Révision: probabilités & ensembles

(CIV8530 - Fiabilité des structures et systèmes)

Enseignant: James-A. Goulet

Département des génies civil, géologique et des mines
Polytechnique Montréal

Chapitre 3 – J.-A. Goulet, (2020), Probabilistic Machine Learning for Civil Engineers, MIT Press

Théorie des probabilités, pourquoi?

"... la *théorie des probabilités* n'est, au fond, que le bon sens réduit au calcul; elle fait apprécier avec exactitude ce que les esprits justes sentent par une sorte d'instinct..."

— Pierre-Simon, marquis de Laplace [1749 – 1827]

[Wikipedia]

Théorie des probabilités, pourquoi?

Pour décrire les phénomènes aléatoires? ⚠ non.

Peu de phénomènes sont intrinsèquement aléatoires

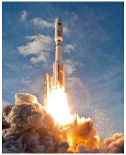
On utilise les probabilités pour **représenter notre connaissance (incertitudes épistémiques)**

- Le module élastique E ne varie pas (localement)
- On ne connait pas exactement sa valeur
- On décrit notre incertitude de E par des probabilités

Le moins on en sait, le plus on devrait utiliser les probabilités

[Google images / Nitipak Samsen]

Qui/qu'est-ce qui utilise les probabilités?



Si cela fonctionne pour eux, cela fonctionne pour nous.

Plan du module #0b

Ensembles

Probabilités

$X-$ **V.A.**

$X-$ **V.A.M.**

$\mathbb{E}(X)$

$g(\mathbf{X})$

Linéarisation

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
- 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

Définitions

Ensemble: Regroupement d'éléments/évènements

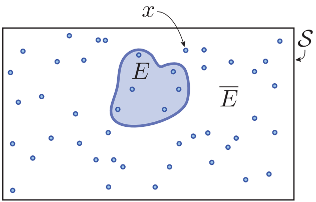
Univers/espace d'échantillonnage ($\mathcal{S}$): Ensemble des résultats possibles (peut être continu/discret & fini/infini)

Évènement élémentaire (x): un seul élément, $x \in \mathcal{S}$

Évènement (E): ensemble d'évènements élémentaires

$E \subseteq \mathcal{S}$	: Sous-ensemble de $\mathcal{S}$
$E = \mathcal{S}$	: Évènement certain
$E = \emptyset$	: Évènement impossible
$\overline{E}$	: Complément de E

Diagramme de Venn



Exemple #1 – État d'une structure après un séisme

Réalité: continuum, espace d'échantillonnage continu

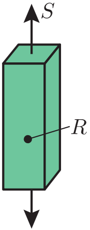
$$\mathcal{S} = \{\text{aucun dégat} \longleftrightarrow \text{éfondrement total}\}$$

Abstraction: espace d'échantillonnage discret

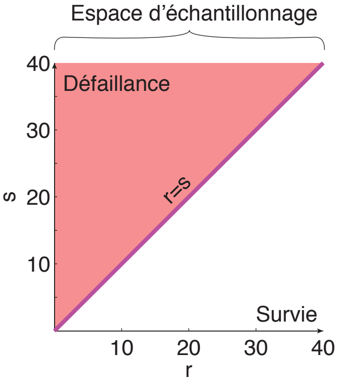
$$\mathcal{S} = \{\text{aucun dégat}, \text{léger dégats}, \text{importants dégats}, \text{éfondrement}\}$$

$$\begin{aligned}
\mathcal{S} &= \{AD, LD, ID, EF\}\\
E_1 &= \{AD, LD\}\\
E_2 &= \{EF\}\\
\overline{E_1} &= \{ID, EF\}\\
\overline{E_2} &= \{AD, LD, ID\}
\end{aligned}$$

Exemple #2 – Défaillance d'une composante en traction



$s \in \mathbb{R}^+$: sollicitation
$r \in \mathbb{R}^+$: résistance
$\mathcal{S} = \mathbb{R}^{2+}$

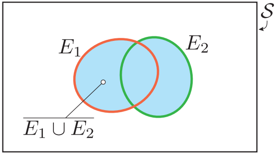


Défaillance : $\{R \leq S\} \equiv \{R - S \leq 0\}$

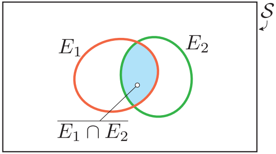
Opérations sur les évènements

Soit des évènements $\{E_1, E_2, \cdots, E_n\} \in \mathcal{S}$

Union (i.e. "ou"): $E_1 \cup E_2$



Intersection (i.e. "et"): $E_1 \cap E_2 \equiv E_1 E_2$

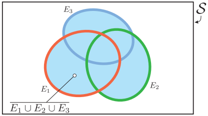


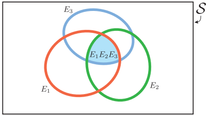
Règles

Commutativité: $E_1 \cup E_2 = E_2 \cup E_1, \quad E_1 E_2 = E_2 E_1$

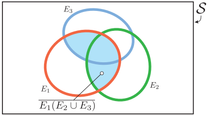
Associativité: $(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3) = E_1 \cup E_2 \cup E_3$

$\cup_{i=1}^{n} E_i = E_1 \cup E_2 \cup \cdots \cup E_n$ $\qquad$ $\cap_{i=1}^{n} E_i = E_1 \cap E_2 \cap \cdots \cap E_n$





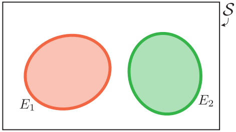
Distributivité: $E_1(E_2 \cup E_3) = (E_1 E_2 \cup E_1 E_3)$



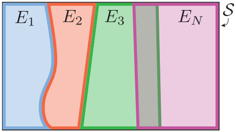
⚠ **Convention:** l'intersection à priorité sur l'union… sinon ( )

Types d'évènements: $\{E_1, E_2, \cdots, E_n\} \in \mathcal{S}$

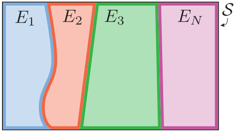
mutuellement exclusifs (ME):
$E_1, E_2, \cdots, E_n$ sont ME
si $E_i E_j = \emptyset, \ \forall i \neq j$

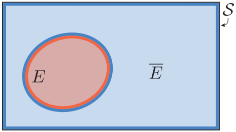


collectivement exhaustifs (CE):
$E_1, E_2, \cdots, E_n$ sont CE
si $\cup_{i=1}^{n} E_i = \mathcal{S}$



mutuellement exclusifs & collectivement exhaustifs (ME & CE) :





Lois de Morgan

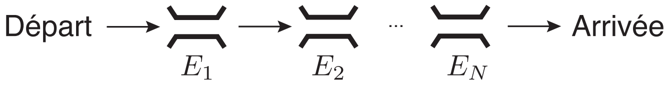
$$\overline{\bigcup_i E_i} = \bigcap_i \overline{E_i}$$

Le complément des unions est égal à l'intersection des compléments

$$\overline{\bigcap_i E_i} = \bigcup_i \overline{E_i}$$

Le complément des intersections est égal à l'union des compléments

Exemple – Lois de Morgan



E_i: Le pont n'est pas ouvert à la circulation

Défaillance: $E_1 \cup E_2 \cup \cdots \cup E_n = \underbrace{\bigcup_{i=1}^{n}}_{\text{“ou”}} E_i$

Survie: $\overline{\bigcup_i E_i} = \underbrace{\bigcap_{i}}_{\text{“et”}} \overline{E_i}$

