)}{\partial x_1} d\mu(x) \right)^2 .$$

- For uniforms on $[0, 1]$ using the orthonormal function obtained from x_1^m, and an integration by part, we obtain:

$$D_1^{\text{tot}} \geq D_1 \geq \frac{2m+1}{m^2} \left(\int (g(1, x_{-1}) - g(x)) dx - w_1^{(m+1)} \right)^2$$

where $w_1^{(m+1)} = \int \frac{\partial g(x)}{\partial x_1} x_1^{m+1} dx$. This improves on the known lower bound which has the same form, with the smaller multiplicative constant $\frac{2m+1}{(m+1)^2}$.

N.B. Better bounds are obtained by adding orth. funct. of the form $\psi_1 \psi_j$.

Results on the application

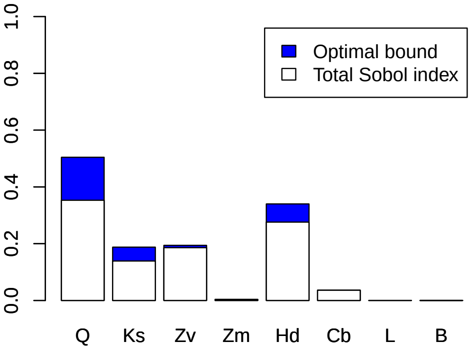


Figure: Results obtained with orth. 1st order pol. tensors $\psi_1, \psi_1\psi_2, \ldots, \psi_1\psi_8$

Results on the application

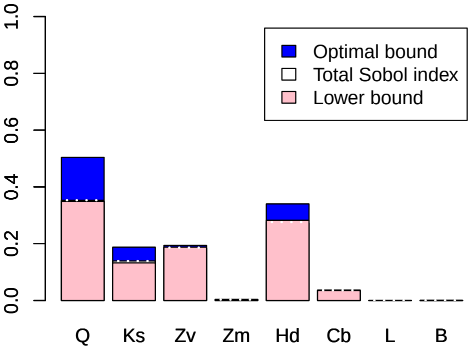


Figure: Results obtained with orth. 1st order pol. tensors $\psi_1, \psi_1\psi_2, \ldots, \psi_1\psi_8$

When using derivatives and other numerical considerations

We must compute squared integrals $\theta = (\int h(x) d\mu(x))^2$, when h has the form:

$$h_{\text{dir}} = g\phi_1, g\phi_1\phi_j, \ldots, \qquad \text{or} \qquad h_{\text{der}} = \frac{\partial g}{\partial x_i}, \frac{\partial g}{\partial x_j} Z_j, \ldots$$

for centered function ϕ_1, ϕ_j, Z_j.

When using derivatives and other numerical considerations

We must compute squared integrals $\theta = (\int h(x) d\mu(x))^2$, when h has the form:

$$h_{\text{dir}} = g\phi_1, g\phi_1\phi_j, \ldots, \qquad \text{or} \qquad h_{\text{der}} = \frac{\partial g}{\partial x_i}, \frac{\partial g}{\partial x_j} Z_j, \ldots$$

for centered function ϕ_1, ϕ_j, Z_j.

The sample estimate $\hat{\theta} = \left(\frac{1}{n} \sum_{i=1}^n h(X^i)\right)^2$, with $X^1, \ldots, X^n$ i.i.d. $\sim \mu$, verifies:

$$\hat{\theta} \approx \mathcal{N}\left(\theta, \frac{4\theta}{n} \mathrm{Var}_\mu(h)\right)$$

Hence, for one squared integral, using the derivative form can reduce estimation error when h_{der} is less variable than h_{dir}.

Partial conclusions

- Lower bounds of a (convex comb. of) ANOVA term g_I can be obtained by *projection onto subspaces of its ANOVA space* $\{g \in L^2(\mu)$ s.t. $g = g_I\}$

  $\rightarrow$ *Illustrated on main and total effects, but very general!*

Partial conclusions

- Lower bounds of a (convex comb. of) ANOVA term g_I can be obtained by *projection onto subspaces of its ANOVA space* $\{g \in L^2(\mu) \text{ s.t. } g = g_I\}$
  $\rightarrow$ *Illustrated on main and total effects, but very general!*
- *Tensors* are used to get *lower bounds as a sum of squared integrals*
  $\rightarrow$ *Chaos polynomials or more general tensors*

