Airbus Group Innovations

Unsupervised parameter estimation in computer experiments

Journées MAS, 28^{th} August 2014

Nabil RACHDI, nabil.rachdi@airbus.com

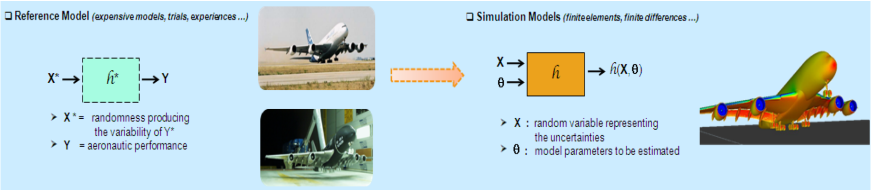
AIRBUS GROUP

Outline

AIRBUS GROUP

Context

- From Real life to Simulated life...

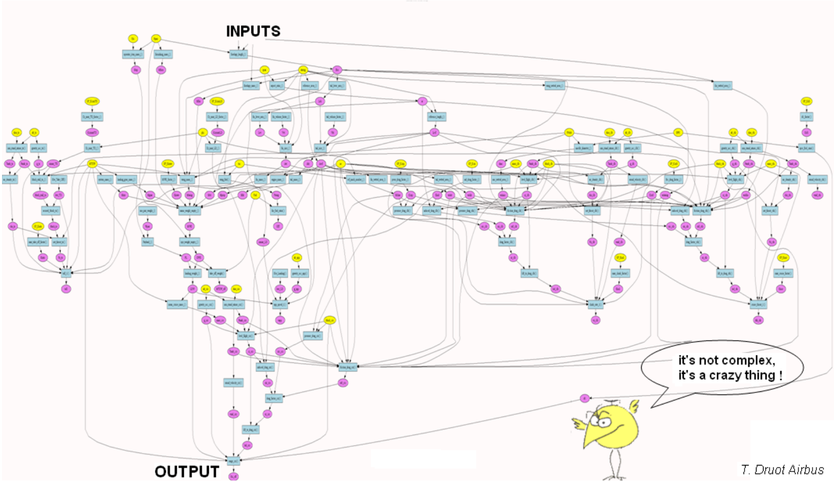


- Y = Variable of Interest (uncertain !)
- ρ^* = Quantity of Interest (*quantile, pdf, exceed. probability* ...)
- **Challenge :**

  From ref. data $Y_1, ..., Y_n$ or $(\mathbf{X}_1^*, Y_1), ..., (\mathbf{X}_n^*, Y_n)$ (n limited !)

  $\longrightarrow$ Choose h and θ to predict ρ^* with simulation model(s) h

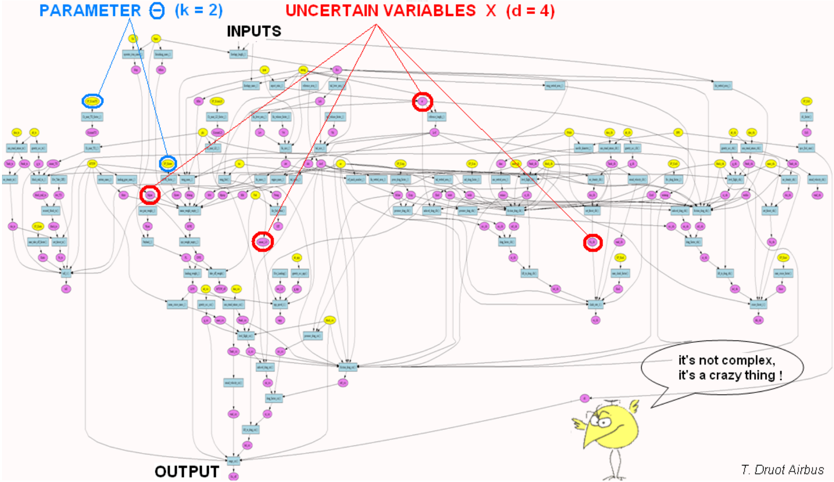
Numerical Simulations under Uncertainties



T. Druot Airbus

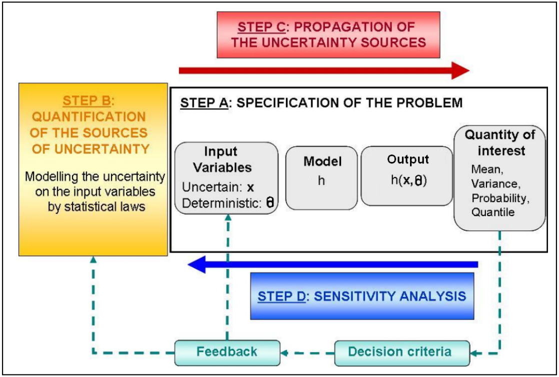
AIRBUS GROUP

Numerical Simulations under Uncertainties

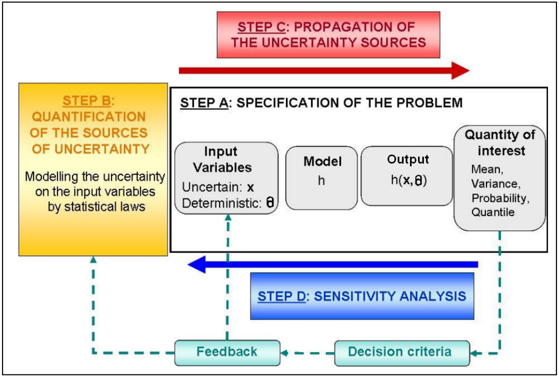


AIRBUS GROUP

Methodology commonly adopted [de Rocquigny et al. (2008)]



Methodology commonly adopted [de Rocquigny et al. (2008)]



⇒ 2 kind of problems:

- **Inverse Problem:** identify the parameter θ (mechanical, thermal...) from a set $Y_1, \ldots, Y_n$
- **Prediction Problem:** estimate θ (tuning parameters, etc.) and simulate with $h(\mathbf{X}, \widehat{\theta})$ under $\mathbf{X}$

Questions ?

- **Inverse problem:** If the "real life" inputs $\mathbf{X}_i^*$'s are not observed ? How to calibrate ?
  (e.g input code $\neq$ experimental conditions etc...)

AIRBUS GROUP

Questions ?

- **Inverse problem:** If the "real life" inputs $\mathbf{X}_i^*$'s are not observed ? How to calibrate ?
  (e.g input code $\neq$ experimental conditions etc...)

  In other words, for each simulation input $\mathbf{X}_i$ we do not have the associated response Y_i, which may be referred to as Unsupervised Learning.

