Probabilités

Julien Lesouple[1] et Jean-Yves Tourneret[2]

(1) Université de Toulouse, ENAC

(2) Université de Toulouse, INP/ENSEEIHT-IRIT

julien.lesouple@enac.fr, jyt@n7.fr

Images optique et radar

Plan du cours

- Chapitre 1 : Eléments de base du calcul des probabilités
  - Triplet de Probabilité (Ω, C, P)
  - Équiprobabilité - Dénombrement
  - Probabilités conditionnelles
  - Indépendance
- Chapitre 2 : Variables aléatoires réelles
- Chapitre 3 : Couples de variables aléatoires réelles
- Chapitre 4 : Vecteurs Gaussiens
- Chapitre 5 : Convergence et théorèmes limites

Bibliographie

- B. Lacaze, M. Maubourguet, C. Mailhes et J.-Y. Tourneret, Probabilités et Statistique appliquées, Cépadues, 1997.
- Athanasios Papoulis and S. Unnikrishna Pillai, Probability, Random Variable and Stochastic Processes, McGraw Hill Higher Education, 4th edition, 2002.

Triplet de Probabilité $(\Omega, \mathcal{C}, P)$

- Ω : Ensemble des résultats d'expérience
- $\mathcal{C}$: Ensemble des événements
  - $\mathcal{C} \subset \mathcal{P}(\Omega)$
  - $\Omega \in \mathcal{C}$ (événement certain)
  - si $A \in \mathcal{C}$ alors $\overline{A} \in \mathcal{C}$ (événement contraire)
  - si $A_i \in \mathcal{C}, i \in I$ (I fini ou infini dénombrable), alors $\cup A_i \in \mathcal{C}$
- P : application probabilité de $\mathcal{C}$ dans $[0, 1]$
  - $P(\Omega) = 1$
  - $P\left(\overline{A}\right) = 1 - P(A)$
  - $P\left(\cup A_i\right) = \sum_{i \in I} P(A_i)$ si les événements A_i sont disjoints.

Propriétés

- Événements
  - $\emptyset \in \mathcal{C}$
  - si $A_i \in \mathcal{C}, i \in I$ (I fini ou infini dénombrable), alors $\cap A_i \in \mathcal{C}$
- Probabilité
  - $P(\emptyset) = 0$
  - si $A \subset B$, alors, $P\left(A\right) \leq P(B)$
  - $P\left(A \cup B\right) = P(A) + P(B) - P(A \cap B)$

Vocabulaire

- si $a \in \Omega$ alors $\{a\}$ est un événement élémentaire
- si $\Omega = \cup_{i \in I} A_i$ avec $A_i \cap A_j = \emptyset$, on dit que $\{A_i\}_{i \in I}$ est un système complet d'événements
- $(\Omega, \mathcal{C})$ espace probabilisable
- $(\Omega, \mathcal{C}, P)$ espace probabilisé
- $\mathcal{C}$ tribu ou σ-algèbre

Équiprobabilité - Dénombrement

- Définition

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}}$$

- Exemples
  - Jet d'un dé
  - Tirages avec remise dans une urne à 2 catégories

$$P(k \text{ succès sur } n \text{ expériences}) = \binom{n}{k} P_s^k (1 - P_s)^{n-k}$$

avec $k = 0, ..., n$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, P_s est la probabilité du succès sur une expérience et n est le nombre d'expériences identiques et indépendantes.

Probabilités conditionnelles

- Définition

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ ou } P(A \cap B) = P(A|B)P(B)$$

- Théorème des probabilités totales

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i)$$

pour tout système complet d'événements $\{A_i\}$.

- Formule de Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Indépendance

- Deux événements
  Deux événements A et B sont indépendants si et ssi

$$P(A \cap B) = P(A)P(B) \text{ ou } P(A|B) = P(A)$$

- Généralisation
  On dit que $\{A_i\}_{i \in I}$ est famille d'événements mutuellement indépendants si et ssi

$$P\left(\cap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i), \; \forall J \subset I$$

☞ **Exercice d'application**

Que faut-il savoir ?

- Probabilité d'une réunion d'événements : $P\left(A \cup B\right) =?$
- Probabilité de l'évènement contraire : $P\left(\overline{A}\right) =?$
- Equiprobabilité : $P(A) =?$
- Loi Binomiale : $P(k \text{ succès sur } n \text{ expériences}) =?$
- Probabilité conditionnelle : $P(A|B) =?$
- Indépendance : $P(A \cap B) =?$
- Formule de Bayes : $P(A|B) =?$

Plan du cours

- Chapitre 1 : Eléments de base du calcul des probabilités
- Chapitre 2 : Variables aléatoires réelles
  - Définition
  - Loi d'une variable aléatoire
  - Fonction de répartition
  - Exemples fondamentaux
  - Espérance mathématique
  - Changements de variables
- Chapitre 3 : Couples de variables aléatoires réelles
- ...

Variable aléatoire réelle

- Définition
  Soient $(\Omega, \mathcal{C}, P)$ un triplet de probabilité qui est associé à l'expérience et $(\Omega', \mathcal{C}')$, avec $\Omega' \subset \mathbb{R}$ un espace probabilisable qui résume les quantités qui nous intéressent. Une variable aléatoire réelle X est une application de Ω dans Ω' qui possède la propriété de mesurabilité :

$$\forall (a, b) \in \mathcal{C}', \{\omega | X(\omega) \in (a, b)\} \in \mathcal{C}.$$

- Exemple : somme des résultats de deux dés

$$X : \begin{array}{rcl} \Omega & \longrightarrow & \Omega' \\ (m, n) & \longmapsto & m + n \end{array}$$

Variable aléatoire discrète

- Loi d'une variable aléatoire discrète
  $\{X(\omega), \omega \in \Omega\}$ est fini ou infini dénombrable. La loi de X est définie par
  - l'ensemble des valeurs possibles de X : $\{x_i, i \in I\}$
  - les probabilités associées $p_i = P[X = x_i]$ avec

$$\sum_{i \in I} p_i = 1 \text{ et } P[X \in \Delta] = \sum_{x_i \in \Delta} p_i$$

- Exemples
  - Jet d'un dé
  - Jet d'une pièce
  - ...

Variables aléatoires continues

- Loi d'une variable aléatoire continue
  $\{X(\omega), \omega \in \Omega\}$ est infini non dénombrable avec $P[X = x_i] = 0, \forall x_i$. La loi de X est définie par
  - l'ensemble des valeurs possibles de X qui est en général une réunion d'intervalles
  - une densité de probabilité p : $\begin{array}{c} \mathbb{R} \to \mathbb{R} \\ x \longmapsto p(x) \end{array}$ telle que

$$p(x) \geq 0, \forall x \in \mathbb{R},$$

$$\int_{\mathbb{R}} p(u)du = 1,$$

$$P[X \in \Delta] = \int_{\Delta} p(u)du.$$

Variables aléatoires continues

- Remarques
  - On peut avoir $p(x) > 1$.
  - $P\left[X \in [x, x+dx[\right] \simeq p(x)dx$ pour dx "petit" [a]
  - lien avec l'histogramme
- Exemples
  - Loi uniforme sur $[a, b]$
  - Loi normale

---

[a] On peut montrer que $\lim\limits_{dx \to 0} \frac{P[X \in [x, x+dx[]}{dx} = p(x)$ en utilisant la fonction de répartition F telle que $F(x) = P[X \leq x]$, qui sera introduite plus tard.