Plan de la section

Interpretation des probabilités

Probabilité de l'évènement E_i: $\boxed{\Pr(E_i)}$

Interpretation Bayesienne:
$\Pr(E_i)$ est une mesure de vraisemblance par rapport aux autres évènements dans $\mathcal{S}$

- Dépend de notre connaissance à priori
- Change lorsque l'on obtient de nouvelles informations

Interpretation Fréquentiste:

$$\Pr(E_i) = \lim_{n\to\infty} \frac{\#x}{n}$$

Axiomes fondamentaux

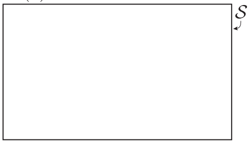
1. $0 \leq \Pr(E_i) \leq 1$

2. $\Pr(\mathcal{S}) = 1$
   (probabilité d'un évènement certain $= 1$)

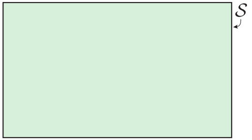
3. si E_1 et E_2 sont mutuellement exclusifs

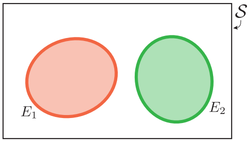
$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$$

$\Pr(\emptyset) = 0$



$\Pr(\mathcal{S}) = 1$



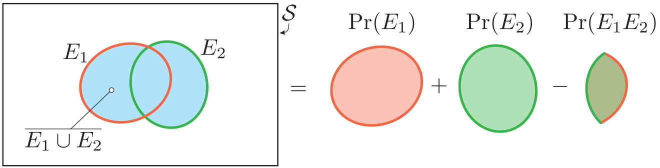


Lois dérivées à partir des axiomes fondamentaux

Probabilité du complément: $\Pr(\overline{E}) = 1 - \Pr(E)$

Probabilité de l'ensemble vide: $\Pr(\emptyset) = 1 - 1 = 0$

Loi de l'addition: $\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 E_2)$

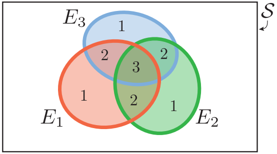


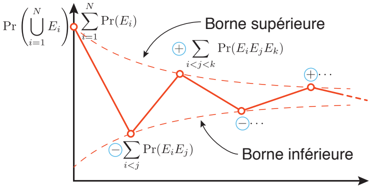
Loi d'inclusion-exclusion:

$$\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n}\Pr(E_i) - \sum_{i<j}\Pr(E_iE_j) + \sum_{i<j<k}\Pr(E_iE_jE_k) - \cdots + (-1)^{n-1}\Pr(E_1E_2\cdots E_n)$$
$$= 1 - \Pr(\overline{E_1}\,\overline{E_2}\cdots\overline{E_n}) \quad (\textbf{Loi de Morgan})$$

Pour $n = 3$

$$\Pr(E_1 \cup E_2 \cup E_3) = \Pr(E_1)+\Pr(E_2)+\Pr(E_3)-\Pr(E_1E_2)-\Pr(E_1E_3)-\Pr(E_2E_3)+\Pr(E_1E_2E_3)$$





⚠ La difficulté d'évaluer la probabilité conjointe $\Pr(E_1E_2\cdots E_N)$ augmente avec N (**mais sa contribution diminue**)

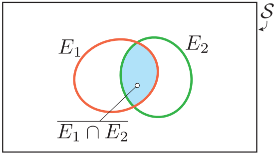
Probabilités conditionnelles

$\Pr(E_1|E_2)$: Probabilité de l'évènement E_1 conditionnelle à la réalisation de l'évènement E_2

$$\Pr(E_1|E_2) = \frac{\Pr(E_1 E_2)}{\Pr(E_2)}, \quad (\Pr(E_2) \neq 0)$$

$$\overbrace{\Pr(E_1 E_2)}^{\text{Probabilité conjointe}} = \overbrace{\Pr(E_1|E_2)}^{\text{Probabilité conditionnelle}} \cdot \overbrace{\Pr(E_2)}^{\text{Probabilité marginale}}$$

$$= \Pr(E_2|E_1) \cdot \Pr(E_1)$$



Si E_1 et E_2 sont statistiquement indépendants ($E_1 \perp\!\!\!\perp E_2$)

$$\Pr(E_1|E_2) = \Pr(E_1) \rightarrow \Pr(E_1 E_2) = \Pr(E_1) \cdot \Pr(E_2)$$

Loi de la dérivation en chaine

$$\begin{aligned}
\Pr(E_1 E_2 \cdots E_n) &= \Pr(E_1|E_2 \cdots E_n)\Pr(E_2 \cdots E_n)\\
&= \Pr(E_1|E_2 \cdots E_n)\Pr(E_2|E_3 \cdots E_n)\Pr(E_3 \cdots E_n)\\
&= \Pr(E_1|E_2 \cdots E_n)\Pr(E_2|E_3 \cdots E_n)\cdots\Pr(E_{n-1}|E_n)\Pr(E_n)
\end{aligned}$$

- n termes
- $n-1$ probabilités conditionnelles
- 1 probabilité marginale

Exemple:

$$Pr(E_1E_2E_3) = \overbrace{\Pr(E_1|E_2E_3)\underbrace{\Pr(E_2|E_3)\Pr(E_3)}_{=\Pr(E_2E_3)}}^{=Pr(E_1E_2E_3)}$$

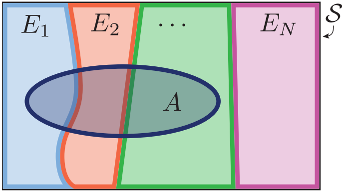
Loi de probabilité totale

Soit un évènement A appartenant à l'espace d'échantillonnage $\mathcal{S}$ (i.e. $A \in \mathcal{S}$)

Soit $\{E_1, E_2, E_3, \cdots, E_n\} \in \mathcal{S}$ un ensemble d'éléments mutuellement exclusifs et collectivement exhaustifs (ME&CE) (i.e. $E_iE_j = \emptyset,\ \forall i \neq j,\ \cup_{i=1}^{n} E_i = \mathcal{S}$)

$$\Pr(A) = \sum_{i=1}^{n} \underbrace{\Pr(A|E_i)\Pr(E_i)}_{=\Pr(AE_i)}$$

(Marginalisation)



Loi de probabilité totale (cont.)

RAPPEL: (ME&CE, $\cup_{i=1}^{n} E_i = \mathcal{S}$, )

Loi de l'addition:

$$\Pr(E_1 \cup E_2 | A) = \Pr(E_1|A) + \Pr(E_2|A) - \Pr(E_1 E_2|A)$$

Loi de la multiplication:

$$\Pr(E_1 E_2|A) = \Pr(E_1|E_2 A)\Pr(E_2|A)$$

$$\Pr(A|B) = \sum_{i=1}^{n} \underbrace{\Pr(A|E_i B)\Pr(E_i|B)}_{=\Pr(AE_i|B)}$$

Théorème de Bayes

Soit $E_1, E_2, E_3, \cdots, E_n$ un ensemble d'éléments mutuellement exclusifs et collectivement exhaustifs (ME&CE)

$$\left.\begin{array}{rcl} \Pr(AE_i) & = & \Pr(A|E_i)\Pr(E_i) \\ & = & \Pr(E_i|A)\Pr(A) \end{array}\right\} \rightarrow \Pr(E_i|A)\Pr(A) = \Pr(A|E_i)\Pr(E_i)$$