AIRBUS GROUP

Link with Statistical Learning

Classical learning areas (see Hastie et al [7], Massart [8])

- **Unsupervised learning**: We observe $\mathbf{X}_1^*, ..., \mathbf{X}_n^*$ i.i.d $\mathbb{P}_{\mathbf{X}}^*$ (unknown) and we look for a probabilistic feature of $P_{\mathbf{X}}^*$
- **Semi-supervised learning** With $l < n$, we observe $(\mathbf{X}_i^*, \mathbf{Y}_i^*)_{i \leq l} + \mathbf{X}_{l+1}^*, ..., \mathbf{X}_n^*$ and we look for a map $g : \mathcal{X}^* \rightarrow \mathcal{Y}^*$
- **Supervised/inductive learning**: We observe $(\mathbf{X}_1^*, \mathbf{Y}_1^*), ..., (\mathbf{X}_n^*, \mathbf{Y}_n^*)$ and we look for a map $g : \mathcal{X}^* \rightarrow \mathcal{Y}^*$

Our learning context

- If the $\mathbf{X}_i^*$'s are observed ?
  Data at disposal:
  $(\mathbf{X}_1^*, \mathbf{Y}_1^*), ..., (\mathbf{X}_n^*, \mathbf{Y}_n^*) \ + \ (\mathbf{X}_1, h(\mathbf{X}_1, \theta)), ..., (\mathbf{X}_m, h(\mathbf{X}_m, \theta)), \quad m >> n$
  The framework $\mathbf{Y}_1^*, ..., \mathbf{Y}_n^* + \mathbf{X}_1, ..., \mathbf{X}_m$ *may be seen between Supervised and Semi-supervised learning...*
- If the $\mathbf{X}_i^*$'s are NOT observed ?
  Data at disposal: $\mathbf{Y}_1^*, ..., \mathbf{Y}_n^* \ + \ h(\mathbf{X}_1, \theta), ..., h(\mathbf{X}_m, \theta), \quad m >> n$
  The framework $\mathbf{Y}_1^*, ..., \mathbf{Y}_n^* + \mathbf{X}_1, ..., \mathbf{X}_m$ *may be seen between Unsupervised and Semi-supervised learning...*

AIRBUS GROUP

Other Questions ?

- **Prediction problem:** even if they are observed, should we always use regression models for predicting some quantity of interest ?

AIRBUS GROUP

Other Questions ?

- **Prediction problem:** even if they are observed, should we always use regression models for predicting some quantity of interest ?

- for instance, what is the meaning of

$$\mathbb{P}(h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg}) > s) \quad \text{or} \quad pdf_{h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg})} \text{ ?}$$

where $\widehat{\boldsymbol{\theta}}_{reg}$ is the mean-squares estimator of the model $Y_i = h(\mathbf{X}^i, \theta) + \varepsilon_i$

AIRBUS GROUP

Other Questions ?

- **Prediction problem:** even if they are observed, should we always use regression models for predicting some quantity of interest ?

- for instance, what is the meaning of

$$\mathbb{P}(h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg}) > s) \quad \text{or} \quad pdf_{h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg})} \, ?$$

where $\widehat{\boldsymbol{\theta}}_{reg}$ is the mean-squares estimator of the model $Y_i = h(\mathbf{X}^i, \theta) + \varepsilon_i$

... "*duality*" between estimation procedure and target prediction ...

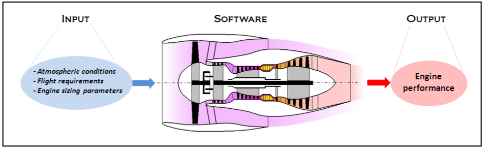
Example 1: Inverse Problem

N. Rachdi, J-C. Fort, T. Klein [2]

- **Fuel Mass data:**

Reference Fuel Masses [kg]							
7918	7671	7719	7839	7912	7963	7693	7815
7872	7679	8013	7935	7794	8045	7671	7985
7755	7658	7684	7658	7690	7700	7876	7769
8058	7710	7746	7698	7666	7749	7764	7667

- **Model (noisy simulator):**



- **Goal:** Identify SFC($=\theta$) (Specific Fuel Consumption) under uncertainties $\mathbf{X}$

Rq: We do not have at disposal the inputs providing the Fuel Mass data

AIRBUS GROUP

Example 1: Inverse Problem

- **Idea:** Minimize the "distance" between the distribution of Fuel Mass reference data Y_i and the distribution of the noisy computer code $h(\mathbf{X}, \theta)$ ($\mathbf{X}$ = uncertainties, θ=SFC)
- **Kullback-Leibler minimization:**

$$KL(f_1, f_2) = \int_{\mathcal{Y}} \log\left(\frac{f_1}{f_2}\right) f_1$$

  Set f = density of Y, $f_{\boldsymbol{\theta}}$ = density of $h(\mathbf{X}, \boldsymbol{\theta})$
- **Goal:** Find $\boldsymbol{\theta}$ that minimizes $KL(f, f_{\boldsymbol{\theta}})$.

AIRBUS GROUP

Example 1: Inverse Problem

- **Idea:** Minimize the "distance" between the distribution of Fuel Mass reference data Y_i and the distribution of the noisy computer code $h(\mathbf{X}, \theta)$ ($\mathbf{X}$ = uncertainties, θ=SFC)
- **Kullback-Leibler minimization:**

$$KL(f_1, f_2) = \int_{\mathcal{Y}} \log\left(\frac{f_1}{f_2}\right) f_1$$

  Set f = density of Y, $f_{\boldsymbol{\theta}}$ = density of $h(\mathbf{X}, \boldsymbol{\theta})$
- **Goal:** Find $\boldsymbol{\theta}$ that minimizes $KL(f, f_{\boldsymbol{\theta}})$.
- **2 Difficulties**
  - f is unknown $\rightarrow$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^{n} \delta_{Y_i}$
  - $f_{\boldsymbol{\theta}}$ intractable $\rightarrow$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_{\boldsymbol{\theta}}^m = \frac{1}{m}\sum_{j=1}^{m} K_{b_m}(\cdot - h(\mathbf{X}_j, \boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^{x}\right)$

AIRBUS GROUP

Example 1: Inverse Problem

- **Idea:** Minimize the "distance" between the distribution of Fuel Mass reference data Y_i and the distribution of the noisy computer code $h(\mathbf{X},\theta)$ ($\mathbf{X}$ = uncertainties, θ=SFC)
- **Kullback-Leibler minimization:**

$$KL(f_1,f_2) = \int_{\mathcal{Y}} \log\left(\frac{f_1}{f_2}\right) f_1$$

Set f = density of Y, $f_{\boldsymbol{\theta}}$ = density of $h(\mathbf{X},\boldsymbol{\theta})$