Variable aléatoire mixte

- Loi d'une variable aléatoire mixte
  $\{X(\omega), \omega \in \Omega)\} = E \cup \{x_i, \in I\}$ est la réunion de deux ensembles, le premier E est infini non dénombrable avec $P[X = x] = 0, \forall x \in E$, le deuxième est fini ou infini dénombrable avec $p_i = P[X = x_i] > 0$. La loi de X est définie par
  - $\{x_i, \in I\}$ avec $p_i = P[X = x_i] > 0$
  - E et une densité de probabilité p telle que

$$p(x) \geq 0, \forall x \in \mathbb{R}$$
$$\int_{\mathbb{R}} p(u)du + \sum_{i \in I} p_i = 1$$
$$P[X \in \Delta] = \int_{\Delta} p(u)du + \sum_{x_i \in \Delta} p_i$$

- Exemple : Tension aux bornes d'un voltmètre

Exemples Fondamentaux de Lois Discrètes

- Loi de Bernoulli : $X \sim \mathcal{B}e(p)$

$$P[X = 1] = p \text{ et } P[X = 0] = q = 1 - p$$

Lancer d'une pièce, "Succès ou Echec", ...

- Loi binomiale : $X \sim \mathcal{B}(n, p)$

$$P[X = k] = \binom{n}{k} p^k q^{n-k}, \quad k = 0, ..., n$$

Probabilité d'avoir k succès sur n expériences, $X = \sum_{i=1}^{n} X_i$ où X_i suit une loi de Bernoulli, ...

- Loi de Poisson : $X \sim \mathcal{P}(\lambda)$

$$P[X = k] = \frac{\lambda^k}{k!} \exp(-\lambda), \quad k \in \mathbb{N}$$

Loi du nombre d'arrivées pendant un temps donné

Exemples Fondamentaux de Lois Continues

- Loi uniforme : $X \sim \mathcal{U}\left([a,b]\right)$

$$p(x) = \frac{1}{b-a}, \quad x \in [a,b]$$

- Loi normale ou Gaussienne : $X \sim \mathcal{N}(m, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \quad x \in \mathbb{R}$$

- Loi gamma : $X \sim \mathcal{G}a(\alpha, \beta)$

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x > 0$$

Pour $\alpha = 1$, on a la loi exponentielle

LOIS DE PROBABILITÉ DISCRÈTES

$\mathbf{m}$: moyenne $\quad \sigma^2$: variance $\quad$ **F. C. :** fonction caractéristique

$p_k = P[X = k] \qquad p_{1,\dots,m} = P[X_1 = k_1, \dots, X_m = k_m]$

LOI	Probabilités	m	σ^2	F. C.
Uniforme	$p_k = \frac{1}{n}$ $k \in \{1, \dots, n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\dfrac{e^{it}\left(1 - e^{itn}\right)}{n\left(1 - e^{it}\right)}$
Bernoulli	$p_1 = P[X = 1] = p$ $p_0 = P[X = 0] = q$ $p \in [0,1] \quad q = 1 - p$	p	pq	$pe^{it} + q$
Binomiale $\mathcal{B}(n,p)$	$p_k = \frac{n!}{k!(n-k)!} p^k q^{n-k}$ $p \in [0,1] \quad q = 1 - p$ $k \in \{0, 1, \dots, n\}$	np	npq	$\left(pe^{it} + q\right)^n$
Binomiale négative	$p_k = \frac{(n+k-1)!}{(n-1)!k!} p^n q^k$ $p \in [0,1] \quad q = 1 - p$ $k \in \mathbb{N}$	$n\dfrac{q}{p}$	$n\dfrac{q}{p^2}$	$\left(\dfrac{p}{1 - qe^{it}}\right)^n$
Multinomiale	$p_{1,\dots,m} = \frac{n!}{k_1! \dots k_m!} p_1^{k_1} \dots p_m^{k_m}$ $p_j \in [0,1] \quad q_j = 1 - p_j$ $k_j \in \{0, 1, \dots, n\}$ $\sum_{j=1}^m k_j = n \quad \sum_{j=1}^m p_j = 1$	np_j	Variance : $np_j q_j$ Covariance : $-np_j p_k$	$\left(\sum_{j=1}^m p_j e^{it}\right)^n$
Poisson $\mathcal{P}(\lambda)$	$p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$ $\lambda > 0 \quad k \in \mathbb{N}$	λ	λ	$\exp\left[\lambda\left(e^{it} - 1\right)\right]$
Géométrique	$p_k = pq^{k-1}$ $p \in [0,1] \quad q = 1 - p$ $k \in \mathbb{N}^*$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{pe^{it}}{1 - qe^{it}}$

LOIS DE PROBABILITÉ CONTINUES

$\mathbf{m}$: moyenne σ^2 : variance **F. C. :** fonction caractéristique

LOI	Densité de probabilité	m	σ^2	F. C.
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\Gamma(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Inverse gamma IG(θ, ν)	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\frac{\theta}{x}} \frac{1}{x^{\nu+1}}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{\theta}{\nu-1}$ si $\nu > 1$	$\frac{\theta^2}{(\nu-1)^2(\nu-2)}$ si $\nu > 2$	(*)
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}$	0	2	$\dfrac{1}{1+t^2}$
Normale $\mathcal{N}(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Khi$_2$ χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\lvert t \rvert}$
Beta $B(a, b)$	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0$ $x \in]0, 1[$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

Fonction de répartition

- Définition

$$F : \begin{array}{c} \mathbb{R} \rightarrow [0,1] \\ x \longmapsto F(x) = P[X \leq x] \end{array}$$

- Propriétés
  - F croissante
  - $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$
  - F caractérise une loi de probabilité
  - Si X est une va discrète, le graphe de F est une fonction en escaliers
  - Si X est une va continue, F est continue et $F(x) = \int_{-\infty}^{x} p(u)du$, i.e., $p(x) = F'(x)$

Espérance mathématique

- Définition

$$E[\alpha(X)] = \begin{cases} X \text{ va discrète : } & \sum_{i\in I} \alpha(x_i)p_i \\ X \text{ va continue : } & \int_{\mathbb{R}} \alpha(u)p(u)du \\ X \text{ va mixte : } & \sum_{i\in I} \alpha(x_i)p_i + \int_{\mathbb{R}} \alpha(u)p(u)du \end{cases}$$

- Propriétés
  - Constante : $E(\text{cste}) = \text{cste}$
  - Linéarité : $E(aX+b) = aE(X)+b$
- Exemples
  - Moments non centrés : $E(X^n)$ ($n=1$: moyenne)
  - Moments centrés : $E\left([X-E(X)]^n\right)$ ($n=2$: variance)
  - Fonction caractéristique : $\phi_X(t) = E\left[\exp(itX)\right]$

Exemples simples

- Variables aléatoires discrètes

$$E[X] = \sum_{i \in I} x_i P[X = x_i], \quad E\left[X^2\right] = \sum_{i \in I} x_i^2 P[X = x_i]$$

$$E\left[e^{jtX}\right] = \sum_{i \in I} e^{jtx_i} P[X = x_i]$$

- Variables aléatoires continues

$$E[X] = \int_{\mathbb{R}} u p(u) du, \quad E\left[X^2\right] = \int_{\mathbb{R}} u^2 p(u) du$$

$$E\left[e^{jtX}\right] = \int_{\mathbb{R}} e^{jtu} p(u) du$$

Propriétés

- Variance
  - var$(X) = E\left([X - E(X)]^2\right) = E(X^2) - E(X)^2$
  - Ecart Type : $\sqrt{\text{variance}}$
  - var$(aX + b) = a^2$var(X)
- Fonction caractéristique
  - Caractérise une loi de probabilité
  - Cas continu

$$\phi_X(t) = \int_{\mathbb{R}} e^{itu} p(u) du$$

est la transformée de Fourier de p.

- Exemples de calculs

Changements de variables

- Problème

  Étant donnée une variable aléatoire réelle X de loi connue, on cherche à déterminer la loi de $Y = g(X)$ où g est une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

- Variables aléatoires discrètes
  - Définition

$$P[Y = y_j] = \sum_{i | y_j = g(x_i)} p[X = x_i]$$

  - Exemple

    $Y = (X - 2)^2$ avec $X \sim \mathcal{P}(\lambda)$

Changements de va continues

- g bijective
  - Théorème : si X est une va continue à valeurs dans un ouvert $O_X \subset \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ application bijective de O_X dans un ouvert $O_Y \subset \mathbb{R}$ différentiable ainsi que son inverse g^{-1}, alors $Y = g(X)$ est une va continue de densité

$$p_Y(y) = p_X\left[g^{-1}(y)\right]\left|\frac{dx}{dy}\right|.$$

    où $\frac{dx}{dy}$ est le Jacobien de la transformation.
  - Exemple 1 : $Y = 1/X$ avec $X \sim \mathcal{E}(1)$.
  - Idée de preuve et preuve
  - Exemple 2 : $Y = aX + b$ avec $X \sim \mathcal{N}(m, \sigma^2)$.

