EDF

CentraleSupélec

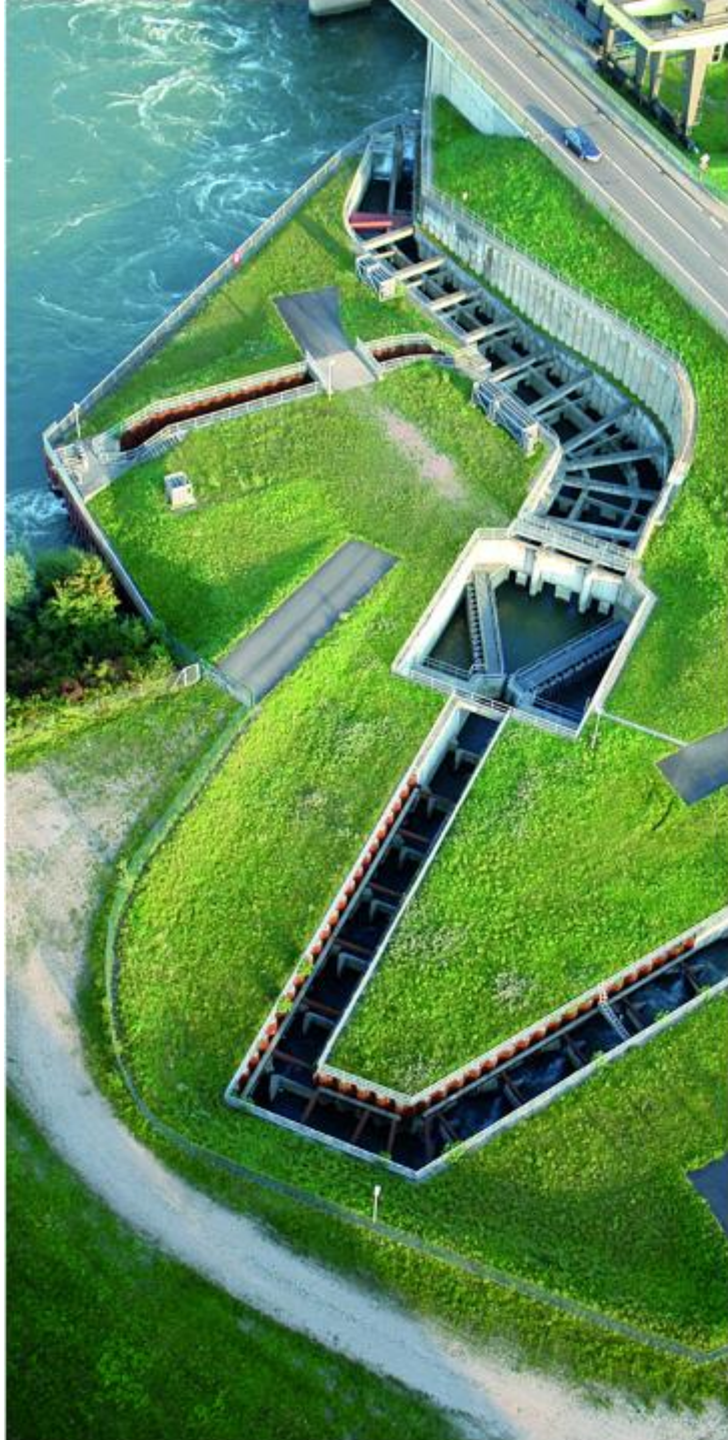
Perturbed-Law based sensitivity Indices for sensitivity analysis in structural reliability

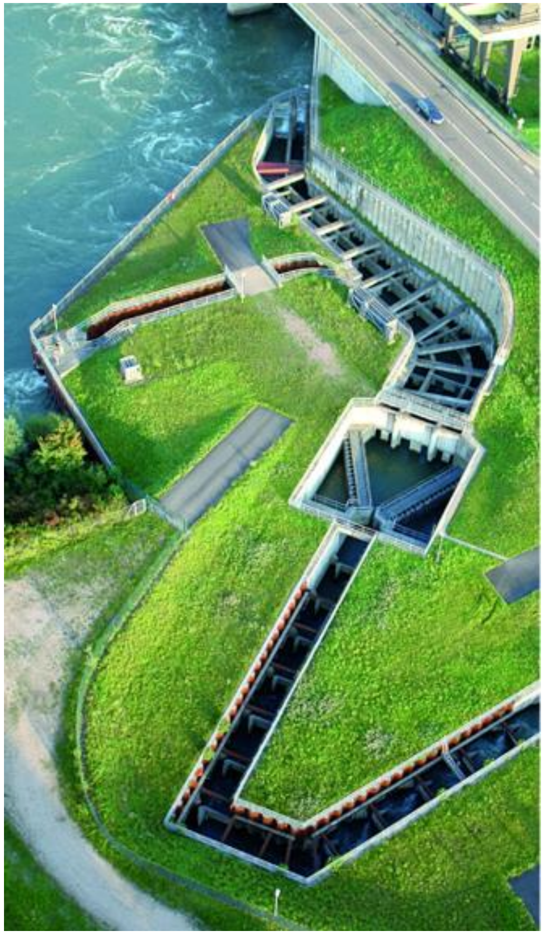
8th International Conference

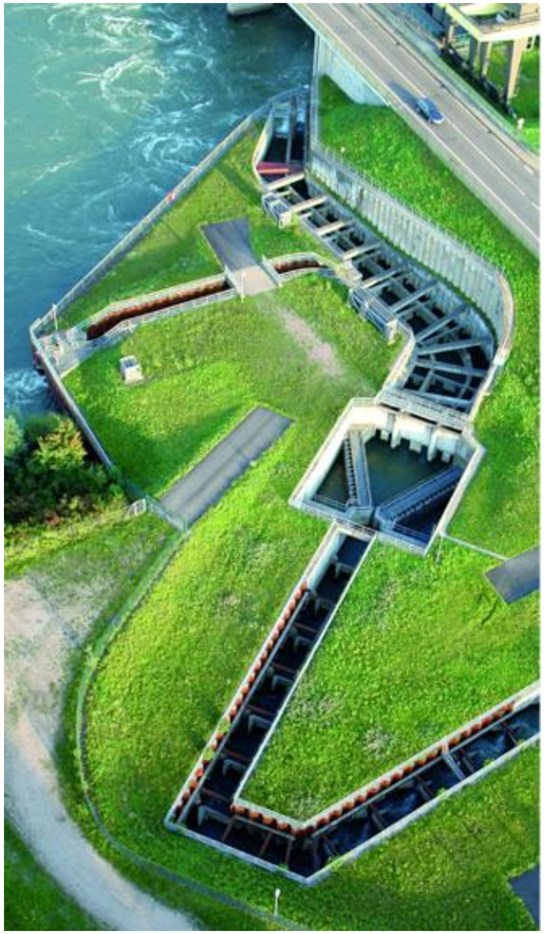
Roman SUEUR[1], Nicolas BOUSQUET[1], Bertrand, IOOSS[1], Julien BECT[2]

[1] *Département de Management des Risques Industriels EDF R&D*

[2] *Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506) CentraleSupélec, CNRS*







1. Context : sensitivity analysis on a failure probability

CONTEXT OF STRUCTURAL RELIABILITY

We want to study a computer code G which :

- **is deterministic**
- **is a « costly » numerical model (CPU time, memory,…)**
- **has d input variables**
- **allows calculating the value $G(\boldsymbol{X})$ of the reliability criterion for a given set of input values $\boldsymbol{X} = (X_1, \dots, X_d)$**

The input variables are uncertain, hence we denote

- $\mathbb{X} \subset \mathbb{R}^d$ **the domain of variation of the random vector $\boldsymbol{X}$**
- $f = \prod_{i=1}^{d} f_i$ **the probability density function of $\boldsymbol{X}$**
  - each f_i is the density of X_i, the i-th marginal of $\boldsymbol{X}$
  - the uncertain input variables $X_1, \dots X_d$ are considered independent

CONTEXT OF STRUCTURAL RELIABILITY

In particular, we get interested in the following event

$$\{G(\boldsymbol{X}) < 0\}$$

► **binary quantity of interest**

We aim at calculating a « failure probability » :

$$P_f = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) d\boldsymbol{x}$$

Problem : how to study the influence of the input variables on P_f ?

Usually, P_f is low ($< 10^{-3}$)

CONTEXT OF STRUCTURAL RELIABILITY

$$P_f = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) d\boldsymbol{x}$$

Problem : how to study the influence of the input variables on P_f ?

- Many techniques aim at detecting most influencial inputs on a quantity of interest on $Y = G(\boldsymbol{X})$ (usually assumed continuous with respect to each X_i)
- The alternative approach we're choosing here consists in quantifying the impact of the lack of knowledge about the law of $\boldsymbol{X}$

A FIRST IDEA : SOBOL' INDICES FOR $x \mapsto \mathbb{I}_{\{G(x)<0\}}$

► **good method to distinguish influential and non-influential variables**

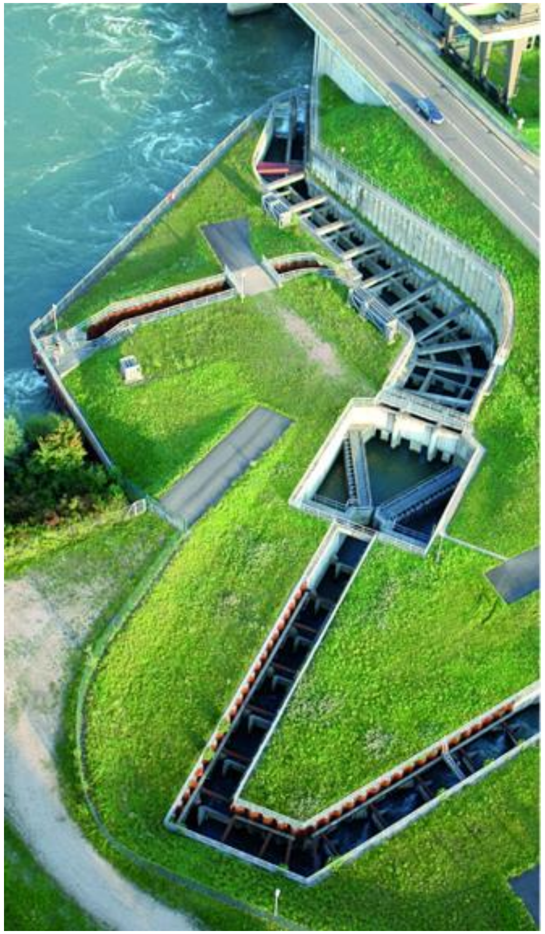
In general :