- **Goal:** Find $\boldsymbol{\theta}$ that minimizes $KL(f,f_{\boldsymbol{\theta}})$.
- **2 Difficulties**
  - f is unknown $\rightarrow$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$
  - $f_{\boldsymbol{\theta}}$ intractable $\rightarrow$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_{\boldsymbol{\theta}}^m = \frac{1}{m}\sum_{j=1}^m K_{b_m}(\cdot - h(\mathbf{X}_j,\boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\right)$
- **Estimator**

$$\widehat{\boldsymbol{\theta}}_{KL} = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, KL(f^n,f_{\boldsymbol{\theta}}^m) = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}} -\frac{1}{n}\sum_{i=1}^n \log(f_{\boldsymbol{\theta}}^m)(Y_i)$$

AIRBUS GROUP

Example 1: Inverse Problem

- **Idea:** Minimize the "distance" between the distribution of Fuel Mass reference data Y_i and the distribution of the noisy computer code $h(\mathbf{X}, \theta)$ ($\mathbf{X}$ = uncertainties, θ=SFC)
- **Kullback-Leibler minimization:**

$$KL(f_1, f_2) = \int_{\mathcal{Y}} \log\left(\frac{f_1}{f_2}\right) f_1$$

Set f = density of Y, $f_{\boldsymbol{\theta}}$ = density of $h(\mathbf{X}, \boldsymbol{\theta})$

- **Goal:** Find $\boldsymbol{\theta}$ that minimizes $KL(f, f_{\boldsymbol{\theta}})$.
- **2 Difficulties**
  - f is unknown $\rightarrow$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$
  - $f_{\boldsymbol{\theta}}$ intractable $\rightarrow$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_{\boldsymbol{\theta}}^m = \frac{1}{m}\sum_{j=1}^m K_{b_m}(\cdot - h(\mathbf{X}_j, \boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\right)$
- **Estimator**
  Remark: This Estimator doesn't depend on the (unknown) $\mathbf{X}_i$'s providing the Y_i's !

Example 1: Inverse Problem

- **Idea:** Minimize the "distance" between the distribution of Fuel Mass reference data Y_i and the distribution of the noisy computer code $h(\mathbf{X}, \theta)$ ($\mathbf{X}$ = uncertainties, θ=SFC)
- **Kullback-Leibler minimization:**

$$KL(f_1, f_2) = \int_{\mathcal{Y}} \log\left(\frac{f_1}{f_2}\right) f_1$$

Set f = density of Y, $f_{\boldsymbol{\theta}}$ = density of $h(\mathbf{X}, \boldsymbol{\theta})$

- **Goal:** Find $\boldsymbol{\theta}$ that minimizes $KL(f, f_{\boldsymbol{\theta}})$.

- **2 Difficulties**
  - f is unknown $\to$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$
  - $f_{\boldsymbol{\theta}}$ intractable $\to$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_{\boldsymbol{\theta}}^m = \frac{1}{m}\sum_{j=1}^m K_{b_m}(\cdot - h(\mathbf{X}_j, \boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^{\mathrm{x}}\right)$
- **Estimator**

$$\widehat{\boldsymbol{\theta}}_{KL} = \underset{\boldsymbol{\theta}\in\Theta}{\mathrm{Argmin}}\, KL(f^n, f_{\boldsymbol{\theta}}^m) = \underset{\boldsymbol{\theta}\in\Theta}{\mathrm{Argmin}} -\frac{1}{n}\sum_{i=1}^n \log(f_{\boldsymbol{\theta}}^m)(Y_i)$$

$$\widehat{SFC} = \widehat{\boldsymbol{\theta}}_{KL}$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f_{MS}}$)

N. Rachdi, J-C. Fort, T. Klein [1]

Suppose that $\mathbf{X}^* = \mathbf{X}$ and that $(\mathbf{X}_1, Y_1), ..., (\mathbf{X}_n, Y_n)$ are available.

Goal: Estimate the pdf of Y from a computer code $h(\mathbf{X}, \theta)$ where $\mathbf{X} \sim P^{\mathrm{x}}$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$)

N. Rachdi, J-C. Fort, T. Klein [1]

Suppose that $\mathbf{X}^* = \mathbf{X}$ and that $(\mathbf{X}_1, Y_1), ..., (\mathbf{X}_n, Y_n)$ are available.

Goal: Estimate the pdf of Y from a computer code $h(\mathbf{X}, \boldsymbol{\theta})$ where $\mathbf{X} \sim P^{\mathrm{x}}$

- **Mean-Squares minimization**

$$\widehat{\boldsymbol{\theta}}_{MS} = \underset{\boldsymbol{\theta} \in \Theta}{\mathrm{Argmin}} \frac{1}{n} \sum_{i=1}^{n} (Y_i - h(\mathbf{X}_i, \boldsymbol{\theta}))^2$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$)

N. Rachdi, J-C. Fort, T. Klein [1]

Suppose that $\mathbf{X}^* = \mathbf{X}$ and that $(\mathbf{X}_1, Y_1), ..., (\mathbf{X}_n, Y_n)$ are available.

Goal: Estimate the pdf of Y from a computer code $h(\mathbf{X}, \boldsymbol{\theta})$ where $\mathbf{X} \sim P^{\mathrm{x}}$

- **Mean-Squares minimization**

$$\widehat{\boldsymbol{\theta}}_{MS} = \underset{\boldsymbol{\theta} \in \Theta}{\mathrm{Argmin}} \frac{1}{n} \sum_{i=1}^{n} (Y_i - h(\mathbf{X}_i, \boldsymbol{\theta}))^2$$

- **Prediction**
  Compute the probability density of $h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{MS})$ under $\mathbf{X} \sim P^{\mathrm{x}}$

$$\boxed{\to \widehat{f}_{MS}}$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_M$)

Other Estimation Procedures...

- **Mean-Squares minimization (version 2)**

$$\widehat{\boldsymbol{\theta}}_M = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{Argmin}} \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \frac{1}{m}\sum_{j=1}^{m} h(\mathbf{X}_j, \boldsymbol{\theta})\right)^2$$

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_M$)

Other Estimation Procedures...

- **Mean-Squares minimization (version 2)**

$$\widehat{\boldsymbol{\theta}}_M = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}} \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \frac{1}{m}\sum_{j=1}^{m} h(\mathbf{X}_j, \boldsymbol{\theta})\right)^2$$

Rq : This version of mean squares minimizes the distance between the "expectations", whereas the previous estimator $\widehat{\boldsymbol{\theta}}_{MS}$ minimizes the distance between "conditional expectations".

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_M$)

Other Estimation Procedures...

