Global sensitivity analysis

O. Roustant, INSA Toulouse

Outline

- Motivation
- The Morris method
- The Sobol-Hoeffding (ANOVA) decomposition
- Ressources

MOTIVATION

What we expect from sensitivity analysis

- $y = f(x_1, \ldots, x_d)$

- Questions :
  - Which variables x_i are the most influential on y ?
  - Which one(s) have no influence ?
  - Are there interactions ?
  - Can we quantify the effects ?
    - x_i alone,
    - x_i with others

Local sensitivity analysis

- Computation of finite differences:

  $[f(x_1, \ldots, x_{i-1}, x_{i0}+h, \ldots, x_d) - f(x_1, \ldots, x_{i-1}, x_{i0}, \ldots, x_d)] / h$

- Limitations:
  - Gives a good idea of the effect of x_i ...

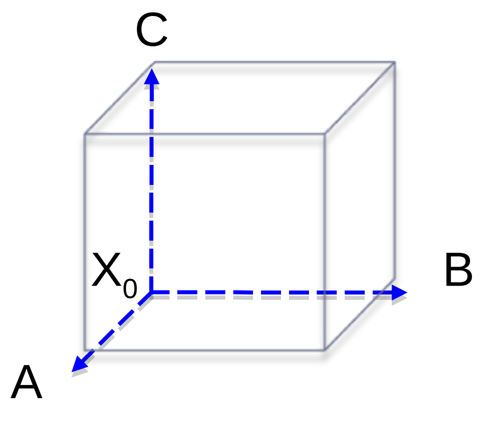
    ...at the neighborhood of x_{i0}
  - Restricted to main effects

The Morris Method

OAT
(The method)

- OAT design (One-At-a-Time)
  - Choose a nominal value X_0
  - Move one coordinate at-a-time, and compute a finite difference:

    $\Delta_i(X_0, h) = [\ f(X_0 + h\ E_i) – f(X_0)\ ] / h$



OAT

(The drawbacks)

- Strengths (?)
  - Easy to run
  - Cheap (d+1 function evaluations)
- Drawbacks (many!)
  - Correct interpretation only for a 1st order polynomial!
  - This is a pure local sensitivity analysis
    - High dependence to the nominal point
  - Poor exploration of the space
    - Included in the unit sphere $x_1^2 + \ldots + x_d^2 = 1$!!!

The Morris method

- Idea (From local to global)
  - Do several OAT with different initial points & random paths

- Description
  - Choose a number of repetitions r
  - Simulate a starting point X_0^*, and a permutation s*
  - Compute the finite differences successively:

    $\Delta_{s^*(1)}(h) = [\ f(X_0^* + h\ E_{s^*(1)}) - f(X_0^*)\ ]\ /\ h$

    $\Delta_{s^*(2)}(h) = [\ f(X_0^* + h\ E_{s^*(1)} + h\ E_{s^*(2)}) - f(X_0^* + h\ E_{s^*(1)})\ ]\ /\ h$