Partial conclusions

- Lower bounds of a (convex comb. of) ANOVA term g_I can be obtained by *projection onto subspaces of its ANOVA space* $\{g \in L^2(\mu)$ s.t. $g = g_I\}$
  $\rightarrow$ *Illustrated on main and total effects, but very general!*
- *Tensors* are used to get *lower bounds as a sum of squared integrals*
  $\rightarrow$ *Chaos polynomials or more general tensors*
- *Integration by part* modify lower bounds into *derivative-based forms*
  $\rightarrow$ *Specific choices of subspaces remove weights* for specific pdfs

Partial conclusions

- Lower bounds of a (convex comb. of) ANOVA term g_I can be obtained by *projection onto subspaces of its ANOVA space* $\{g \in L^2(\mu)$ s.t. $g = g_I\}$
  $\rightarrow$ *Illustrated on main and total effects, but very general!*
- *Tensors* are used to get *lower bounds as a sum of squared integrals*
  $\rightarrow$ *Chaos polynomials or more general tensors*
- *Integration by part* modify lower bounds into *derivative-based forms*
  $\rightarrow$ *Specific choices of subspaces remove weights* for specific pdfs
- Using derivative-based inequalities may be useful when the derivative is less variable than the function itself.

Part IV

Tail dependograph

Joint work with C. Mercadier

Multivariate dependence

Denote F a multivariate cdf,

$$F(x) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$$

Assume that F is in the domain of attraction of a max-stable distribution H i.e. there exist vector sequences $a_n > 0, b_n$ s.t. for indep. samples $X^1, \dots, X^n$ of F

$$\mathbb{P}\left(\frac{\max_{k=1}^n(X_1^k) - b_{n,1}}{a_{n,1}} \leq x_1, \dots, \frac{\max_{k=1}^n(X_d^k) - b_{n,d}}{a_{n,d}} \leq x_d\right) \underset{n\to\infty}{\to} H(x)$$

Multivariate dependence

Denote F a multivariate cdf,

$$F(x) = \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d)$$

Assume that F is in the domain of attraction of a max-stable distribution H i.e. there exist vector sequences $a_n > 0, b_n$ s.t. for indep. samples $X^1, \ldots, X^n$ of F

$$\mathbb{P}\left(\frac{\max_{k=1}^n(X_1^k) - b_{n,1}}{a_{n,1}} \leq x_1, \ldots, \frac{\max_{k=1}^n(X_d^k) - b_{n,d}}{a_{n,d}} \leq x_d\right) \underset{n\to\infty}{\to} H(x)$$

- In the univariate case, H is a generalized extreme value distribution, summarizing the three types Fréchet, Weibull, Gumbel

Multivariate dependence

Denote F a multivariate cdf,

$$F(x) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$$

Assume that F is in the domain of attraction of a max-stable distribution H i.e. there exist vector sequences $a_n > 0, b_n$ s.t. for indep. samples $X^1, \dots, X^n$ of F

$$\mathbb{P}\left(\frac{\max_{k=1}^n (X_1^k) - b_{n,1}}{a_{n,1}} \leq x_1, \dots, \frac{\max_{k=1}^n (X_d^k) - b_{n,d}}{a_{n,d}} \leq x_d \right) \underset{n \to \infty}{\to} H(x)$$

- In the univariate case, H is a generalized extreme value distribution, summarizing the three types Fréchet, Weibull, Gumbel
- In the multivariate case, the margins are gevd, and *the multivariate dependence is characterized by a multivariate function*
  - extreme value copula, *stable tail dependence function*, ...

Multivariate dependence: stable tail dependence function

Stable tail dependence function (stdf) ℓ

$$-\log H(x) = \ell(-\log H_1(x_1), \dots, -\log H_d(x_d))$$

Properties (see e.g. [de Haan and Ferreira, 2006])

- ℓ is continuous, convex and homogeneous of order 1
  $\rightarrow$ *we can restrict it on* $[0,1]^d$
- $\underbrace{\max(u_1, \dots, u_d)}_{\text{Asymptotic dependence}} \leq \ell(u) \leq \underbrace{u_1 + \cdots + u_d}_{\text{Asymptotic independence}}$
- $\ell(u) = \lim_{z \to +\infty} z \left(1 - F\left(F_1^{-1}(u_1/z), \dots, F_d^{-1}(u_d/z)\right)\right)$