- Weak first order indices
- Strong total indices

assesses that no variable is influential on its own, and that most variables contribute to the failure probability when interacting with the others

► **in most structural reliability cases, this is an already known information : the equipment fails when all variable take extreme values at the same time**

+ this information is obtained at a high computational cost



1. Context : sensitivity analysis on a failure probability
2. Principle of PLI indices

PLI INDICES : THE PRINCIPLE

For a particular random input X_i which's density is f_i, we introduce a δ perturbation, leading to a perturbed density $f_{i\delta}$

The failure probability becomes :

$$P_{i\delta} = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) \frac{f_{i\delta}(x_i)}{f_i(x_i)} d\boldsymbol{x}$$

The corresponding PLI indice is defined by :

$$S_{i\delta} = \left(\frac{P_{i\delta}}{P_f} - 1\right) \mathbb{I}_{\{P_{i\delta} > P_f\}} + \left(1 - \frac{P_f}{P_{i\delta}}\right) \mathbb{I}_{\{P_{i\delta} < P_f\}}$$

Introduced in *Lemaître et al. (2015)*

PLI INDICES : THE PRINCIPLE

$$P_{i\delta} = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) \frac{f_{i\delta}(x_i)}{f_i(x_i)} d\boldsymbol{x}$$

The corresponding PLI indice is defined by :

$$S_{i\delta} = \left(\frac{P_{i\delta}}{P_f} - 1\right) \mathbb{I}_{\{P_{i\delta}>P_f\}} + \left(1 - \frac{P_f}{P_{i\delta}}\right) \mathbb{I}_{\{P_{i\delta}<P_f\}}$$

PLI INDICES : THE PRINCIPLE

$$P_{i\delta} = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) \frac{f_{i\delta}(x_i)}{f_i(x_i)} d\boldsymbol{x}$$

The corresponding PLI indice is defined by :

$$S_{i\delta} = \left(\frac{P_{i\delta}}{P_f} - 1\right) \mathbb{I}_{\{P_{i\delta} > P_f\}} + \left(1 - \frac{P_f}{P_{i\delta}}\right) \mathbb{I}_{\{P_{i\delta} < P_f\}}$$

- $S_{i\delta} = 0$ when $P_{i\delta} = P_f$ e.g. when perturbating f_i has no impact on the failure probability

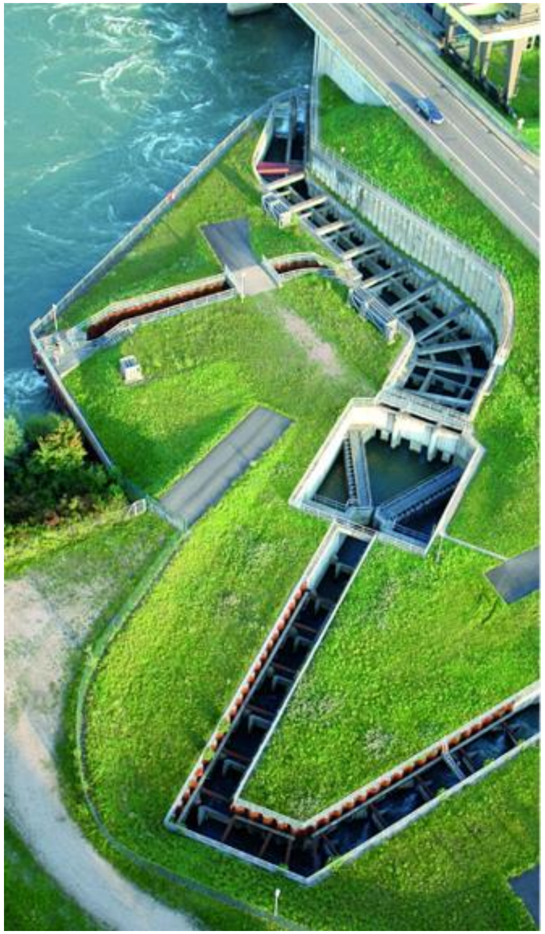
PLI INDICES : THE PRINCIPLE

$$P_{i\delta} = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) \frac{f_{i\delta}(x_i)}{f_i(x_i)} d\boldsymbol{x}$$

The corresponding PLI indice is defined by :

$$S_{i\delta} = \left(\frac{P_{i\delta}}{P_f} - 1\right) \mathbb{I}_{\{P_{i\delta}>P_f\}} + \left(1 - \frac{P_f}{P_{i\delta}}\right) \mathbb{I}_{\{P_{i\delta}<P_f\}}$$

- $S_{i\delta} = 0$ when $P_{i\delta} = P_f$ e.g. when berturbating f_i has no impact on the failure probability
- The sign of $S_{i\delta}$ indicates how the perturbation modifies the failure probability (symetrical behaviour for increase or decrease of the output probability)

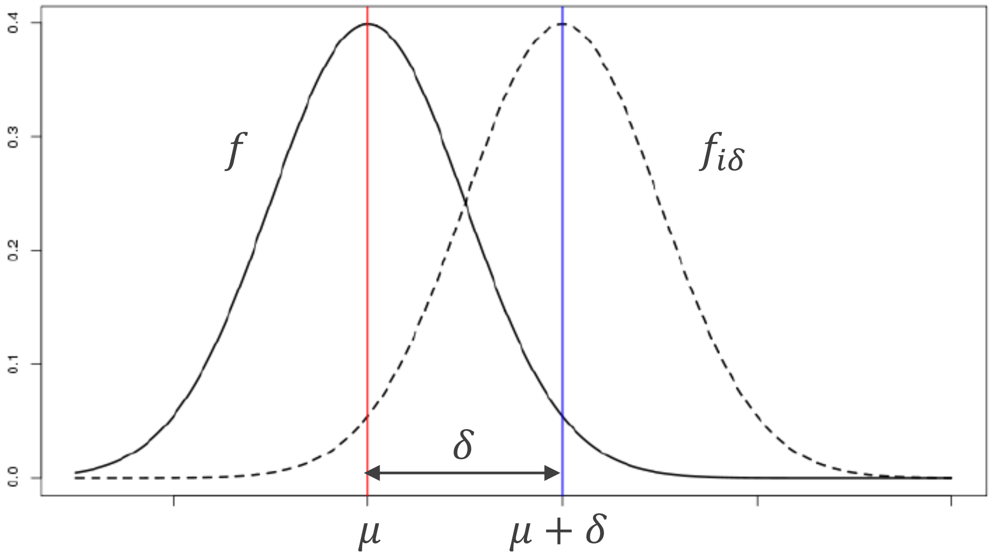


1. Context : sensitivity analysis on a failure probability
2. Principle of PLI indices
3. **Density perturbation**

CentraleSupélec

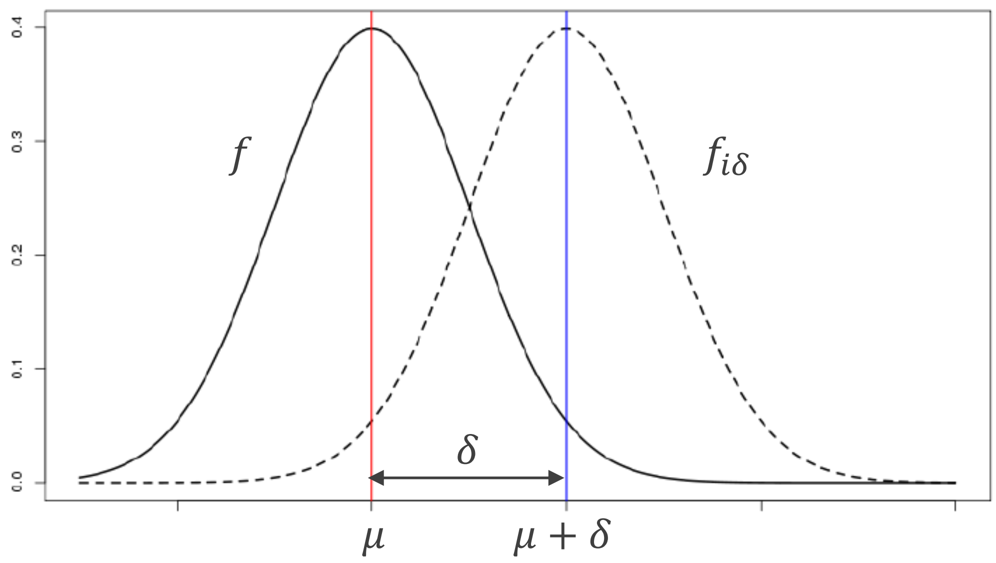
HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **What if the mean of X_i was not μ but $\mu + \delta$?**



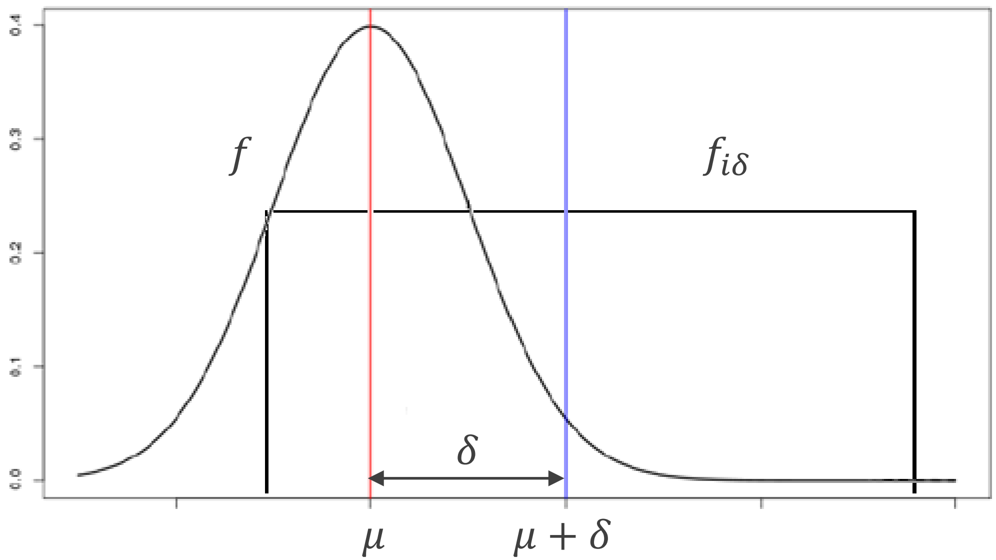
HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **How to define $f_{i\delta}$ with the constraint $\int_{\mathbb{X}_i} x_i f_{i\delta}(x_i)dx_i = \mu + \delta$?**