Changements de va continues

- g bijective par morceaux
  On suppose que g est différentiable sur chaque morceau ainsi que son inverse.
  - Méthode : On ajoute la contribution de chaque bijection.
  - Exemple : $Y = X^2$ avec $X \sim \mathcal{N}(0, 1)$.
- g non bijective et non bijective par morceaux
  - Détermination de la fonction de répartition
  - Détermination de la fonction caractéristique

Que faut-il savoir ?

- Loi d'une variable aléatoire discrète : ?
- Loi d'une variable aléatoire continue : ?
- Appartenance à un intervalle : $P[X \in \Delta] =?$
- Signification d'une densité : $P\left[X \in [x, x+dx[\right] \simeq ?$
- Fonction de répartition : $F(x) =?$
- Espérance mathématique : $E[X] =?$, $E[X^2] =?$
- Variance : $\text{Var}[X] =?$, Ecart-type : ?
- Relations utiles : $E[aX+b] =?$, $\text{Var}[aX+b] =?$
- Fonction caractéristique : $\phi(t) =?$
- Changement de variables : ?

Plan du cours

- Chapitre 1 : Eléments de base du calcul des probabilités
- Chapitre 2 : Variables aléatoires réelles
- Chapitre 3 : Couples de variables aléatoires réelles
  - Définition
  - Fonction de répartition
  - Lois marginales, lois conditionnelles, indépendance
  - Espérances mathématiques
  - Changements de variables
- Chapitre 4 : Vecteurs Gaussiens
- Chapitre 5 : Convergence et théorèmes limites

Couple de va réelles

- Définition
  Soit (Ω, C, P) un espace probabilisé et (Ω', C') un espace probabilisable avec $\Omega' \subset \mathbb{R}^2$ et C' construit à partir des réunions et intersections finies ou dénombrables des pavés $(a, b) \times (c, d)$ de $\mathbb{R}^2$. Un couple (X, Y) de variables aléatoires réelles est une application mesurable de Ω dans Ω'.
- notation
  On notera $P\left[(X, Y) \in \Delta\right], \Delta \subset \mathbb{R}^2$, la probabilité que le couple (X, Y) prenne ses valeurs dans Δ.

Loi d'un couple de va

- Variables aléatoires discrètes
  La loi du couple (X, Y) est définie par l'ensemble des valeurs possibles du couple (qui est un ensemble fini ou dénombrable) noté $\{(x_i, y_j), i \in I, j \in J\}$ et par les probabilités associées $p_{ij} = P[X = x_i, Y = y_j]$, $i \in I, j \in J$ telles que $p_{ij} > 0$ et $\sum_{i,j} p_{ij} = 1$.
- Variables aléatoires continues
  La loi du couple (X, Y) est définie par l'ensemble des valeurs possibles du couple (qui est un ensemble infini non dénombrable), en général une réunion d'intervalles de $\mathbb{R}^2$, et par une densité de probabilité $p(x, y)$ telle que

$$p(x, y) \geq 0, \text{ et } \int\int_{\mathbb{R}^2} p(x, y)dxdy = 1.$$

Loi d'un couple de va

- Variables aléatoires discrètes et continues
  La loi du couple (X, Y), où X est discrète à valeurs dans $\{x_i, i \in I\}$ et Y est continue, est définie par le domaine de définition du couple et par un ensemble de densités de probabilités $p_i(y), i \in I$ tel que

$$p_i(y) \geq 0, \text{ et } \sum_{i \in I} \int_{\mathbb{R}} p_i(y) dy = 1.$$

Propriétés

- Couples de va discrètes

$$P\left[(X,Y)\in\Delta\right]=\sum_{(i,j)|(x_i,y_j)\in\Delta}P[X=x_i,Y=y_j],\quad \Delta\subset\mathbb{R}^2$$

- Couples de va continues

$$P\left[(X,Y)\in\Delta\right]=\int\int_{\Delta}p(u,v)dudv,\quad \Delta\subset\mathbb{R}^2$$

Remarque : signification de $p(u,v)$

- Couples de va discrètes et continues

$$P\left[(X,Y)\in\Delta\times\Delta'\right]=\sum_{i|x_i\in\Delta}\int_{\Delta'}p_i(v)dv,\quad \Delta\times\Delta'\subset\mathbb{R}^2$$

Fonction de répartition

- Définition

$$F : \begin{array}{l} \mathbb{R}^2 \to [0,1] \\ (x,y) \longmapsto F(x,y) = P[X \leq x, Y \leq y] \end{array}$$

- Propriétés
  - C'est une fonction étagée lorsque (X,Y) est un couple de va discrètes
  - C'est une fonction continue lorsque (X,Y) est un couple de va continues avec

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u,v)dudv \text{ d'où } p(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

Lois marginales

- Cas discret

$$P[X = x_i] = p_{i.} = \sum_{j \in J} p_{ij}$$

$$P[Y = y_j] = p_{.j} = \sum_{i \in I} p_{ij}$$

- Cas continu

densité de X : $p(x, .) = \int_{\mathbb{R}} p(x, y) dy$

densité de Y : $p(., y) = \int_{\mathbb{R}} p(x, y) dx$

Lois marginales

- Cas discret

$$p_{00} = \frac{1}{2}, p_{01} = \frac{1}{6}, p_{10} = \frac{1}{6}, p_{11} = \frac{1}{6}.$$

Lois de X et de Y ?

- Cas continu

$$p(x, y) = \begin{cases} \theta^2 e^{-\theta x} \text{ si } x > y > 0 \\ 0 \text{ sinon} \end{cases}$$

Montrer que $X \sim \Gamma(\theta, 2)$ et que $Y \sim \Gamma(\theta, 1)$.

Lois conditionnelles

Les lois conditionnelles d'un couple (X, Y) sont les lois de $X|\, Y = y$ et de $Y|\, X = x$.

- Cas discret

$$P\left[X = x_i|\, Y = y_j\right] = \frac{p_{ij}}{p_{.j}}$$

$$P\left[Y = y_j|\, X = x_i\right] = \frac{p_{ij}}{p_{i.}}$$

- Cas continu

densité de $X|\, Y = y \qquad p(x|\, y) = \dfrac{p(x,y)}{p(.,y)}$

densité de $Y|\, X = x \qquad p(y|\, x) = \dfrac{p(x,y)}{p(x,.)}$

Théorème de Bayes

- Cas discret

$$P\left[X = x_i \middle| Y = y_j\right] = \frac{P\left[Y = y_j \middle| X = x_i\right] P\left[X = x_i\right]}{P\left[Y = y_j\right]}$$

- Cas continu

$$p(x | y) = \frac{p(y | x)p(x, .)}{p(., y)}$$

Indépendance

Les variables aléatoires X et Y sont indépendantes si

$$P\left[X \in \Delta, Y \in \Delta'\right] = P\left[X \in \Delta\right] P\left[Y \in \Delta'\right], \forall \Delta, \forall \Delta'$$

- Cas discret

$$p_{ij} = p_{i.}p_{.j} \qquad \forall i \in I, \forall j \in J$$

- Cas continu

$$p(x, y) = p(x, .)p(., y) \; \forall x, \forall y$$

ou

$$p(x|\, y) = p(x, .), \; \forall x, \forall y$$

Propriété

si X et Y sont des variables aléatoires indépendantes et α et β sont des applications continues de $\mathbb{R}$ dans $\mathbb{R}$, alors $\alpha(X)$ et $\beta(Y)$ sont des variables aléatoires indépendantes. La réciproque est vraie si α et β sont des applications bijectives. Par contre, dans le cas où α et β ne sont pas bijectives, la réciproque est fausse. On vérifiera par exemple que le couple (X, Y) de densité

$$p(x,y) = \begin{cases} \frac{1}{4}(1+xy) \text{ si } |x|<1 \text{ et } |y|<1 \\ 0 \text{ sinon} \end{cases}$$

est tel que X^2 et Y^2 sont indépendantes alors que X et Y ne le sont pas.

