A New Estimator for Quantile-Oriented Sensitivity Indices

Thomas Browne

EDF R&D-MRI Chatou - Université Paris 5

Supervisors: J-C. Fort (Paris 5)

Advisors: B. Iooss & L. Le Gratiet (EDF R&D-MRI Chatou)

December 1st, 2016, Université du Tampon, La Réunion

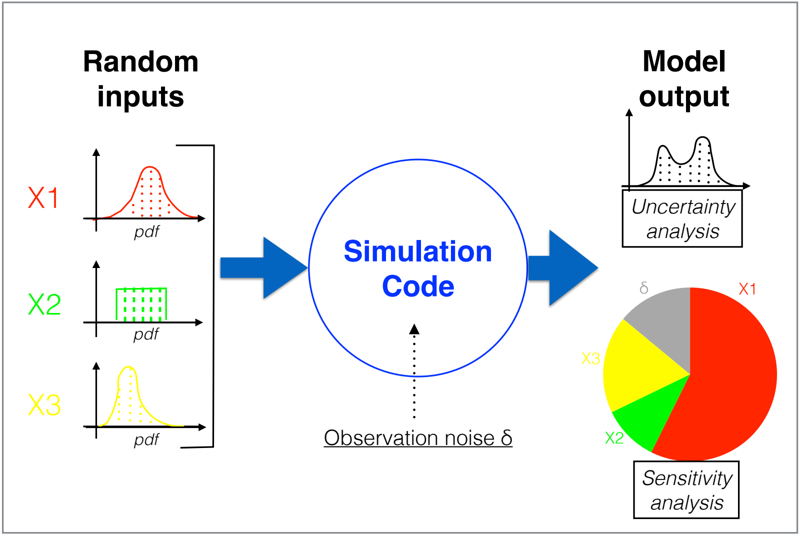
Sensitivity Analysis - Introduction

- Numerical code g.
- Random inputs $(X_1, \ldots, X_d) \sim (f_1, \ldots, f_d)$ independent.
- Random output $Y \in \mathbb{R}$ such that

$$\boxed{Y = g(X_1, \ldots, X_d)}$$

- Main goal : for $i \in \{1, \ldots, d\}$, how much does the uncertainty on X_i propagate through g ?

Sensitivity Analysis - Schema



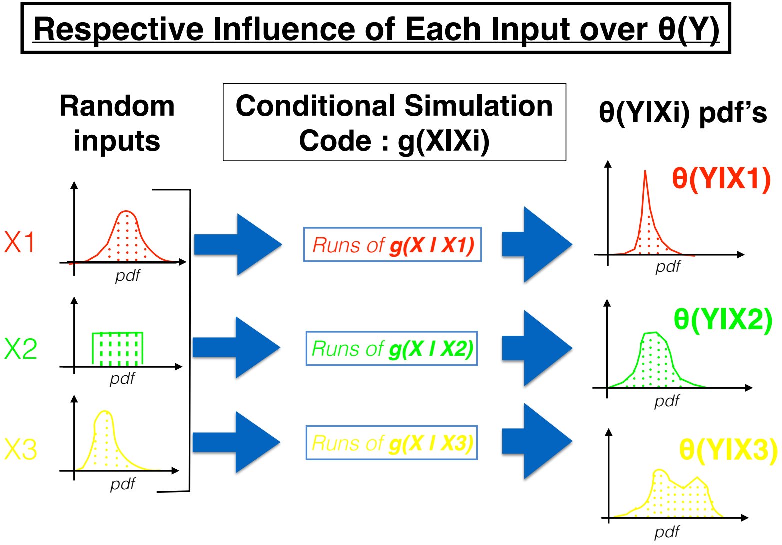
Goal-Oriented Sensitivity Analysis

- In practice, Y's distribution might not need to be fully known.
- Choice of a probability feature $\theta(Y)$ (mean, quantiles etc ...) which may be relevant for the study.
- Goal-Oriented Sensitivity Analysis (*GOSA*) [N. Rachdi, 2011] :

  $\longrightarrow$ For $i \in \{1, \dots, d\}$, quantification of X_i's influence over $\theta(Y)$.