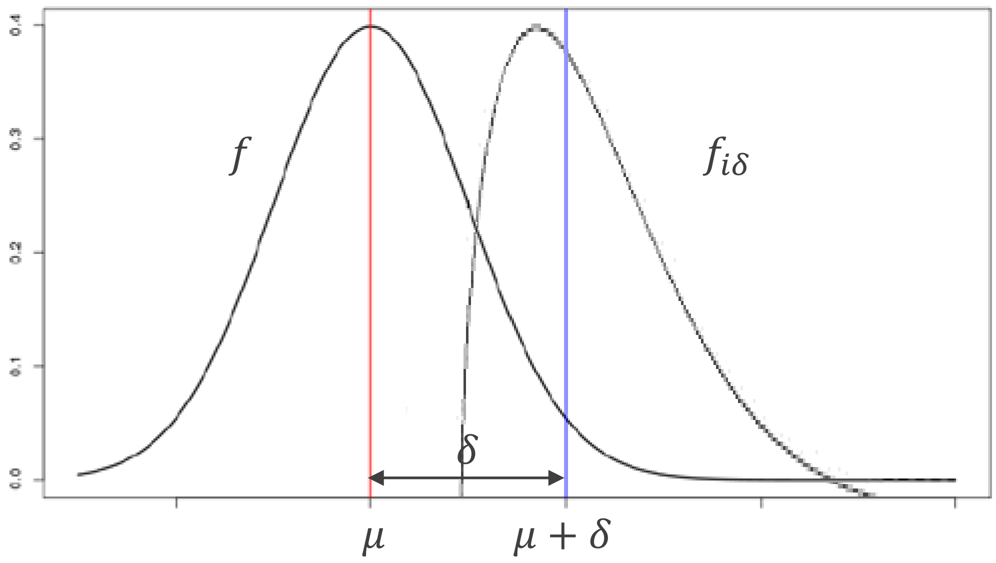
HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **How to define $f_{i\delta}$ with the constraint $\int_{\mathbb{X}_i} x_i f_{i\delta}(x_i) dx_i = \mu + \delta$?**



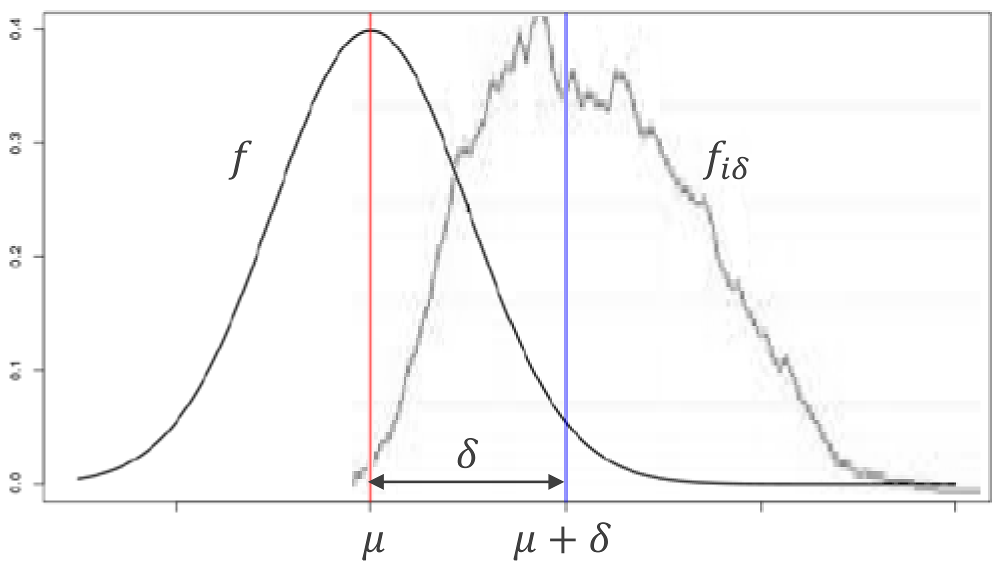
HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **How to define $f_{i\delta}$ with the constraint $\int_{\mathbb{X}_i} x_i f_{i\delta}(x_i) dx_i = \mu + \delta$?**



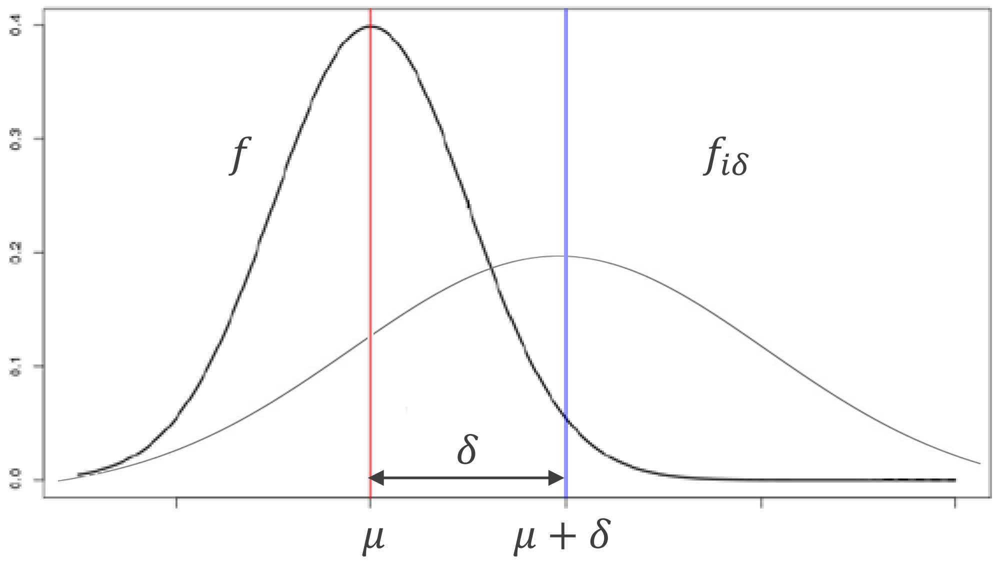
HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **How to define $f_{i\delta}$ with the constraint $\int_{\mathbb{X}_i} x_i f_{i\delta}(x_i) dx_i = \mu + \delta$?**



HOW TO DEFINE A DENSITY PERTURBATION ?

- **Let's assume that the X_i input variable has a normal distribution $X_i \sim \mathcal{N}(\mu, \sigma^2)$**
- **How to define $f_{i\delta}$ with the constraint $\int_{\mathbb{X}_i} x_i f_{i\delta}(x_i) dx_i = \mu + \delta$?**



HOW TO DEFINE A DENSITY PERTURBATION ?

- We suggest to define the perturbed density $f_{i\delta}$ as **the closest one from the initial f_i** in the sense of the entropy, **under the constraint of perturbation**

- Recall that :

$$KL(\pi_1, \pi_2) = \int_{-\infty}^{+\infty} \pi_1(x) \log\left(\frac{\pi_1(x)}{\pi_2(x)}\right) dx$$

- So we can give a general formal definition for $f_{i\delta}$ the following way :

$$f_{i\delta} = \underset{\substack{\pi \\ s.t.\ \mathbb{E}_\pi[g_k]=\delta_k \\ k=1,\dots,K}}{\operatorname{argmin}} KL(\pi, f_i)$$

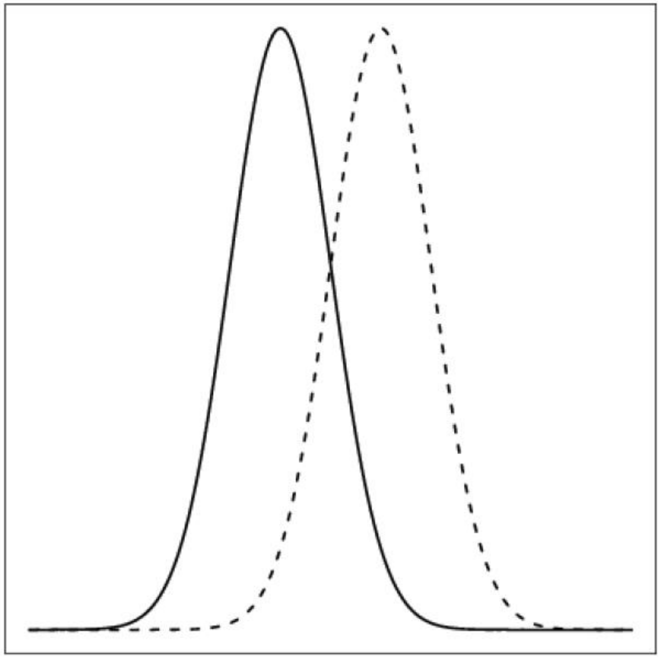
where :

- $g_1, \dots, g_K$ are K linear constraints on the modified density
- and $\delta_1, \dots, \delta_K$ are the values for the perturbed parameters