- **Mean-Squares minimization (version 2)**

$$\widehat{\boldsymbol{\theta}}_M = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}} \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \frac{1}{m}\sum_{j=1}^{m} h(\mathbf{X}_j, \boldsymbol{\theta})\right)^2$$

Rq : This version of mean squares minimizes the distance between the "expectations", whereas the previous estimator $\widehat{\boldsymbol{\theta}}_{MS}$ minimizes the distance between "conditional expectations".

- **Prediction**
  Compute the probability density of $h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_M)$ under $\mathbf{X} \sim P^x$

$$\boxed{\rightarrow \widehat{f}_M}$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_{M}$, $\widehat{f}_{KL}$)

Other Estimation Procedures...

- **Kullback-Leibler minimization** $KL(f_1, f_2) = \int_{\mathcal{Y}} \log(\frac{f_1}{f_2})\, f_1$
  - f = density of Y, f_θ = density of $h(\mathbf{X}, \theta)$
  - Goal: Find θ that minimizes $KL(f, f_\theta)$.
- **2 Difficulties**
  - f is unknown $\rightarrow$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^{n} \delta_{Y_i}$
  - f_θ intractable $\rightarrow$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_\theta^m = \frac{1}{m}\sum_{j=1}^{m} K_{b_m}(\cdot - h(\mathbf{X}_j, \theta)), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\right)$
- **Estimator**

$$\widehat{\theta}_{KL} = \underset{\theta\in\Theta}{\text{Argmin}}\, KL(f^n, f_\theta^m) = \underset{\theta\in\Theta}{\text{Argmin}} - \frac{1}{n}\sum_{i=1}^{n} \log(f_\theta^m)(Y_i)$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_M$, $\widehat{f}_{KL}$)

Other Estimation Procedures...

- **Kullback-Leibler minimization** $KL(f_1, f_2) = \int_{\mathcal{Y}} \log(\frac{f_1}{f_2})\, f_1$
  - f = density of Y, f_θ = density of $h(\mathbf{X}, \theta)$
  - Goal: Find θ that minimizes $KL(f, f_\theta)$.
- **2 Difficulties**
  - f is unknown $\rightarrow$ replaced by $f^n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$
  - f_θ intractable $\rightarrow$ replaced by a **simulation density** (Kernel, projection, etc...) $\left(f_\theta^m = \frac{1}{m}\sum_{j=1}^m K_{b_m}(\cdot - h(\mathbf{X}_j, \theta)), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\right)$
- **Estimator**

$$\widehat{\theta}_{KL} = \underset{\theta\in\Theta}{\text{Argmin}}\, KL(f^n, f_\theta^m) = \underset{\theta\in\Theta}{\text{Argmin}} -\frac{1}{n}\sum_{i=1}^n \log(f_\theta^m)(Y_i)$$

- **Prediction**
  Compute the probability density of $h(\mathbf{X}, \widehat{\theta}_{KL})$ under $\mathbf{X} \sim P^x$

$$\boxed{\rightarrow \widehat{f}_{KL}}$$

AIRBUS GROUP

Example 2: Density prediction ($\widehat{f}_{MS}$, $\widehat{f}_{M}$, $\widehat{f}_{KL}$)

Question ?

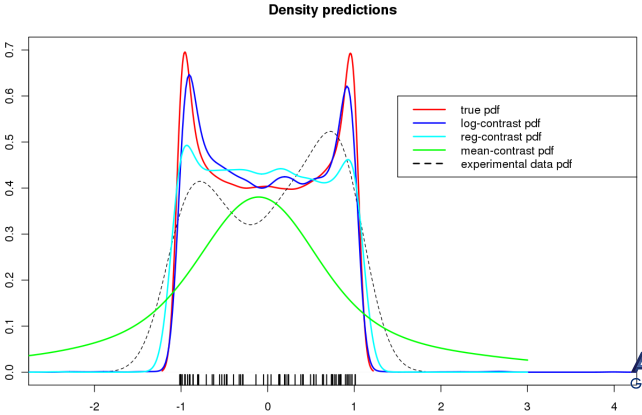
What is the "best" estimator of the density f of Y,

$\widehat{f}_{MS}$, $\widehat{f}_{M}$ or $\widehat{f}_{KL}$?

Toy application

- $Y = \sin(X^*) + 0.01\,\varepsilon, \quad X^*, \varepsilon \sim \mathcal{N}(0,1)$ independents
- $h(X, \boldsymbol{\theta}) = \theta_1 + \theta_2\, X + \theta_3\, X^3,\ X \sim P^x = \mathcal{N}(0,1)$
- $n = 50$ and $m = 10^3$

true pdf, $\widehat{f}_{MS}$, $\widehat{f}_{M}$, $\widehat{f}_{KL}$



Issues

- **Inverse problem:**
  Formalize Stochastic Inverse Problems in a Statistical Learning framework
- **Prediction problem:**
  Define "adapted" estimation procedures (learning algorithms) for a computer code based prediction

AIRBUS GROUP

General Framework

- **Reference data** : set $\mathbf{X}^* = \mathbf{X}$ (i.e " phenomenon causes = code inputs ")

$$Z_1 = (\mathbf{X}_1, Y_1), ..., Z_n = (\mathbf{X}_n, Y_n)$$

with (unknown) dist. $Q^{\mathbf{z}}$ and denote by Q the marginal dist. of Y
$\longrightarrow \mathbf{X}_1, ..., \mathbf{X}_n$ may be unobserved

- **Model** : $\{\mathbf{x} \in \mathcal{X} \mapsto h(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{Y}, \quad \boldsymbol{\theta} \in \Theta\}$
  - mathematical model : $h(\mathbf{x}, \boldsymbol{\theta}) = \sum_{l=1}^{l=q} \phi(\mathbf{x})\, \boldsymbol{\theta}$ etc ...
  - physical/simulation model : $h(\mathbf{x}, \boldsymbol{\theta})$ is the result of a computer code

- **Uncertainty** : Equip $\mathcal{X}$ with a prob. measure $P^{\mathbf{x}}$: $\mathbf{X} \in (\mathcal{X}, P^{\mathbf{x}})$
  - stochastic codes, Monte-Carlo codes, uncertain variables etc...