- *GOSA* $=$ Quantification of $\theta(Y \mid X_i)$'s variability.

GOSA - Schema



Respective Influence of Each Input over θ(Y)
Random inputs
Conditional Simulation Code : g(XIXi)
θ(YIXi) pdf's
X1
X2
X3
pdf
pdf
pdf
Runs of g(X I X1)
Runs of g(X I X2)
Runs of g(X I X3)
θ(YIX1)
θ(YIX2)
θ(YIX3)
pdf
pdf
pdf

Contrast Functions

Use contrast functions to quantify $\theta(Y \mid X_i)$'s variability

- Simple contrasts : $\forall (y, \theta) \in \mathbb{R}^2 \quad \varphi(y, \theta) \geq 0$

  quantify a "distance" between two real components.

- Mean Contrasts : for Y real rand var, $\mathbb{E}_Y[\varphi(Y, \theta)]$.

- Y's feature : $\theta(Y) := \arg\min_{\theta \in \mathbb{R}} \mathbb{E}_Y[\varphi(Y, \theta)]$.

Contrast Functions : Mean and Quantiles

- If $\varphi(y,\theta) = m(y,\theta) = \mid y - \theta \mid^2$:

  Mean Contrast : $\mathbb{E}_Y[\mid Y - \theta \mid^2]$,

  $\longrightarrow \theta(Y) = \mathbb{E}[Y]$.

- If, for $\alpha \in]0;1[$, $\varphi(y,\theta) = c_\alpha(y,\theta) = (y-\theta)(\alpha - 1_{y \leq \theta})$:

  Mean Contrast : $\mathbb{E}_Y[(Y-\theta)(\alpha - 1_{Y \leq \theta})]$,

  $\longrightarrow \theta(Y) = q^\alpha(Y)$, α-quantile de Y.

- We focus on $\varphi = c_\alpha$: $\theta(Y) = q^\alpha(Y)$.

  Remark : $\min_\theta \mathbb{E}\left[c_\alpha\left(Y,\theta\right) \mid X_i = x\right] = \mathbb{E}\left[c_\alpha\left(Y, q^\alpha(Y \mid X_i)\right) \mid X_i = x\right]$.

Sensitivity Analysis with Respect to a Contrast

Need to quantify the variability of $\theta(Y \mid X_i)$!

- Sensitivity indices based on contrasts [Fort et al., 2013]

$$\begin{aligned} S_{\varphi}^{Xi}(Y) &= \min_{\theta \in \mathbb{R}} \mathbb{E}\left[\varphi\left(Y,\theta\right)\right] - \mathbb{E}\left[\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\varphi\left(Y,\theta\right) \mid X_i\right]\right] \\ &= \mathbb{E}\left[\varphi\left(Y,\theta(Y)\right)\right] - \mathbb{E}_{X_i}\left[\varphi\left(Y,\theta(Y \mid X_i)\right)\right]. \end{aligned}$$

$\longrightarrow$ quantifies the influence of the input X_i on $\theta(Y)$.

- If $\varphi(y,\theta) = (y-\theta)^2$, $S_{\varphi}^{Xi}(Y)$ is the Sobol index !

Indices' Properties

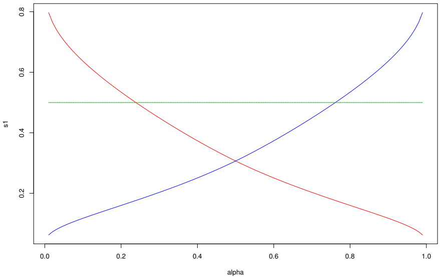
$$\boxed{S_{\varphi}^{Xi}(Y) = \min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_{\alpha}\left(Y,\theta\right)\right] - \mathbb{E}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_{\alpha}\left(Y,\theta\right) \mid X_i\right]\right]}$$