PERTURBATION EXAMPLES

- **Perturbation of the mean**

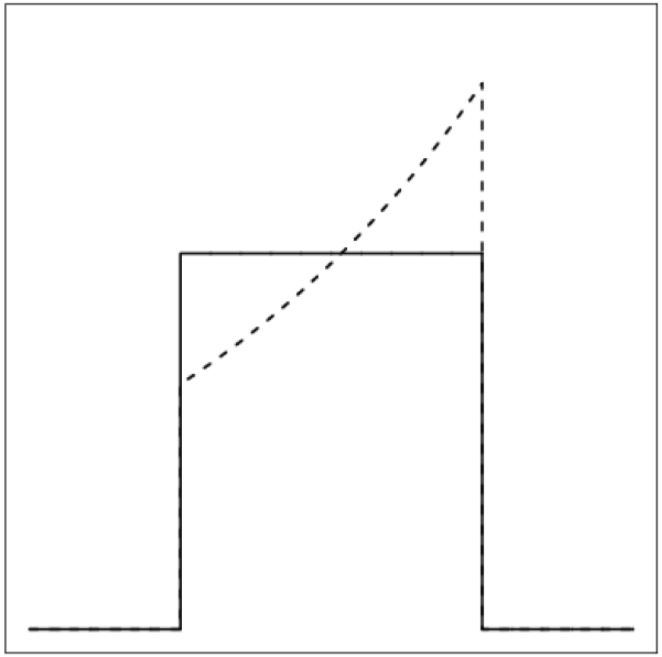
$$\mathbb{E}[X_i] = \mu + \delta$$



PERTURBATION EXAMPLES

- **Perturbation of the mean**

$$\mathbb{E}[X_i] = \mu + \delta$$

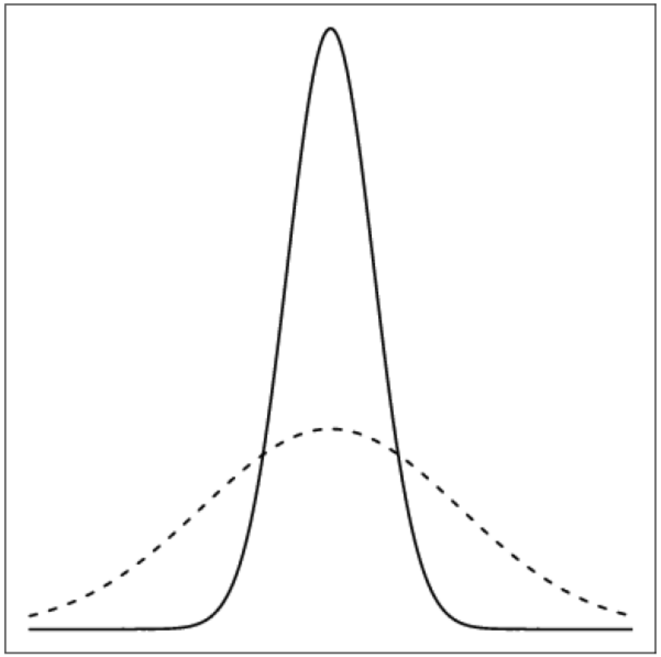


PERTURBATION EXAMPLES

- **Perturbation of the variance**

$$\mathbb{E}[X_i] = \mu$$

$$\mathrm{Var}[X_i] = \mathbb{E}\left[X_i^2\right] - \mu^2 = \sigma^2 + \mu$$

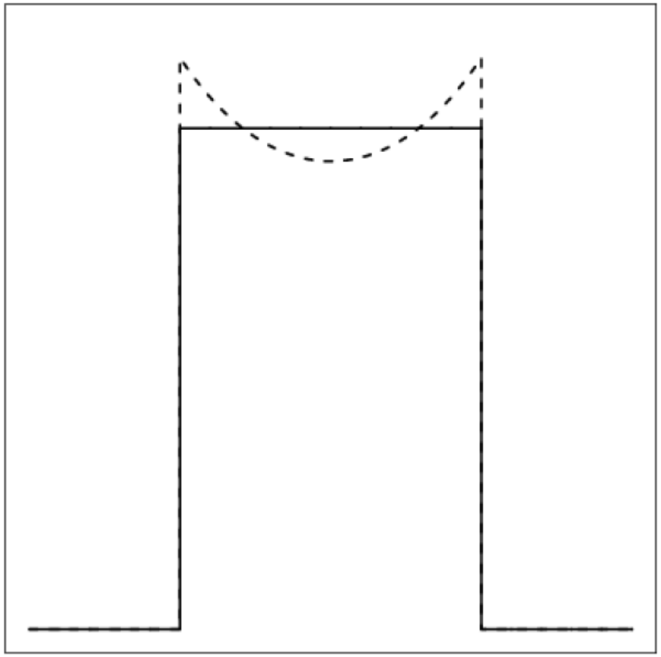


PERTURBATION EXAMPLES

- **Perturbation of the variance**

$$\mathbb{E}[X_i] = \mu$$

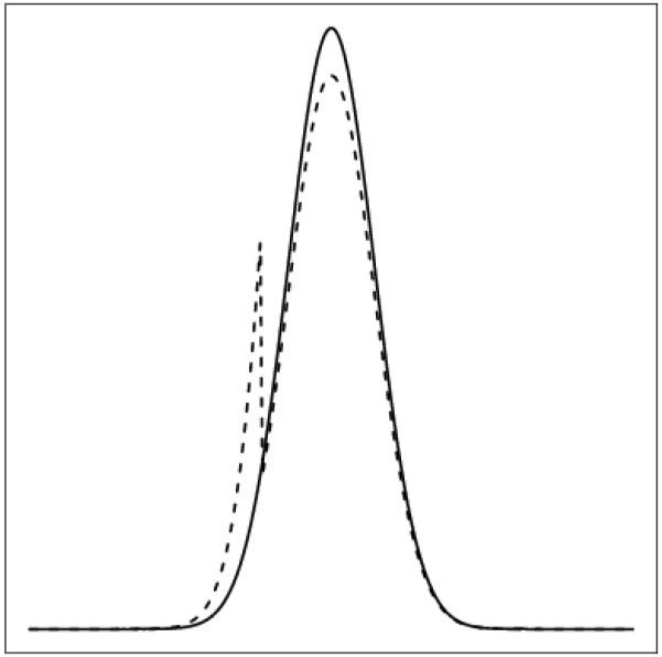
$$\text{Var}[X_i] = \mathbb{E}\left[X_i^2\right] - \mu^2 = \sigma^2 + \mu$$



PERTURBATION EXAMPLES

- **Perturbation of a α-quantile**

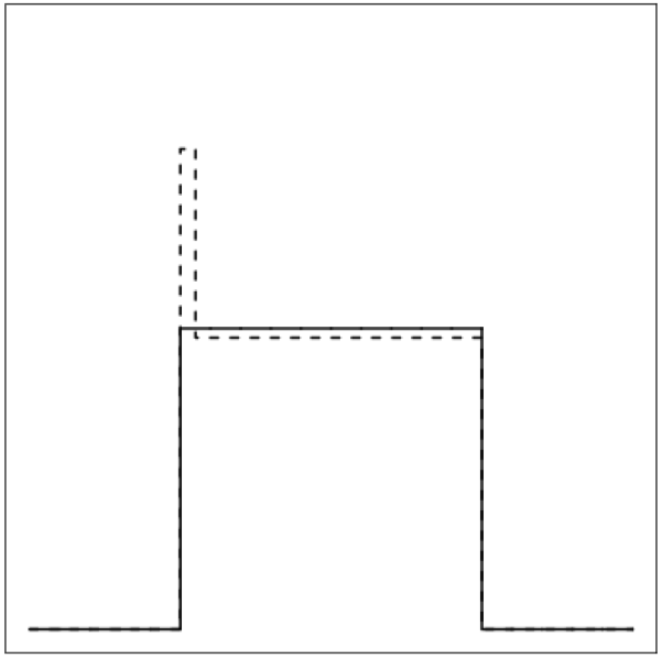
$$q_\alpha(X_i) = F_i^{-1}(\alpha) = q + \delta$$

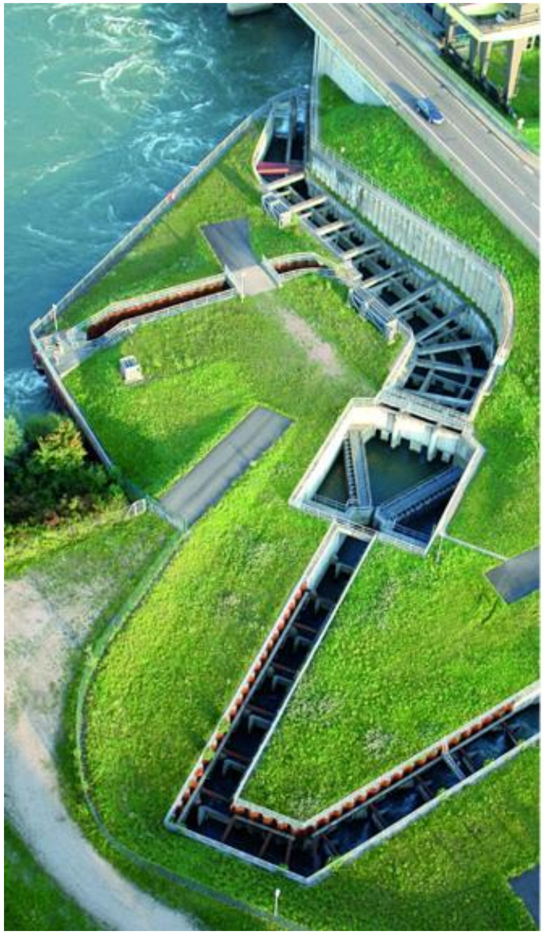


PERTURBATION EXAMPLES

- **Perturbation of a α-quantile**

$$q_\alpha(X_i) = F_i^{-1}(\alpha) = q + \delta$$





1. Context : sensitivity analysis on a failure probability
2. Principle of PLI indices
3. Density perturbation
4. **Estimation of perturbed probability**

REVERSE IMPORTANCE SAMPLING

Lemaître et al. (2015)

- **We use a Monte-Carlo framework**

$$\hat{P}_f^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}}$$

REVERSE IMPORTANCE SAMPLING

Lemaître et al. (2015)

- **We use a Monte-Carlo framework**

$$\hat{P}_f^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}}$$

- **We can estimate $P_{i\delta}$ using the « reverse » importance sampling method**

$$\hat{P}_{i\delta}^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}} \frac{f_{i\delta}\left(x_i^{(n)}\right)}{f_i\left(x_i^{(n)}\right)}$$

REVERSE IMPORTANCE SAMPLING

Lemaître et al. (2015)

- **We use a Monte-Carlo framework**

$$\hat{P}_f^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}}$$

- **We can estimate $P_{i\delta}$ using the « reverse » importance sampling method**

$$\hat{P}_{i\delta}^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}} \frac{f_{i\delta}\left(x_i^{(n)}\right)}{f_i\left(x_i^{(n)}\right)}$$

► $\hat{P}_f^N$ **and** $\hat{P}_{i\delta}^N$ **can be estimated using the same sample**

no need for additional runs of the code G

REVERSE IMPORTANCE SAMPLING

Lemaître et al. (2015)

- **We use a Monte-Carlo framework**

$$\hat{P}_f^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}}$$

- **We can estimate $P_{i\delta}$ using the « reverse » importance sampling method**

$$\hat{P}_{i\delta}^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{G(\boldsymbol{x}^{(n)})<0\}} \frac{f_{i\delta}\left(x_i^{(n)}\right)}{f_i\left(x_i^{(n)}\right)}$$

► $\hat{P}_f^N$ **and** $\hat{P}_{i\delta}^N$ **can be estimated using the same sample**

no need for additional runs of the code G

- **For a couple $\left(\hat{P}_f^N, \hat{P}_{i\delta}^N\right)$, the plug-in estimator of the $S_{i\delta}$ indice is given by :**

$$\hat{S}_{i\delta}^N = \left(\frac{\hat{P}_{i\delta}^N}{\hat{P}_f^N} - 1\right)\mathbb{I}_{\left\{\hat{P}_{i\delta}^N > \hat{P}_f^N\right\}} + \left(1 - \frac{\hat{P}_f^N}{\hat{P}_{i\delta}^N}\right)\mathbb{I}_{\left\{\hat{P}_{i\delta}^N < \hat{P}_f^N\right\}}$$