Espérance mathématique

- Définition

$$E[\alpha(X,Y)] = \begin{cases} X \text{ et } Y \text{ va discrètes : } & \sum_{i,j \in I \times J} \alpha(x_i, y_j) p_{ij} \\ X \text{ et } Y \text{ va continues : } & \int_{\mathbb{R}^2} \alpha(u,v) p(u,v) du dv \end{cases}$$

- Propriétés
  - Constante : $E(\text{cste}) = \text{cste}$
  - Linéarité :
    $E\left[a\alpha(X,Y) + b\beta(X,Y)\right] = aE\left[\alpha(X,Y)\right] + bE\left[\beta(X,Y)\right]$
  - Définition cohérente (cas continu) :

$$E\left[\alpha(X)\right] = \int_{\mathbb{R}^2} \alpha(u) p(u,v) du dv = \int_{\mathbb{R}} \alpha(u) p(u,.) du$$

  - Indépendance : si X et Y sont indépendantes, alors

$$E\left[\alpha(X)\beta(Y)\right] = E\left[\alpha(X)\right] E\left[\beta(Y)\right], \ \forall \alpha \forall \beta$$

Exemples

- Moments centrés et non centrés

$$m_{ij} = E\left(X^i Y^j\right),\ i \in \mathbb{N}, j \in \mathbb{N}$$
$$\mu_{ij} = E\left([X - E(X)]^i [Y - E(Y)]^j\right),\ i \in \mathbb{N}, j \in \mathbb{N}$$

- Covariance et matrice de covariance

$$\text{cov}(X,Y) = E\left([X - E(X)][Y - E(Y)]\right) = E\left(XY\right) - E(X)E(Y)$$
$$E\left[\boldsymbol{V}\boldsymbol{V}^T\right] = \begin{pmatrix} \text{var}X & \text{cov}(X,Y) \\ \text{cov}(X,Y) & \text{var}Y \end{pmatrix},\ \boldsymbol{V} = \begin{pmatrix} X - E[X] \\ Y - E[Y] \end{pmatrix}$$

- Fonction caractéristique

$$\phi_{X,Y}(u_1,u_2) = E\left[\exp(i\boldsymbol{u}^T\boldsymbol{W})\right],\ \boldsymbol{u} = (u_1,u_2)^T, \boldsymbol{W} = (X,Y)^T.$$

Coefficient de Corrélation

- Définition

$$r(X,Y) = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y},$$

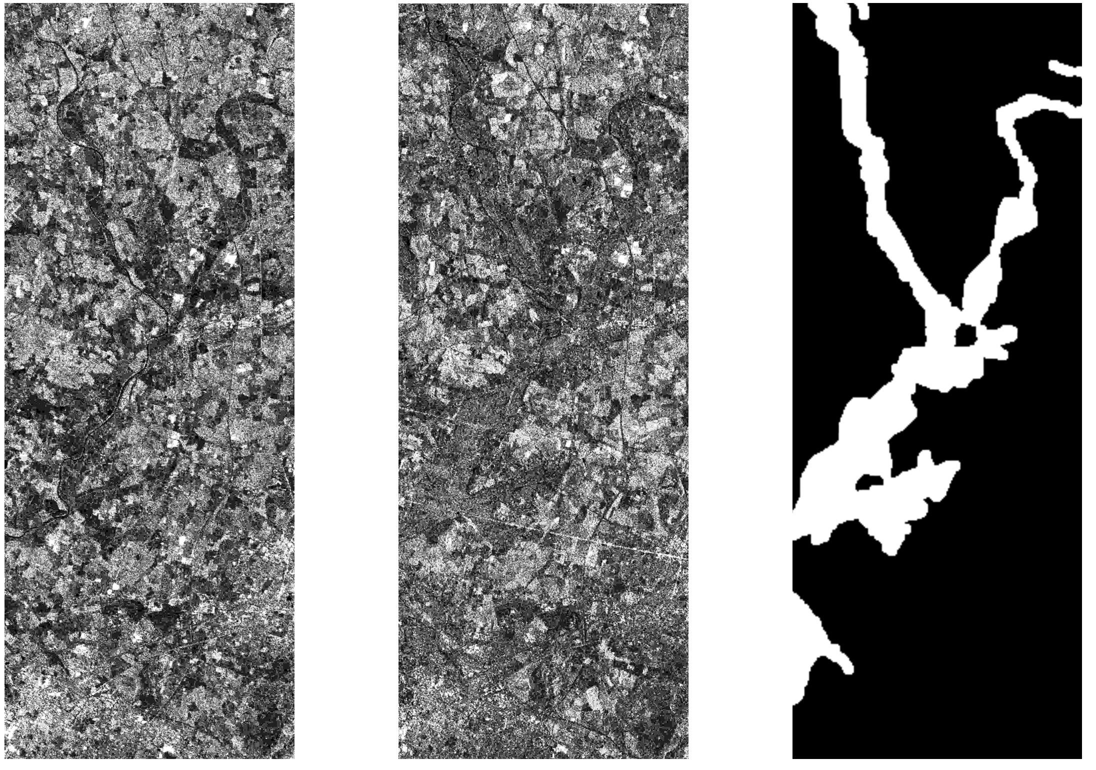
  où σ_X et σ_Y sont les écart-types des va X et Y.

- Propriétés
  - $-1 \leq r(X,Y) \leq 1$
  - $r(X,Y) = \pm 1$ si et ssi X et Y sont reliées par une relation affine
  - si X et Y sont des va indépendantes, alors $r(X,Y) = 0$ mais la réciproque est fausse

- Conclusion

  $r(X,Y)$ est une mesure imparfaite mais très pratique du lien entre les va X et Y.