Fonction de vraisemblance

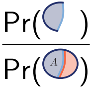
Probabilité à priori

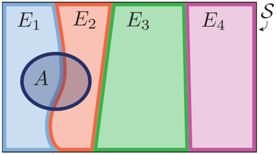
$$\underbrace{\Pr(E_i|A)}_{\text{Probabilité à posteriori}} = \frac{\Pr(A|E_i)\Pr(E_i)}{\Pr(A)}$$

constante de normalisation

Théorème de Bayes:

$$\Pr(E_i|A) = \frac{\Pr(A|E_i)\Pr(E_i)}{\Pr(A)} =$$





Constante de normalisation:

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A|E_i)\Pr(E_i)$$

⚠ typiquement, $\Pr(A)$ est difficile à évaluer

Exemple – Probabilités conditionnelles & Bayes

Soit un séisme:

$$\mathcal{S}_{\text{intensité}} = \{\text{Léger (L)}, \text{Modéré (M)}, \text{Important (I)}\}$$

et une structure pouvant être dans l'un des deux états suivants:

$$\mathcal{S}_{\text{état}} = \{\text{endommagé } (D), \neg\,\text{endommagé } (\overline{D})\}$$

On cherche Pr(D) sachant que

$$\begin{array}{rcl} \Pr(D|L) & = & 0.01 \\ \Pr(D|M) & = & 0.10 \\ \Pr(D|I) & = & 0.60 \end{array} \quad \& \quad \left.\begin{array}{rcl} \Pr(L) & = & 0.90 \\ \Pr(M) & = & 0.08 \\ \Pr(I) & = & 0.02 \end{array}\right\} \sum = 1$$

$$\begin{aligned} \Pr(D) &= \overbrace{\Pr(D|L)\Pr(L)}^{\Pr(DL)} + \overbrace{\Pr(D|M)\Pr(M)}^{\Pr(DM)} + \overbrace{\Pr(D|I)\Pr(I)}^{\Pr(DI)} \\ &= 0.01 \times 0.90 + 0.10 \times 0.08 + 0.6 \times 0.02 \\ &= 0.029 \end{aligned}$$

À l'inverse: si l'on **observe des dommages** sur une structure (état $:= D$), quelle est la **probabilité de chaque intensité**?

$$\Pr(I|D) = \frac{\Pr(D|I)\Pr(I)}{\Pr(D)} = \frac{0.60 \times 0.02}{0.029} = 0.41$$

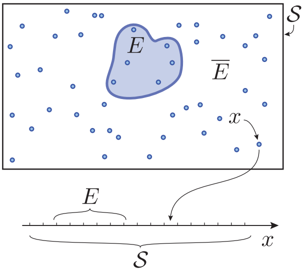
$$\Pr(M|D) = \frac{\Pr(D|M)\Pr(M)}{\Pr(D)} = \frac{0.10 \times 0.08}{0.029} = 0.28$$

$$\Pr(L|D) = \frac{\Pr(D|L)\Pr(L)}{\Pr(D)} = \frac{0.01 \times 0.90}{0.029} = 0.31$$

Plan de la section

Variables aléatoires (V.A.)

Représentation de l'espace d'échantillonnage $\mathcal{S}$ sur une ligne



Notation:

X : Variable aléatoire
x : Réalisation d'une V.A.

La probabilité que la **variable aléatoire** X prenne une **valeur** x est décrite soit par une **fonction de masse (PMF)** ou une **densité de probabilité (PDF)**

Variables aléatoires discrètes

Fonction de masse (PMF): $p_X(x) = \Pr(X = x)$

Propriétés:

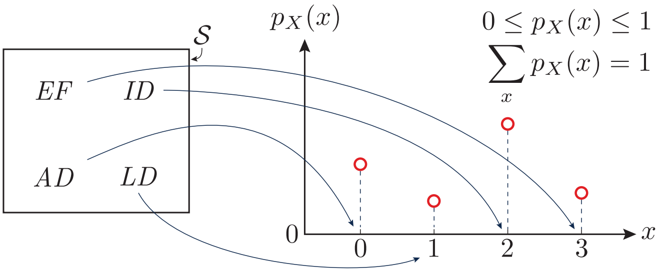
$$0 \leq p_X(x) \leq 1$$

$$\sum_x p_X(x) = 1$$

L'ensemble des valeurs possibles de x constitue un ensemble mutuellement exclusif et collectivement exhaustifs (ME&CE)

Exemple – endommagement d'une structure

$$\mathcal{S} = \left\{ \begin{array}{r} \text{aucun dégat (AD)} \\ \text{léger dégats (LD)} \\ \text{importants dégats (ID)} \\ \text{éfondrement (EF)} \end{array} \right\}$$



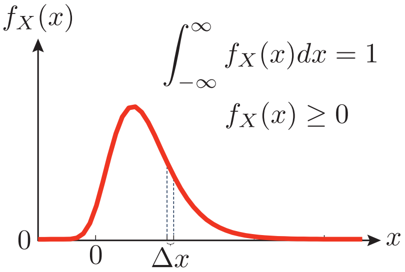
Dommages légers ou importants: $\{1 \leq x \leq 2\}$

Variables aléatoires continues

Densité de probabilité (PDF): $f_X(x)$

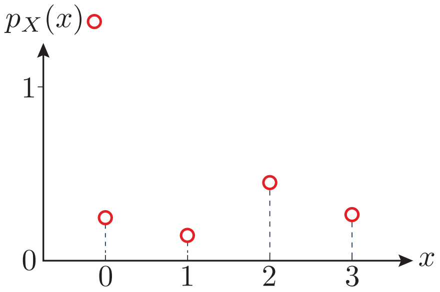
$$\Pr(x < X \leq x + \Delta x) = f_X(x)\Delta x$$

$$\Pr(X = x) = 0$$



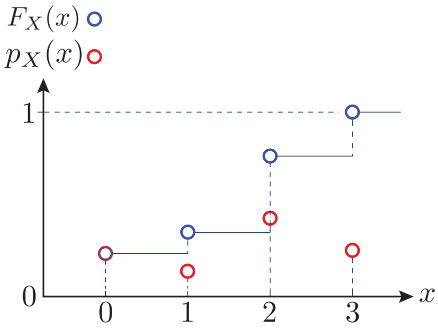
Densité de probabilité cumulative/fonction de répartition (CDF)

$$F_X(x) = \Pr(X \leq x) = \sum_{x' \leq x} p_X(x'), \quad \text{(V.A. discrète)}$$



Densité de probabilité cumulative/fonction de répartition (CDF)

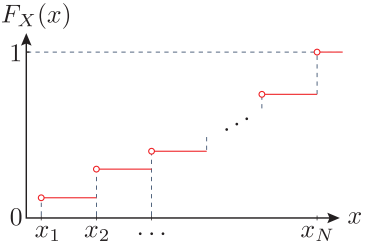
$$F_X(x) = \Pr(X \leq x) = \sum_{x' \leq x} p_X(x'), \text{ (V.A. discrète)}$$

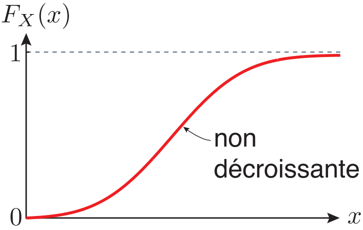


Densité de probabilité cumulative (CDF)

$$F_X(x) = \Pr(X \leq x) = \sum_{x' \leq x} p_X(x'), \quad \text{(V.A. discrète)}$$

$$= \int_{-\infty}^{x'} f_X(x')dx, \quad \text{(V.A. continue)}$$





$$F_X(-\infty) = 0, \quad F_X(+\infty) = 1$$

PDF $\leftrightarrow$ CDF

Pour les V.A. continues

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx \leftrightarrow \boxed{f_X(x) = \frac{dF_X(x)}{dx}}$$