Asymptotic independence and tail dependograph

Let A, B a partition of $\{1, \ldots, d\}$

X_A and X_B are *asymptotically independent*

$$
\begin{aligned}
&\Leftrightarrow \quad H(x) \text{ if of the form } H(x) = H_A(x_A)H_B(x_B) \\
&\Leftrightarrow \quad \ell(u) \text{ if of the form } \ell(u) = \ell(u_A) + \ell(u_B) \\
&\Leftrightarrow \quad \forall i \in A, \forall j \in B, \quad \ell_{i,j}^{\text{tot}} \equiv 0
\end{aligned}
$$

Asymptotic independence and tail dependograph

Let A, B a partition of $\{1, \dots, d\}$

X_A and X_B are *asymptotically independent*

$$\begin{aligned}
&\Leftrightarrow \quad H(x) \text{ if of the form } H(x) = H_A(x_A)H_B(x_B) \\
&\Leftrightarrow \quad \ell(u) \text{ if of the form } \ell(u) = \ell(u_A) + \ell(u_B) \\
&\Leftrightarrow \quad \forall i \in A, \forall j \in B, \quad \ell_{i,j}^{\text{tot}} \equiv 0
\end{aligned}$$

Thus

$$X_A \underset{\infty}{\perp\!\!\!\perp} X_B \quad \textit{if} \quad \underbrace{\textit{the FANOVA graph of } \ell}_{\text{“tail dependograph”}} \textit{ is partitioned by A and B}$$

Asymptotic independence and extremal coefficients

The extremal coefficients $\theta_I(\ell)$ are defined by

$$\mathbb{P}\left(X_j \leq F_j^{-1}(p),\ \text{for all } j \in I\right) = p^{\theta_I(\ell)}$$

Equivalently $\theta_I(\ell) = \ell(1_I)$

Asymptotic independence and extremal coefficients

The extremal coefficients $\theta_I(\ell)$ are defined by

$$\mathbb{P}\left(X_j \leq F_j^{-1}(p),\ \text{for all } j \in I\right) = p^{\theta_I(\ell)}$$

Equivalently $\theta_I(\ell) = \ell(1_I)$, and in particular

$$\underbrace{1}_{\text{Asymptotic dependence}} \leq \theta_I(\ell) \leq \underbrace{|I|}_{\text{Asymptotic independence}}$$

Asymptotic independence and extremal coefficients

The extremal coefficients $\theta_I(\ell)$ are defined by

$$\mathbb{P}\left(X_j \leq F_j^{-1}(p),\ \text{for all } j \in I\right) = p^{\theta_I(\ell)}$$

Equivalently $\theta_I(\ell) = \ell(1_I)$, and in particular

$$\underbrace{1}_{\text{Asymptotic dependence}} \leq \theta_I(\ell) \leq \underbrace{|I|}_{\text{Asymptotic independence}}$$

Hence,

$$X_i \underset{\infty}{\perp\!\!\!\perp} X_j \quad \text{if} \quad \theta_{i,j}(\ell) = 2$$

Illustration: Revealing asymptotic dependence for asymmetric models

Consider a 4-dim. random vector X with standard Gumbel margins, and s.t.d.f. built as a mixture of independence and logistic:

$$\begin{aligned}\ell(u) &= (1-w)(u_1+u_2) + w\left(u_1^{1/\alpha} + u_2^{1/\alpha}\right)^{\alpha} \\ &+ (1-w')(u_3+u_4) + w'\left(u_3^{1/\alpha'} + u_4^{1/\alpha'}\right)^{\alpha'},\end{aligned}$$

with asymetric parameters: $(w, \alpha) = (0.2, 0.2), \quad (w', \alpha') = (0.8, 0.83)$.

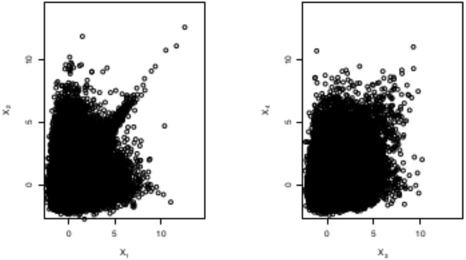
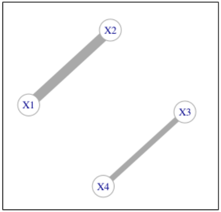


