Module #11
Estimation bayesienne des paramètres $f(\theta|\mathcal{D})$

(CIV8530 - Fiabilité des structures et systèmes)

Enseignant: James-A. Goulet

Département des génies civil, géologique et des mines
Polytechnique Montréal

Chapitre 7 - A. Der Kiureghian, (2015), *Probabilistic Methods for Engineering Risk Analysis*. Department of Civil and Environmental Engineering, UC Berkeley

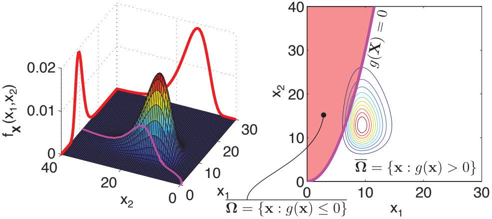
Probabilité de défaillance

$\mathbf{X} = \{X_1, X_2, \cdots, X_n\}$: ensemble de variables aléatoires

$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta_f})$: densité de probabilité conjointe

$\boldsymbol{\theta_f}$: Paramètres de $f_{\mathbf{X}}(\mathbf{x})$

$g(\mathbf{X})$: fonction d'état limite



Comment estimer la densité de probabilité $f(\theta)$ à l'aide de mesures?

Plan du module #11

Organisation de la matière

Théorie probabilité
- 0 Révision algèbre & probabilités
- 1 Lois de probabilités
- 3 PDF multivariés

Fiabilité des structures & systèmes
- Introduction
- 2 Formulation fiabilité composantes $g_1(\mathbf{X})$
- 9 Formulation fiabilité systèmes

Estimation – p_f échantillonnage
- 4 Monte Carlo
- 12 Échantillonnage par importance

Estimation – p_f analytique
- 5 FOSM
- 6 FORM
- 8 SORM

Utilisation des résultats
- 7 Analyse de sensibilité
- 10 Incertitudes aléatoires & épistémiques
- 11 Données empiriques
- 13 Prise de décision
- 14 Métamodèles & modèles empiriques

Plan de la section

Catégorisation des incertitudes

Les incertitudes sont habituellement regroupées en deux catégories:

Aléatoires (i.e. variabilité intrinsèque): Les incertitudes aléatoires sont considérées comme étant irréductibles.
Davantage d'informations, de mesures, ou d'observations <u>ne peuvent</u> les réduire.

Épistémiques (i.e. manque de connaissances): Les incertitudes épistémiques sont associées à un manque de connaissances.
Davantage d'informations, de mesures, ou d'observations peuvent les réduire.

Les variables X sont habituellement affectées par les deux types d'incertitudes.

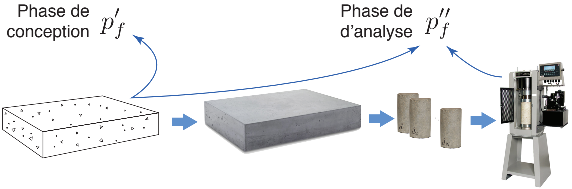
Évolution temporelle des incertitudes

Phase de conception: Les variables sont considérées comme étant aléatoires.

- Variabilité intrinsèque
- Incertitude quant à au processus de construction

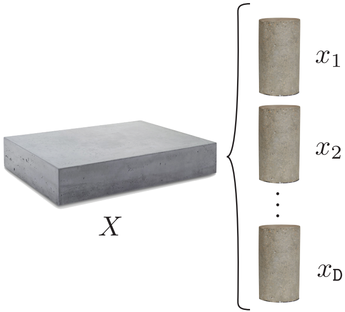
Phase d'analyse (structure existante): Les variables sont habituellement affectées par les deux types d'incertitudes.

- Variabilité intrinsèque
- Manque de connaissance quant au processus de construction



Plan de la section

Observations de quantités incertaines



- X: variable aléatoire, $f_X(x;\boldsymbol{\theta})$
- $\boldsymbol{\theta}$: Paramètres du PDF de X
- $\{x_1, x_2, \cdots, x_n\}$: réalisations de X

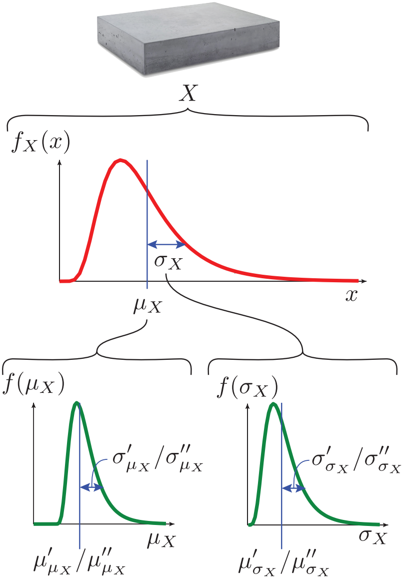
Que fait-on lorsque les paramètres θ sont inconnus?

X : Variable aléatoire

$\underbrace{\mu_X, \sigma_X}_{\boldsymbol{\theta}}$: Paramètres de $f_X(x)$

$\mu_{\mu_X}, \sigma_{\mu_X}$: hyper-paramètres de μ_X

$\mu_{\sigma_X}, \sigma_{\sigma_X}$: hyper-paramètres de σ_X



Qualificatifs des observations

On cherche à obtenir $f(\boldsymbol{\theta}|\mathcal{D})$, $\mathcal{D} = \{y_1, y_2, \cdots, y_n\}$

Il y a trois qualificatifs pour les observations (Y)

1. Types:
   - **Directes** : $Y = X$
   - **Indirectes**: $Y = h(X)$
2. Incertitudes:
   - **Exactes** : $Y = X$
   - **Inexactes**: $\underbrace{Y = X + V}_{\text{additive}}, \quad \underbrace{Y = X \cdot V}_{\text{multiplicative}}$
3. Bornes:
   - **Supérieure** : $Y \geq X$
   - **Inférieure**: $Y \leq X$
   - $\emptyset$: $Y = X$