Pour les V.A. discrètes

$$p_X(x_i) = F_X(x_i) - F_X(x_{i-1})$$

Les **densités de probabilités communes** seront vues au module #1

Loi normale univariée - $X \sim \mathcal{N}(\mu, \sigma^2)$ [.m]

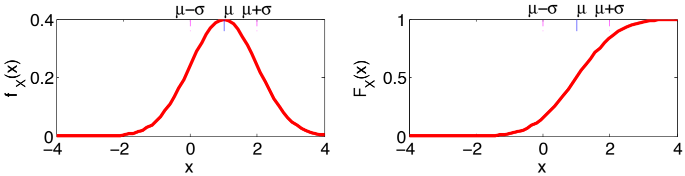
Densité de probabilité (PDF) pour 1 variable aléatoire X

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Densité de probabilité cumulative (CDF)

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

$\mu = 1$
$\sigma = 1$



Probabilités conditionnelles & V.A. [ ]

Pour un évènement E et une variable aléatoire X; la probabilité de E conditionnelle à la réalisation de l'évènement $\{X = x\}$ est:

$$\begin{aligned}
\Pr(E|\{X=x\}) &= \frac{\Pr(E \cap \{X=x\})}{\Pr(\{X=x\})} \\
\Pr(E|X=x) &= \frac{\Pr(E \cap X=x)}{\Pr(X=x)} \\
&= \frac{\Pr(X=x|E)\Pr(E)}{\Pr(X=x)} \\
\Pr(E|x) &= \frac{\Pr(x|E)\Pr(E)}{\Pr(x)}
\end{aligned}$$

Rappel:

$$\boxed{\Pr(E_1|E_2) = \frac{\Pr(E_1E_2)}{\Pr(E_2)}}$$

Probabilités conditionnelles – évènements

Si X est une variable aléatoire **discrète**

$$\boxed{E|x} \quad \rightarrow \quad \Pr(E|x) = \frac{p_{X|E}(x|E)\Pr(E)}{p_X(x)}$$

$$\boxed{X|E} \quad \rightarrow \quad \Pr(X|E) = p_{X|E}(x|E) = \frac{\Pr(E|x)p_X(x)}{\Pr(E)}$$

Si X est une variable aléatoire **continue**

$$\boxed{E|x} \quad \rightarrow \quad \Pr(E|x) = \frac{f_{X|E}(x|E)\Pr(E)}{f_X(x)}$$

$$\boxed{X|E} \quad \rightarrow \quad f_{X|E}(x|E) = \frac{\Pr(E|x)f_X(x)}{\Pr(E)}$$

Plan de la section

X– **V.A.M.**

Variables aléatoires multiples

Soit $\mathbf{X} = [X_1, X_2, \cdots, X_n]^\intercal$ { vecteur (colonne) de variable aléatoires

Soit $\mathbf{x} = [x_1, x_2, \cdots, x_n]^\intercal$ { vecteur (colonne) représentant une réalisation de $\mathbf{X}$

$\mathbf{X}$ décrit l'occurrence simultanée de plusieurs phénomènes

Densités de probabilité multivariées discrètes (PMF)

Définition:

- $p_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 = x_1 \cap X_2 = x_2 \cap \cdots \cap X_n = x_n)$
- $0 \leq p_{\mathbf{X}}(\mathbf{x}) \leq 1$

Marginalisation:

- $\displaystyle\sum_{x_n} p_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = p_{X_1,\cdots,X_{n-1}}(x_1,\cdots,x_{n-1})$
- $\displaystyle\sum_{x_1}\cdots\sum_{x_n} p_{\mathbf{X}}(\mathbf{x}) = 1$

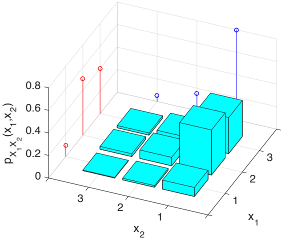
$p_{X_i}(x_i)$ **: Fonction de masse marginale**

Exemple – marginalisation [.m]

Soit deux V.A. discrètes $X_1 \perp\!\!\!\perp X_2$

$$p_{X_1}(x_1)\begin{cases} p_{X_1}(1) = 0.1 \\ p_{X_1}(2) = 0.5 \\ p_{X_1}(3) = 0.4 \end{cases}$$

$$p_{X_2}(x_2)\begin{cases} p_{X_2}(1) = 0.8 \\ p_{X_2}(2) = 0.15 \\ p_{X_2}(3) = 0.05 \end{cases}$$



$$p_{X_1X_2}(x_1,x_2) = p_{X_1}(x_1)\cdot p_{X_2}(x_2)$$

$p_{X_1X_2}(x_1,x_2) =$

	$x_1 = 1$	$x_1 = 2$	$x_1 = 3$
$x_2 = 1$	0.08	0.4	0.32
$x_2 = 2$	0.015	0.075	0.06
$x_2 = 3$	0.005	0.025	0.02
$\sum_{i=1}^{3} p_{X_1X_2}(x_1,i)$	0.1	0.5	0.4

Densités de probabilité multivariées continues (PDF)

Définition:

- $f_{\mathbf{X}}(\mathbf{x})\Delta\mathbf{x} = \Pr(x_1 < X_1 \leq x_1 + \Delta x_1 \cap \cdots \cap x_n < X_n \leq x_n + \Delta x_n)$
- $0 \leq f_{\mathbf{X}}(\mathbf{x})$ ⚠ peut être > 1

Marginalisation:

- $\int_{-\infty}^{\infty} f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)dx_n = f_{X_1,\cdots,X_{n-1}}(x_1,\cdots,x_{n-1})$
- $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1$

$f_{X_i}(x_i)$: **Densité de probabilité marginale**

Loi normale bi-variée (PDF) [.m]

Densité de probabilité (PDF) pour 2 variables aléatoires X_1, X_2

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_{X_1}}{\sigma_{X_1}}\right)^2+\left(\frac{x_2-\mu_{X_2}}{\sigma_{X_2}}\right)^2-2\rho\left(\frac{x_1-\mu_{X_1}}{\sigma_{X_1}}\right)\left(\frac{x_2-\mu_{X_2}}{\sigma_{X_2}}\right)\right]\right\}$$

5 paramètres: $\mu_{X_1}, \sigma_{X_1}, \mu_{X_2}, \sigma_{X_2}, \rho$

Loi normale bi-variée (PDF) [.m]

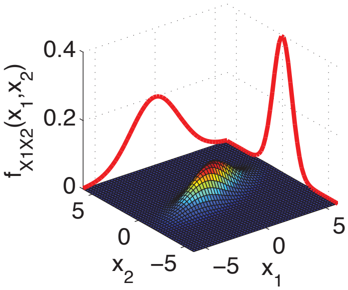
$$\boxed{\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})}$$

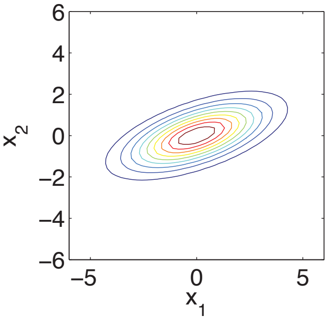
5 paramètres:

$\mu_{X_1}, \sigma_{X_1}, \mu_{X_2}, \sigma_{X_2}, \rho$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{X_1}^2 & \rho\sigma_{X_1}\sigma_{X_2} \\ \rho\sigma_{X_1}\sigma_{X_2} & \sigma_{X_2}^2 \end{bmatrix}$$

$$\mu_{X_1} = 0$$
$$\sigma_{X_1} = 2$$
$$\mu_{X_2} = 0$$
$$\sigma_{X_2} = 1$$
$$\rho = 0.6$$