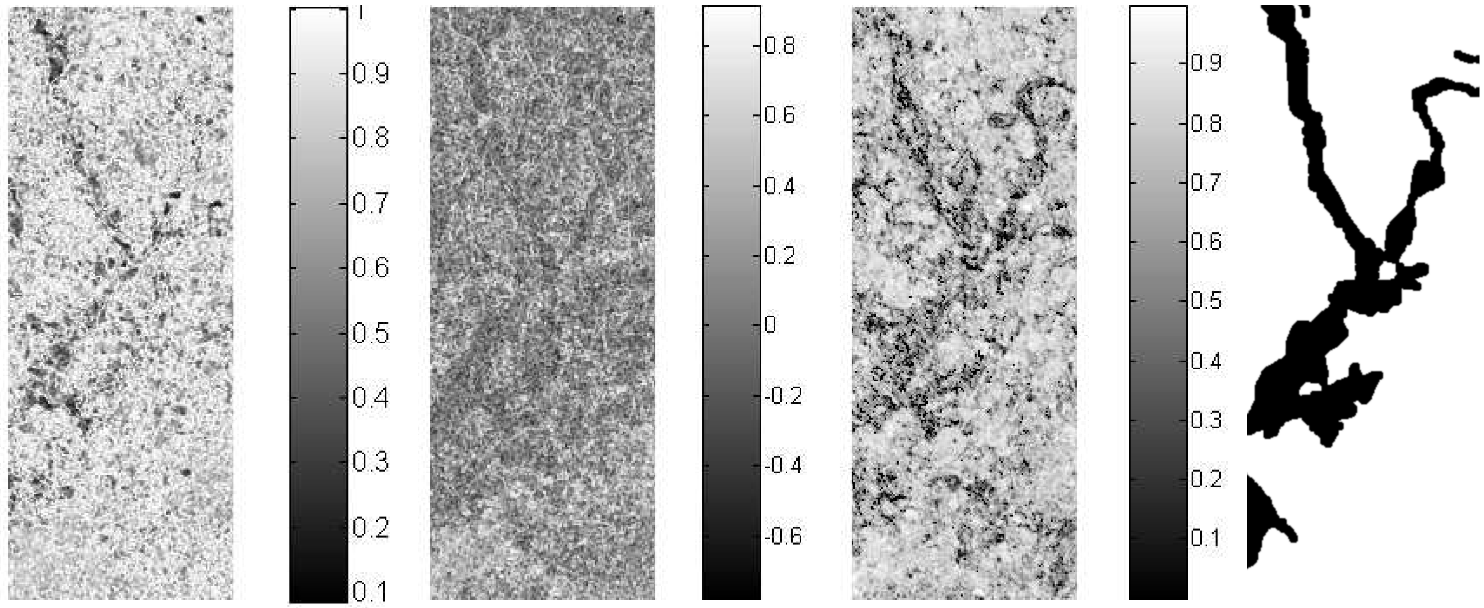
Données réelles : images de Gloucester



(a) Avant (b) Après (c) Masque

Images radar ERS (3049 × 1170 pixels) de Gloucester, Angleterre, avant et après une inondation

Application aux images de Gloucester : cartes de changement



(a) Ratio Edge (b) Correlation Moment (c) Correlation MV (d) Masque

Cartes de changement pour les images de Gloucester obtenues pour une fenêtre d'estimation de taille $n = 15 \times 15$

Espérance conditionnelle

Théorème

$$E\left[\alpha(X,Y)\right] = E_X\left[E_Y\left[\alpha(X,Y)|X\right]\right]$$

Exemples

- Exemple 1 : soit X et Y deux variables aléatoires indépendantes telles que $X \sim \mathcal{N}(0,1)$ et $P[Y=1] = p \in]0,1[= 1 - P[Y=-1]$. Déterminer la fonction caractéristique de $Z = XY$ et en déduire la loi de Z.
- Exemple 2 : déterminer $E[Y_N]$ lorsque

$$Y_N = \sum_{i=1}^{N} X_i$$

où $P[X_i = 1] = p$, $P[X_i = 0] = q = 1 - p$ et N est une va de loi de Poisson de paramètre λ indépendante des va X_i.

Changements de variables

- Problème
  Étant donné un couple de variables aléatoires réelles (X, Y) de loi connue, on cherche à déterminer la loi de $(U, V) = g(X, Y)$ où g est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ et U et V sont deux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$.
- Variables aléatoires discrètes
  - Définition

$$P[(U, V) = (u_k, v_l)] = \sum_{i,j|g(x_i,y_j)=(u_k,v_l)} p[X = x_i, Y = y_j]$$

  - Exemple
    voir TD

Changements de va continues de $\mathbb{R}^2 \to \mathbb{R}^2$

- Théorème pour g bijective
  si (X, Y) est un couple de va continues à valeurs dans un ouvert $O \subset \mathbb{R}^2$ et $g : \mathbb{R}^2 \to \mathbb{R}^2$ est une application bijective de O dans un ouvert $\Delta \subset \mathbb{R}^2$ continument différentiable ainsi que son inverse g^{-1}, alors $(U, V) = g(X, Y)$ est un couple de va continues de densité

$$p_{U,V}(u, v) = p_{X,Y}\left[g^{-1}(u, v)\right] |\text{det}(J)|,$$

  où J est la matrice Jacobienne définie par

$$J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

Changements de va continues de $\mathbb{R}^2 \to \mathbb{R}^2$

- Exemples
  - Exemple 1
    $X \sim \mathcal{U}(0,1)$, $Y \sim \mathcal{U}(0,1)$, X et Y indépendantes.
    Quelle est la loi du couple (U,V) avec $U = X + Y$ et $V = X$?
  - Exemple 2
    $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$, X et Y va indépendantes.
    Quelle est la loi de (R,Θ) avec $X = R\cos\Theta$ et $Y = R\sin\Theta$?
- Généralisation
  Si g est bijective par morceaux, on ajoute les contributions de chaque morceau.

Changements de va continues de $\mathbb{R}^2 \to \mathbb{R}$

- Problème
  Si (X, Y) est un couple de va continues de loi connue et $g : \mathbb{R}^2 \to \mathbb{R}$, on cherche la loi de $U = g(X, Y)$.
- Solution 1
  - Variable intermédiaire : on introduit une va $V = h(X, Y)$ (e.g., $V = X$ ou $V = Y$), on cherche la loi du couple (U, V), puis la loi marginale de U
  - Exemple : $X \sim \mathcal{U}(0, 1)$, $Y \sim \mathcal{U}(0, 1)$, X et Y indépendantes. Quelle est la loi de $U = X + Y$?

Changements de va continues de $\mathbb{R}^2 \to \mathbb{R}$

- Problème
  Si (X, Y) est un couple de va continues de loi connue et $g : \mathbb{R}^2 \to \mathbb{R}$, on cherche la loi de $U = g(X, Y)$.
- Solution 2
  - Calcul de la fonction de répartition de U

$$\begin{aligned} P[U < u] &= P[g(X, Y) < u] \\ &= P[(X, Y) \in \Delta_u] \\ &= \int_{\Delta_u} p(x, y) dx dy. \end{aligned}$$

  - Exemple : $X \sim \mathcal{U}(0, 1)$, $Y \sim \mathcal{U}(0, 1)$, X et Y indépendantes. Quelle est la loi de $U = X + Y$?

Changements de va continues de $\mathbb{R}^2 \to \mathbb{R}$

- Problème
  Si (X, Y) est un couple de va continues de loi connue et $g : \mathbb{R}^2 \to \mathbb{R}$, on cherche la loi de $U = g(X, Y)$.
- Solution 3 : Cas particulier de $U = X + Y$, X et Y indépendantes
  - Calcul de la fonction caractéristique de U.
  - Exemple : $X \sim \mathcal{N}(m_1, \sigma_1^2)$, $Y \sim \mathcal{N}(m_2, \sigma_2^2)$, X et Y indépendantes. Quelle est la loi de $U = X + Y$?

Que faut-il savoir ?

- Loi d'un couple de va discrètes et continues : $?$
- Appartenance à un intervalle : $P[(X,Y) \in \Delta] =?$
- Comment calculer les lois marginales d'un couple ?
- Comment calculer les lois conditionnelles d'un couple ?
- Indépendance de deux variables aléatoires ?
- Espérance mathématique : $E[XY] =?$
- Covariance : $\text{cov}(X,Y) =?$
- Coeff. de corrélation : $r(X,Y) =?$, $r(X,Y) \in ?$ Intérêt ?
- Espérances conditionnelles : ?
- Trois méthodes de changements de variables : ?

Plan du cours

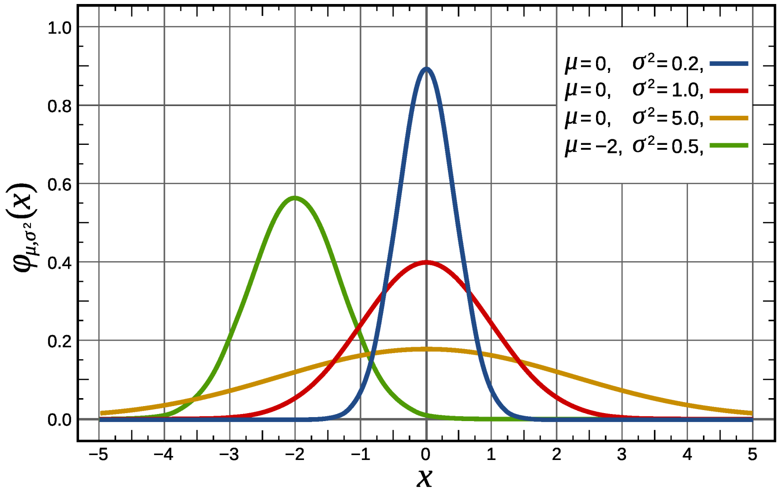
- Chapitre 1 : Eléments de base du calcul des probabilités
- Chapitre 2 : Variables aléatoires réelles
- Chapitre 3 : Couples de variables aléatoires réelles
- Chapitre 4 : Vecteurs Gaussiens
  - Définition
  - Transformation affine
  - Lois marginales, lois conditionnelles, indépendance
  - Lois du chi2, de Student et de Fisher
- Chapitre 5 : Convergence et théorèmes limites