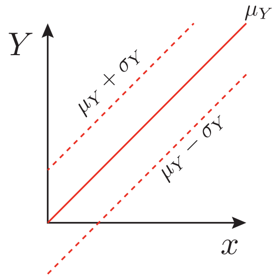
Incertitudes d'observations

Il y a deux principaux modèles d'incertitudes pour observations **inexactes**

Erreurs additives

$$Y = X + V : \{X, V, Y\} \sim \mathcal{N}(\cdot,\cdot)$$

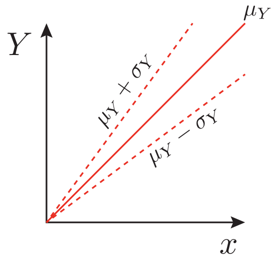
$$V \sim \mathcal{N}(0, \sigma_V)$$



Erreurs multiplicatives

$$Y = X \cdot V : \{X, V, Y\} \sim \ln\mathcal{N}(\cdot,\cdot)$$

$$V \sim \ln\mathcal{N}(0, \sigma_{\ln V}) \approx \mathcal{N}(1, \delta_V)$$



Plan de la section

Bayes

Notation

$f_X(x) \equiv f(x;\boldsymbol{\theta})$: Densité de probabilité d'une variable

$f_{\mathbf{Y}}(\mathbf{y})$: Distribution conjointe des observations $\mathbf{Y} = [Y_1, \cdots, Y_n]^\intercal$

$$\underbrace{Y_i = X + V_i}_{\text{additive}} \quad | \quad \underbrace{Y_i = X \cdot V_i}_{\text{multiplicative}}$$

où V est une variable aléatoire représentant l'erreur de mesure.
Si la mesure est *non biaisé* :

$$\underbrace{V_i \sim \mathcal{N}(0, \sigma_V^2), V_i \perp\!\!\!\perp V_j, \forall i \neq j}_{\text{additive}} \quad | \quad \underbrace{V_i \sim \ln\mathcal{N}(0, \sigma_{\ln V}), V_i \perp\!\!\!\perp V_j, \forall i \neq j}_{\text{multiplicative}}$$

On cherche à obtenir $f(x|\mathcal{D})$ (**"PDF prédictif à postériori"**) à partir de $f(\boldsymbol{\theta}|\mathcal{D})$ (**"PDF à postériori"**)

$$\boxed{f''(x) \equiv f(x|\mathcal{D}) = \int f(x;\boldsymbol{\theta}) f(\boldsymbol{\theta}|\mathcal{D}) d\boldsymbol{\theta}}$$

Théorème de Bayes

$f(\boldsymbol{\theta}|\mathcal{D})$ est obtenue à partir de la densité conjointe $f(\boldsymbol{\theta},\mathcal{D})$

$$f(\boldsymbol{\theta},\mathcal{D}) = f(\boldsymbol{\theta}|\mathcal{D})f(\mathcal{D}) = f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})$$

$$\boxed{f''(\boldsymbol{\theta}) \equiv f(\boldsymbol{\theta}|\mathcal{D}) = \frac{f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathcal{D})}}$$ **Théorème de Bayes**

$f''(\boldsymbol{\theta}) \equiv f(\boldsymbol{\theta}|\mathcal{D})$: Densité de probabilité à posteriori

$f'(\boldsymbol{\theta}) \equiv f(\boldsymbol{\theta})$: Densité de probabilité à priori

$f(\mathcal{D}|\boldsymbol{\theta}) \equiv \mathcal{L}(\boldsymbol{\theta}|\mathcal{D})$: Fonction de vraisemblance

$f(\mathcal{D}) = \Pr(\mathcal{D})$: Constante de normalisation

$$\int f(\boldsymbol{\theta}|\mathcal{D})d\boldsymbol{\theta} \equiv 1, \quad \text{donc} \quad f(\mathcal{D}) = \int f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta}$$

$f(\boldsymbol{\theta}) \equiv f'(\boldsymbol{\theta})$: Densité de probabilité à priori

Connaissances subjectives: Connaissances à priori du problème, contraintes physiques, jugement d'expert, etc.

Connaissances diffuses: i.e. probabilité égale pour toutes les possibilités. e.g.: $X \sim \mathcal{N}(\mu, \sigma^2)$, $f'(\mu, \sigma) = 1$

Apprentissage itératif: La densité de probabilité à postériori à l'étape $i-1$ correspond à la connaissance à priori à l'étape i
e.g.: $f'_i(\boldsymbol{\theta}) = f''_{i-1}(\boldsymbol{\theta})$

$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance

$f(\mathcal{D}|\boldsymbol{\theta}) \equiv f(\mathbf{Y}=\mathbf{y}|\boldsymbol{\theta}) \equiv f(\mathbf{y}|\boldsymbol{\theta}) \equiv \mathcal{L}(\boldsymbol{\theta}|\mathbf{y})$: Vraisemblance d'observer y_1 en fonction de $\boldsymbol{\theta}$

Soit une observation $\mathcal{D} = \{y_1\}$ étant une réalisation de $\underbrace{Y = X + V}_{\text{additive}}$, où $V \sim \mathcal{N}(0, \sigma_V^2)$ et $X \sim \mathcal{N}(\underbrace{\mu_X, \sigma_X^2}_{\boldsymbol{\theta}})$

$$f(y|\boldsymbol{\theta}) = \mathcal{N}(\mu_X, \sigma_X^2 + \sigma_V^2)$$

Soit une observation $\mathcal{D} = \{y_1\}$ étant une réalisation de $\overbrace{Y = X \cdot V}^{\text{multiplicative}}$, où $V \sim \ln\mathcal{N}(0, \sigma_{\ln V}^2)$ et $X \sim \ln\mathcal{N}(\underbrace{\mu_{\ln X}, \sigma_{\ln X}^2}_{\boldsymbol{\theta}})$

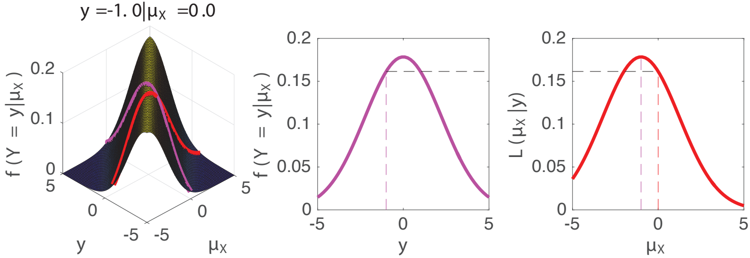
$$f(y|\boldsymbol{\theta}) = \ln\mathcal{N}(\mu_{\ln X}, \sigma_{\ln X}^2 + \sigma_{\ln V}^2)$$

$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $\overbrace{Y = X + V}^{\text{additive}}$.

$$f(x) = \mathcal{N}(\mu_X, 2^2),\ \theta = \mu_X,\ V_i \sim \mathcal{N}(0, 1^2)$$

$$f(y|\mu_X) \equiv \mathcal{L}(\mu_X|y) = \frac{1}{\sqrt{2\pi}\cdot\sqrt{1^2+2^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_X}{\sqrt{1^2+2^2}}\right)^2\right]$$