Fonctions de répartitions multivariées continues (CDF)

Définition:

- $F_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 \leq x_1 \cap \cdots \cap X_n \leq x_n)$
- $0 \leq F_{\mathbf{X}}(\mathbf{x}) \leq 1$
- $F_{X_1,\cdots,X_{n-1},X_n}(x_1,\cdots,x_{n-1},-\infty) = 0$
- $F_{X_1,\cdots,X_{n-1},X_n}(+\infty,\cdots,+\infty,+\infty) = 1$

Marginalisation: $(F_{X_1X_2}(x_1,+\infty) = F_{X_1}(x_1))$

- $F_{X_1,\cdots,X_{n-1},X_n}(x_1,\cdots,x_{n-1},+\infty) = F_{X_1,\cdots,X_{n-1}}(x_1,\cdots,x_{n-1})$

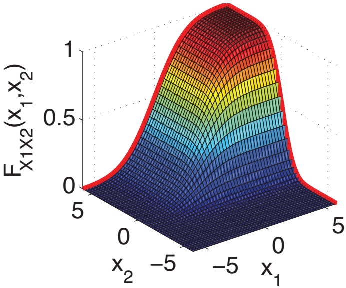
$$\boxed{F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}, \qquad f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = \frac{\partial^n F_{\mathbf{X}}(\mathbf{x})}{\partial x_1\cdots\partial x_n}}$$

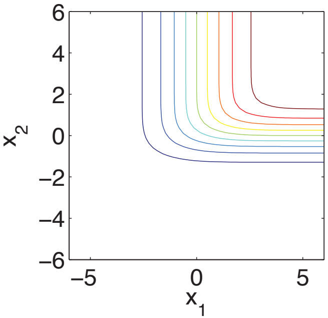
Loi normale bi-variée (CDF) [.m]

Densité de probabilité cumulative (CDF) pour 2 variables aléatoires X_1, X_2

$$F_{X_1,X_2}(x_1,x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1,X_2}(x_1,x_2)\partial x \partial y$$

$$\mu_{X_1} = 0$$
$$\sigma_{X_1} = 2$$
$$\mu_{X_2} = 0$$
$$\sigma_{X_2} = 1$$
$$\rho = 0.6$$





Probabilités conditionnelles – discrètes

Si X_i est une variable aléatoire **discrète**

$$p_{X_1X_2}(x_1, x_2) = p_{X_1|X_2}(x_1|x_2)p_{X_2}(x_2) = p_{X_2|X_1}(x_2|x_1)p_{X_1}(x_1)$$

$$\boxed{X_1|x_2} \rightarrow p_{X_1|X_2}(x_1|x_2) = \frac{p_{X_2|X_1}(x_2|x_1)p_{X_1}(x_1)}{p_{X_2}(x_2)}$$

$$\boxed{X_2|x_1} \rightarrow p_{X_2|X_1}(x_2|x_1) = \frac{p_{X_1|X_2}(x_1|x_2)p_{X_2}(x_2)}{p_{X_1}(x_1)}$$

Probabilités conditionnelles – discrètes

X_1 et X_2 sont statistiquement indépendantes ($\perp\!\!\!\perp$) si

$$p_{X_1|X_2}(x_1|x_2) = p_{X_1}(x_1)$$

si $X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_n$

$$p_{X_1\cdots X_n}(x_1,\cdots,x_n) = p_{X_1}(x_1)p_{X_2}(x_2)\cdots p_{X_n}(x_n)$$

Cas général: X_1 et X_2 **pas** statistiquement indépendantes

$\rightarrow$ **dérivation en chaine**

$$p_{X_1\cdots X_n}(x_1,\cdots,x_n) = p_{X_1|X_2\cdots X_n}(x_1|x_2\cdots,x_n)\cdots p_{X_{n-1}|X_n}(x_{n-1}|x_n)p_{X_n}(x_n)$$

e.g. $p_{X_1X_2X_3}(x_1,x_2,x_3) = p_{X_1|X_2X_3}(x_1|x_2,x_3)\cdot p_{X_2|X_3}(x_2|x_3)\cdot p_{X_3}(x_3)$

Probabilités conditionnelles – continues

Exactement la même chose que pour les V.A. discrètes

$$p_{\mathbf{X}}(\mathbf{x}) \rightarrow f_{\mathbf{X}}(\mathbf{x})$$

Plan de la section

Espérance mathématique (expected value)

Espérance mathématique: Somme pondérée de toutes les valeurs possibles d'une variable aléatoire

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx \quad \text{(V.A. continue)}$$

$$\mathbb{E}[X] = \sum x \cdot p_X(x)dx \quad \text{(V.A. discrète)}$$

L'espérance est une **opération linéaire**

$$\mathbb{E}[g(\mathbf{X})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) \cdot f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$$

Moments

Moment d'ordre m: $\mathbb{E}[X^m]$

$$\mathbb{E}[X^m] = \int_{-\infty}^{\infty} x^m f_X(x)dx$$

Pour $m = 1$: $\mathbb{E}[X] = \mu_X$ (**espérance / centre de gravité**)

Pour $m = 2$: $\mathbb{E}[X^2]$ (espérance des carrés)

Moments centrés d'ordre m: $\mathbb{E}[(X - \mu_X)^m]$

$$\mathbb{E}[(X - \mu_X)^m] = \int_{-\infty}^{\infty} (x - \mu_X)^m f_X(x)dx$$

Pour $m = 1$: $\mathbb{E}[(X - \mu_X)^1] = 0$

Pour $m = 2$: $\mathbb{E}[(X - \mu_X)^2] = \sigma_X^2 = \text{var}[X]$ (**variance / inertie**)

Variance – $\mathbb{E}[(X-\mu_X)^2]$

$$\mathbb{E}[(X-\mu_X)^2] = \sigma_X^2 = \text{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

var$[X]$: La variance mesure la dispersion de la densité de probabilité. Concept analogue à l'inertie d'une poutre.

σ_X: Écart-type

$\delta_X = \frac{\sigma_X}{\mu_X}$: coefficient de variation (C.O.V.).
Mesure adimensionnelle de la dispersion (⚠ $\mu_X \neq 0$)

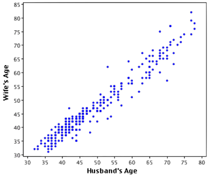
Covariance – $\mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)]$

Soit deux variables aléatoires X_i, X_j

$$\text{cov}[X_i X_j] = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$$

ρ_{ij}: coefficient de corrélation.
(Quantifie la **dépendance linéaire** entre X_i et X_j)

$$\rho_{ij} = \frac{\text{cov}[X_i X_j]}{\sigma_i \sigma_j}, \quad -1 \leq \rho_{ij} \leq +1$$



- $X_i \perp\!\!\!\perp X_j \implies \rho_{ij} = 0$
- $\rho_{ij} = 0 \not\Longrightarrow X_i \perp\!\!\!\perp X_j$ ⚠
- correlation $\not\Longleftrightarrow$ causalité ⚠

[onlinestatbook.com, Wikipedia]

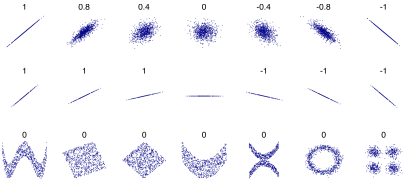
Covariance – $\mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)]$

Soit deux variables aléatoires X_i, X_j

$$\text{cov}[X_iX_j] = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathbb{E}[X_iX_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$$