- $S_{\varphi}^{X_i}(Y) \geq 0$.
- We divide $S_{\varphi}^{X_i}(Y)$ by $\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_{\alpha}\left(Y,\theta\right)\right]$ so that $0 \leq S_{\varphi}^{X_i}(Y) \leq 1$.
- We proved :

$$S_{c_{\alpha}}^{X_i}(Y) = 0 \Leftrightarrow q^{\alpha}(Y \mid X_i) = q^{\alpha}(Y) \ a.s.$$

$$S_{c_{\alpha}}^{X_i}(Y) = 1 \Leftrightarrow (Y \mid X_i = x) = constant(x) \ a.s.$$

Sensitivity Analysis with Respect to a Contrast



- $Y = X_1 - X_2$ with $X_1 \sim Exp(1)$ and $X_2 \sim Exp(1)$ independent.
- $S_m^{X_1} = S_m^{X_2} = 0.5$ (Sobol indices).
- Both inputs are influential on the mean $\mathbb{E}[Y]$!
- $S_{c_\alpha}^{X_1}$: X_1's influence on Y's α-quantile.
- $S_{c_\alpha}^{X_2}$: X_2's influence on Y's α-quantile.
- Sensitivity changes regarding the level of quantile α.

Estimate for the Quantile-Oriented Index

Goal : from a iid n-sample $(X_i^1, Y^1), \ldots, (X_i^n, Y^n)$, estimate of

$$S_{\varphi}^{X_i}(Y) = \min_{\theta \in \mathbb{R}} \mathbb{E}\left[c_\alpha(Y, \theta)\right] - \mathbb{E}\left[\min_{\theta \in \mathbb{R}} \mathbb{E}\left[c_\alpha(Y, \theta) \mid X_i\right]\right].$$

1st term estimation :

$$\min_{\theta} \frac{1}{n} \sum_{j=1}^{n} c_\alpha(Y^j, \theta) = \frac{1}{n} \sum_{j=1}^{n} c_\alpha\left(Y^j, \hat{q}^\alpha(Y)\right),$$

where $\hat{q}^\alpha(Y)$ is the classical empirical quantile estimator

$\longrightarrow$ this estimator converges *a.s.*

Estimate for the Second Term

Second term : $\mathbb{E}_{X_i}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i\right]\right].$

- Several issues :

  -Double expectation

  -Conditional expectation

  -Minimization problem.

- Requirements : Be able to estimate $\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i = x\right]$, for x any possible realization of X_i.
- Idea :
  $\mathbb{E}_{X_i}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i\right]\right] \simeq \frac{1}{m}\sum_{k=1}^{m} \min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i = X_i'^{k}\right]$,
  where $\left(X_i'^{1},\ldots,X_i'^{m}\right) \sim f_i$ iid, $m \in \mathbb{N}$.

Kernel regression

Second term : $\mathbb{E}_{X_i}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i\right]\right].$

- **Quantile Regression** [Fan et al., 1994], for x any possible realization of X_i :

$$\boxed{\arg\min_{\theta} \sum_{j=1}^{n} c_\alpha\left(Y^j,\theta\right) K_{h(n)}\left(X_i^j - x\right) \xrightarrow[n\to\infty]{\mathbb{P}} \arg\min_{\theta} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i = x\right]},$$

where f_i is the *pdf* of X_i with a compact support, K a 2-order positive kernel and $(h(n))_{n\in\mathbb{N}}$ a bandwidth sequence such that $h(n) \to 0$ while $n \times h(n) \to \infty$.

Estimate for the Second Term

- We prove :

$$\min_{\theta} \frac{1}{n.f_i(x)} \sum_{j=1}^{n} c_\alpha\left(Y^j, \theta\right) K_{h(n)}\left(X_i^j - x\right) \xrightarrow[n\to\infty]{\mathbb{P}} \min_{\theta} \mathbb{E}\left[c_\alpha(Y, \theta) \mid X_i = x\right].$$

Remark : this function is easy to minimize (convex and piecewise linear) !