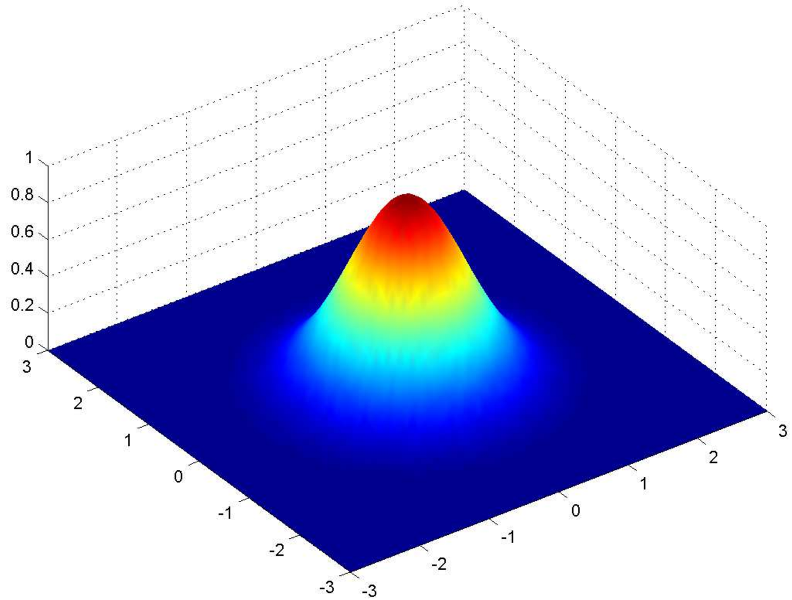
Vecteurs Gaussiens

- Définition
  On dit que $\boldsymbol{X} = (X_1, ..., X_n)^T$ suit une loi normale à n dimensions et on notera $\boldsymbol{X} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma})$, si la densité de probabilité de $\boldsymbol{X}$ s'écrit

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{m})\right]$$

  où $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{m} \in \mathbb{R}^n$ et $\boldsymbol{\Sigma} \in \mathcal{M}_n(\mathbb{R})$ est une matrice symétrique définie positive.
- Page Wikipedia Loi Normale Multidimensionnelle





Vecteurs Gaussiens

- Cas particuliers
  - $n = 1$
  - $\boldsymbol{\Sigma}$ diagonale
- Autres définitions
  - à partir de sa fonction caractéristique définie p. 59
  - à partir de la propriété “un vecteur $\boldsymbol{X} = (X_1, ..., X_n)^T$ est gaussien si toutes les combinaisons linéaires $\sum_{i=1}^{n} a_i X_i$ sont gaussienne”

  Les deux définitions ci-dessus n’imposent pas que la matrice $\boldsymbol{\Sigma}$ soit inversible. Si $\boldsymbol{\Sigma}$ n’est pas inversible, le vecteur $\boldsymbol{X}$ n’a pas de densité.

Vecteur Gaussien

- Exercice

$$p(x,y) \propto \exp\left(-x^2 - \frac{3}{2}y^2 - xy + 4x + 7y\right)$$

Quelle est la loi de (X,Y) ?

Signification de m et Σ

- Fonction caractéristique

$$\phi_{\boldsymbol{X}}(\boldsymbol{u}) = E\left(e^{i\boldsymbol{u}^T\boldsymbol{X}}\right) = \exp\left(i\boldsymbol{u}^T\boldsymbol{m} - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}\boldsymbol{u}\right)$$

- Fonction génératrice des moments

$$M_{\boldsymbol{X}}(\boldsymbol{u}) = E\left(e^{\boldsymbol{u}^T\boldsymbol{X}}\right) = \exp\left(\boldsymbol{u}^T\boldsymbol{m} + \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}\boldsymbol{u}\right)$$

  Propriétés
  - Soient X et Y deux variables aléatoires réelles telles qu'il existe $r > 0$ tel que $\forall u \in]-r, +r[, M_X(u) = M_Y(u)$ alors X et Y ont la même loi
  - Contrairement à la fonction caractéristique, la fonction génératrice des moments n'est pas toujours définie (e.g., loi de Cauchy)
  - Si $M_X(u)$ existe dans un voisinage de l'origine alors X admet des moments $E[X^k] < \infty, \forall k \in \mathbb{N}$
  - Ne pas confondre avec la fonction génératrice des probabilités $G_X(s) = E[s^X]$.
- $\boldsymbol{m}$ et $\boldsymbol{\Sigma}$
  - $\boldsymbol{m}$ est le vecteur moyenne
  - $\boldsymbol{\Sigma}$ est la matrice de covariance

Cas Bivarié

- Fonction génératrice des moments

$$M_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\left[u_1 m_1 + u_2 m_2 + \frac{1}{2}\Sigma_{11} u_1^2 + \frac{1}{2}\Sigma_{22} u_2^2 + \Sigma_{12} u_1 u_2\right]$$

- Dérivées partielles

$$\frac{\partial M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_1} = M_{\boldsymbol{X}}(\boldsymbol{u})(m_1 + \Sigma_{12} u_2 + \Sigma_{11} u_1)$$

$$\frac{\partial M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_2} = M_{\boldsymbol{X}}(\boldsymbol{u})(m_2 + \Sigma_{12} u_1 + \Sigma_{22} u_2)$$

$$\frac{\partial^2 M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_1 \partial u_2} = M_{\boldsymbol{X}}(\boldsymbol{u})(m_1 + \Sigma_{12} u_2 + \Sigma_{11} u_1)(m_2 + \Sigma_{12} u_1 + \Sigma_{22} u_2) + \Sigma_{12} M_{\boldsymbol{X}}(\boldsymbol{u})$$

- Moments

$$\left.\frac{\partial M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_1}\right|_{u=0} = E[X_1] = m_1, \quad \left.\frac{\partial M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_2}\right|_{u=0} = E[X_2] = m_2$$

$$\left.\frac{\partial^2 M_{\boldsymbol{X}}(\boldsymbol{u})}{\partial u_1 \partial u_2}\right|_{u=0} = E[X_1 X_2] = m_1 m_2 + \Sigma_{12}$$

Transformation affine

- Problème : Soit $\boldsymbol{X} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma})$ un vecteur Gaussien. Quelle est la loi de $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$, où $\boldsymbol{Y}$ est un vecteur aléatoire de $\mathbb{R}^p$, $\boldsymbol{b} \in \mathbb{R}^p$ et $\boldsymbol{A}$ est une matrice de taille $p \times n$ avec $p \leq n$?
- Idée : on calcule la fonction génératrice de $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$

$$M_{\boldsymbol{Y}}(\boldsymbol{v}) = \exp\left[\boldsymbol{v}^T(\boldsymbol{A}\boldsymbol{m} + \boldsymbol{b}) + \frac{1}{2}\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T v\right]$$

- Conclusion

$$\boldsymbol{Y} \sim \mathcal{N}_p\left(\boldsymbol{A}\boldsymbol{m} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T\right)$$

si $\boldsymbol{A}$ est de rang p (i.e., de rang maximal).

Lois marginales

- Hypothèses

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma}), \ \boldsymbol{m} = \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_2 \end{pmatrix}$$

- Problème

  Quelle est la loi de $\boldsymbol{X}_1$?

- Conclusion

$$\boldsymbol{X}_1 \sim \mathcal{N}_p\left(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1\right)$$

  où p est la dimension de $\boldsymbol{X}_1$.