ρ_{ij}: coefficient de corrélation.
(Quantifie la **dépendance linéaire** entre X_i et X_j)

$$\rho_{ij} = \frac{\text{cov}[X_iX_j]}{\sigma_i\sigma_j}, \qquad -1 \leq \rho_{ij} \leq +1$$



- $X_i \perp\!\!\!\perp X_j \implies \rho_{ij} = 0$
- $\rho_{ij} = 0 \;\not\Longrightarrow\; X_i \perp\!\!\!\perp X_j$ ⚠
- correlation $\not\Longleftrightarrow$ causalité ⚠

[onlinestatbook.com, Wikipedia]

Notation matricielle

Soit n variables aléatoires $\mathbf{X} = [X_1, X_2, \cdots, X_n]^\intercal$ où,
$\mathbf{M_X} = [\mu_1, \mu_2, \cdots, \mu_n]^\intercal$ est un vecteur contenant les moyennes,
$\mathbf{D_{XX}}$ est la matrice contenant les écarts-types, $\mathbf{R_{XX}}$ est la matrice de corrélation, et $\mathbf{\Sigma_{XX}}$ est la matrice de covariance.

$$\mathbf{D_{XX}} = \begin{bmatrix} \sigma_{X_1} & 0 & 0 & 0 \\ & \sigma_{X_2} & 0 & 0 \\ & & \ddots & 0 \\ \text{sym.} & & & \sigma_{X_n} \end{bmatrix}, \quad \mathbf{R_{XX}} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ & 1 & \cdots & \rho_{2n} \\ & & \vdots & \rho_{n-1n} \\ \text{sym.} & & & 1 \end{bmatrix}$$

$$\mathbf{\Sigma_{XX}} = \mathbf{D_{XX} R_{XX} D_{XX}} = \begin{bmatrix} \sigma^2_{X_1} & \rho_{12}\sigma_{X_1}\sigma_{X_2} & \cdots & \rho_{1n}\sigma_{X_1}\sigma_{X_n} \\ & \sigma^2_{X_2} & \cdots & \rho_{2n}\sigma_{X_2}\sigma_{X_n} \\ & & \vdots & \vdots \\ \text{sym.} & & & \sigma^2_{X_n} \end{bmatrix}$$

Plan de la section

Fonctions de variables aléatoires – $g(\mathbf{X})$

Soit $\mathbf{X} = [X_1, X_2, \cdots, X_n]^\intercal$ définit par le PDF $f_{\mathbf{X}}(\mathbf{x})$, et soit $\mathbf{Y} = [Y_1, Y_2, \cdots, Y_m]^\intercal$ obtenu à partir d'une fonction

$$\mathbf{Y} = \mathbf{g}(\mathbf{X}) = \begin{bmatrix} g_1(\mathbf{X}) \\ g_2(\mathbf{X}) \\ \vdots \\ g_m(\mathbf{X}) \end{bmatrix}$$

Question: déterminer la densité de probabilité $f_{\mathbf{Y}}(\mathbf{y})$,

3 cas:

1. $m = n = 1$
2. $m = n > 1$
3. $m = 1, n > 1$

On s'intéresse aux **fonctions linéaires**
i.e. $Y_k = g_k(X) = aX_k + b$

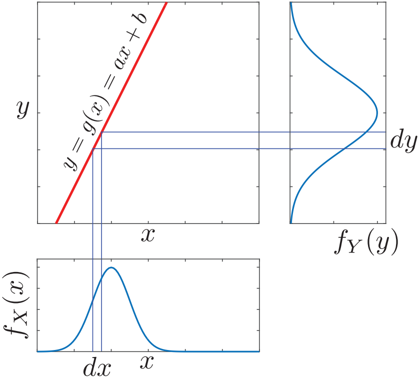
$Y = g(X)$: Cas 1, $m = n = 1$

$$f_Y(y)dy = f_X(x)dx$$

$$f_Y(y) = \underbrace{f_X(x)\left|\frac{dx}{dy}\right|}_{f_X(x)\geq 0}$$

Pour $Y = aX + b$

$$\left|\frac{dy}{dx}\right| = |a|$$



$$\{X : \mu_X, \sigma_X\} \rightarrow \{Y : a\mu_X + b, |a|\sigma_X\}$$

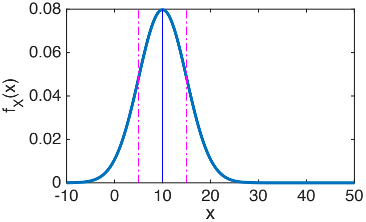
Exemple: Cas 1, $Y = 2X$

Variable aléatoire:

$$X \sim \mathcal{N}(\underbrace{10}_{\mu_X}, \underbrace{5^2}_{\sigma_X^2})$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]$$

$$\frac{dy}{dx} = 2$$



Transformation $X \rightarrow Y$:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = f_X(x)\left|\frac{dy}{dx}\right|^{-1} = f_X(\underbrace{\tfrac{y}{2}}_{x=y/2}) \cdot \frac{1}{2}$$

$$\boxed{f_Y(y) = \frac{1}{2\sqrt{2\pi}\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{\frac{y}{2} - \mu_X}{\sigma_X}\right)^2\right], \quad Y \sim \mathcal{N}(\underbrace{2\mu_X}_{\mu_Y}, (\underbrace{2\sigma_X}_{\sigma_Y})^2)}$$

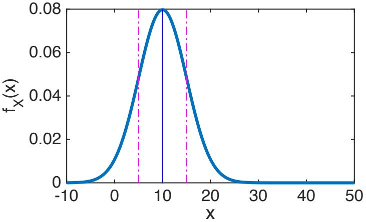
Exemple: Cas 1, $Y = 2X$

Variable aléatoire:

$$X \sim \mathcal{N}(\underbrace{10}_{\mu_X}, \underbrace{5^2}_{\sigma_X^2})$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]$$

$$\frac{dy}{dx} = 2$$



Transformation $X \to Y$:

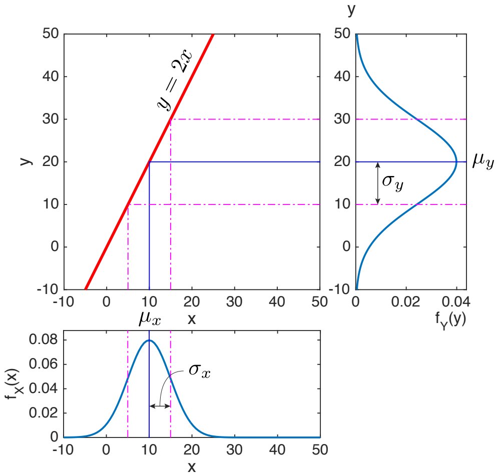
$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = f_X(x)\left|\frac{dy}{dx}\right|^{-1} = f_X(\underbrace{\tfrac{y}{2}}_{x=y/2}) \cdot \frac{1}{2}$$

$$\boxed{f_Y(y) = \frac{1}{\sqrt{2\pi}2\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{y-2\mu_X}{2\sigma_X}\right)^2\right], \quad Y \sim \mathcal{N}(\underbrace{2\mu_X}_{\mu_Y}, (\underbrace{2\sigma_X}_{\sigma_Y})^2)}$$