ASYMPTOTIC LAW OF THE COUPLE $\left(\hat{P}_f^N, \hat{P}_{i\delta}^N\right)$

Lemaître et al. (2015)

Proposition

if $P_{i\delta} \neq P_f$ and under usual conditions :

i. $\mathrm{Supp}(f_{i\delta}) \subseteq \mathrm{Supp}(f_i)$

ii. $\int_{\mathrm{Supp}(f_{i\delta})} \frac{f_{i\delta}^2(x_i)}{f_i(x_i)} dx_i < +\infty$

$$\sqrt{N}\left(\hat{S}_{i\delta}^N - S_{i\delta}\right) \underset{N \to \infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(0, \boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{i\delta} \boldsymbol{\Delta}\right)$$

with :

$$\boldsymbol{\Delta} = \begin{pmatrix} \frac{\partial s}{\partial P_f}\left(P_f, P_{i\delta}\right) \\ \frac{\partial s}{\partial P_{i\delta}}\left(P_f, P_{i\delta}\right) \end{pmatrix} \text{ for } P_{i\delta} \neq P_f \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{i\delta} = \begin{pmatrix} \hat{P}_f^N\left(1 - \hat{P}_f^N\right) & \hat{P}_{i\delta}^N\left(1 - \hat{P}_f^N\right) \\ \hat{P}_{i\delta}^N\left(1 - \hat{P}_f^N\right) & \hat{\sigma}^2{}_{i\delta,N} \end{pmatrix}$$

ASYMPTOTIC LAW OF THE COUPLE $\left(\widehat{P}_f^N, \widehat{P}_{i\delta}^N\right)$

Lemaître et al. (2015)

Proposition

$$\sqrt{N}\left(\hat{S}_{i\delta}^N - S_{i\delta}\right) \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(0, \boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{i\delta} \boldsymbol{\Delta}\right)$$

with :

$$\boldsymbol{\Delta} = \begin{pmatrix} \frac{\partial s}{\partial P_f}\left(P_f, P_{i\delta}\right) \\ \frac{\partial s}{\partial P_{i\delta}}\left(P_f, P_{i\delta}\right) \end{pmatrix} \text{ for } P_{i\delta} \neq P_f \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{i\delta} = \begin{pmatrix} \widehat{P}_f^N\left(1-\widehat{P}_f^N\right) & \widehat{P}_{i\delta}^N\left(1-\widehat{P}_f^N\right) \\ \widehat{P}_{i\delta}^N\left(1-\widehat{P}_f^N\right) & \hat{\sigma}^2{}_{i\delta,N} \end{pmatrix}$$

► In spite of the use of a single set of data for both $\widehat{P}_f^N$ and $\widehat{P}_{i\delta}^N$:

ASYMPTOTIC LAW OF THE COUPLE $\left(\widehat{P}_f^N, \widehat{P}_{i\delta}^N\right)$

Lemaître et al. (2015)

Proposition

$$\sqrt{N}\left(\hat{S}_{i\delta}^N - S_{i\delta}\right) \underset{N \to \infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(0, \boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{i\delta} \boldsymbol{\Delta}\right)$$

with :

$$\boldsymbol{\Delta} = \begin{pmatrix} \frac{\partial s}{\partial P_f}\left(P_f, P_{i\delta}\right) \\ \frac{\partial s}{\partial P_{i\delta}}\left(P_f, P_{i\delta}\right) \end{pmatrix} \text{ for } P_{i\delta} \neq P_f \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{i\delta} = \begin{pmatrix} \widehat{P}_f^N\left(1 - \widehat{P}_f^N\right) & \widehat{P}_{i\delta}^N\left(1 - \widehat{P}_f^N\right) \\ \widehat{P}_{i\delta}^N\left(1 - \widehat{P}_f^N\right) & \hat{\sigma}^2{}_{i\delta,N} \end{pmatrix}$$

► In spite of the use of a single set of data for both $\widehat{P}_f^N$ and $\widehat{P}_{i\delta}^N$:

- $\hat{S}_{i\delta}^N$ **converges to the true value** $S_{i\delta}$

ASYMPTOTIC LAW OF THE COUPLE $\left(\widehat{P}_f^N, \widehat{P}_{i\delta}^N\right)$

Lemaître et al. (2015)

Proposition

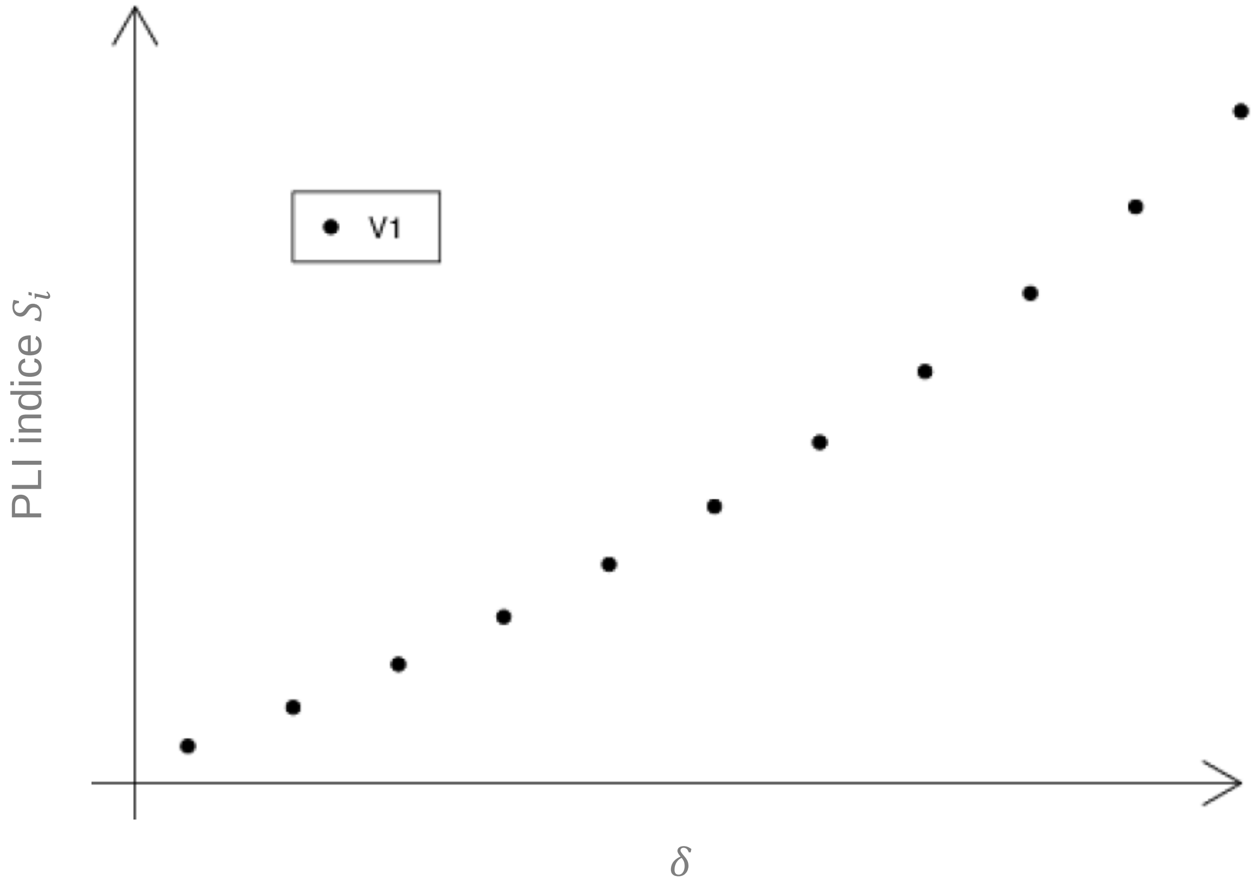
$$\sqrt{N}\left(\hat{S}_{i\delta}^N - S_{i\delta}\right) \underset{N \to \infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(0, \boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{i\delta} \boldsymbol{\Delta}\right)$$

with :

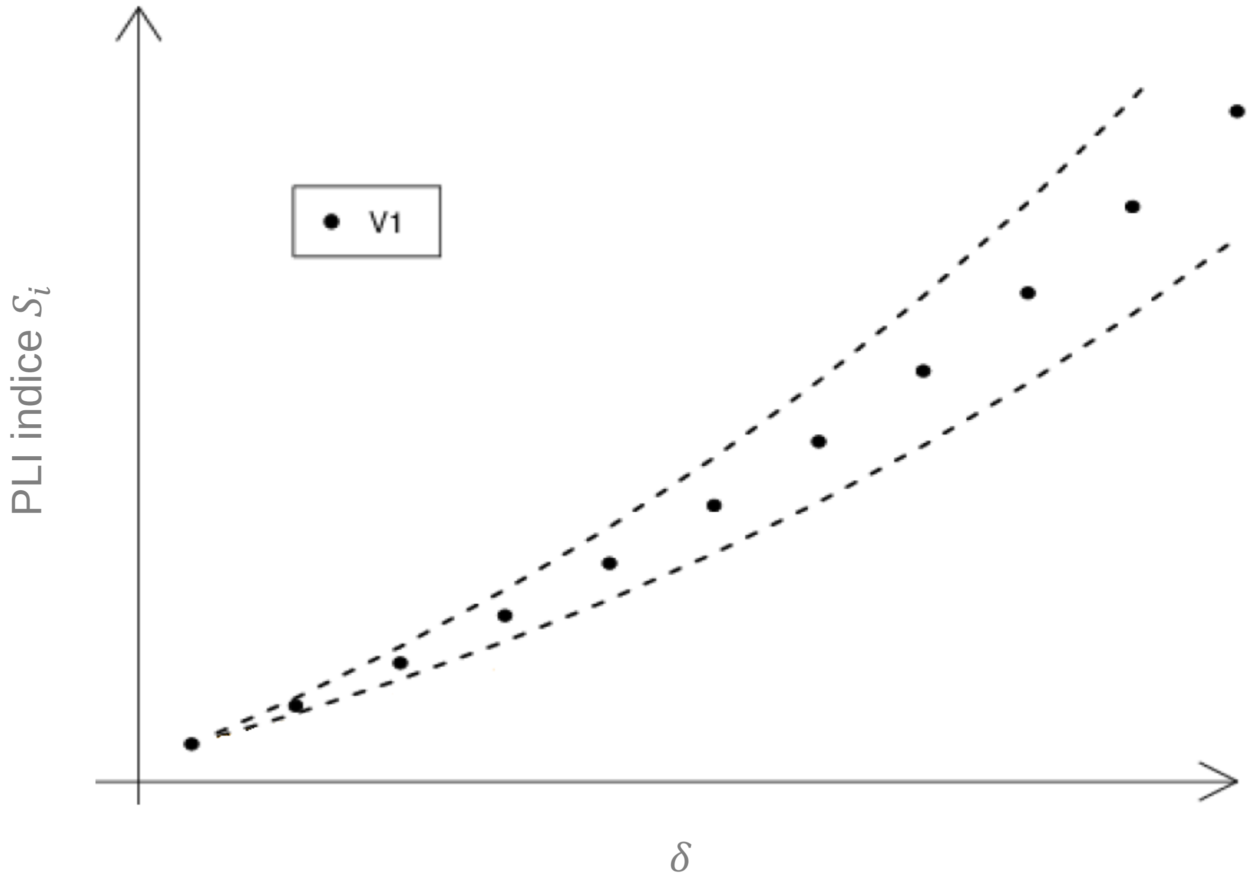
$$\boldsymbol{\Delta} = \begin{pmatrix} \frac{\partial s}{\partial P_f}\left(P_f, P_{i\delta}\right) \\ \frac{\partial s}{\partial P_{i\delta}}\left(P_f, P_{i\delta}\right) \end{pmatrix} \text{ for } P_{i\delta} \neq P_f \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{i\delta} = \begin{pmatrix} \widehat{P}_f^N\left(1 - \widehat{P}_f^N\right) & \widehat{P}_{i\delta}^N\left(1 - \widehat{P}_f^N\right) \\ \widehat{P}_{i\delta}^N\left(1 - \widehat{P}_f^N\right) & \hat{\sigma}^2{}_{i\delta,N} \end{pmatrix}$$