- We define the estimator for the second term :

$$\frac{1}{m} \sum_{k=1}^{m} \min_{\theta} \frac{1}{n.f_i(X_i'^k)} \sum_{j=1}^{n} c_\alpha\left(Y^j, \theta\right) K_{h(n)}\left(X_i^j - X_i'^k\right)$$

- We hope to have :

$$\boxed{\hat{S}_{c_\alpha}^{X^i}(Y) \xrightarrow[n,m\to\infty]{\mathbb{P}} S_{c_\alpha}^{X^i}(Y)}$$

Consistency : back up

- So far, we only have, for any $h > 0$:

$$\frac{1}{m}\sum_{k=1}^{m}\min_{\theta}\frac{1}{n.f_i(X_i^k)}\sum_{j=1}^{n}c_\alpha\left(Y^j,\theta\right)K_h\left(X_i^j - X_i^k\right)$$

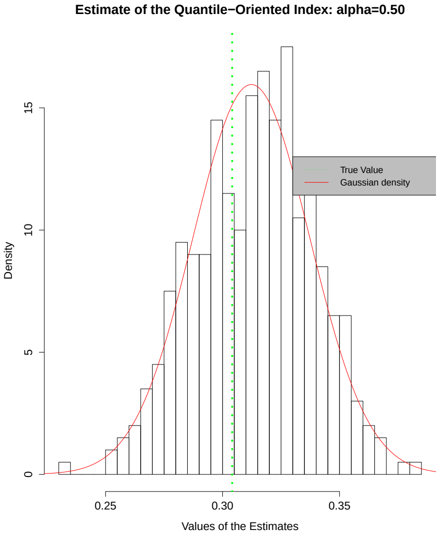
$$\xrightarrow[n,m\to\infty]{\mathbb{P}} \mathbb{E}\left[c_\alpha\left(Y, q^\alpha(Y \mid X_i')\right)K_h\left(X_i - X_i'\right)\right],$$

with X_i and X_i' *iid*.

- While we have :

$$\mathbb{E}\left[c_\alpha\left(Y, q^\alpha(Y \mid X_i')\right)K_h\left(X_i - X_i'\right)\right] \xrightarrow[h\to 0]{} \mathbb{E}\left[\min_{\theta\in\mathbb{R}}\mathbb{E}\left[c_\alpha\left(Y,\theta\right) \mid X_i\right]\right].$$

Applications to the first example

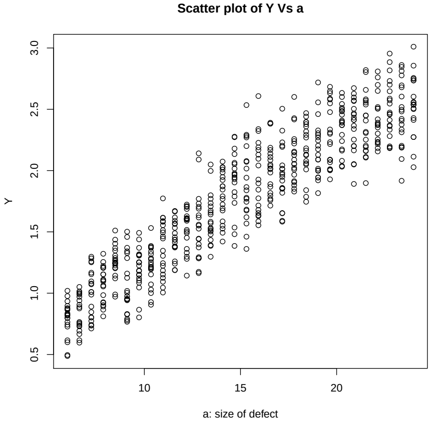


- $Y = X_1 - X_2$ with $X_1 \sim Exp(1)$ and $X_2 \sim Exp(1)$ independent.
- Estimation of $S^{X^1}_{c_\alpha}(Y)$, $\alpha = 0.5$.
- We already know : $S^{X^1}_{c_\alpha}(Y) \simeq 0.30$
- 10^3 repetitions of the experiments : histogram of $\hat{S}^{X^1}_{c_\alpha}(Y)$, with $n = 500$.
- *pdf* of $\mathcal{N}\left(1/10^3 \sum \hat{S}^{X^1}_{c_\alpha}(Y), \mathrm{Var}\left(\hat{S}^{X^1}_{c_\alpha}(Y)\right)\right)$.
- CLT ?