Illustration: Revealing asymptotic dependence for asymmetric models



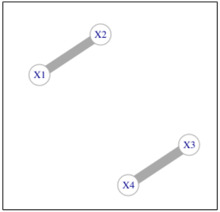


Figure: Tail dependograph (left) and graph representing $2 - \theta_{i,j}$ (right)

$\rightarrow$ Both indices recover the asympt. indep. between (X_1, X_2) and (X_3, X_4)

$\rightarrow$ Asymmetry in tail dependence is more visible on tail dependograph

Inference

The formula $\ell(u) = \lim_{z \to +\infty} z \left(1 - F\left(F_1^{-1}(u_1/z), \dots, F_d^{-1}(u_d/z)\right)\right)$ leads to the natural estimator ([Huang, 1992])

$$
\begin{aligned}
\hat{\ell}_{k,n}(u) &= \frac{n}{k}\left(1 - \frac{1}{n}\sum_{s=1}^{n} \mathbf{1}\left\{X_s^{(1)} < X^{(1)}_{n-[ku_1]+1,n}, \dots, X_s^{(d)} < X^{(d)}_{n-[ku_d]+1,n}\right\}\right) \\
&= \frac{n}{k}\left(1 - \frac{1}{n}\sum_{s=1}^{n} \mathbf{1}\left\{u_1 < \tilde{R}_s^{(1)}, \dots, u_d < \tilde{R}_s^{(d)}\right\}\right) \\
&= \frac{n}{k} - \frac{1}{k}\sum_{s=1}^{n} \underbrace{\prod_{t=1}^{d} \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\}}_{\text{separable function}}
\end{aligned}
$$

with:

- $X^{(t)}_{1,n}, \dots, X^{(t)}_{n,n}$: sorted data (asc. order) for coordinate t
- $\tilde{R}_s^{(t)} := \frac{n - R_s^{(t)} + 1}{k}$, where $R_s^{(t)}$ is the rank of $X_s^{(t)}$ among $X_1^{(t)}, \dots, X_n^{(t)}$.

Inference

Let $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ a measure on $[0,1]^d$ (without special link with F).

As a sum of separable functions, *the whole Sobol-Hoeffding decomposition of the stdf estimator can be computed in closed form*, and in particular

$$\hat{\ell}^{\text{tot}}_{k,n;\{i,j\}}(u) = -\frac{1}{k}\sum_{s=1}^{n}\prod_{t=1}^{d}\left(\mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} - \mathbf{1}_{\{t\in\{i,j\}\}}\mu_t\left(\tilde{R}_s^{(t)}\right)\right)$$

and the tail dependograph as well

$$D^{\text{tot}}_{\{i,j\}}(\hat{\ell}_{k,n}) = \frac{1}{k^2}\sum_{s=1}^{n}\sum_{s'=1}^{n}\prod_{t=1}^{d}\left(\mu_t\left(\tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)}\right) - \mathbf{1}_{\{t\in\{i,j\}\}}\mu_t\left(\tilde{R}_s^{(t)}\right)\mu_t\left(\tilde{R}_{s'}^{(t)}\right)\right).$$

Inference

As the terms of S.-H. decomposition are obtained by linear operation,

inference properties of the stdf transfer to its ANOVA terms...

Inference

As the terms of S.-H. decomposition are obtained by linear operation,

inference properties of the stdf transfer to its ANOVA terms...

Consider the usual assumptions for stdf inference, with corresponding valid sequences $k = k(n)$. Then, for all $I \subseteq \{1, \dots, d\}$,

- $\sup_{u_I \in [0,1]^{|I|}} |\hat{\ell}_{k,n;I}(u_I) - \ell_I(u_I)| \xrightarrow{\mathbb{P}} 0$.
- $\sqrt{k} \left\{ \hat{\ell}_{k,n;I}(u_I) - \ell_I(u_I) \right\} \xrightarrow{d} Y_{\ell;I}(u_I)$

where $Y_{\ell;I}$ is some Gaussian process.

Inference

As the terms of S.-H. decomposition are obtained by linear operation,

inference properties of the stdf transfer to its ANOVA terms...

Consider the usual assumptions for stdf inference, with corresponding valid sequences $k = k(n)$. Then, for all $I \subseteq \{1, \dots, d\}$,

- $\sup_{u_I \in [0,1]^{|I|}} |\hat{\ell}_{k,n;I}(u_I) - \ell_I(u_I)| \xrightarrow{\mathbb{P}} 0$.
- $\sqrt{k}\left\{\hat{\ell}_{k,n;I}(u_I) - \ell_I(u_I)\right\} \xrightarrow{d} Y_{\ell;I}(u_I)$

where $Y_{\ell;I}$ is some Gaussian process.