- In spite of the use of a single set of data for both $\widehat{P}_f^N$ and $\widehat{P}_{i\delta}^N$:
  - **$\hat{S}_{i\delta}^N$ converges to the true value $S_{i\delta}$**
  - **We are able to control the estimation error**

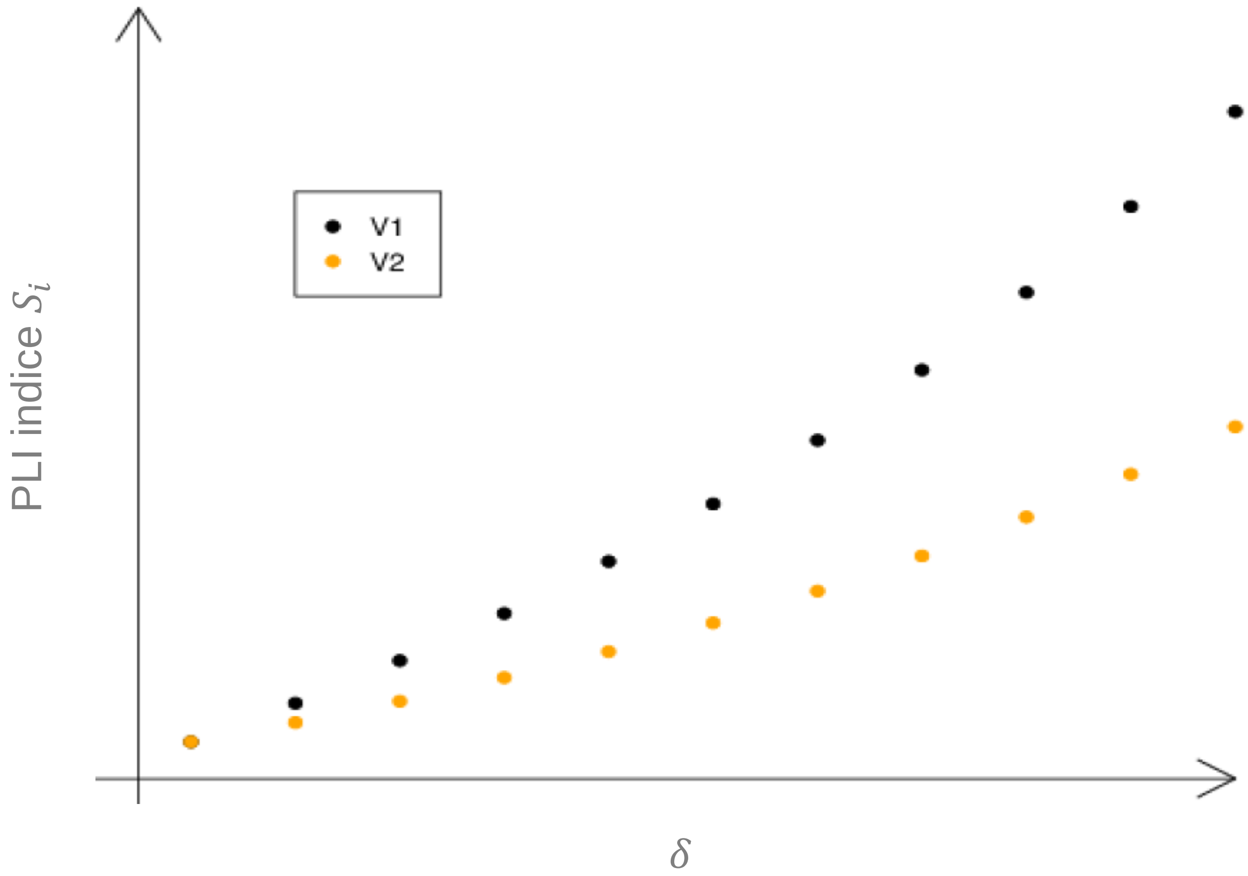
ILLUSTRATION



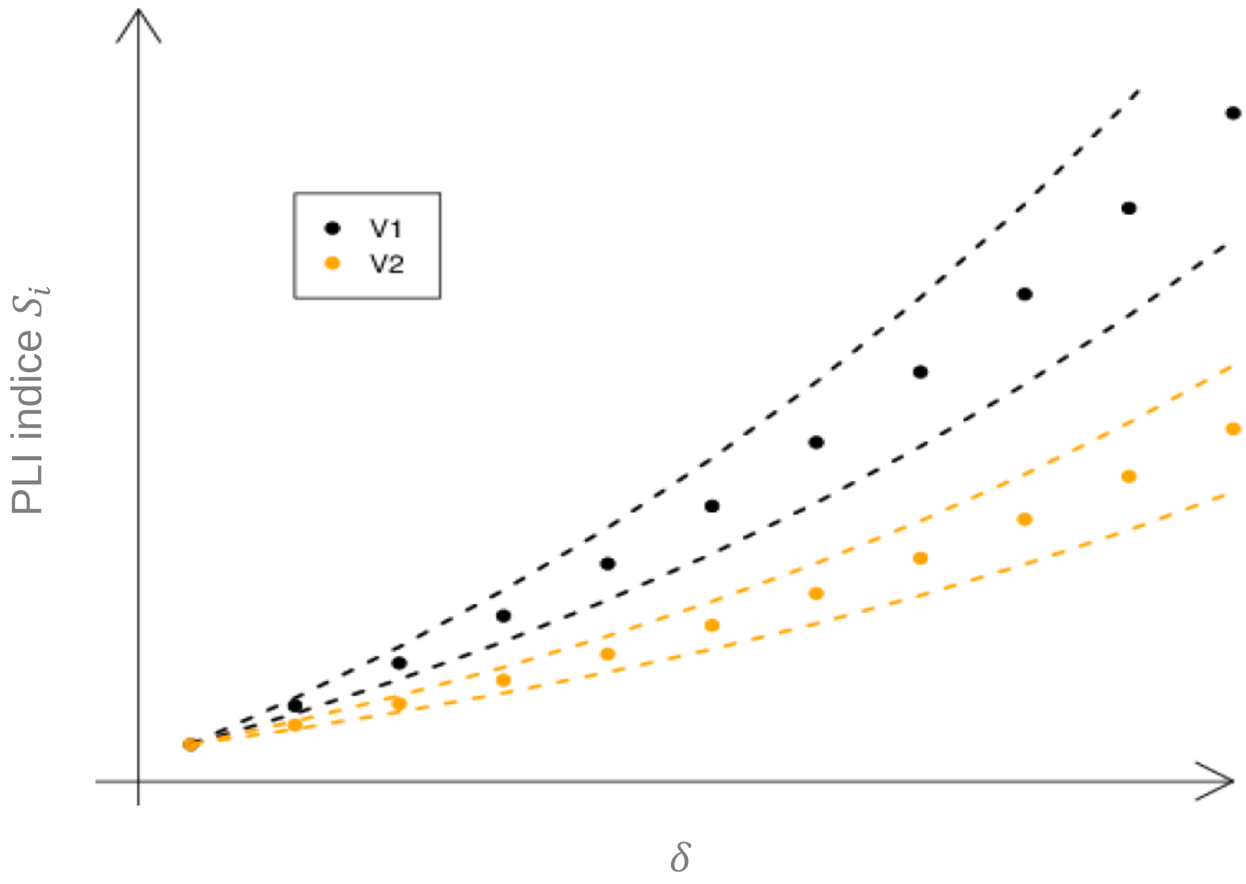
ILLUSTRATION

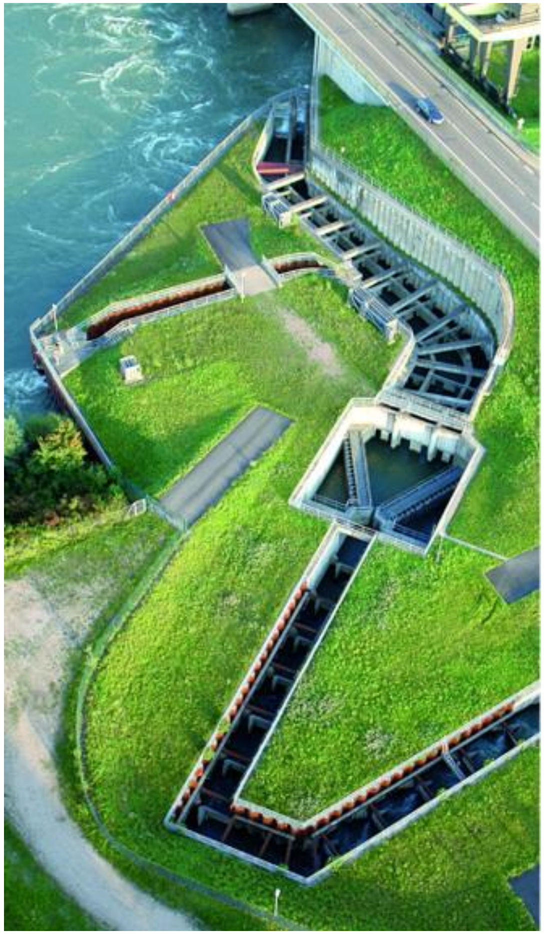


ILLUSTRATION



ILLUSTRATION





1. Context : sensitivity analysis on a failure probability
2. Principle of PLI indices
3. Density perturbation
4. Estimation of perturbed probability
5. **Bayesian approach using GP**

WHY A BAYESIAN APPROACH ?