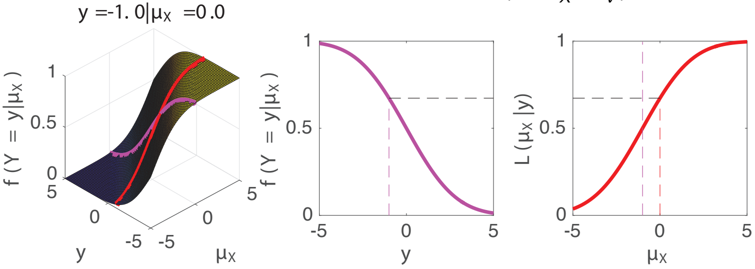


$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance (borne inf.) [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $Y < X + V$.

$$f(x) = \mathcal{N}(\mu_X, 2^2),\ \theta = \mu_X,\ V_i \sim \mathcal{N}(0, 1^2)$$

$$f(X+V > y|\mu_X) = \underbrace{\int_y^{\infty} \frac{1}{\sqrt{2\pi}\cdot\sqrt{1^2+2^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_X}{\sqrt{1^2+2^2}}\right)^2\right] dy}_{1-\text{CDF}\mathcal{N}(\mu_X,\sigma_X^2+\sigma_V^2)}$$

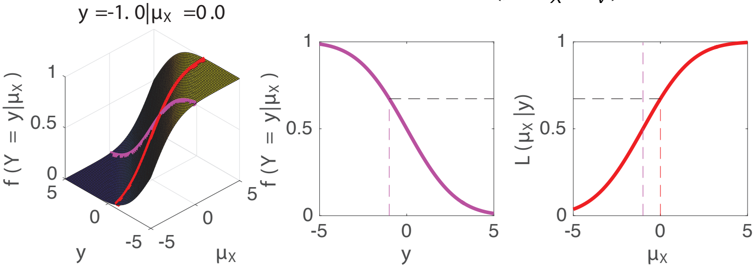


$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance (borne inf.)

Soit une observation $y_1 = -1$ étant une réalisation de $Y < X + V$.

$$f(x) = \mathcal{N}(\mu_X, 2^2),\ \theta = \mu_X,\ V_i \sim \mathcal{N}(0, 1^2)$$

$$f(X + V > y|\mu_X) = \underbrace{1 - \left(\frac{1}{2} + \frac{1}{2}\text{erf}\left[\frac{y - \mu_X}{\sqrt{2}\sqrt{1^2 + 2^2}}\right]\right)}_{1-\text{CDF}\mathcal{N}(\mu_X,\sigma_X^2+\sigma_V^2)}$$

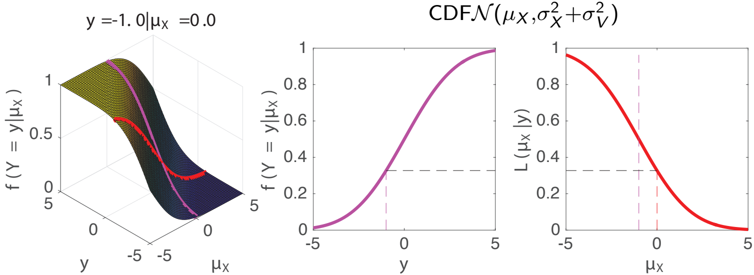


$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance (borne sup.) [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $Y > X + V$.

$$f(x) = \mathcal{N}(\mu_X, 2^2), \;\; \theta = \mu, \;\; V_i \sim \mathcal{N}(0, 1^2)$$

$$f(X+V < y|\mu_X) = \underbrace{\int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\cdot\sqrt{1^2+2^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_X}{\sqrt{1^2+2^2}}\right)^2\right] dy}_{\text{CDF}\mathcal{N}(\mu_X,\sigma_X^2+\sigma_V^2)}$$

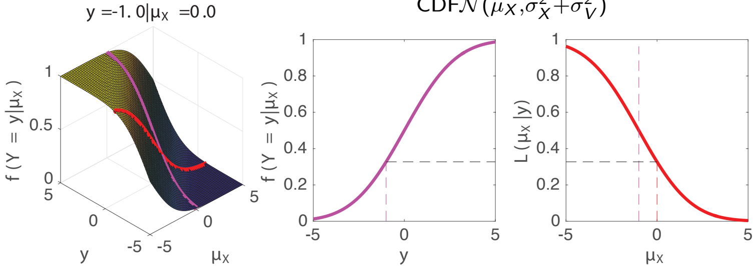


$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance (borne sup.) [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $Y > X + V$.

$$f(x) = \mathcal{N}(\mu_X, 2^2),\ \theta = \mu,\ V_i \sim \mathcal{N}(0, 1^2)$$

$$f(X+V<y|\mu_X) = \underbrace{\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left[\frac{y-\mu_X}{\sqrt{2}\sqrt{1^2+2^2}}\right]}_{\mathrm{CDF}\mathcal{N}(\mu_X, \sigma_X^2+\sigma_V^2)}$$



$f(\mathcal{D}|\boldsymbol{\theta})$: Fonction de vraisemblance

Note 1 – terminologie:

$f(Y=y|\theta)$ est appelé vraisemblance car il n'est pas nécessaire que

$$\int_{\theta} f(Y=y|\theta)=1$$

Note 2 – N observations:

Si $y_i \perp\!\!\!\perp y_j, \forall i \neq j$ (i.e. les observations sont indépendantes)

$$\begin{aligned} f(Y_1=y_1,\cdots,Y_N=y_N|\theta) &= \overbrace{\prod_{i=1}^{N} f(Y_i=y_i|\theta)}^{\approx 0, N\gg 1} \\ &= \exp\left(\sum_{i=1}^{N}\ln(f(Y_i=y_i|\theta))\right) \end{aligned}$$

$f(\mathcal{D})$: Constante de normalisation

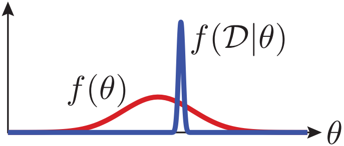
$$\Pr(\mathcal{D}) = f(\mathcal{D}) = \int f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta} = \mathbb{E}_{\boldsymbol{\theta}}[f(\mathcal{D}|\boldsymbol{\theta})]$$

$$\widehat{\Pr(\mathcal{D})}_{\mathrm{MC}} = \frac{1}{m}\sum_{i=1}^{m}\frac{f(\mathcal{D}|\boldsymbol{\theta}_i)\cdot f(\boldsymbol{\theta}_i)}{h(\boldsymbol{\theta}_i)},\ \boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) : \text{densité d'échantillonnage}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\frac{f(\mathcal{D}|\boldsymbol{\theta}_i)\cdot f(\boldsymbol{\theta}_i)}{\frac{1}{b-a}},\ \boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) = \mathcal{U}(a,b)$$

$$= \frac{1}{m}\sum_{i=1}^{m} f(\mathcal{D}|\boldsymbol{\theta}_i),\ \boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})$$

Attention: l'intégration Monte Carlo avec $h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})$ n'est plus une option lorsque N est grand.