    …

The Morris method

- Path visualization for d=2 and d=3
  - See [Badea and Bolagno, 2008], page 18

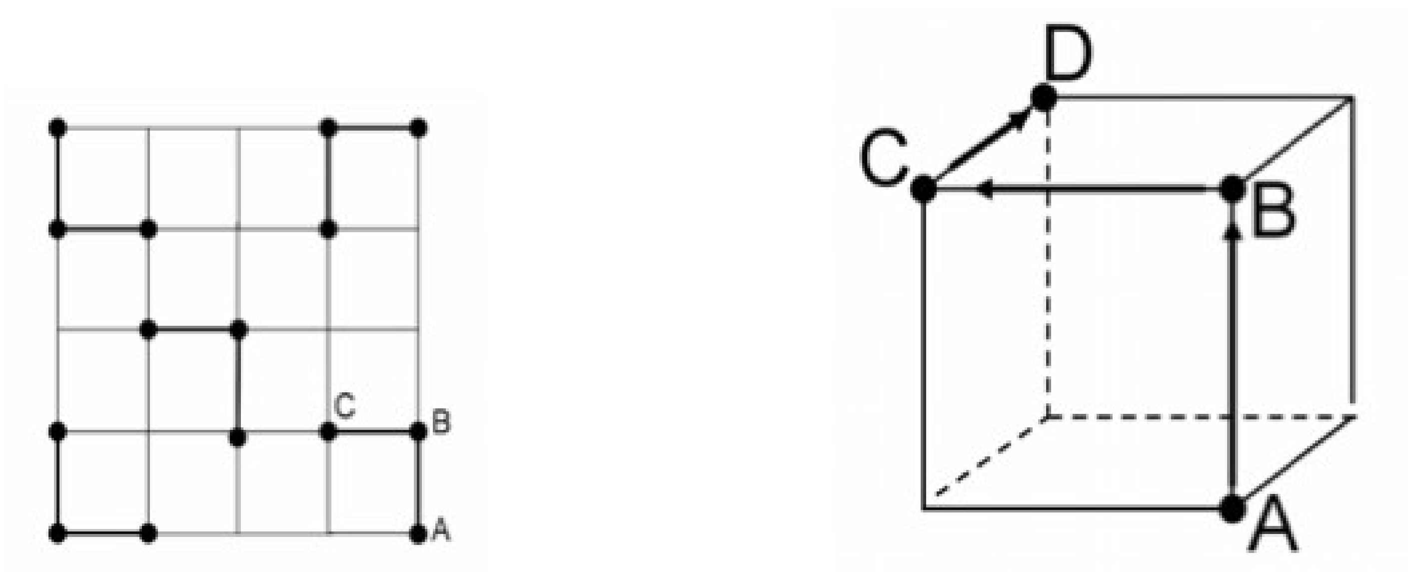


Figure 4.3.1.1.- Example of Morris' OAT designs in 2 (left) and 3 (right) dimensions

The Morris method

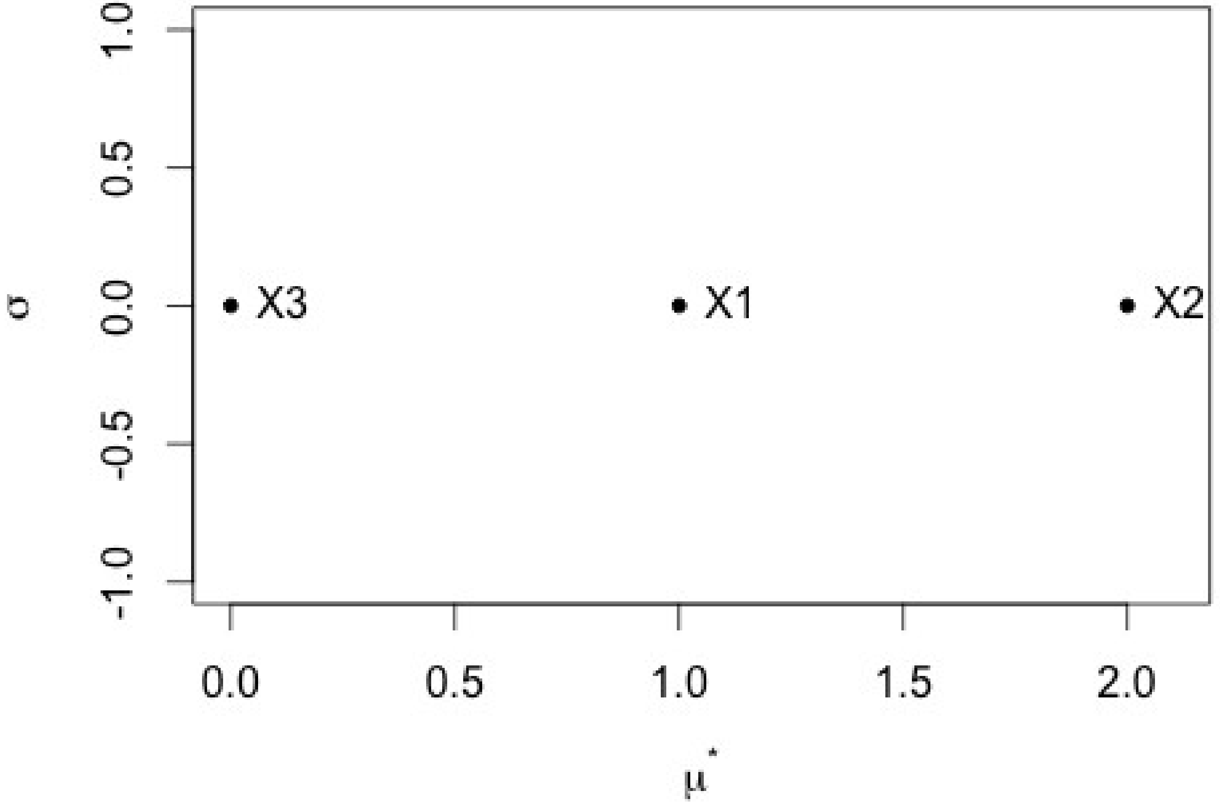
- Outputs:
  - μ_i^* : The average of the r absolute values $|\Delta_i(h)|$
  - σ_i : The standard deviation of the r values $\Delta_i(h)$