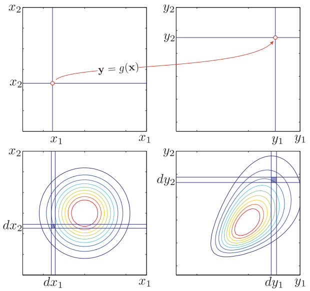
Exemple: $Y = 2X$ (cont.) [.m]

$$X \sim \mathcal{N}(\underbrace{10}_{\mu_X}, \underbrace{5^2}_{\sigma_X^2})$$

$$Y \sim \mathcal{N}(\underbrace{2\mu_X}_{\mu_Y}, (\underbrace{2\sigma_X}_{\sigma_Y})^2)$$



$\mathbf{Y} = \mathbf{g}(\mathbf{X})$: Cas 2, $m = n > 1$



$$f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = f_{\mathbf{Y}}(\mathbf{y})d\mathbf{y}$$

$$f_{\mathbf{X}}(\mathbf{x})\frac{d\mathbf{x}}{d\mathbf{y}} = f_{\mathbf{Y}}(\mathbf{y})$$

$$\boxed{J_{k,i} = J_{y_k,x_i} = \frac{\partial y_k}{\partial x_i}}$$

Jacobien ($\mathbf{J}_{\mathbf{y},\mathbf{x}}$): taille du "voisinage" de $d\mathbf{y}$ par rapport à $d\mathbf{x}$

$$\mathbf{J}_{\mathbf{y},\mathbf{x}} = \begin{bmatrix} J_{y_1,x_1} & \cdots & J_{y_1,x_n} \\ \vdots & \ddots & \vdots \\ J_{y_m,x_1} & \cdots & J_{y_m,x_n} \end{bmatrix} = \begin{bmatrix} \nabla g_1(\mathbf{x}) \\ \vdots \\ \nabla g_n(\mathbf{x}) \end{bmatrix}$$

$$\frac{d\mathbf{x}}{d\mathbf{y}} = |\det \mathbf{J}_{\mathbf{y},\mathbf{x}}|^{-1}$$

$$\frac{d\mathbf{y}}{d\mathbf{x}} = |\det \mathbf{J}_{\mathbf{y},\mathbf{x}}|$$

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{x})\,|\det \mathbf{J}_{\mathbf{y},\mathbf{x}}|^{-1}$$

$\mathbf{Y} = \mathbf{g}(\mathbf{X})$: Cas 2, $m = n > 1$ (cont.)

Soit une fonction linéaire $\mathbf{Y} = \mathbf{g}(\mathbf{X}) = \mathbf{AX} + \mathbf{b}$ où

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad \mathbf{M_X} = \begin{bmatrix} \mu_{X_1} \\ \vdots \\ \mu_{X_n} \end{bmatrix}, \quad \mathbf{\Sigma_{XX}} = \begin{bmatrix} \sigma^2_{X_1} & \cdots & \rho_{1n}\sigma_{X_1}\sigma_{X_n} \\ & \vdots & \vdots \\ \text{sym.} & & \sigma^2_{X_n} \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} \end{bmatrix}_{n\times 1}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \end{bmatrix}_{n\times n}}_{\mathbf{A}=\mathbf{J_{y,x}}} \times \underbrace{\begin{bmatrix} \end{bmatrix}_{n\times 1}}_{\mathbf{X}} + \underbrace{\begin{bmatrix} \end{bmatrix}_{n\times 1}}_{\mathbf{b}}$$

$$\boxed{\begin{aligned} \mathbf{M}_Y &= \mathbf{g}(\mathbf{M}_X) = \mathbf{AM}_X + \mathbf{b} \\ \mathbf{\Sigma}_{YY} &= \mathbf{A\Sigma}_{XX}\mathbf{A}^\intercal = \mathbf{J_{y,x}}\mathbf{\Sigma}_{XX}\mathbf{J_{y,x}}^\intercal \end{aligned}}$$

$Y = g(\mathbf{X})$: Cas 3, $m = 1$, $n > 1$

Soit une fonction linéaire $Y = g(\mathbf{X}) = \mathbf{AX} + b$ où

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad \mathbf{M_X} = \begin{bmatrix} \mu_{X_1} \\ \vdots \\ \mu_{X_n} \end{bmatrix}, \quad \mathbf{\Sigma_{XX}} = \begin{bmatrix} \sigma^2_{X_1} & \cdots & \rho_{1n}\sigma_{X_1}\sigma_{X_n} \\ & \vdots & \vdots \\ \text{sym.} & & \sigma^2_{X_n} \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} \quad \end{bmatrix}_{1\times 1}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \qquad\qquad \end{bmatrix}_{1\times N}}_{\mathbf{A}=\nabla g(\mathbf{x})} \times \underbrace{\begin{bmatrix} \quad \\ \\ \\ \end{bmatrix}_{N\times 1}}_{\mathbf{X}} + \underbrace{\begin{bmatrix} \quad \end{bmatrix}_{1\times 1}}_{b}$$

$$\boxed{\begin{aligned} \mu_Y &= g(\mathbf{M_X}) = \mathbf{AM}_X + b \\ \sigma^2_Y &= \nabla g(\mathbf{X})\mathbf{\Sigma_{XX}}\nabla g(\mathbf{X})^\intercal \\ &= \mathbf{A\Sigma_{XX}A}^\intercal \end{aligned}}$$

où $\underbrace{\nabla g(\mathbf{x}) = \left[\frac{\partial g(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial g(\mathbf{x})}{\partial x_n}\right]_{\mathbf{x}}}_{\text{Gradient}}$

Fonctions non-linéaires: $Y = \mathbf{g}(\mathbf{X})$

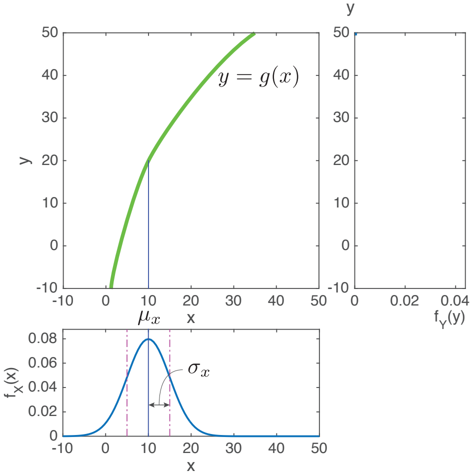
Que fait-on si **$\mathbf{g}(\mathbf{X})$ est non-linéaire?** **On linéarise.**

$$Y = \mathbf{g}(\mathbf{X}) \approx \mathbf{A}\mathbf{X} + \mathbf{b}$$

Plan de la section

Linéarisation

On remplace $g(X)$ par une fonction linéaire (e.g. une droite)

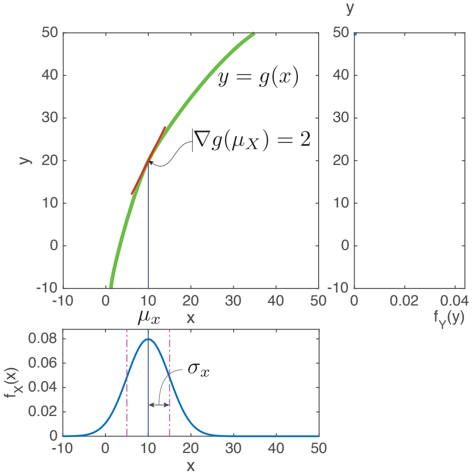


Linéarisation autour de la moyenne μ_x
⚠ **contenu de probabilité élevé**

$$\underbrace{\nabla g(\mu_x) = \left[\frac{dg(x)}{dx}\right]_{x=\mu_x}}_{\text{Gradient}}$$

Linéarisation

On remplace $g(X)$ par une fonction linéaire (e.g. une droite)