$f(\mathcal{D})$: Constante de normalisation [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $Y = X + V$.
$f(x) = \mathcal{N}(\mu_X, 2^2)$, $\theta = \mu_X$, $V \sim \mathcal{N}(0, 1^2)$, $f(\mu_X) = \mathcal{N}(1, 5^2)$

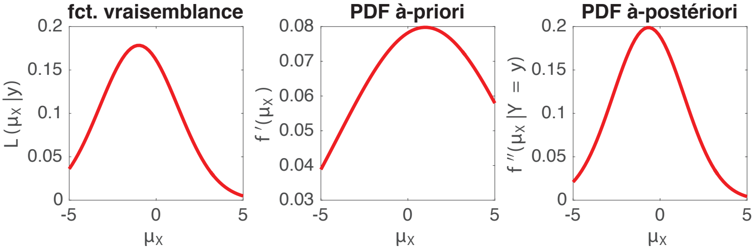
$$\widehat{\Pr(\mathcal{D})}_{\text{MC}} = \frac{1}{m}\sum_{i=1}^{m} f(\mathcal{D}|\boldsymbol{\theta}_i) = 0.0681 \; (m = 10^6)$$

.

$f(\boldsymbol{\theta}|\mathcal{D})$: Probabilité à posteriori [.m]

Soit une observation $y_1 = -1$ étant une réalisation de $Y = X + V$.
$f(x) = \mathcal{N}(\mu_X, 2^2)$, $\theta = \mu_X$, $V \sim \mathcal{N}(0, 1^2), f(\mu_X) = \mathcal{N}(1, 5^2)$

$$f(\boldsymbol{\theta}|\mathcal{D}) = \frac{f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathcal{D})}$$



$f(\boldsymbol{\theta}|\mathcal{D})$: Probabilité à posteriori

$$f(\boldsymbol{\theta}|\mathcal{D}) = \frac{f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathcal{D})}$$

$f(\boldsymbol{\theta}|\mathcal{D}) \rightarrow$ souvent pas de formulation analytique

Solution:

$$f(\boldsymbol{\theta}|\mathcal{D}) \approx \mathcal{N}(\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}], \text{var}[\boldsymbol{\theta}|\mathcal{D}])$$

$$\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}] = \int \boldsymbol{\theta} \cdot f(\boldsymbol{\theta}|\mathcal{D})d\boldsymbol{\theta}$$

$$\text{var}[\boldsymbol{\theta}|\mathcal{D}] = \int (\boldsymbol{\theta} - \mathbb{E}[\boldsymbol{\theta}|\mathcal{D}])^2 \cdot f(\boldsymbol{\theta}|\mathcal{D})d\boldsymbol{\theta}$$

Échantillonnage Monte-Carlo, $f(\boldsymbol{\theta}|\mathcal{D})$

$$\begin{aligned}
\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}] &= \int \boldsymbol{\theta} \cdot f(\boldsymbol{\theta}|\mathcal{D}) d\boldsymbol{\theta} \\
&\approx \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{\theta}_i \cdot \frac{f(\mathcal{D}|\boldsymbol{\theta}_i)\cdot f(\boldsymbol{\theta}_i)}{f(\mathcal{D})} \cdot \frac{1}{h(\boldsymbol{\theta}_i)} \\
&\approx \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{\theta}_i \cdot \frac{f(\mathcal{D}|\boldsymbol{\theta}_i)}{f(\mathcal{D})}, \;\; \boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})
\end{aligned}$$

$$\begin{aligned}
\text{var}[\boldsymbol{\theta}|\mathcal{D}] &= \int (\boldsymbol{\theta} - \mathbb{E}[\boldsymbol{\theta}|\mathcal{D}])^2 \cdot f(\boldsymbol{\theta}|\mathcal{D}) d\boldsymbol{\theta} \\
&\approx \frac{1}{m}\sum_{i=1}^{m} (\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}|\mathcal{D}])^2 \cdot \frac{f(\mathcal{D}|\boldsymbol{\theta}_i)\cdot f(\boldsymbol{\theta}_i)}{f(\mathcal{D})} \cdot \frac{1}{h(\boldsymbol{\theta}_i)} \\
&\approx \frac{1}{m}\sum_{i=1}^{m} (\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}|\mathcal{D}])^2 \cdot \frac{f(\mathcal{D}|\boldsymbol{\theta}_i)}{f(\mathcal{D})}, \;\; \boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})
\end{aligned}$$

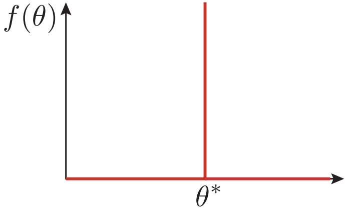
$f(\boldsymbol{\theta}|\mathcal{D})$: Probabilité à posteriori – effet de N

Lorsque le nombre d'observations indépendantes $N \to \infty$

$$f(\boldsymbol{\theta}|\mathcal{D}) = \frac{f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathcal{D})} \to \underbrace{\delta(\overbrace{\boldsymbol{\theta}^*}^{\text{valeur réelle}})}_{\text{Dirac delta PDF}}$$

$$\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}] \to \underbrace{\boldsymbol{\theta}^*}_{\text{valeur réelle}}$$

$$\text{var}[\boldsymbol{\theta}|\mathcal{D}] \to \underbrace{0}_{\text{valeur réelle}}$$



Résumé – Échantillonnage Monte-Carlo

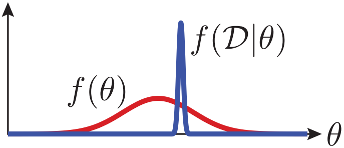
Pour N observations et m échantillons : $\boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})$

$$f(\mathcal{D}) \approx \frac{1}{m}\sum_{i=1}^{m}\left[\prod_{j=1}^{N} f(y_j|\boldsymbol{\theta}_i)\right]$$

$$\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}] \approx \frac{1}{m}\sum_{i=1}^{m}\left[\boldsymbol{\theta}_i \cdot \frac{\prod_{j=1}^{N} f(y_j|\boldsymbol{\theta}_i)}{f(\mathcal{D})}\right]$$

$$\text{var}[\boldsymbol{\theta}|\mathcal{D}] \approx \frac{1}{m}\sum_{i=1}^{m}\left[(\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}|\mathcal{D}])^2 \cdot \frac{\prod_{j=1}^{N} f(y_j|\boldsymbol{\theta}_i)}{f(\mathcal{D})}\right]$$

Attention: l'intégration Monte Carlo avec $h(\boldsymbol{\theta}) = f(\boldsymbol{\theta})$ n'est plus une option lorsque N est grand.