- Interpretation:
  - If μ_i^* is large, then X_i is influent
  - If X_i has no interaction with other variables AND IF the output is linear / X_i, then $\sigma_i = 0$

The Morris method

- Illustration with the toy function defined over $[-0.5, 0.5]^3$:

$$f(x_1, x_2, x_3) = x_1 - 2x_2 + b_{12}x_1x_2 + b_{11}x_1^2$$

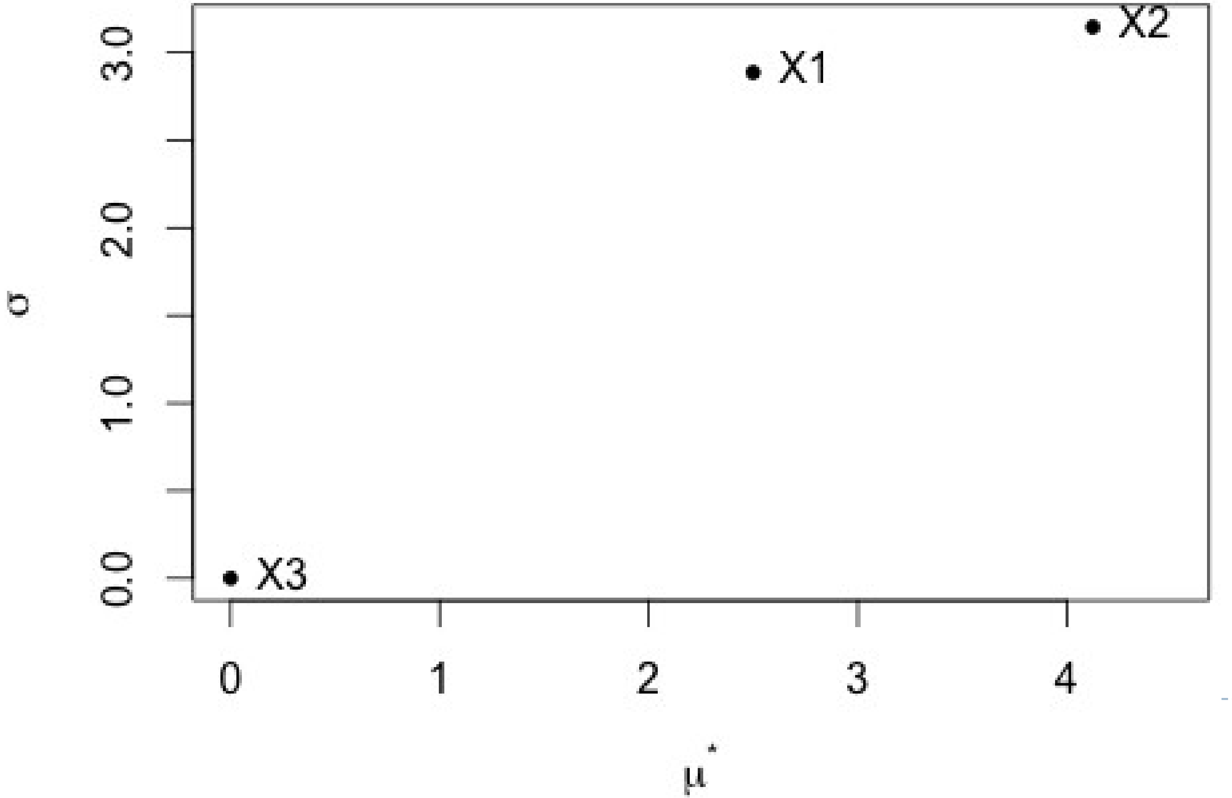


$b_{12} = 0$
$b_{11} = 0$

The Morris method

- Illustration with the toy function defined over $[-0.5, 0.5]^3$:

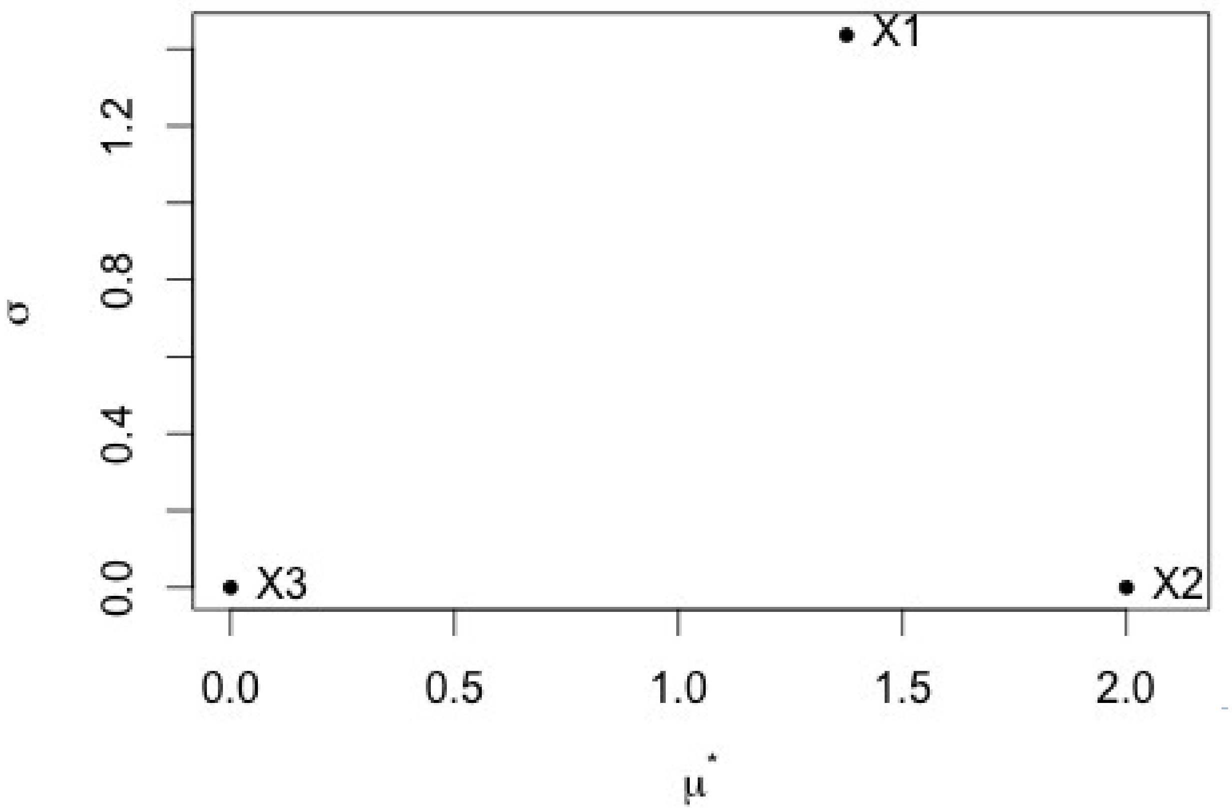
$$f(x_1, x_2, x_3) = x_1 - 2x_2 + b_{12}x_1x_2 + b_{11}x_1^2$$



$b_{12} = 10$
$b_{11} = 0$

The Morris method

- Question (with the same function)



b_{12} = ??
b_{11} = ??

The Morris method

- Question: What do you obtain, still over $[-0.5, 0.5]^2$ with

$$f(x_1, x_2) = |x_1|$$

??

Conclusions

- (Please...) **NEVER USE OAT !**

- **The MORRIS method is a much better alternative**:
  - Has the same advantages of OAT
    - Easy to run, easy to interpret
    - Low cost: r*(d+1) function evaluations
  - The drawbacks of OAT are avoided
    - No dependence to the starting point
    - Interactions & Non-linearities are accepted
    - Domain exploration

The Sobol-Hoeffding decomposition

Sobol-Hoeffding decomposition
(Efron and Stein, 1981, Hoeffding 1948, Sobol 1993)

- Assume that $X_1, \dots, X_d$ are independent random variables. Let f be a function defined on D in R^d . Then f is uniquely decomposed as:

$$f(\boldsymbol{X}) = \mu_0 + \sum_{i=1}^{d} \mu_i(X_i) + \sum_{i<j} \mu_{ij}(X_i, X_j) + \cdots + \mu_{1,\dots,d}(X_1, \dots, X_d)$$

with centering conditions: $\mathrm{E}(\mu_I(X_I)) = 0, \quad I \subseteq \{1, \dots, d\}$
and non-simplification conditions, implying orthogonality:

$$\mathrm{E}(\mu_{ii'}(X_i X_{i'}) \mid X_i) = \mathrm{E}(\mu_{ii'i''}(X_i X_{i'} X_{i''}) \mid X_i X_{i'}) = \cdots = 0.$$

Sobol-Hoeffding decomposition
(main effects, interactions)