Difficulties in the estimation of $P_{i\delta}$ and P_f :

- The computer code G is costly (CPU time, memory,…)
- P_f is low ($< 10^{-3}$), sometimes $\approx 10^{-6}$ or 10^{-8}

Hence, we will use :

- an importance sampling method to generate more $\boldsymbol{x}^{(n)}$ in the failure domain $\mathcal{D}_f$ than 1 out of $1/P_f$ (on average)
- a GP meta-model to reduce the number of code runs required for the estimation

WHY A BAYESIAN APPROACH ?

Hence, we will use :

- an importance sampling method to generate more $\boldsymbol{x}^{(n)}$ in the failure domain $\mathcal{D}_f$ than 1 out of $1/P_f$ (on average)
- a GP meta-model to reduce the number of code runs required for the estimation

Two important remarks about IS estimation :

- Desirable property : IS allows unbiased estimation of $P_{i\delta}$ and P_f
- The closer the instrumental density $q(\boldsymbol{x})$ from the optimal one

$$q^*(\boldsymbol{x}) = \frac{\mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) d\boldsymbol{x}}{P_f},$$

the more accurate (low variance) the estimation

WHY A BAYESIAN APPROACH ?

Hence, we will use :

- an importance sampling method to generate more $\boldsymbol{x}^{(n)}$ in the failure domain $\mathcal{D}_f$ than 1 out of $1/P_f$ (on average)
- a GP meta-model to reduce the number of code runs required for the estimation

1st possibility : replace G (costly) by a meta-model $\tilde{G}$ in the expression of $\hat{P}_f^N$

$$\hat{P}_f^N = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{\{\tilde{G}(\boldsymbol{x}^{(n)})<0\}}$$

► In general, this estimator is biased because of the direct use of $\tilde{G}$

2nd possibility : use $\tilde{G}$ only to define a relevant instrumental density and calculate $\hat{P}_f^N$ using the true G

► allows to take advantage of the regularity of G, without losing the unbiased estimation property

BAYESIAN IMPORTANCE SAMPLING

A two-stages procedure :

1) **Build a surrogate model for G (or $G \times f$) using a $n_0 < n$ budget of code runs and deduce an instrumental density we hope to be close from q^***

2) **Calculate $\hat{P}_f^N$ (and $\hat{P}_{i\delta}^N$) by importance sampling, the $n_1 = n - n_0$ remaining budget, using the instrumental density obtained at the previous stage**

BAYESIAN IMPORTANCE SAMPLING

NIS algorithm : approximation of $G \times f$ by a kernel estimator (Zhang, 1996) or a « Linear Blend Frequency Polygon » (Neddermeyer, 2009)

More recently : procedures using Bayesian approximation for G with gaussian process a priori (Dubourg, 2011), (Auffray *et al.*, 2012), (Dubourg *et al.*, 2013), (Bect & Vazquez, 2013)

BAYESIAN IMPORTANCE SAMPLING

We make the hypothesis the code G is the realization of a GP - Notations are :

$\mathbb{P}_0^G$ a priori law on G

$\mathbb{P}_{n_0}^G$ a posteriori law after the first stage

$\mathbb{P}^q$ law of the $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n_1)}$ (iid) generated for the estimation of $\hat{P}_f^N$

$\mathbb{P}_{n_0}^{G,q}$ joint a posteriori law of $G, \boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n_1)}$

Then, for a cost function $L: \mathbb{R}^2 \rightarrow \mathbb{R}$, we look for :

$$q^*\left(\mathbb{P}_{n_0}^G\right) = \underset{q}{\operatorname{argmin}}\, \mathbb{E}_{n_0}^{G,q}\left[L\left(P_f, \hat{P}_f^{IS,n_1}\right)\right]$$

OPTIMAL DENSITY FOR A QUADRATIC COST FUNCTION

Let's assume that :

$$L : (a, b) \to (a - b)^2$$

and G admits moments of order 2 under $\mathbb{P}_{n_0}^G$

Then the optimal density is :

$$q^*\left(\mathbb{P}_{n_0}^G\right) \propto f \times \sqrt{\mathbb{E}_{n_0}^G[G^2]}$$

We denote $g_{n_0}(\boldsymbol{x}) = \sqrt{\mathbb{E}_{n_0}^G[G(\boldsymbol{x})^2]}$, and $q_{n_0}^* = q^*\left(\mathbb{P}_{n_0}^G\right)$

BAYESIAN ESTIMATOR OF P_f

The constant $Z = \int_{\mathbb{X}} g_{n_0} f$ can be estimated without any call to G

► We estimate it by simple MC of size m (large) : $\hat{Z}_m$

The corresponding estimator of P_f

$$\hat{P}_f^{BIS,n_0,n_1} = \hat{Z}_m \frac{1}{n_1} \sum_{n=1}^{n_1} \frac{G\left(\boldsymbol{X}^{(n)}\right)}{g_{n_0}(\boldsymbol{X}^{(n)}) f(\boldsymbol{X}^{(n)})} f\left(\boldsymbol{X}^{(n)}\right) = \hat{Z}_m \frac{1}{n_1} \sum_{n=1}^{n_1} \frac{G\left(\boldsymbol{X}^{(n)}\right)}{g_{n_0}(\boldsymbol{X}^{(n)})}$$

► Unbiased since $\hat{Z}_m$ is an unbiased estimator of Z, independent on the $\boldsymbol{X}^{(n)}$

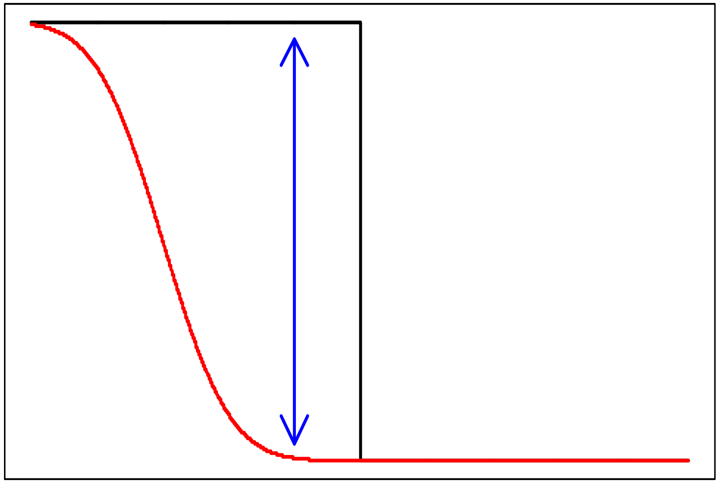
PROBLEM : IF $q^*_{n_0}$ UNDER-ESTIMATES THE OPTIMAL DENSITY $q^*(h)$ ON SOME PARTS OF $\mathbb{X}$

For a perfect model, when $g_{n_0}(\boldsymbol{x}) = G(\boldsymbol{x})$, the theoretical variance is null :

$$\mathrm{Var}\left(\hat{P}_f^{BIS,n_0,n_1}\right) = \frac{1}{n_1}\left(Z\int_{\mathbb{X}} \frac{G^2(\boldsymbol{x})f(\boldsymbol{x})}{g_{n_0}(\boldsymbol{x})}d\boldsymbol{x} - {P_f}^2\right)$$

If g_{n_0} is « biased » the variance tends to $+\infty$

PROBLEM : IF $q^*_{n_0}$ UNDER-ESTIMATES THE OPTIMAL DENSITY $q^*(h)$ ON SOME PARTS OF $\mathbb{X}$



In dimension 1, $\mathbb{I}^2_{\{G(\boldsymbol{x})<0\}}$ in black, $g_{n_0}(\boldsymbol{x})$ in red

SOLUTION : A « DEFENSIVE » SAMPLING

Review of the techniques allowing to control this problem by Owen & Zhou (2000)

Technique of the « defensive mixture » by Hesterberg (1995) :

We add a ε quantity to bound the variance

$$q_\varepsilon(\boldsymbol{x}) = \frac{\left(g_{n_0}(\boldsymbol{x}) + \varepsilon\right) f(\boldsymbol{x})}{Z + \varepsilon}$$

$$= \frac{g_{n_0}(\boldsymbol{x}) f(\boldsymbol{x})}{Z + \varepsilon} + \frac{\varepsilon}{Z + \varepsilon} f(\boldsymbol{x})$$

$$= c_{\boldsymbol{\varepsilon}} f(\boldsymbol{x}) + (1 - c_\varepsilon) q^*(\boldsymbol{x})$$

SOLUTION : A « DEFENSIVE » SAMPLING

The estimator remains unbiased :

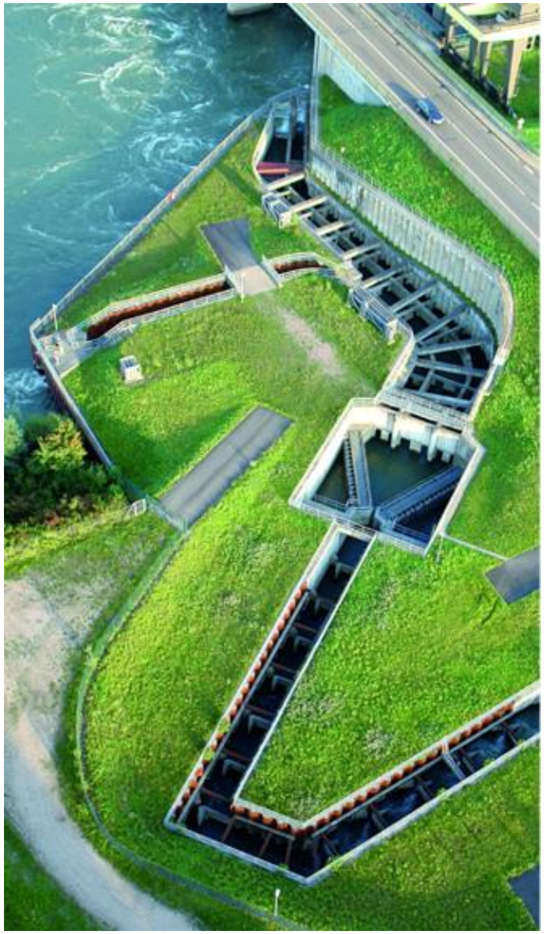
$$\widehat{P}_f^{BIS,n_0,n_1,\varepsilon} = \frac{Z+\varepsilon}{n_1}\sum_{\boldsymbol{n}=1}^{n_1}\frac{G\left(\boldsymbol{X}^{(n)}\right)}{g_{n_0}(\boldsymbol{X}^{(n)})+\varepsilon}, \qquad \boldsymbol{X}^{(n)} \sim q_\varepsilon,$$

with a variance :

$$\mathrm{Var}\left(\widehat{P}_f^{BIS,n_0,n_1,\varepsilon}\right) = \frac{1}{n_1}\left((Z+\varepsilon)\int_{\mathbb{X}}\frac{G^2(\boldsymbol{x})f(\boldsymbol{x})}{g_{n_0}(\boldsymbol{x})+\varepsilon}d\boldsymbol{x} - {P_f}^2\right)$$

$$\leq \frac{1}{n_1}\left(\frac{Z+\varepsilon}{\varepsilon}P_f - {P_f}^2\right)$$

which has a finite value



1. Context : sensitivity analysis on a failure probability
2. Principle of PLI indices
3. Density perturbation
4. Estimation of perturbed probability
5. Bayesian approach using GP
6. **Conclusion**