Plan de la section

Exemples

Exemple #11.1 – Estimation des paramètres [ ]

Réalité (inconnue!, $\mu_R = 42\,MPa$, $\sigma_R = 2\,MPa$)

$$R \sim \ln\mathcal{N}(\underbrace{\mu_{\ln R}}_{\mu_R=?}, \underbrace{\sigma_{\ln R}}_{\sigma_R=?})\ \text{MPa}$$

(Note: la résistance (R) du béton à une variabilité intrinsèque)

On cherche:

$$\boxed{\overbrace{\color{red}{f''(\mu_R,\sigma_R)}}^{\text{PDF à postériori}}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

Exemple #11.1 - Nomenclature

$$R \sim \ln\mathcal{N}(\underbrace{\mu_{\ln R}}_{\mu_R=?}, \underbrace{\sigma_{\ln R}}_{\sigma_R=?})\ \text{MPa}$$

R: *Variable aléatoire* associée à la résistance

$\{\mu_R, \sigma_R\}$: *Paramètres* décrivant la densité de probabilité de R, i.e. moyenne et écart-type.

$\{\mu'_{\mu_R}, \sigma'_{\mu_R}, \mu'_{\sigma_R}, \sigma'_{\sigma_R}\}$: *Hyper-paramètres* décrivant les densités de probabilités à priori de μ_R et σ_R, i.e. moyennes et écarts-types.

$\{\mu''_{\mu_R}, \sigma''_{\mu_R}, \mu''_{\sigma_R}, \sigma''_{\sigma_R}\}$: *Hyper-paramètres* décrivant les densités de probabilités à posteriori de μ_R et σ_R, i.e. moyennes et écarts-types.

$$\overbrace{f''(\mu_R,\sigma_R)}^{\text{PDF à postériori}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

Connaissances à priori: $f'(\mu_R,\sigma_R) = f'(\mu_R)f'(\sigma_R)$, $(\mu_R \perp\!\!\!\perp \sigma_R)$

$$\underbrace{f'(\mu_R) = \ln\mathcal{N}(\mu'_{\ln\mu_R},\sigma'_{\ln\mu_R})}_{\text{Informations à priori}}, \quad \underbrace{f'(\sigma_R) = \ln\mathcal{N}(\mu'_{\ln\sigma_R},\sigma'_{\ln\sigma_R})}_{\text{Informations à priori}}$$

$$\mu'_{\mu_R} = 45\,\text{MPa} \qquad \mu'_{\sigma_R} = 5\,\text{MPa}$$
$$\sigma'_{\mu_R} = 7\,\text{MPa} \qquad \sigma'_{\sigma_R} = 2\,\text{MPa}$$

Rappel, densité de probabilité lognormale:

$$\sigma_{\ln X_i} = \sqrt{\ln\left(1+\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)^2\right)}, \quad \mu_{\ln X_i} = \ln\mu_{X_i} - \frac{\sigma^2_{\ln X_i}}{2}$$

$$\overbrace{f''(\mu_R,\sigma_R)}^{\text{PDF à postériori}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

Cas #1: La capacité de la presse > R

Erreur de mesure: $\ln Y = \ln R + V,\ V \sim \mathcal{N}(0, 0.01^2)$ MPa

Fonction de vraisemblance: $f(y|\boldsymbol{\theta}) = \ln\mathcal{N}(\mu_{\ln R}, \sigma^2_{\ln R} + \underbrace{\sigma^2_V}_{=0.01^2})$

$$f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R) = \underbrace{\prod_{i=1}^{N}\frac{1}{y_i\sqrt{2\pi}\sqrt{\sigma^2_{\ln R}+0.01^2}}\exp\left[-\frac{1}{2}\left(\frac{\ln y_i-\mu_{\ln R}}{\sqrt{\sigma^2_{\ln R}+0.01^2}}\right)^2\right]}_{\text{PDF Lognormal}}$$

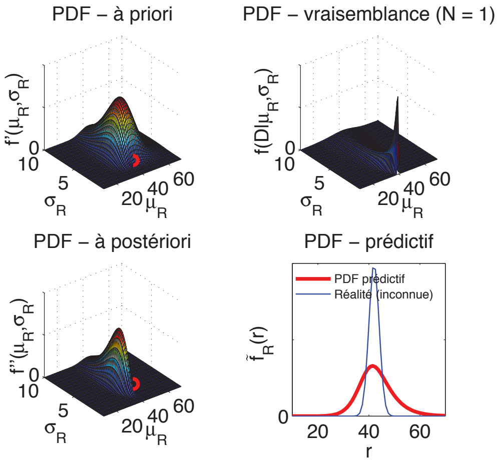
$$\underbrace{f''(\mu_R,\sigma_R)}_{\text{PDF à postériori}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

$$f(Y_1=y_1,\cdots,Y_N=y_N)=\Pr(Y_1=y_1,\cdots,Y_N=y_N)=$$

$$\int_{\mu_R}\int_{\sigma_R} f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)f'(\mu_R,\sigma_R)d\sigma_R d\mu_R$$

Exemple #11.1 - Résultats [ ]

$y_1 = 39.7$ MPa

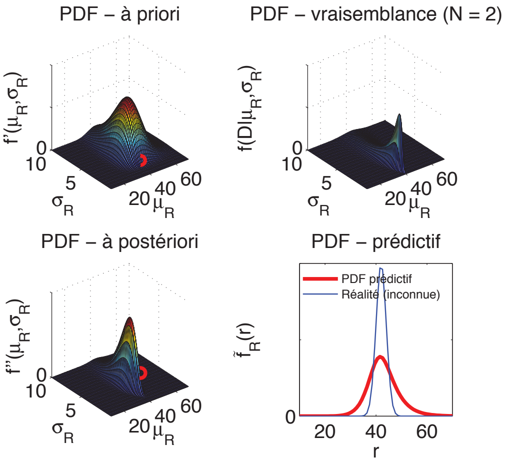


Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Exemple #11.1 - Résultats [.m]