- The terms are obtained recursively:
  - Mean, Main effects

$$\mu_0 = E(f(X)) \qquad \mu_i(X_i) = E(f(X)|X_i) - \mu_0$$

  - 2nd order interactions

$$\mu_{ij}(X_i, X_j) = E(f(X)|X_i, X_j) - \mu_i(X_i) - \mu_j(X_j) - \mu_0$$

  - And more generally:

$$\mu_I(X_I) = E(f(X)|X_I) - \sum_{I' \subsetneq I} \mu_{I'}(X_{I'})$$

An example

Example. Ishigami function, with uniform measure on $[-\pi,\pi]^3$.
With $a=1/2$, $b=\pi^4/5$, we have:

$$f(\mathbf{x}) = \sin(x_1) + 7\sin^2(x_2) + 0.1(x_3)^4\sin(x_1)$$

$$= \underbrace{7a}_{\mu_0} + \underbrace{\sin(x_1)(1+0.1b)}_{\mu_1(x_1)} + \underbrace{7(\sin^2(x_2)-a)}_{\mu_2(x_2)} + \underbrace{0.1[(x_3)^4-b]\sin(x_1)}_{\mu_{1,3}(x_1,x_3)}$$

μ_0

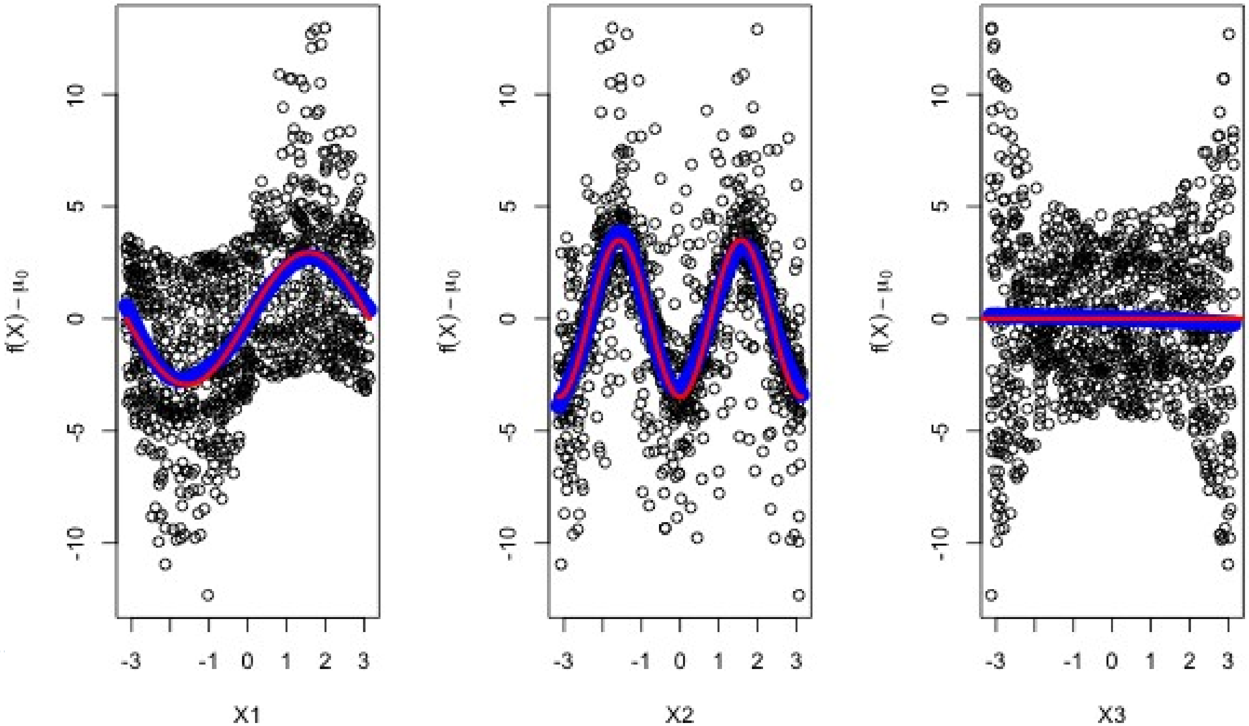
$\mu_1(x_1)$ *(main effect)*

$\mu_2(x_2)$ *(main effect)*

$\mu_{1,3}(x_1,x_3)$ *(2nd order interaction)*

An example

- Estimation of the main effects:
  - Simulate a sample of size N for $X=(X_1, X_2, X_3)$ and compute $z=f(X)$
  - Estimate the global mean μ_0 by the mean of z
  - Estimate the conditional expectation $E(f(X) \mid X_i) - \mu_0$
    - Use a smoother: Local polynomials, smoothing splines, etc.