... and hence to the tail dependograph

- $D_I(\hat{\ell}_{k,n}) \xrightarrow{\mathbb{P}} D_I(\ell)$
- If $D_I(\ell) > 0$, then $D_I(\hat{\ell}_{k,n})$ is asympt. normal with rate $\sqrt{k}$
- If $D_I(\ell) = 0$, then $D_I(\hat{\ell}_{k,n})$ is asympt. χ^2 type with rate k

(The same is true for D_I^{tot})

Inference

(A piece of intuition about asymptotic distribution)

$$\hat{\ell}_{k,n;I}(u_I) \quad = \quad \ell_I(u_I) + \frac{1}{\sqrt{k}} Y_{\ell,I}(u_I) + \dots$$

Inference

(A piece of intuition about asymptotic distribution)

$$\hat{\ell}_{k,n;I}(u_I) \quad = \quad \ell_I(u_I) + \frac{1}{\sqrt{k}} Y_{\ell,I}(u_I) + \dots$$

- If $\ell_I \not\equiv 0$,

$$\underbrace{\int \hat{\ell}^2_{k,n;I}(u_I) d\mu(u)}_{D_I(\hat{\ell}_{k,n})} = \underbrace{\int \ell_I^2(u_I) d\mu(u)}_{D_I(\ell)} + \frac{1}{\sqrt{k}} \underbrace{\int 2 Y_{\ell,I}(u_I)\ell_I(u_I) d\mu(u)}_{\text{a Gaussian r.v.}} + \dots$$

Inference

(A piece of intuition about asymptotic distribution)

$$\hat{\ell}_{k,n;I}(u_I) \quad = \quad \ell_I(u_I) + \frac{1}{\sqrt{k}} Y_{\ell,I}(u_I) + \dots$$

- If $\ell_I \not\equiv 0$,

$$\underbrace{\int \hat{\ell}^2_{k,n;I}(u_I) d\mu(u)}_{D_I(\hat{\ell}_{k,n})} = \underbrace{\int \ell_I^2(u_I) d\mu(u)}_{D_I(\ell)} + \frac{1}{\sqrt{k}} \underbrace{\int 2 Y_{\ell,I}(u_I) \ell_I(u_I) d\mu(u)}_{\text{a Gaussian r.v.}} + \dots$$

- If $\ell_I \equiv 0$,

$$\underbrace{\int \hat{\ell}^2_{k,n;I}(u_I) d\mu(u)}_{D_I(\hat{\ell}_{k,n})} = \underbrace{0}_{D_I(\ell)} + \frac{1}{k} \underbrace{\int Y^2_{\ell,I}(u_I) d\mu(u)}_{\text{a } \chi^2 \text{ type r.v.}} + \dots$$

Application on real data

Dataset: yearly maxima temperatures at 21 French cities during 1946 – 2000.

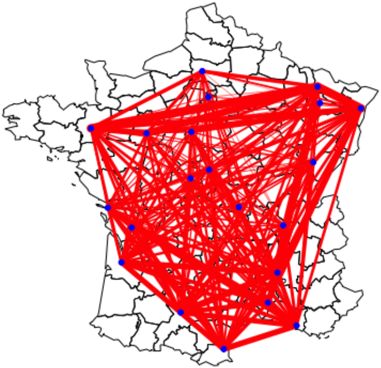


Figure: Estimated tail dependograph: complete, 30 largest values, 9 largest

Application on real data

Dataset: yearly maxima temperatures at 21 French cities during 1946 – 2000.

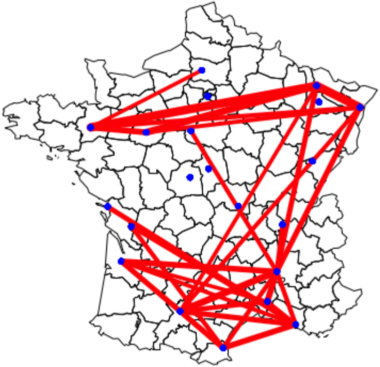


Figure: Estimated tail dependograph: complete, 30 largest values, 9 largest

Application on real data

Dataset: yearly maxima temperatures at 21 French cities during 1946 – 2000.