$y_2 = 42.0$ MPa

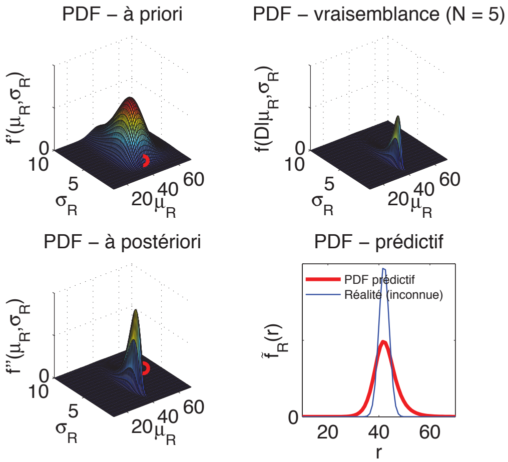


Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Exemple #11.1 - Résultats [ ]

$N = 5$

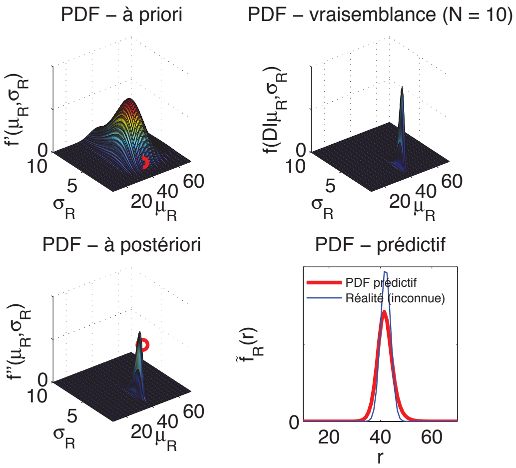


Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Exemple #11.1 - Résultats [.m]

$N = 10$



Densité de probabilité prédictive

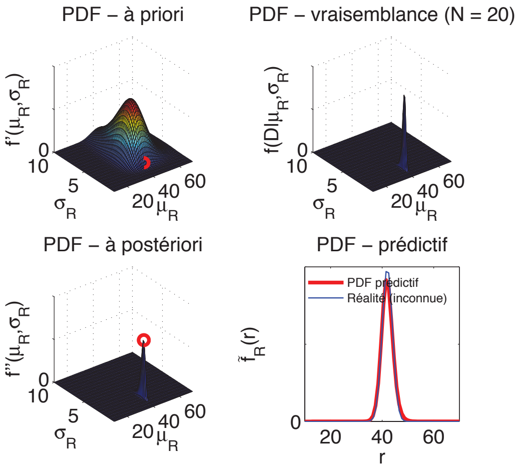
$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Exemple #11.1 - Résultats [.m]

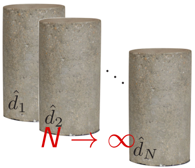
$N = 20$

Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f''(\boldsymbol{\theta}) d\boldsymbol{\theta}$$



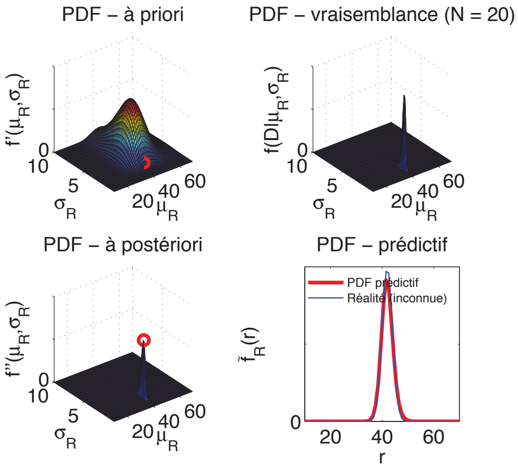
Exemple #11.1 - Résultats [.m]



$$f(\boldsymbol{\theta}|\mathcal{D}) \to \underbrace{\delta(\boldsymbol{\theta}^*)}_{\text{Dirac delta PDF}}$$

$$\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}] \to \underbrace{\boldsymbol{\theta}^*}_{\text{valeur réelle}}$$

$$\tilde{f}_R(r) \to \underbrace{f_R(r)}_{\text{PDF réel}}$$



$$\underbrace{f''(\mu_R,\sigma_R)}^{\text{PDF à postériori}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

Cas #2: La capacité de la presse $= 35\,\text{MPa} < R$

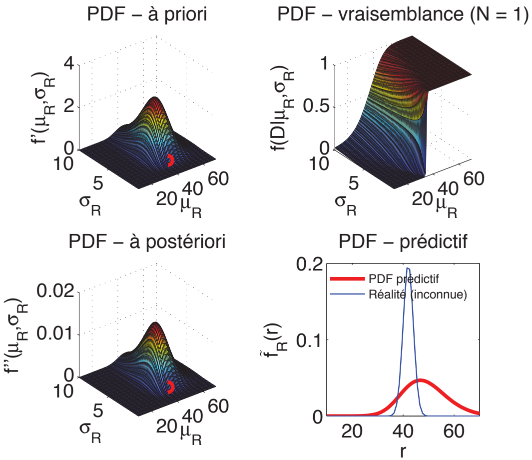
Erreur de mesure: $\ln Y < \ln R + V,\ \ V \sim \mathcal{N}(0, 0.01^2)\,\text{MPa}$

Fonction de vraisemblance: $f(y|\boldsymbol{\theta}) = \ln\mathcal{N}(\mu_{\ln R}, \sigma^2_{\ln R} + \underbrace{\sigma_V^2}_{=0.01^2})$

$$f(Y_1 > y_1,\cdots,Y_N > y_N|\mu_R,\sigma_R) = \prod_{i=1}^{N} 1 - \underbrace{\left[\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{\ln y_i - \mu_{\ln R}}{\sqrt{2}\sqrt{\sigma^2_{\ln R}+0.01^2}}\right)\right]}_{\text{CDF log Normal}}$$

Exemple #11.1 - Résultats [.m]