Numerical experiments : CIVA Example

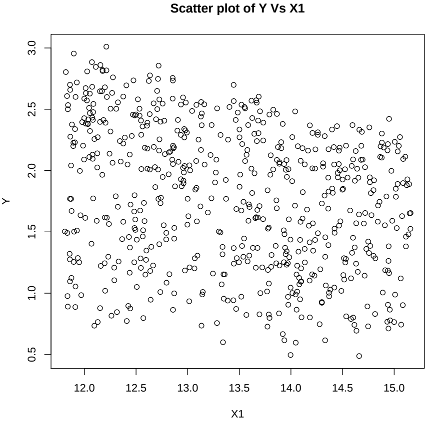
- Non-destructive examination (*NDE*) for defect detection.
- A ultrasonic wave is sent through a structure to analyze.
- One measures Y, the amplitude of the output signal.
- Deterministic simulator g : $Y = g(a, X_1, \ldots, X_6)$.
- n=600 *iid* simulations.

Numerical experiments : CIVA Example



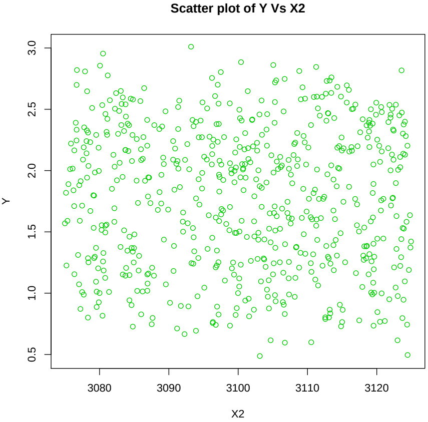
- Defect detection : wave control through a structure to study.
- Y Vs a, the size of defect.

Numerical experiments : CIVA Example



- Defect detection : wave control through a structure to study.
- Y Vs X_1, the thickness of the structure.

Numerical experiments : CIVA Example



- Defect detection : wave control through a structure to study.
- Y Vs X_2, the speed of transverse waves.

Numerical experiments : CIVA Example

α	a	X_1	X_2	X_3	X_4	X_5	X_6
0.1	0.89	0.10	0.12	0.10	0.14	0.12	0.15
0.25	0.90	0.11	0.13	0.11	0.13	0.10	0.13
0.5	0.93	0.13	0.13	0.10	0.13	0.10	0.12
0.75	0.92	0.16	0.13	0.10	0.14	0.09	0.11
0.9	0.87	0.21	0.14	0.11	0.13	0.11	0.12

TABLE: List of the values of the estimates for $S^{\cdot}_{c_\alpha}(Y)$, for all the values, at different level of quantiles, $\alpha \in]0,1[$.

- a is far more influential over any zone of the distribution.
- The other inputs are poorly influential.
- The influence of X_1 rises with α.

Conclusions

- Relevant information for Sensitivity Analysis - useful alternative to Sobol indices !
  $\Longrightarrow$ global methodology for Sensitivity Analysis.
- Do not sum up to 1.
- *CLT* for the convergence of $\hat{S}^{X}_{c_\alpha}$?
- Determination of an Optimal Bandwidth.
- How to calibrate m and n for a given time budget ?

N. Rachdi
Statistical Learning and Computer Experiments
PhD thesis, Université Paul Sabatier, France, 2011.

J. C. Fort, T. Klein, N. Rachdi
New sensitivity analysis subordinated to a contrast
Communication in Statistics : Theory and Methods, *In press*, 2013.

J. Fan, T. Hu and Y. K. Truong
Robust Non-Parametric Function Estimation
Scandinavian Journal of Statistics, Vol. 21, No. 4, pp. 433-446, 1994.

H. J. Kushner, D. S. Clark
Stochastic Approximations for Constrained and Unconstrained Systems
Springer, Berlin, 1978.

Da Veiga S.
Global Sensitivity Analysis with Dependence Measures
Journal of Statistical Computation and Simulation, Vol. 85, N. 7, pp. 1283–1305, 2015.