- **Stochastic Output**: $h(\mathbf{X}, \boldsymbol{\theta})$ supposed known through input/output simulations
  - for instance $\mathbf{x} \mapsto h(\mathbf{x}, \boldsymbol{\theta})$ is has an analytical form but too complicated to compute the distribution $h(\mathbf{X}, \boldsymbol{\theta})$
  - or $\mathbf{x} \mapsto h(\mathbf{x}, \boldsymbol{\theta})$ is an input/output simulation code

AIRBUS GROUP

From Loss function to Contrast function

- **Loss function:** Given an action set $\mathcal{A}$ and an output set $\mathcal{Y}$ (for us $\mathcal{A} = \mathcal{Y}$)

$$\begin{array}{rcl} \ell : \mathcal{Y} \times \mathcal{Y} & \longrightarrow & \mathbb{R} \\ (a, y) & \longmapsto & \ell(a, y) \end{array}$$

  $\rightarrow$ here think $a \in \mathcal{A}$ as: $a = h(\mathbf{x}, \boldsymbol{\theta})$
  $\rightarrow$ ex: the square loss writes $\ell(h(\mathbf{x}, \boldsymbol{\theta}), y) = (h(\mathbf{x}, \boldsymbol{\theta}) - y)^2$

- **Towards Contrast functions:** For instance in the case of the square loss, we define the associated "contrast" function as

$$\ell(h(\mathbf{x}, \boldsymbol{\theta}), y) = (h(\mathbf{x}, \boldsymbol{\theta}) - y)^2 = \Psi(h(\cdot, \boldsymbol{\theta}), (\mathbf{x}, y))$$

- **Definition:** Denote by $\mathcal{F}$ some feature space, a contrast Ψ is defined as

$$\begin{array}{rcl} \Psi : \mathcal{F} & \longrightarrow & L_1(Q^{\mathbf{z}}) \\ \rho & \longmapsto & \Psi(\rho, \cdot) : (\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y} \longmapsto \Psi(\rho, (\mathbf{x}, y)) \end{array}$$

  In the example before, if we consider $\mathcal{F} = \{\rho : \mathcal{X} \rightarrow \mathcal{Y}, \|\rho\|_{L_2(P^{\mathbf{x}})} < \infty\}$, we may define $F = \{\rho_{\boldsymbol{\theta}} : \mathbf{x} \mapsto h(\mathbf{x}, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$.
  We will call F as (computer code based) **Model**

AIRBUS GROUP

From Loss function to Contrast function

- **Loss function:** Given an action set $\mathcal{A}$ and an output set $\mathcal{Y}$ (for us $\mathcal{A} = \mathcal{Y}$)

$$\begin{array}{rcl} \ell : \mathcal{Y} \times \mathcal{Y} & \longrightarrow & \mathbb{R} \\ (a, y) & \longmapsto & \ell(a, y) \end{array}$$

  $\rightarrow$ here think $a \in \mathcal{A}$ as: $a = h(\mathbf{x}, \boldsymbol{\theta})$
  $\rightarrow$ ex: the square loss writes $\ell(h(\mathbf{x}, \boldsymbol{\theta}), y) = (h(\mathbf{x}, \boldsymbol{\theta}) - y)^2$

- **Towards Contrast functions:** For instance in the case of the square loss, we define the associated "contrast" function as

$$\ell(h(\mathbf{x}, \boldsymbol{\theta}), y) = (h(\mathbf{x}, \boldsymbol{\theta}) - y)^2 = \Psi(h(\cdot, \boldsymbol{\theta}), (\mathbf{x}, y))$$

- **Definition:** Denote by $\mathcal{F}$ some feature space, a contrast Ψ is defined as

$$\begin{array}{rcl} \Psi : \mathcal{F} & \longrightarrow & L_1(Q^{\mathbf{z}}) \\ \rho & \longmapsto & \Psi(\rho, \cdot) : (\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y} \longmapsto \Psi(\rho, (\mathbf{x}, y)) \end{array}$$

  In the example before, if we consider $\mathcal{F} = \{\rho : \mathcal{X} \to \mathcal{Y}, \|\rho\|_{L_2(P^{\mathbf{x}})} < \infty\}$, we may define $F = \{\rho_{\boldsymbol{\theta}} : \mathbf{x} \mapsto h(\mathbf{x}, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$.
  We will call F as (computer code based) **Model**

The contrast function emphasizes the quantity of interest in $\mathcal{F}$ involved

AIRBUS GROUP

Notion of Risk

- **Risk with Loss function:**

$$\mathcal{R}(f) = \mathbb{E}[\ell(f(X), Y)] \underset{e.g}{=} \mathbb{E}(f(X) - Y)^2$$

- **Risk with Contrast function, Ψ-Risk** :

$$\mathcal{R}_\Psi(\rho) := \mathbb{E}\,\Psi\,(\rho\,,\,(\mathbf{X}, Y))$$

- **Target** :

$$\rho^* = \underset{\rho \in \mathcal{F}}{\text{Argmin}}\,\mathcal{R}_\Psi(\rho) \quad \text{if it exists}$$

- **Interpretation:**
  In Computer Experiments framework, the "target" defined before will be the "quantity of interest" (QoI) depending on the contrast considered

Examples of contrasts and associated QoI

(Abuse of notation: we will write $\Psi(\rho, y)$ a contrast function which does not depend on the joint data $(\mathbf{x}, y)$)

- $\mathcal{F} = \{\rho : \mathcal{X} \to \mathcal{Y}, \, \|\rho\|_{L_2(P^{\mathbf{x}})} < \infty\}$

  regression-contrast: $\Psi(\rho, (\mathbf{x}, y)) = (y - \rho(\mathbf{x}))^2 \quad \to \rho^*(\cdot) = \mathbb{E}(Y|\mathbf{X} = \cdot)$

- $\mathcal{F} = \mathbb{R}$

  mean-contrast: $\Psi(\rho, y) = (y - \rho)^2 \quad \to \rho^* = \mathbb{E}(Y)$

  prob-contrast: $\Psi(\rho, y) = (\mathbb{1}_{y \geq s} - \rho)^2 \quad \to \rho^* = \mathbb{P}(Y \geq s)$

  (α)quantile-contrast: $\Psi(\rho, y) = (y - \rho)(\alpha - \mathbf{1}_{y \leq \rho}) \quad \to \rho^* = q_\alpha(Y)$

- $\mathcal{F} = \{\text{density functions on } \mathcal{Y}\}$

  (log)pdf-contrast: $\Psi(\rho, y) = -\log \rho(y) \quad \to \rho^* = pdf_Y$

  (L_2)pdf-contrast: $\Psi(\rho, y) = ||\rho||_2^2 - 2\,\rho(y) \quad \to \rho^* = pdf_Y$

AIRBUS GROUP

Examples of contrasts and associated QoI

(Abuse of notation: we will write $\Psi(\rho, y)$ a contrast function which does not depend on the joint data $(\mathbf{x}, y)$)

- $\mathcal{F} = \{\rho : \mathcal{X} \to \mathcal{Y}, \|\rho\|_{L_2(P^x)} < \infty\}$

  regression-contrast: $\Psi(\rho, (\mathbf{x}, y)) = (y - \rho(\mathbf{x}))^2 \quad \to \rho^*(\cdot) = \mathbb{E}(Y|\mathbf{X} = \cdot)$