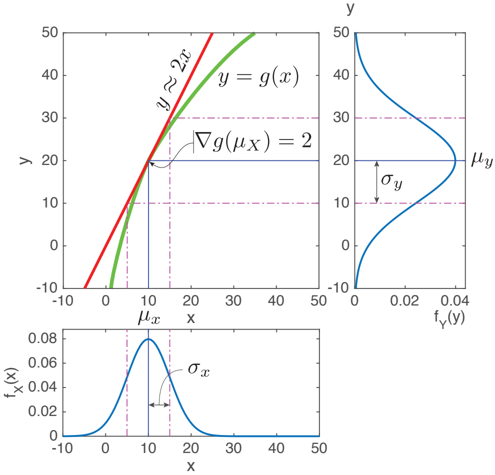


Linéarisation autour de la moyenne μ_x
⚠ **contenu de probabilité élevé**

$$\underbrace{\nabla g(\mu_x) = \left[\frac{dg(x)}{dx}\right]_{x=\mu_x}}_{\text{Gradient}}$$

Linéarisation

On remplace $g(X)$ par une fonction linéaire (e.g. une droite)



Linéarisation autour de la moyenne μ_x
⚠ **contenu de probabilité élevé**

$$\underbrace{\nabla g(\mu_x) = \left[\frac{dg(x)}{dx}\right]_{x=\mu_x}}_{\text{Gradient}}$$

Linéarisation – Série de Taylor développée autour de $\mathbf{M_X}$

Soit une fonction **non-linéaire** $Y = \mathbf{g(X)} \approx \mathbf{AX} + \mathbf{b}$

$$Y \approx \underbrace{\overbrace{\underbrace{g(\mathbf{M_X}) + \overbrace{\nabla g(\mathbf{M_X})}^{\text{Gradient}}(\mathbf{X}-\mathbf{M_X})}_{\text{approximation du } 1^{er} \text{ ordre}} + \frac{1}{2}(\mathbf{X}-\mathbf{M_X})^\intercal \overbrace{\mathbf{H}(\mathbf{M_X})}^{\text{Hessien}}(\mathbf{X}-\mathbf{M_X})}^{\text{approximation du } 2^{eme} \text{ ordre}} + \cdots}_{\text{approximation du } N^{eme} \text{ ordre}}$$

Note: Pour une fonction non-linéaire, le gradient/Jacobien varie en fonction du point où il est évalué $\rightarrow \nabla\mathbf{G}(\mathbf{M_X})$ / $\mathbf{J_{Y,X}}(\mathbf{M_X})$

Linéarisation – Approximation du premier ordre

Approximation du premier ordre centré sur la moyenne $\equiv$ linéarisation de $g(\mathbf{X})$ à $\mathbf{M_X}$

$$\begin{aligned} Y = g(\mathbf{X}) &\cong g(\mathbf{M_X}) + \nabla g(\mathbf{M_X})(\mathbf{X} - \mathbf{M_X}) \\ \mu_Y &\cong g(\mathbf{M_X}) \\ \sigma_Y^2 &\cong \nabla g(\mathbf{M_X})\mathbf{\Sigma_{XX}}\nabla g(\mathbf{M_X})^{\intercal} \end{aligned}$$

où

$$\underbrace{\nabla g(\mathbf{M_X}) = \left[\frac{\partial g(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial g(\mathbf{x})}{\partial x_n}\right]_{\mathbf{x}=\mathbf{M_X}}}_{\text{Gradient}}$$

Importance des variables, $\text{Imp}(X_i \to Y)$

$$\text{Imp}(X_i \to Y) \propto |a_i|\sigma_{X_i}$$

$$\underbrace{\left[\begin{array}{ccc} a_1 & \cdots & a_n \end{array}\right]_{1\times n}}_{\mathbf{A}=[\frac{\partial y}{\partial x_i}]=\nabla g(\mathbf{x})} \quad \mathbf{\Sigma_{XX}} = \left[\begin{array}{ccc} \sigma^2_{X_1} & \cdots & \rho_{1n}\sigma_{X_1}\sigma_{X_n} \\ & \vdots & \vdots \\ \text{sym.} & & \sigma^2_{X_n} \end{array}\right]$$

Résumé

Probabilités:

- Les probabilités **représentent notre connaissance**
- Le moins on en sait, le plus on devrait utiliser les probabilités

Interpretation Bayesienne: $\Pr(E_i)$ est une mesure de vraisemblance par rapport aux autres évènements dans $\mathcal{S}$

Lois/opération évènements: $\cap, \cup, \subset, \subseteq, \in$

Axiomes fondamentaux:

1. $0 \leq \Pr(E_i) \leq 1$
2. $\Pr(\mathcal{S}) = 1$
3. Si E_1 et E_2 sont mutuellement exclusifs $\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$

Loi d'inclusion-exclusion: $\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \ldots$

Théorème de Bayes: $\Pr(E_i|A) = \frac{\Pr(A|E_i)\Pr(E_i)}{\Pr(A)}$

Densités de probabilités: PDF, CDF, PMF, CMF

Normale multivariée: $\mu_1, \sigma_1, \mu_2, \sigma_2, \rho_{12}$

Densités de probabilité multivariées:

- $0 \leq f_{\mathbf{X}}(\mathbf{x})$
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 1$

Probabilités conditionnelles:

- si $p_{X_1|X_2}(x_1|x_2) = p_{X_1}(x_1)$, $X_1 \perp\!\!\!\perp X_2$
- si $X_1 \perp\!\!\!\perp X_2$ $p_{X_1 X_2}(x_1, x_2) = p_{X_1}(x_1) p_{X_2}(x_2)$

Cas général: $X_1 \not\!\perp\!\!\!\perp X_2 \rightarrow$ **dérivation en chaine**

Espérance & Variance:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx \quad (\text{V.A. continue})$$

$$\mathbb{E}[(X - \mu_X)^2] = \sigma_X^2 = \text{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Notation matricielle: $\mathbf{\Sigma_{XX}} = \mathbf{D_{XX} R_{XX} D_{XX}}$

Fonctions de variables aléatoires: $Y = g(\mathbf{X})$

$$f_Y(y) dy = f_X(x) dx$$

$$\begin{aligned} \mathbf{M}_Y &= \mathbf{g}(\mathbf{M}_X) = \mathbf{A}\mathbf{M}_X + \mathbf{b} \\ \mathbf{\Sigma}_{YY} &= \mathbf{A}\mathbf{\Sigma}_{XX}\mathbf{A}^{\mathsf{T}} = \mathbf{J_{y,x}}\mathbf{\Sigma}_{XX}\mathbf{J_{y,x}}^{\mathsf{T}} \end{aligned}$$

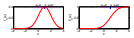
Linéarisation – Approximation du premier ordre

$$\begin{aligned} Y = g(\mathbf{X}) &\cong g(\mathbf{M_X}) + \nabla g(\mathbf{M_X})(\mathbf{X} - \mathbf{M_X}) \\ \mu_Y &\cong g(\mathbf{M_X}) \\ \sigma_Y^2 &\cong \nabla g(\mathbf{M_X}) \mathbf{\Sigma_{XX}} \nabla g(\mathbf{M_X})^{\mathsf{T}} \end{aligned}$$

Importance des variables

$$\text{Imp}(X_i \rightarrow Y) \propto |a_i| \sigma_{X_i}$$

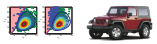
Théorie probabilité

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

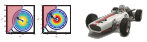
Fiabilité des structures & systèmes

- Introduction
- 2 Formulation fiabilité composantes 
- 9 Formulation fiabilité systèmes 

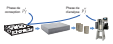
Estimation – p_f échantillonnage

- 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