CONCLUSION

- **We defined a new type of sensitivity indices called PLI adapted to some reliability quantities of interest such as probailities of failure** $P_f = \int_{\mathbb{X}} \mathbb{I}_{\{G(\boldsymbol{x})<0\}} f(\boldsymbol{x}) d\boldsymbol{x}$
- **To this aim we formalized a way of perturbing a law of probability with respect to one of its parameters**
- **We developped an estimation technique of our indices** $S_{i\delta}$ **taking into account our two main constraints :**
  - G is costly
  - P_f is low $\approx 10^{-6}$ or 10^{-8}

CONCLUSION

Perspectives

- **Reach the same kind of theoretical/practical results for PLI indices over quantiles**
- **Take dependences into account**

REFERENCES

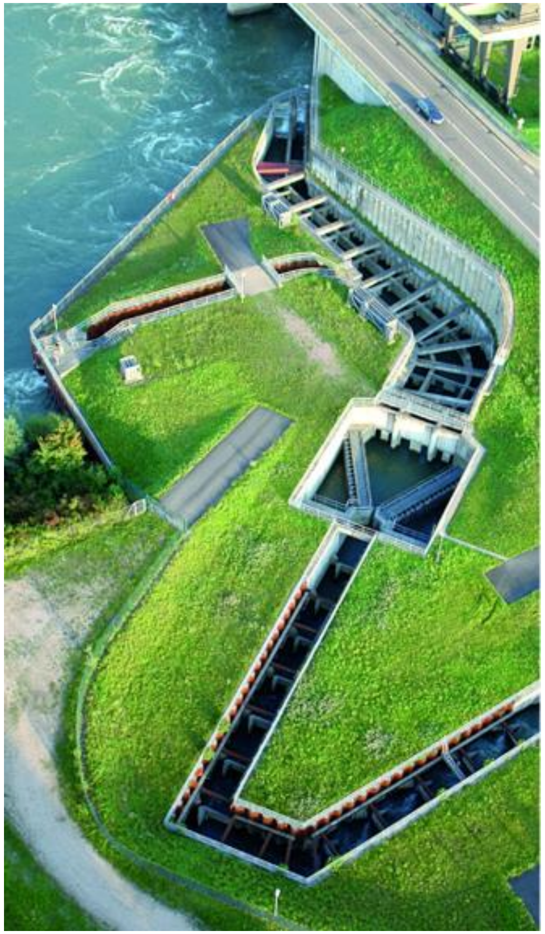
P. Lemaître. Sensitivity analysis in structural reliability, Thèse de l'Université Bordeaux I, 2014

P. Lemaître, E. Sergienko, A. Arnaud, N. Bousquet, F. Gamboa and B. Iooss. Density modification based reliability sensitivity analysis. Journal of Statistical Computation and Simulation, 85 :1200-1223, 205

J. Bect, R. Sueur, A. Gerossier, L. Mongellaz, S. Petit and E. Vazquez (2015), Echantillonnage preferentiel et meta-modeles : methodes bayesiennes optimale et defensive, 47emes Journees de Statistique de la SFdS, Lille, France

MERCI DE VOTRE ATTENTION

ANNEXES



Expression of the variance of $\widehat{P}_f^{IS,N}$

Demonstration of the optimality of a density $\propto f g_{n_0}$

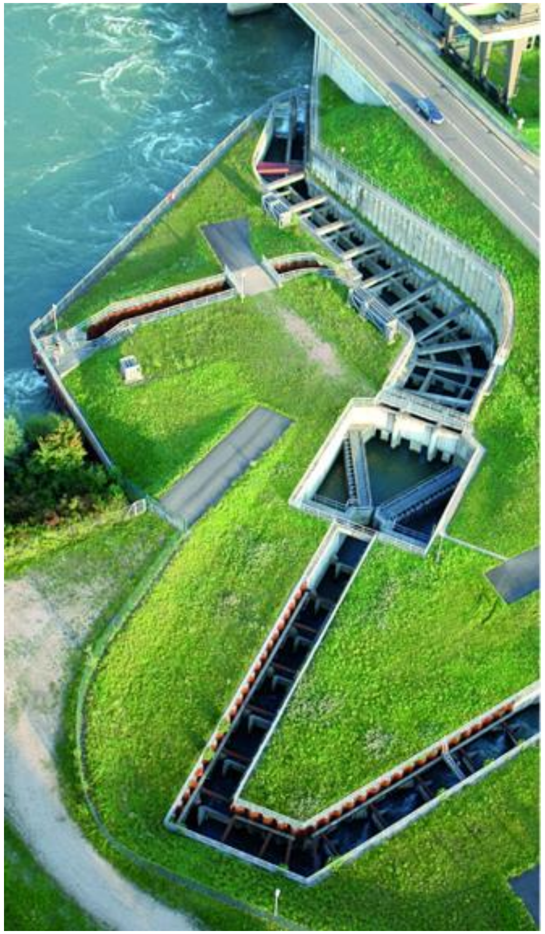
EXPRESSION OF THE VARIANCE OF $\widehat{P}_f^{IS,N}$

The variance of $\widehat{P}_f^{IS,N}$ is :

$$\mathrm{Var}\left(\widehat{P}_f^{IS,N}\right) = \frac{1}{N}\left(\int_{\mathbb{X}} \frac{G^2 f^2}{q} - {P_f}^2\right) = \frac{{P_f}^2}{N}\left(\int_{\mathbb{X}} \frac{\left(q - q^*(G)\right)^2}{q}\right)$$

CentraleSupélec

ANNEXES



Expression of the variance of $\widehat{P}_f^{IS,N}$

Demonstration of the optimality of a density $\propto f g_{n_0}$

DEMONSTRATION OF THE OPTIMALITY OF A DENSITY $\propto f g_{n_0}$

We have the following formula for the variance :

$$\mathrm{Var}\left(\widehat{P}_f^{IS,n_1}\right) = \frac{1}{n_1}\left(\int_{\mathbb{X}} \frac{G^2 f^2}{q} - P_f^2\right) = \frac{P_f^2}{n_1}\left(\int_{\mathbb{X}} \frac{\left(q - q^*(G)\right)^2}{q}\right) \qquad (\lozenge)$$

Moreover, by conditionning by the GP G, we get :

$$\mathbb{E}_{n_0}^{G,q}\left[\left(\widehat{P}_f^{IS,n_1} - P_f\right)^2\right] = \mathbb{E}_{n_0}^{G}\left[\mathbb{E}_{n_0}^{q}\left[\left(\widehat{P}_f^{IS,n_1} - P_f\right)^2\right]\right] \qquad (\circ)$$

CentraleSupélec

DEMONSTRATION OF THE OPTIMALITY OF A DENSITY $\propto f g_{n_0}$

Injecting (◊) into (○), we obtain :

$$\mathbb{E}_{n_0}^{G,q}\left[\left(\widehat{P}_f^{IS,n_1} - P_f\right)^2\right] = \mathbb{E}_{n_0}^{G}\left[\frac{1}{n_1}\left(\int_{\mathbb{X}} \frac{G^2 f^2}{q} - P_f^2\right)\right]$$

$$= \frac{1}{n_1}\left(\int_{\mathbb{X}} \frac{\mathbb{E}_{n_0}^{G}[G^2] f^2}{q} - \mathbb{E}_{n_0}^{G}\left[P_f^2\right]\right)$$

$$= \frac{1}{n_1}\left(\int_{\mathbb{X}} \frac{g_{n_0}^2 f^2}{q} - \mathbb{E}_{n_0}^{G}\left[P_f^2\right]\right) \qquad \textbf{(♣)}$$

DEMONSTRATION OF THE OPTIMALITY OF A DENSITY $\propto f g_{n_0}$

Applying Cauchy-Schwarz inequality to the first term of (♣), appears that :

$$\int_{\mathbb{X}} \frac{g_{n_0}^2 f^2}{q} = \int_{\mathbb{X}} \left(\frac{g_{n_0} f}{q}\right)^2 q \geq \left(\int_{\mathbb{X}} \frac{g_{n_0} f}{q} q\right)^2 = \left(\int_{\mathbb{X}} g_{n_0} f\right)^2 \textbf{(♠)}$$

DEMONSTRATION OF THE OPTIMALITY OF A DENSITY $\propto f g_{n_0}$

Injecting (♠) in (♣), and as this term is the only one that depends on the density q, we show that :

$$\mathbb{E}_{\boldsymbol{n_0}}^{\boldsymbol{G,q}}\left[\left(\widehat{P}_f^{IS,n_1}-\boldsymbol{\alpha}\right)^2\right]=\frac{\mathbf{1}}{\boldsymbol{n_1}}\left(\int_{\mathbb{X}}\frac{\boldsymbol{g_{n_0}^2 f^2}}{\boldsymbol{q}}-\mathbb{E}_{n_0}^{G}\left[P_f^2\right]\right)$$

$$\geq\frac{\mathbf{1}}{\boldsymbol{n_1}}\left(\left(\int_{\mathbb{X}}\boldsymbol{g_{n_0}\pi}\right)^{\mathbf{2}}-\mathbb{E}_{n_0}^{G}\left[P_f^2\right]\right)\qquad (☺)$$

DEMONSTRATION OF THE OPTIMALITY OF A DENSITY $\propto f g_{n_0}$

To conclude, one notices that the equality case is obtained in (♠) and hence in (☺) when q is proportional to $g_{n_0} f$, which means for the instrumental density $q^*\left(\mathbb{P}_{n_0}^G\right)$.