$R > y_1 = 35$ MPa



Densité de probabilité prédictive

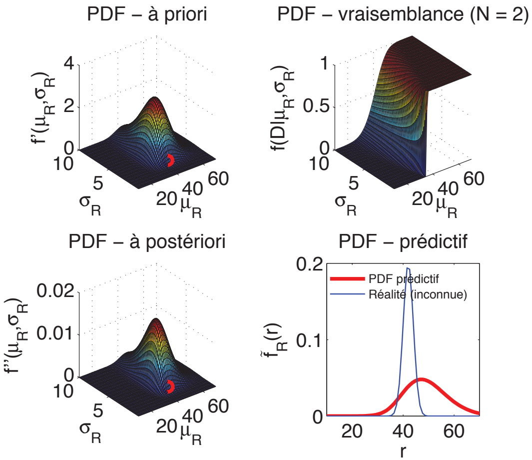
$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Exemple #11.1 - Résultats []

$$R > y_1 = 35 \text{ MPa}$$
$$R > y_2 = 35 \text{ MPa}$$

Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

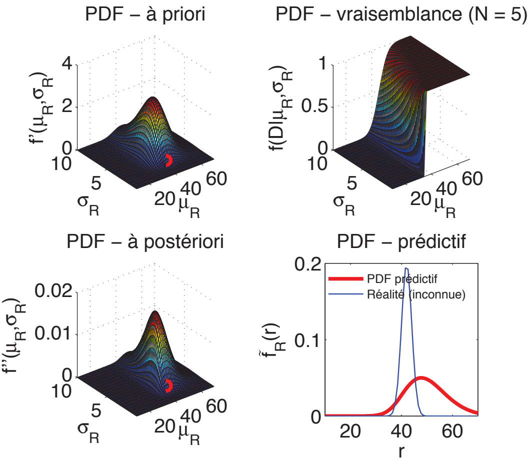


Exemple #11.1 - Résultats [.m]

$N = 5$

Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

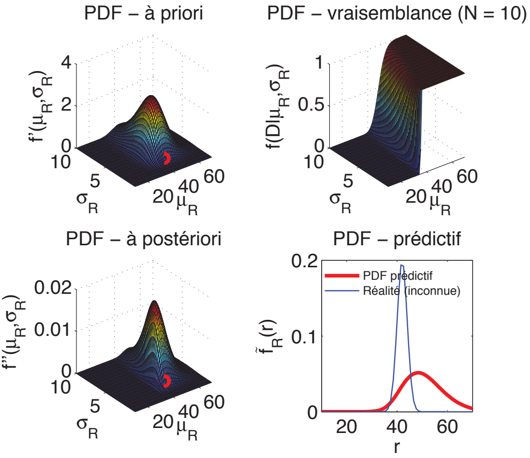


Exemple #11.1 - Résultats [.m]

$N = 10$

Densité de probabilité prédictive

$$\tilde{f}_R(r) = \int_{\boldsymbol{\theta}} f_R(r|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$$



Plan de la section

Constante de normalisation, $f(\mathcal{D})$

$$f(\boldsymbol{\theta}|\mathcal{D}) = \frac{f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathcal{D})}, \quad f(\mathcal{D}) = \int_{\boldsymbol{\theta}} f(\mathcal{D}|\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta}$$

Pour $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_m]^\intercal$, lorsque $m > 3$; calcul analytique de $f(\mathcal{D})$ devient très difficile.
Lorsque $N > 10$ une approche MC de base devient insuffisante.

Solutions:

- Approximation MLE: $\boldsymbol{\theta}^{\{\text{MLE}\}} = \arg\max_{\boldsymbol{\theta}} f(\mathcal{D}|\boldsymbol{\theta})$
- Approximation de Laplace:
  $$f(\boldsymbol{\theta}|D) \approx \mathcal{N}(\boldsymbol{\theta}^{\{\text{MLE}\}}, -\boldsymbol{H}(\ln f(\mathcal{D}|\boldsymbol{\theta}^{\{\text{MLE}\}}))^{-1})$$
- Intégration Monté Carlo avancés, $\boldsymbol{\theta}_i \sim h(\boldsymbol{\theta}) \propto f(\boldsymbol{\theta}|\mathcal{D})$
- **Densité de probabilité à priori conjuguée**

Densité de probabilité à priori conjuguée

Pour certaines densités de probabilités $f(\mathcal{D}|\boldsymbol{\theta})$ et $f(\boldsymbol{\theta})$, la densité à postériori $f(\boldsymbol{\theta}|\mathcal{D})$ fait parti de la même famille que $f(\boldsymbol{\theta})$.

Pour les densités de probabilités à priori conjuguées, la densité de probabilité a posteriori des paramètres $f(\boldsymbol{\theta}|\mathcal{D})$ est décrite par des hyper-paramètres qui peuvent être calculés analytiquement.

Exemples: `http://en.wikipedia.org/wiki/Conjugate_prior`

Hyper-paramètres: Paramètres décrivant la densité de probabilité des paramètres que l'on veut identifier.

Exemple #11.2 [.m]

Supposons que l'on connait
$\sigma_{\ln R} = 0.05, \sigma_V = 0.01$

Réalité (inconnue!, $\mu_R = 42\,MPa$):

$$R \sim \ln\mathcal{N}(\underbrace{\mu_{\ln R}}_{=?}, \underbrace{\sigma_{\ln R}}_{=0.05})\,\text{MPa}$$

(Note: la résistance (R) du béton à une variabilité intrinsèque)

$$f_{\ln R}(\ln r|\mu_R) = \mathcal{N}(\mu_{\ln R}, \sigma^2_{\ln R} + \sigma^2_V)$$

$$f'(\mu_{\ln R}) = \mathcal{N}(\mu'_{\mu_{\ln R}}, \sigma'_{\mu_{\ln R}}); \quad \{\mu'_{\mu_R} = 45\,MPa,\ \sigma'_{\mu_R} = 7\,MPa\}$$

$$f''(\mu_{\ln R}) = \mathcal{N}(\mu''_{\mu_{\ln R}}, \sigma''_{\mu_{\ln R}}); \quad \{\mu''_{\mu_R} = ?,\ \sigma''_{\mu_R} = ?\}$$

Exemple #11.2 - Formulation [.m]

$$f''(\mu_{\ln R}) = \mathcal{N}(\mu''_{\mu_{\ln R}}, \sigma''_{\mu_{\ln R}}); \quad \{\mu''_{\mu_R} = ?, \ \sigma''_{\mu_R} = ?\}$$