- $\mathcal{F} = \mathbb{R}$

  mean-contrast: $\Psi(\rho, y) = (y - \rho)^2 \quad \to \rho^* = \mathbb{E}(Y)$

  prob-contrast: $\Psi(\rho, y) = (1\!\!1_{y \geq s} - \rho)^2 \quad \to \rho^* = \mathbb{P}(Y \geq s)$

  (α)quantile-contrast: $\Psi(\rho, y) = (y - \rho)(\alpha - \mathbf{1}_{y \leq \rho}) \quad \to \rho^* = q_\alpha(Y)$

- $\mathcal{F} = \{\text{density functions on } \mathcal{Y}\}$

  (log)pdf-contrast: $\Psi(\rho, y) = -\log \rho(y) \quad \to \rho^* = pdf_Y$

  (L_2)pdf-contrast: $\Psi(\rho, y) = ||\rho||_2^2 - 2\,\rho(y) \quad \to \rho^* = pdf_Y$

- **In practice** we define a model $F \subset \mathcal{F}$ based on a code $h(\mathbf{X}, \boldsymbol{\theta})$ where $\mathbf{X} \sim P^x$

  $F = \{\rho_{\boldsymbol{\theta}} : \mathbf{x} \mapsto h(\mathbf{x}, \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\} \subset \{\rho : \mathcal{X} \to \mathcal{Y}\}$
  $F = \{\rho_{\boldsymbol{\theta}} = \text{pdf of } h(\mathbf{X}, \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\} \subset \{\text{density functions on } \mathcal{Y}\}$
  Etc.

$$\boxed{F = \{\rho(\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\}}$$

AIRBUS GROUP

Empirical Risk Minimisation

Given a set $Y_1, ..., Y_n$ and a simulation code $h(\mathbf{X}, \boldsymbol{\theta})$, with $\mathbf{X} \sim P^x$.
Consider a contrast $\Psi : \mathcal{F} \to L_1(Q)$ (i.e contrasts only the data y)
and a Model $F = \{\rho(\boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$ provided by the simulation code

- **Goal**: estimate the parameter

$$\boldsymbol{\theta}^*_\Psi = \underset{\boldsymbol{\theta} \in \Theta}{\text{Argmin}}\, \mathcal{R}_\Psi(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta} \in \Theta}{\text{Argmin}}\, \mathbb{E}\, \Psi\left(\rho(\boldsymbol{\theta}),\ Y\right)$$

Empirical Risk Minimisation

Given a set $Y_1, ..., Y_n$ and a simulation code $h(\mathbf{X}, \boldsymbol{\theta})$, with $\mathbf{X} \sim P^{\mathrm{x}}$.
Consider a contrast $\Psi : \mathcal{F} \to L_1(Q)$ (i.e contrasts only the data y)
and a Model $F = \{\rho(\boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$ provided by the simulation code

- **Goal**: estimate the parameter

$$\boldsymbol{\theta}^*_{\Psi} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{Argmin}} \, \mathcal{R}_{\Psi}(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{Argmin}} \, \mathbb{E} \, \Psi\left(\rho(\boldsymbol{\theta}), \, Y\right)$$

But Q is unknown and $\rho(\boldsymbol{\theta})$ is not analytically tractable !

AIRBUS GROUP

Empirical Risk Minimisation

Given a set $Y_1, ..., Y_n$ and a simulation code $h(\mathbf{X}, \boldsymbol{\theta})$, with $\mathbf{X} \sim P^x$.
Consider a contrast $\Psi : \mathcal{F} \to L_1(Q)$ (i.e contrasts only the data y)
and a Model $F = \{\rho(\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$ provided by the simulation code

- **Goal**: estimate the parameter

$$\boldsymbol{\theta}^*_\Psi = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathcal{R}_\Psi(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\ Y\right)$$

  But Q is unknown and $\rho(\boldsymbol{\theta})$ is not analytically tractable !

- **Risk "Empirization"**:

$$\mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\ Y\right) \quad \Rrightarrow \quad \frac{1}{n}\sum_{i=1}^{n} \Psi(\rho^m(\boldsymbol{\theta}), Y_i)$$

  where $\rho^m(\boldsymbol{\theta})$ is a kernel estimate of $\rho(\boldsymbol{\theta})$

$$\rho^m(\boldsymbol{\theta}) = \frac{1}{m}\sum_{j=1}^{m} \kappa(h(\mathbf{X}_j, \boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\,.$$

AIRBUS GROUP

Empirical Risk Minimisation

Given a set $Y_1, ..., Y_n$ and a simulation code $h(\mathbf{X}, \boldsymbol{\theta})$, with $\mathbf{X} \sim P^x$.
Consider a contrast $\Psi : \mathcal{F} \to L_1(Q)$ (i.e contrasts only the data y)
and a Model $F = \{\rho(\boldsymbol{\theta}),\, \boldsymbol{\theta} \in \Theta\} \subset \mathcal{F}$ provided by the simulation code

- **Goal**: estimate the parameter

$$\boldsymbol{\theta}^*_\Psi = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathcal{R}_\Psi(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\, Y\right)$$

  But Q is unknown and $\rho(\boldsymbol{\theta})$ is not analytically tractable !

- **Risk "Empirization"**:

$$\mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\, Y\right) \quad \Rightarrow \quad \frac{1}{n}\sum_{i=1}^{n} \Psi(\rho^m(\boldsymbol{\theta}), Y_i)$$

  where $\rho^m(\boldsymbol{\theta})$ is a kernel estimate of $\rho(\boldsymbol{\theta})$

$$\rho^m(\boldsymbol{\theta}) = \frac{1}{m}\sum_{j=1}^{m} \kappa(h(\mathbf{X}_j, \boldsymbol{\theta})), \quad \mathbf{X}_j \underset{i.i.d}{\sim} P^x\,.$$

  Example: $\mathcal{F} =$ "means", $\kappa(y) = y$
  $\mathcal{F} =$ "densities", $\kappa(y)(\cdot) = \frac{1}{\sqrt{2\,\pi}\,b}\exp((y-\cdot)^2/2\,b^2)$
  etc...