ANOVA
(Sobol indices)

- The name "ANOVA" (ANalysis Of VAriance) comes from the relation on variances implied by orthogonality:

$$D = \mathrm{var}(f(\boldsymbol{X})) = \mathrm{var}(\mu_0) + \sum_{i=1}^{d} \mathrm{var}(\mu_i(X_i)) + \sum_{i<j} \mathrm{var}(\mu_{ij}(X_i, X_j)) + \cdots + \mathrm{var}(\mu_{1,\ldots,d}(X_1, \ldots, X_d))$$

- (unnormalized) Sobol indices:

$$D_I = \mathrm{var}(\mu_I(X_I))$$

FANOVA decomposition
(Total indices)

- The total index of one variable X_i implies all the subsets J containing {i}

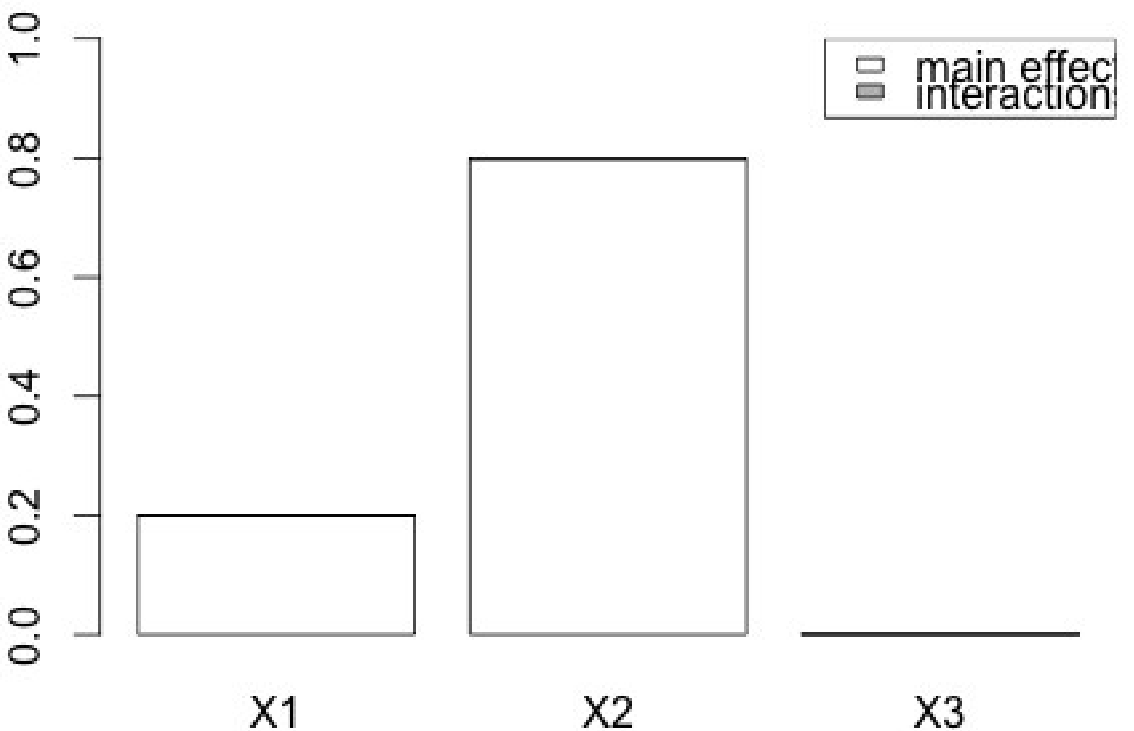
$$D_i^T = \sum_{J \supseteq \{i\}} D_J$$

- Extension for a group of variables X_I: implies all the subsets J that contain at least one element in I