Thank you for your attention

Interpretation of the index

We rewrite the normalized index :

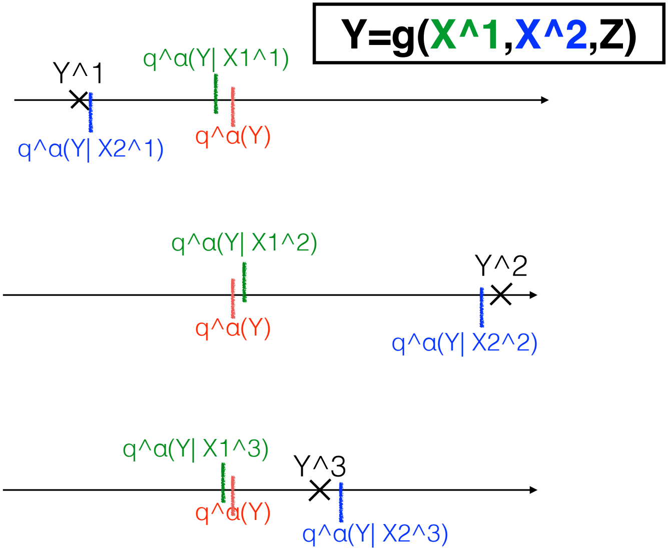
$$S_{c_\alpha}^{X_i}(Y) = 1 - \frac{\mathbb{E}\left[\varphi\left(Y, q^\alpha(Y \mid X_i)\right)\right]}{\mathbb{E}\left[\varphi\left(Y, q^\alpha(Y)\right)\right]}.$$

Therefore :

$$S_{c_\alpha}^{X_i}(Y) \simeq 0 \quad \leftrightarrow \quad \mathbb{E}\left[\varphi\left(Y, q^\alpha(Y \mid X_i)\right)\right] \simeq 0 \quad \leftrightarrow \quad q^\alpha(Y \mid X_i) \simeq q^\alpha(Y),$$

$$S_{c_\alpha}^{X_i}(Y) \simeq 1 \;\leftrightarrow\; \mathbb{E}\left[\varphi\left(Y, q^\alpha(Y \mid X_i)\right)\right] << \mathbb{E}\left[\varphi\left(Y, q^\alpha(Y)\right)\right] \;\leftrightarrow\; Y \simeq q^\alpha(Y \mid X_i).$$

Interpretation of the index



In practice : choice of bandwidth

- Usually, we set : $h^* := \arg\min_{h>0} MSE\left(\hat{S}^{X^i}_{c_\alpha}\right)$,
  but we could not compute $MSE\left(\hat{S}^{X^i}_{c_\alpha}\right)$.

- We found, for $\frac{1}{n-1} \min_\theta \frac{1}{f_i(x)} \sum_{\substack{j=1 \\ j\neq k}}^{n} c_\alpha\left(Y^j, \theta\right) K_h\left(X_i^j - x\right)$:

$$MSE \quad = h_n . \int u^2 K(u) du \int c_\alpha \left(y, q^\alpha(Y \mid X_i = x)\right) \partial_x^2 f(x,y) dy$$

$$+ \frac{\int K^2 . \mathrm{Var}(c_\alpha(Y, q^\alpha(Y|X_i))|X_i = x)}{nh_n} + o(h_n) + o\left(\frac{1}{nh_n}\right).$$

The optimal bandwidth cannot be determined in this case.

In practice : choice of bandwidth

- Idea : instead of optimizing the error over the minimum's estimation, we focus on the minimizer's estimation, *ie* :

$$q^{\alpha}(Y \mid X_i = x) = \arg\min_{\theta} \mathbb{E}\left[c_{\alpha}(Y, \theta) \mid X_i = x_i\right].$$

- In [Yu,1998], regarding the *MSE*, the authors introduced

$$h^{*}(x) \simeq h_{mean}(x) \left[\frac{\alpha(1-\alpha)}{\phi\left(\Phi^{-1}(\alpha)\right)}\right]^{1/5},$$

where $h_{mean}(x)$ is the optimal bandwidth for the estimation of $m(x) = \mathbb{E}[Y \mid X_i = x]$:

$$h_{mean}(x)^5 = \frac{R(K)\sigma^2(x)}{n\mu_2(K)^2 m''(x)^2 f(x)},$$

with $\sigma^2(x) = \mathrm{Var}(Y \mid X_i = x)$.