AIRBUS GROUP

Ψ-Estimator

- **Generic Ψ-estimator**:

$$\boxed{\widehat{\boldsymbol{\theta}}_{\Psi} = \operatorname*{Argmin}_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{n} \Psi(\rho^m(\boldsymbol{\theta}), Y_i)}$$

- **Examples**:

▶ mean-contrast Ψ_{mean}, $\widehat{\boldsymbol{\theta}}_{mean} = \operatorname*{Argmin}_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} (Y_i - h(\mathbf{X}_j, \boldsymbol{\theta})) \right)^2$

▶ log-contrast $\Psi_{\log}$, $\widehat{\boldsymbol{\theta}}_{\log} = \operatorname*{Argmin}_{\boldsymbol{\theta}\in\Theta} - \sum_{i=1}^{n} \log \left(\sum_{j=1}^{m} K_b(Y_i - h(\mathbf{X}_j, \boldsymbol{\theta})) \right)$

▶ L_2-contrast Ψ_{L_2}

$$\widehat{\boldsymbol{\theta}}_{L_2} = \operatorname*{Argmin}_{\boldsymbol{\theta}\in\Theta} \left\{ \left\| \sum_{j=1}^{m} K_b(\,\cdot\, - h(\mathbf{X}_j, \boldsymbol{\theta})) \right\|_2^2 - \frac{2\,m}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} K_b(\,Y_i\, - h(\mathbf{X}_j, \boldsymbol{\theta})) \right\}$$

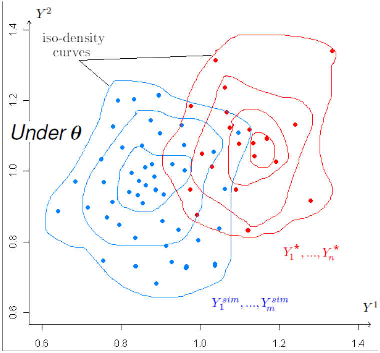
▶ Etc...

AIRBUS GROUP

Contrast minimization: "the way of minimizing"

$$\widehat{\boldsymbol{\theta}}_{\Psi} \quad \overset{plug}{\hookrightarrow} \quad \mathbf{X} \mapsto h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{\Psi}),\ \mathbf{X} \sim P^{\mathbf{x}}$$

- In blue: Simulated data
- In red: Reference data

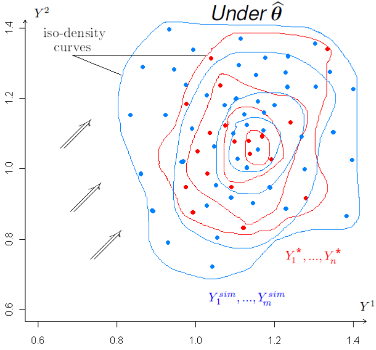


Contrast minimization: "the way of minimizing"

$$\widehat{\boldsymbol{\theta}}_{\Psi} \quad \overset{plug}{\hookrightarrow} \quad \mathbf{X} \mapsto h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{\Psi}),\ \mathbf{X} \sim P^{\mathbf{x}}$$

- In blue: Simulated data
- In red: Reference data

Depending on the contrast used ...



Results

N. Rachdi, J-C. Fort and T. Klein, *Risk bounds for new M-estimation problems*, ESAIM : Probability & Statistics, Volume 17 (2013), p. 740–766

Consistency Results

Under regularity and tightness conditions, in probability

$$\widehat{\boldsymbol{\theta}}_{\Psi}^{n,m} \underset{n,m\to\infty}{\Longrightarrow} \boldsymbol{\theta}_{\Psi}^{*} = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{Argmin}}\, \mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\; Y\right)$$

Results

N. Rachdi, J-C. Fort and T. Klein, *Risk bounds for new M-estimation problems*, ESAIM : Probability & Statistics, Volume 17 (2013), p. 740–766

Consistency Results

Under regularity and tightness conditions, in probability

$$\widehat{\boldsymbol{\theta}}_{\Psi}^{n,m} \underset{n,m\to\infty}{\Longrightarrow} \boldsymbol{\theta}_{\Psi}^{*} = \underset{\boldsymbol{\theta}\in\Theta}{\text{Argmin}}\, \mathbb{E}\,\Psi\left(\rho(\boldsymbol{\theta})\,,\, Y\right)$$

Central Limit Theorem in progress ...

Back to the questions in Introduction ...

- **Inverse problem:** If the $\mathbf{X}_i^*$'s are not observed ? How to calibrate ? (e.g input code $\neq$ experimental conditions etc...)
- **Prediction problem:** even if they are observed, should we always use regression models for predicting some quantity of interest ?
- for instance, what is the meaning of

$$\mathbb{P}(h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg}) > s) \quad \text{or} \quad pdf_{h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg})} \text{ ?}$$

where $\widehat{\boldsymbol{\theta}}_{reg}$ is the mean-squares estimator of the model $Y_i = h(\mathbf{X}^i, \theta) + \varepsilon_i$

... "*duality*" between estimation procedure and target prediction

Back to the questions in Introduction ...

- **Inverse problem:** If the $\mathbf{X}_i^*$'s are not observed ? How to calibrate ?
  (e.g input code $\neq$ experimental conditions etc...)
- **Prediction problem:** even if they are observed, should we always use regression models for predicting some quantity of interest ?
- for instance, what is the meaning of

  $$\mathbb{P}(h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg}) > s) \quad \text{or} \quad pdf_{h(\mathbf{X}, \widehat{\boldsymbol{\theta}}_{reg})} \; ?$$

  where $\widehat{\boldsymbol{\theta}}_{reg}$ is the mean-squares estimator of the model $Y_i = h(\mathbf{X}^i, \theta) + \varepsilon_i$

  ... "*duality*" between estimation procedure and target prediction

Now, we have tools to study that questions ...

Back to the introductive example

- From $(\mathbf{X_1}, Y_1), ..., (\mathbf{X}_n, Y_n)$ and a computer code $h(\mathbf{X}, \boldsymbol{\theta})$, we built the density predictions $\widehat{f}_{MS}$, $\widehat{f}_M$, $\widehat{f}_{KL}$

AIRBUS GROUP

Back to the introductive example

- From $(\mathbf{X}_1, Y_1), \ldots, (\mathbf{X}_n, Y_n)$ and a computer code $h(\mathbf{X}, \boldsymbol{\theta})$, we built the density predictions $\widehat{f}_{MS}$, $\widehat{f}_{M}$, $\widehat{f}_{KL}$

- We used 3 contrasts (Ψ_{reg}, Ψ_{mean} and Ψ_{pdf}) for parameter estimation

$$Y = \sin(X) + 0.01\varepsilon, \quad (X, \varepsilon \sim \mathcal{N}(0,1) \quad iid)$$

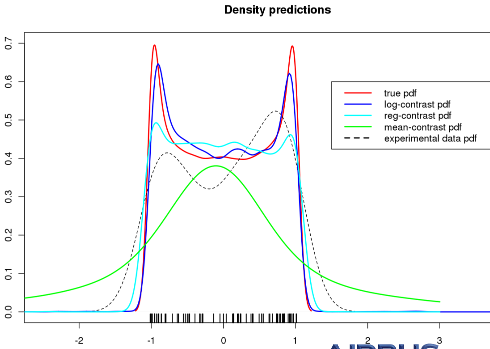
→ **Quantity of Interest:** pdf(Y)