$$D_I^T = \sum_{\substack{J \\ J \cap I \neq \emptyset}} D_J$$

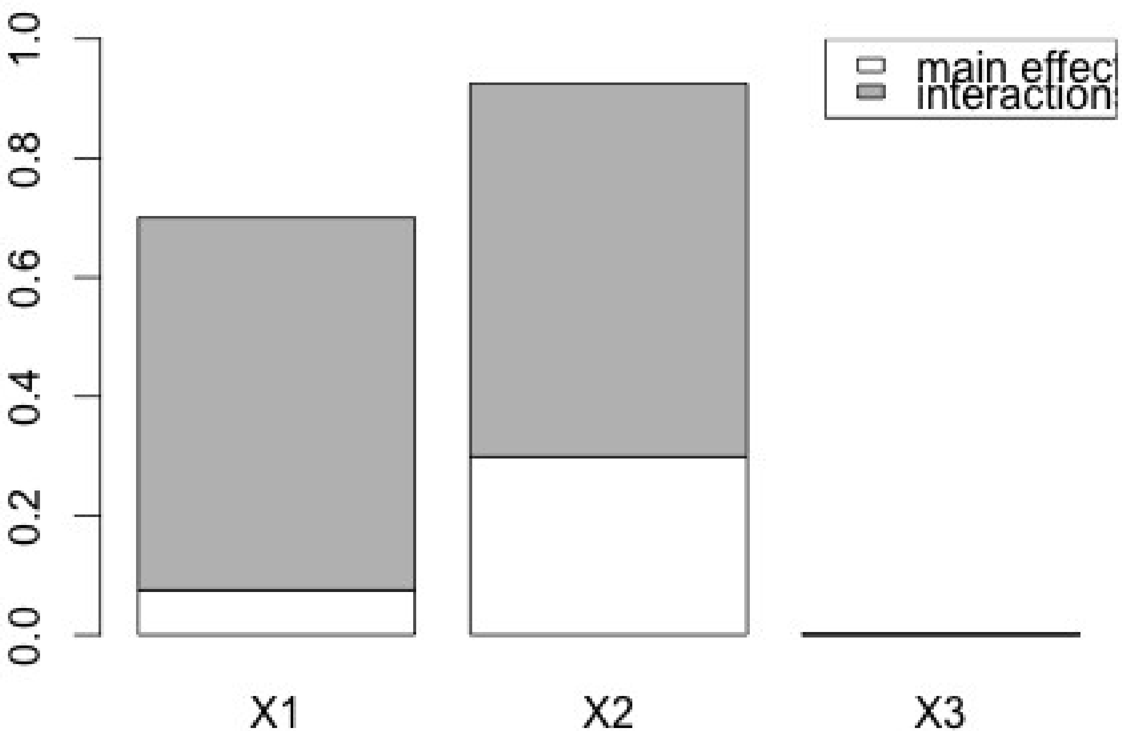
Take again our toy 3D function

- $f(x_1, x_2, x_3) = x_1 - 2x_2 + b_{12}x_1x_2 + b_{11}x_1^2$
  with x_1, x_2, x_3 i.i.d. from $U([-0.5, 0.5])$



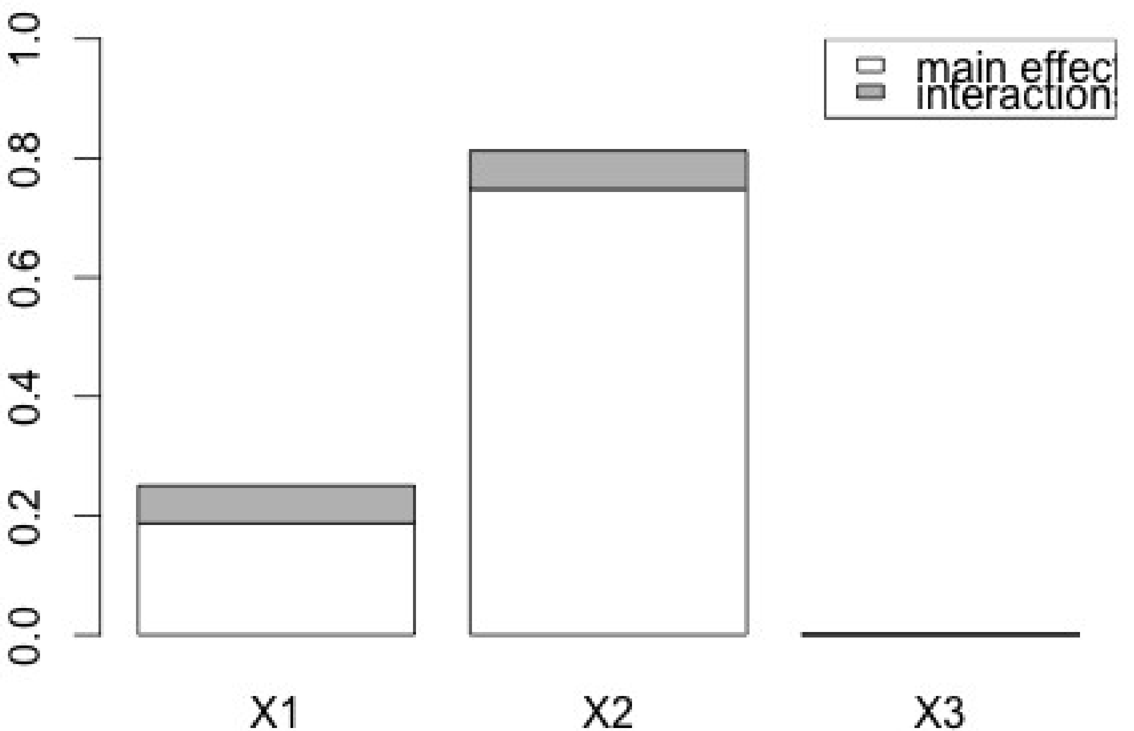
$b_{12} = 0$
$b_{11} = 0$

- $f(x_1, x_2, x_3) = x_1 - 2x_2 + b_{12}x_1x_2 + b_{11}x_1^2$ with x_1, x_2, x_3 i.i.d. from $U([-0.5, 0.5])$