Indépendance

- Hypothèses

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma}),\ \boldsymbol{m} = \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_2 \end{pmatrix}$$

- Conclusion

  $\boldsymbol{X}_1$ et $\boldsymbol{X}_2$ sont des vecteurs indépendants si et ssi $\boldsymbol{\Sigma}_{12} = 0$.

  Preuve

  - en utilisant l'expression de la densité d'un vecteur gaussien
  - en utilisant le fait que $\boldsymbol{X}_1$ et $\boldsymbol{X}_2$ sont des vecteurs indépendants si et ssi la fonction génératrice (ou caractéristique) de $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ est le produit des fonctions génératrices (ou caractéristiques) de $\boldsymbol{X}_1$ et de $\boldsymbol{X}_2$.

Lois conditionnelles

- Hypothèses

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma}),\ \boldsymbol{m} = \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_2 \end{pmatrix}$$

- Problème

  Quelle est la loi de $\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{a}$?

- Conclusion

$$\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{a} \sim \mathcal{N}_p \left(\boldsymbol{m}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{a} - \boldsymbol{m}_2), \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_{12}^T \right)$$

où p est la dimension de $\boldsymbol{X}_1$.

Loi du chi2

- Définition
  Si $X_1, ..., X_n$ sont n va indépendantes de loi $\mathcal{N}(0,1)$, alors $Y = \sum_{i=1}^n X_i^2 \sim \chi_n^2$ suit une loi du chi2 à n degrés de liberté.
- Propriétés
  - Densité de probabilité : $p_n(y) = \frac{y^{\frac{n}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\mathbb{I}_{\mathbb{R}^+}(y)$
  - Fonction caractéristique : $\phi_n(t) = (1-2it)^{-\frac{n}{2}}$
  - Moyenne et variance : $E(Y) = n$ et var$(Y) = 2n$
  - Additivité : si $Y \sim \chi_n^2$, $Z \sim \chi_m^2$, Y et Z ind. alors

$$Y + Z \sim \chi_{n+m}^2$$

Loi de Student

- Définition

  Si $X \sim \mathcal{N}(0,1)$, $Y \sim \chi^2_n$, X et Y indépendantes, alors

$$Z = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$$

- Propriétés
  - Densité de probabilité

$$p_n(z) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{z^2}{n}\right)^{-\frac{n+1}{2}}$$

  - Moyenne et variance (pour $n > 2$)

$$E(Z) = 0 \text{ et } \operatorname{var}(Z) = \frac{n}{n-2}$$

  - pour $n = 1$, on a une loi de Cauchy

Loi de Fisher

- Définition
  Si $X \sim \chi^2_n$, $Y \sim \chi^2_m$, X et Y indépendantes, alors

$$Z = \frac{X/n}{Y/m} \sim f_{n,m}$$

- Propriétés
  - Densité de probabilité connue (voir livres)
  - Moyenne et variance (pour $m > 4$)

$$E(Z) = \frac{m}{m-2} \text{ et } \mathrm{var}(Z) = \frac{2m^2(n+m-2)}{n(m-4)(m-2)^2}$$

Théorème de Cochran

- Hypothèses
  Soit $\boldsymbol{X}$ un vecteur Gaussien de loi $\mathcal{N}_n(\boldsymbol{m}, \sigma^2 \boldsymbol{I}_n)$, où $\boldsymbol{m} \in \mathbb{R}^n$, $\sigma^2 > 0$ et $\boldsymbol{I}_n$ la matrice identité de taille $n \times n$.
  Soient p sous-espaces vectoriels orthogonaux $E_1, ..., E_p$ de dimensions $d_1, ..., d_p$ tels que $\mathbb{R}^n = E_1 \oplus ... \oplus E_p$ et $\boldsymbol{Y}_k = \boldsymbol{P}_k \boldsymbol{X}$ la projection orthogonale de $\boldsymbol{X}$ sur E_k ($\boldsymbol{P}_k$ matrice de projection orthogonale sur E_k).
- Conclusions
  - Les vecteurs $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_p$ sont indépendants et $\boldsymbol{Y}_k \sim \mathcal{N}_n(\boldsymbol{P}_k \boldsymbol{m}, \sigma^2 \boldsymbol{P}_k)$
  - Les variables aléatoires $Z_k = \|\boldsymbol{Y}_k - \boldsymbol{P}_k \boldsymbol{m}\|^2$ sont indépendantes et $\frac{Z_k}{\sigma^2} \sim \chi^2_{d_k}$.

Preuve

- Vecteurs $\boldsymbol{Y}_k$

$$\begin{pmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_1 \\ \vdots \\ \boldsymbol{P}_p \end{pmatrix} \boldsymbol{X} = \boldsymbol{A}\boldsymbol{X}$$

  avec $\boldsymbol{P}_k = \boldsymbol{P}_k^T = \boldsymbol{P}_k^2$ et $\boldsymbol{P}_i \boldsymbol{P}_j = \boldsymbol{0}$ for $i \neq j$.

- Variables Z_k
  Si $\boldsymbol{e}_{k,1}, ..., \boldsymbol{e}_{k,d_k}$ est une base orthonormée de E_k, alors le vecteur $\boldsymbol{Y}_k - \boldsymbol{P}_k \boldsymbol{m}$ s'écrit $\boldsymbol{Y}_k - \boldsymbol{P}_k \boldsymbol{m} = \sum_{i=1}^{d_k} \tilde{y}_{k,i} \boldsymbol{e}_{k,i}$ avec $\tilde{y}_{k,i}$ projection de $\boldsymbol{X} - \boldsymbol{m}$ sur $\boldsymbol{e}_{k,i}$ donc $\tilde{y}_{k,i} = \boldsymbol{e}_{k,i}^T (\boldsymbol{X} - \boldsymbol{m}) \sim \mathcal{N}(0, \sigma^2)$. Comme les vecteurs $\boldsymbol{e}_{k,1}, ..., \boldsymbol{e}_{k,d_k}$ sont orthogonaux, les variables $\tilde{y}_{k,1}, ..., \tilde{y}_{k,d_k}$ sont indépendantes, donc

$$\sum_{i=1}^{d_k} \frac{\tilde{y}_{k,i}^2}{\sigma^2} \sim \chi_{d_k}^2.$$

  Remarque : $E(\tilde{y}_{k,i}^2) = \boldsymbol{e}_{k,i}^T (\sigma^2 I_n) \boldsymbol{e}_{k,i} = \sigma^2$.

Statistiques des échantillons gaussiens

- Théorème
  Soit $\boldsymbol{X} = (X_1, ..., X_n)^T$ un échantillon de variables aléatoires réelles indépendante de même loi $\mathcal{N}(m, \sigma^2)$. Alors la moyenne empirique $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ et la variance empirique $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$ sont des variables aléatoires indépendantes telles que

$$\bar{X} \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right) \text{ et } \frac{n-1}{\sigma^2}S^2 = \sum_{i=1}^n \left(X_i - \bar{X}\right)^2 \sim \chi^2_{n-1}$$

- Preuve
  Utiliser le fait que $Y_1 = \bar{X}\mathbf{1}$ est la projection orthogonale de $\boldsymbol{X}$ sur F engendré par $\mathbf{1} = (1, ..., 1)^T$ (car $\langle \boldsymbol{X} - \bar{X}\mathbf{1}, \mathbf{1} \rangle = 0$) et que $Y_2 = \boldsymbol{X} - \bar{X}\mathbf{1}$ est la projection de $\boldsymbol{X}$ sur $F^{\perp}$.

Que faut-il savoir ?

$$\boldsymbol{X} \sim \mathcal{N}_n(\boldsymbol{m}, \boldsymbol{\Sigma})$$

- Signification de $\boldsymbol{m}$ et de $\boldsymbol{\Sigma}$?
- Transformation affine ($\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$) d'un vecteur gaussien ? Condition sur la matrice $\boldsymbol{A}$ associée ?
- Lois marginales d'un vecteur gaussien ?
- Indépendance de deux sous vecteurs d'un vecteur gaussien ? Application : théorème de Cochran (pour ceux qui veulent aller en MoDIA).
- loi de $Y = \sum_{i=1}^{n} X_i^2$ lorsque $X_1, ..., X_n$ sont n va indépendantes de loi $\mathcal{N}(0, 1)$?