→ **model:** $h(x, \boldsymbol{\theta}) = \theta_1 + \theta_2\, x + \theta_3\, x^3$

→ **data:** $n = 50$ $((X_i, Y_i))$, $m = 10^3$ (X_j)

→ **output simulations** under $X_j \sim P^x = \mathcal{N}(0,1)$ using

- $h(\cdot, \widehat{\boldsymbol{\theta}}_{pdf})$ (blue)
- $h(\cdot, \widehat{\boldsymbol{\theta}}_{reg})$ (cyan)
- $h(\cdot, \widehat{\boldsymbol{\theta}}_{mean})$ (green)



AIRBUS GROUP

Back to the introductive example

- From $(\mathbf{X}_1, Y_1), \ldots, (\mathbf{X}_n, Y_n)$ and a computer code $h(\mathbf{X}, \boldsymbol{\theta})$, we built the density predictions $\widehat{f}_{MS}$, $\widehat{f}_M$, $\widehat{f}_{KL}$
- We used 3 contrasts (Ψ_{reg}, Ψ_{mean} and Ψ_{pdf}) for parameter estimation

$$Y = \sin(X) + 0.01\varepsilon, \quad (X, \varepsilon \sim \mathcal{N}(0,1) \quad iid)$$

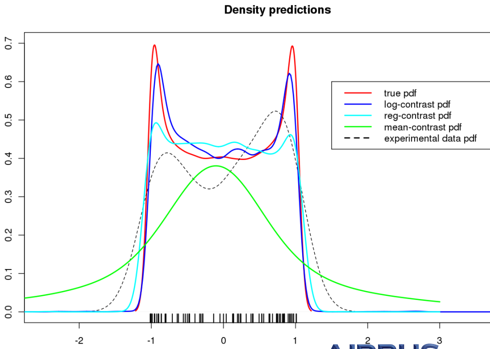
→ **Quantity of Interest**: pdf(Y)

→ **model**: $h(x, \boldsymbol{\theta}) = \theta_1 + \theta_2\, x + \theta_3\, x^3$

→ **data**: $n = 50$ $((X_i, Y_i))$, $m = 10^3$ (X_j)

→ **output simulations** under $X_j \sim P^x = \mathcal{N}(0,1)$ using

- $h(\cdot, \widehat{\boldsymbol{\theta}}_{pdf})$ (blue)
- $h(\cdot, \widehat{\boldsymbol{\theta}}_{reg})$ (cyan)
- $h(\cdot, \widehat{\boldsymbol{\theta}}_{mean})$ (green)



AIRBUS GROUP

Other example: Conditional expectation

- $Y = \sin(X) + \varepsilon,\ X, \varepsilon \sim \mathcal{N}(0,1)$ independents
- $h(X, \boldsymbol{\theta}) = \theta_{\mathbf{1}} + \theta_{\mathbf{2}}\, X + \theta_{\mathbf{3}}\, X^{\mathbf{3}},\ X \sim P^{x} = \mathcal{N}(0,1)$
- $n = 50$ $((X_i, Y_i))$ and $m = 10^{\mathbf{3}}$ (X_j)

QoI: $\rho^* = \mathbb{E}(Y/X = \cdot)(= \sin(\cdot))$
--→ $\Psi^* = \Psi_{reg}$ ("adapted" contrast)

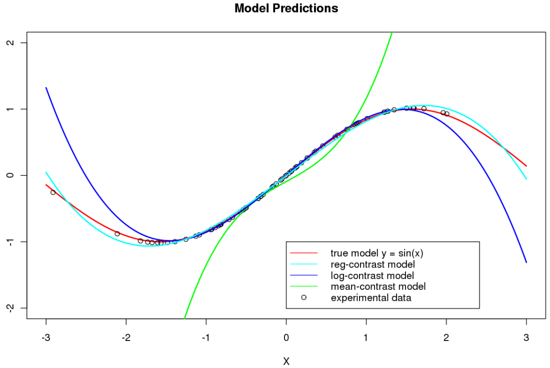
Other example: Conditional expectation

- $Y = \sin(X) + \varepsilon,\ X, \varepsilon \sim \mathcal{N}(0,1)$ independents
- $h(X, \boldsymbol{\theta}) = \theta_1 + \theta_2\, X + \theta_3\, X^3,\ X \sim P^x = \mathcal{N}(0,1)$
- $n = 50$ $((X_i, Y_i))$ and $m = 10^3$ (X_j)

QoI: $\rho^* = \mathbb{E}(Y/X = \cdot)(= \sin(\cdot))$
$\dashrightarrow \Psi^* = \Psi_{reg}$ ("adapted" contrast)

■ **consider 3 Ψ-estimators:** $\widehat{\boldsymbol{\theta}}_{\Psi^*}$, $\widehat{\boldsymbol{\theta}}_{\Psi_{\mathbf{log}}}$ and $\widehat{\boldsymbol{\theta}}_{\Psi_{mean}}$

	$\widehat{\boldsymbol{\theta}}_\Psi$	$\mathcal{R}_{\Psi^*}(\widehat{\boldsymbol{\theta}}_\Psi)$
$\Psi = \Psi^*$	$(-0.0049, 0.9259, -0.1048)$	0.064
$\Psi = \Psi_{\mathbf{log}}$	$(0.0057, 1.025, -0.163)$	0.36
$\Psi = \Psi_{mean}$	$(-0.0924, 0.6607, 0.5965)$	6.18



AIRBUS GROUP

Thank you for your attention !

N. Rachdi, J-C Fort, T. Klein (2013), Risk bounds for new M-estimation problems, ESAIM : Probability & Statistics - doi: 10.1051/ps/2012025

N. Rachdi, J-C Fort, T. Klein (2012), Stochastic Inverse Problem with Noisy Simulator, Ann. Fac. Sci. Toulouse, S. 6, 21 no. 3 (2012), p. 593-622

J-C Fort, T. Klein, N. Rachdi (2014), New sensitivity indices subordinated to a contrast, Communication in Statistics: Theory and Methods, to appear

N. Rachdi (2011), Statistical Learning and Computer Experiments, PhD thesis from University Paul Sabatier of Toulouse.

F. Mangeant (2011), Joined initiative around uncertainty management, Annals of Telecommunications.

E. de Rocquigny, N. Devictor, S. Tarantola - Eds (2008), Uncertainty in Industrial Practice, Wiley Verlag.

T. Hastie, R. Tibshirani, J. Friedman (2008) - The elements of statistical learning - Springer Series in Statistics.

P. Massart (2007), Concentration inequalities and model selection: Ecole d'Ete de Probabilites de Saint-Flour XXXIII-2003. Springer Verlag.

AIRBUS GROUP