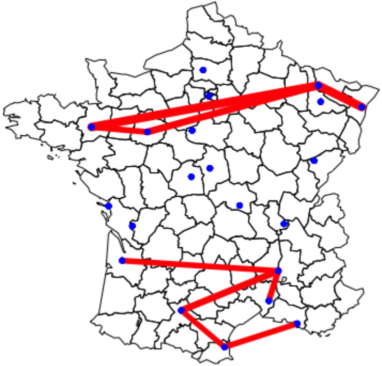


Figure: Estimated tail dependograph: complete, 30 largest values, 9 largest

Some conclusions

Tail dependograph is a graphical tool to investigate multivariate independence.

- Asymptotic independence is visible by partitions in the graph
- Asymetric seems to be better visible, compared to extremal coefficients

Some conclusions

Tail dependograph is a graphical tool to investigate multivariate independence.

- Asymptotic independence is visible by partitions in the graph
- Asymetric seems to be better visible, compared to extremal coefficients
- A natural estimator can be computed analytically
- Inference properties of the stdf transfer to the tail dependograph

Acknowledgements

Part of this research was conducted within the frame of the Chair in Applied Mathematics OQUAIDO, gathering partners in technological research (BRGM, CEA, IFPEN, IRSN, Safran, Storengy) and academia (CNRS, Ecole Centrale de Lyon, Mines Saint-Etienne, University of Grenoble, University of Nice, University of Toulouse) around advanced methods for Computer Experiments.

Part V

References - Thank you for your attention!

de Haan, L. and Ferreira, A. (2006).
Extreme value theory. An introduction.
Springer Series in Operations Research and Financial Engineering. New York, NY: Springer.

Efron, B. and Stein, C. (1981).
The jackknife estimate of variance.
The Annals of Statistics, 9(3):586–596.

Fruth, J., Roustant, O., and Kuhnt, S. (2014).
Total interaction index: A variance-based sensitivity index for second-order interaction screening.
Journal of Statistical Planning and Inference, 147:212 – 223.

Hoeffding, W. (1948).
A class of statistics with asymptotically normal distribution.
Ann. Math. Statist., 19(3):293–325.

Huang, X. (1992).
Statistics of bivariate extremes.
PhD Thesis, Erasmus University Rotterdam, Tinbergen Institute Research series No. 22.

Iooss, B. (2011).
Revue sur l'analyse de sensibilité globale de modèles numériques.
Journal de la Société Francaise de Statistique, 152:1–23.

Iooss, B. and Lemaitre, P. (2015).
A review on global sensitivity analysis methods.
In Meloni, C. and Dellino, G., editors, *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, pages 101–122. Springer.

Kucherenko, S. and Iooss, B. (2017).
Derivative-Based Global Sensitivity Measures, pages 1241–1263.
Springer International Publishing, Cham.

Kuo, F., Sloan, I., Wasilkowski, G., and H.Woźniakowski (2010).
On decompositions of multivariate functions.
Mathematics of computation, 79(270):953–966.

Lamboni, M., Iooss, B., Popelin, A.-L., and Gamboa, F. (2013).
Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests.
Mathematics and Computers in Simulation, 87:45–54.

Mercadier, C. and Roustant, O. (2018).
The hoeffding-sobol decomposition in extreme value theory. exploring the asymptotic dependence structure.
https://hal.archives-ouvertes.fr/hal-01649596.

Muehlenstaedt, T., Roustant, O., Carraro, L., and Kuhnt, S. (2012).
Data-driven kriging models based on fanova-decomposition.
Statistics and Computing, 22(3):723–738.

Petit, S., Zaoui, F., Popelin, A.-L., Goeury, C., and Goutal, N. (2016).
Couplage entre indices à base de dérivées et mode adjoint pour l'analyse de sensibilité globale. Application sur le code Mascaret.
Preprint, https://hal.archives-ouvertes.fr/hal-01373535.

Roustant, O., Barthe, F., and Iooss, B. (2017).
Poincaré inequalities on intervals - application to sensitivity analysis.
Electron. J. Statist., 11(2):3081–3119.

Roustant, O., Fruth, J., Iooss, B., and Kuhnt, S. (2014).
Crossed-derivative based sensitivity measures for interaction screening.
Mathematics and Computers in Simulation, 105:105 – 118.

Sobol, I. (1993).

Sensitivity estimates for non linear mathematical models.
Mathematical Modelling and Computational Experiments, 1:407–414.