Formulation analytique pour $\mu''_{\mu_{\ln R}}, \sigma''_{\mu_{\ln R}}$

$$\mu''_{\mu_{\ln R}} = \frac{\frac{\sum_{i=1}^{N} \ln y_i}{N} \sigma'^2_{\mu_{\ln R}} + \mu'_{\mu_{\ln R}} \left(\frac{\sigma^2_{\ln R} + \sigma^2_V}{N}\right)}{\sigma'^2_{\mu_{\ln R}} + \frac{\sigma^2_{\ln R} + \sigma^2_V}{N}}$$

$$\sigma''^2_{\mu_{\ln R}} = \frac{\sigma'^2_{\mu_{\ln R}} \left(\frac{\sigma^2_{\ln R} + \sigma^2_V}{N}\right)}{\sigma'^2_{\mu_{\ln R}} + \frac{\sigma^2_{\ln R} + \sigma^2_V}{N}}$$

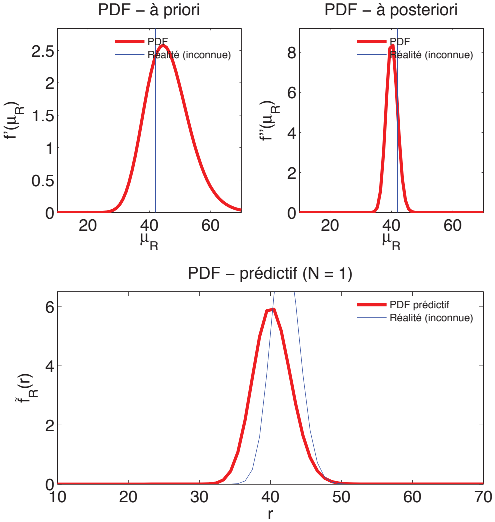
Densité de probabilité prédictive

$$\tilde{f}_{\ln R}(\ln r) = \mathcal{N}(\mu''_{\mu_{\ln R}}, \sigma^2_{\ln R} + \sigma''^2_{\mu_{\ln R}})$$

$$\tilde{f}_R(r) = \ln\mathcal{N}(\mu''_{\mu_{\ln R}}, \sigma_{\ln R} + \sigma''_{\mu_{\ln R}})$$

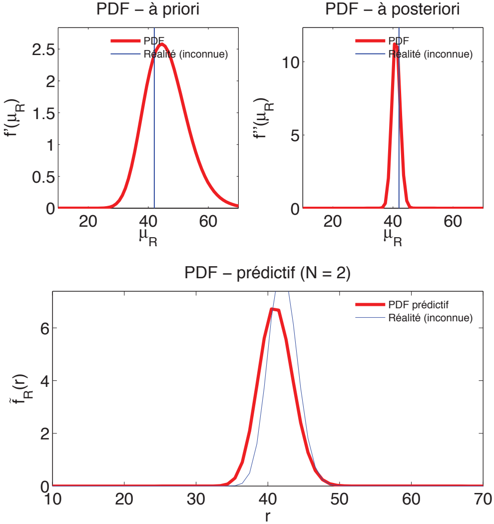
Exemple #11.2 - Résultats [.m]

$y_1 = 39.7$ MPa



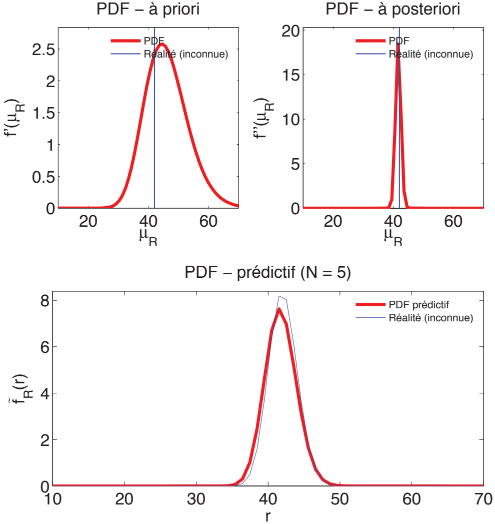
Exemple #11.2 - Résultats [.m]

$y_1 = 39.7$ MPa
$y_2 = 42.0$ MPa



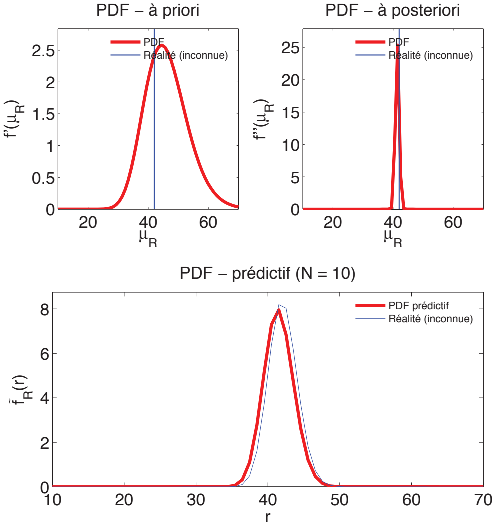
Exemple #11.2 - Résultats [.m]

$N = 5$



Exemple #11.2 - Résultats [.m]

$N = 10$



Résumé

- L'estimation Bayesienne permet de tirer un maximum d'information à partir d'observations afin de réduire les incertitudes épistémiques (⚠ les incertitudes de type aléatoires sont irréductibles)

$$\overbrace{f''(\mu_R,\sigma_R)}^{\text{PDF à postériori}} = \frac{\overbrace{f(Y_1=y_1,\cdots,Y_N=y_N|\mu_R,\sigma_R)}^{\text{fct. de vraisemblance}}\cdot\overbrace{f'(\mu_R,\sigma_R)}^{\text{PDF à priori}}}{\underbrace{f(Y_1=y_1,\cdots,Y_N=y_N)}_{\text{cte. de normalisation}}}$$

- Les observations peuvent être des mesures directes ou des bornes (supérieures ou inférieures)

- L'estimation des paramètres peut se faire à partir d'une intégration numérique, intégration Monté Carlo, ou à l'aide des densités de probabilité à priori conjuguée.

Théorie probabilité

- 0 Révision algèbre & probabilités
- 1 Lois de probabilités
- 3 PDF multivariés

Fiabilité des structures & systèmes

- Introduction
- 2 Formulation fiabilité composantes $g_1(\mathbf{X})$
- 9 Formulation fiabilité systèmes

Estimation – p_f échantillonnage

- 4 Monte Carlo
- 12 Échantillonnage par importance

Estimation – p_f analytique

- 5 FOSM
- 6 FORM
- 8 SORM

Utilisation des résultats

- 7 Analyse de sensibilité
- 10 Incertitudes aléatoires & épistémiques
- 11 Données empiriques
- 13 Prise de décision
- 14 Métamodèles & modèles empiriques