$b_{12} = 10$
$b_{11} = 0$

- $f(x_1, x_2, x_3) = x_1 - 2x_2 + b_{12}x_1x_2 + b_{11}x_1^2$ with x_1, x_2, x_3 i.i.d. from $U([-0.5, 0.5])$



$b_{12} = 2$

$b_{11} = 0$

Question

- Same function, with the same assumptions

??

$b_{12} = 0$
$b_{11} = 3$

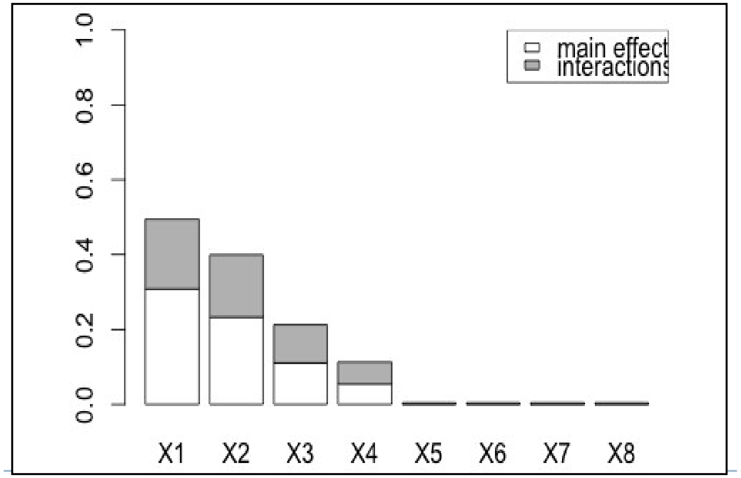
Total indices and screening

- If $D_i^T=0$, the variable X_i is removed (no terms containing X_i)
  - Remark: A condition is required on the probability measure

$$g(X_1,\ldots,X_d)=\prod_{k=1}^{d}\frac{|4X_k-2|+a_k}{1+a_k}$$

a = (0, 1, 4.5, 9, 99, 99, 99, 99)

[package sensitivity]



Total indices of the g-Sobol function:

X_5, X_6, X_7, X_8 can be removed

Sobol-Hoeffding decomposition
(History and aliases)

- Pioneering work of [Hoeffding, 1948]
- Presented as above by [Efron and Stein, 1981]
- Reintroduced by [Sobol, 1993], in the context of sensitivity analysis, who defined the Sobol indices
  - First alias: Sobol (-Hoeffding) decomposition
- The total effects were introduced by [Homma and Saltelli, 1996]
- In the book of [Saltelli, Chan, Scott, 2000]:
  - Second alias: HDMR decomposition (High Dimensional Model Representation)

RESSOURCES

Books of review articles

- Badea A. et Bolado R., Review of sensitivity analysis methods and experience. PAMINA 6th FPEC Project, European Commission, 2008. http ://www.ip-pamina.eu/downloads/pamina.m2.1.d.4.pdf
- Iooss B. (2011), Revue sur l'analyse de sensibilité globale de modèles numériques, Journal de la Société Française de Statistique, **152** (1), 1-23, http://www.sfds.asso.fr/journal
- Saltelli A., Chan K. et Scott E.M., editeurs (2000), *Sensitivity analysis*. Wiley Series in Probability and Statistics.

More specific articles

- B. Efron, C. Stein (1981), The jackknife estimate of variance, *The Annals of Statistics,* **9**, 586–596.
- Hoeffding, W. (1948). A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics,* **19**, 293–325.
- Kuo F.Y., Sloan I.H., Wasilkowski G.W., Wozniakowski H. (2010), On de-compositions of multivariate functions, *Mathematics of computation*, **79** (270), 953-966.
- Morris M. (1991), Factorial sampling plans for preliminary computational experiments. *Technometrics*, **33**, p. 161-174.
- I. Sobol (1993), Sensitivity estimates for non linear mathematical models, *Mathematical Modelling and Computational Experiments,* **1**, 407– 414.

Software

- Pujol G., Iooss B. et Janon A. (2013), sensitivity: Sensitivity Analysis. R package version 1.7

  http://CRAN.R-project.org/package=sensitivity