Plan du cours

- Chapitre 1 : Eléments de base du calcul des probabilités
- Chapitre 2 : Variables aléatoires réelles
- Chapitre 3 : Couples de variables aléatoires réelles
- Chapitre 4 : Vecteurs Gaussiens
- Chapitre 5 : Convergence et théorèmes limites
  - Convergence (en loi, en probabilité, en moyenne quadratique, presque sure)
  - Théorèmes limites (loi des grands nombres, théorème de la limite centrale)

Convergence

- Problèmes

  Que signifient les relations suivantes ?

$$X_n \underset{n \to +\infty}{\to} X = 0, \quad X_n \underset{n \to +\infty}{\to} X \sim \mathcal{N}(0,1)$$

$$\frac{1}{n}\sum_{k=1}^{n} X_k \underset{n \to +\infty}{\to} X = 0, \quad \frac{1}{n}\sum_{k=1}^{n} X_k \underset{n \to +\infty}{\to} X \sim \mathcal{N}(0,1)$$

Convergence en loi

- Définition

  La suite de va $X_1, ..., X_n$ converge en loi vers la va X si et ssi la suite des fonctions de répartition $F_n(x) = P[X_n < x]$ converge simplement vers $F(x) = P[X < x]$ en tout point x où F est continue.

- Notation

$$X_n \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} X$$

- Exemple

$$P[X_n = 1] = \frac{1}{n} \text{ et } P[X_n = 0] = 1 - \frac{1}{n}$$

- Autre définition

  La suite de va $X_1, ..., X_n$ converge en loi vers la va X si et ssi pour toute fonction f continue bornée sur $\mathbb{R}$

$$E[f(X_n)] \underset{n \to +\infty}{\to} E[f(X)].$$

Convergence en loi

- Propriétés
  - Théorème de Levy
    X_n cv en loi vers X si et ssi ϕ continue en $t = 0$ et
    $$\phi_n(t) = E\left[e^{itX_n}\right] \underset{n \to +\infty}{\to} \phi(t) = E\left[e^{itX}\right], \forall t.$$
  - Si X_n est une suite de va continues de densités $p_n(x)$ et que $p_n(x) \underset{n \to +\infty}{\to} p(x)$ p.p., alors $X_n \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} X$.
  - Si $X_n \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} X$ et $g : \mathbb{R} \to \mathbb{R}$ continue, alors
    $$g(X_n) \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} g(X).$$

Convergence en probabilité

- Définition
  La suite de va $X_1, ..., X_n$ converge en probabilité vers la va X si et ssi $\forall \epsilon > 0$, on a
$$P\left[|X_n - X| > \epsilon\right] \underset{n \to +\infty}{\to} 0.$$
- Notation
$$X_n \underset{n \to +\infty}{\overset{\mathcal{P}}{\to}} X$$
- Exemple
  X_n de densité $p_n(x) = \frac{ne^{-nx}}{(1+e^{-nx})^2}$.
- Propriété
  Si $X_n \underset{n \to +\infty}{\overset{\mathcal{P}}{\to}} X$ et $g : \mathbb{R} \to \mathbb{R}$ continue, alors
$$g(X_n) \underset{n \to +\infty}{\overset{\mathcal{P}}{\to}} g(X).$$

Convergence en moyenne quadratique

- **Définition**
  La suite de va $X_1, ..., X_n$ converge en moyenne quadratique vers la va X si et ssi

$$E\left[(X_n - X)^2\right] \underset{n\to+\infty}{\to} 0.$$

- **Notation**

$$X_n \underset{n\to+\infty}{\overset{\mathcal{MQ}}{\to}} X$$

- **Exemple**

$$P[X_n = n] = \frac{1}{n^p} \text{ et } P[X_n = 0] = 1 - \frac{1}{n^p}$$

avec $p = 2$ et $p = 3$.

Convergence presque sûre

- Définition
  La suite de va $X_1, ..., X_n$ converge presque sûrement vers la va X si et ssi

$$X_n(\omega) \underset{n \to +\infty}{\to} X(\omega), \quad \forall \omega \in A | P(A) = 1.$$

- Notation

$$X_n \underset{n \to +\infty}{\overset{\mathcal{PS}}{\to}} X$$

- Comparaison entre les différents types de convergence

Loi faible des grands nombres

- Loi faible des grands nombres
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E[X_k] = m < +\infty$, alors $\overline{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$ converge en loi (et donc en probabilité) vers m.
- Preuve

$$\varphi_{\overline{X}_n}(t) = E\left[e^{it\frac{1}{n}\sum_{k=1}^n X_k}\right] = E\left[\prod_{k=1}^n e^{i\frac{t}{n}X_k}\right] = \left[\varphi\left(\frac{t}{n}\right)\right]^n$$

Développement de Taylor de ϕ autour de 0

$$\varphi(t) = \varphi(0) + t\varphi'(0) + t\lambda(t) = 1 + itm + t\lambda(t)$$

On en déduit

$$\ln\left[\varphi_{\overline{X}_n}(t)\right] = n\ln\left[1 + i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right] = n\left[i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right]$$

Preuve

d'où

$$\lim_{n\to+\infty} \varphi_{\overline{X}_n}(t) = e^{itm} \qquad \forall t$$

i.e.,

$$\overline{X}_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\to}} m \Leftrightarrow \overline{X}_n \underset{n\to+\infty}{\overset{\mathcal{P}}{\to}} m$$

Loi forte des grands nombres

- Loi forte des grands nombres
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E[X_k] = m < +\infty$ et de variance $\sigma^2 < +\infty$, alors la va $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en moyenne quadratique vers m.

- Preuve

$$E\left[\left(\overline{X} - m\right)^2\right] = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[\left(X_k - m\right)\left(X_l - m\right)\right]$$

Mais

$$E\left[\left(X_k - m\right)\left(X_l - m\right)\right] = \begin{cases} \sigma^2 \text{ si } k = l \\ 0 \text{ si } k \neq l \end{cases}$$

Donc

$$E\left[\left(\overline{X}_n - m\right)^2\right] = \frac{\sigma^2}{n} \underset{n\to+\infty}{\rightarrow} 0$$

i.e.,

$$\overline{X}_n \underset{n\to+\infty}{\overset{\mathcal{MQ}}{\rightarrow}} m$$

Méthode de Monte Carlo

- Hypothèses

  $X_1, ..., X_n$ sont des va indépendantes et de même loi uniforme sur $]0, 1[$ et h est une fonction définie sur $]0, 1[$ telle que $\int_0^1 h(u)du < +\infty$.

- Conclusion

$$\frac{h(X_1) + ... + h(X_n)}{n} \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} I = E[h(X_i)] = \int_0^1 h(u)du$$

  On peut utiliser ce résultat pour approcher l'intégrale I avec des tirages uniformes sur $]0, 1[$.

- Généralisation

$$\frac{h(X_1) + ... + h(X_n)}{n} \underset{n \to +\infty}{\overset{\mathcal{L}}{\to}} \int_J h(u)f(u)du$$

  où $X_1, ..., X_n$ sont des variables indépendantes et de même loi de densité f de support J telles que $E[h(X_i)] < +\infty$.

Échantillonnage d'importance

- Hypothèses
  $X_1, ..., X_n$ sont des va indépendantes et de même loi de densité g (qu'on peut échantillonner simplement et dont le support contient celui de f).
- Conclusion
  Si les hypothèses de la loi des grands nombres sont vérifiées

$$\frac{1}{n}\sum_{i=1}^{n} h(X_i)\frac{f(X_i)}{g(X_i)} \underset{n\to+\infty}{\overset{\mathcal{L}}{\to}} \int_J h(u)f(u)du$$

Échantillonnage d'importance

- Exemple
  Soit f la densité d'une loi de Student à ν degrés de liberté. Calcul de

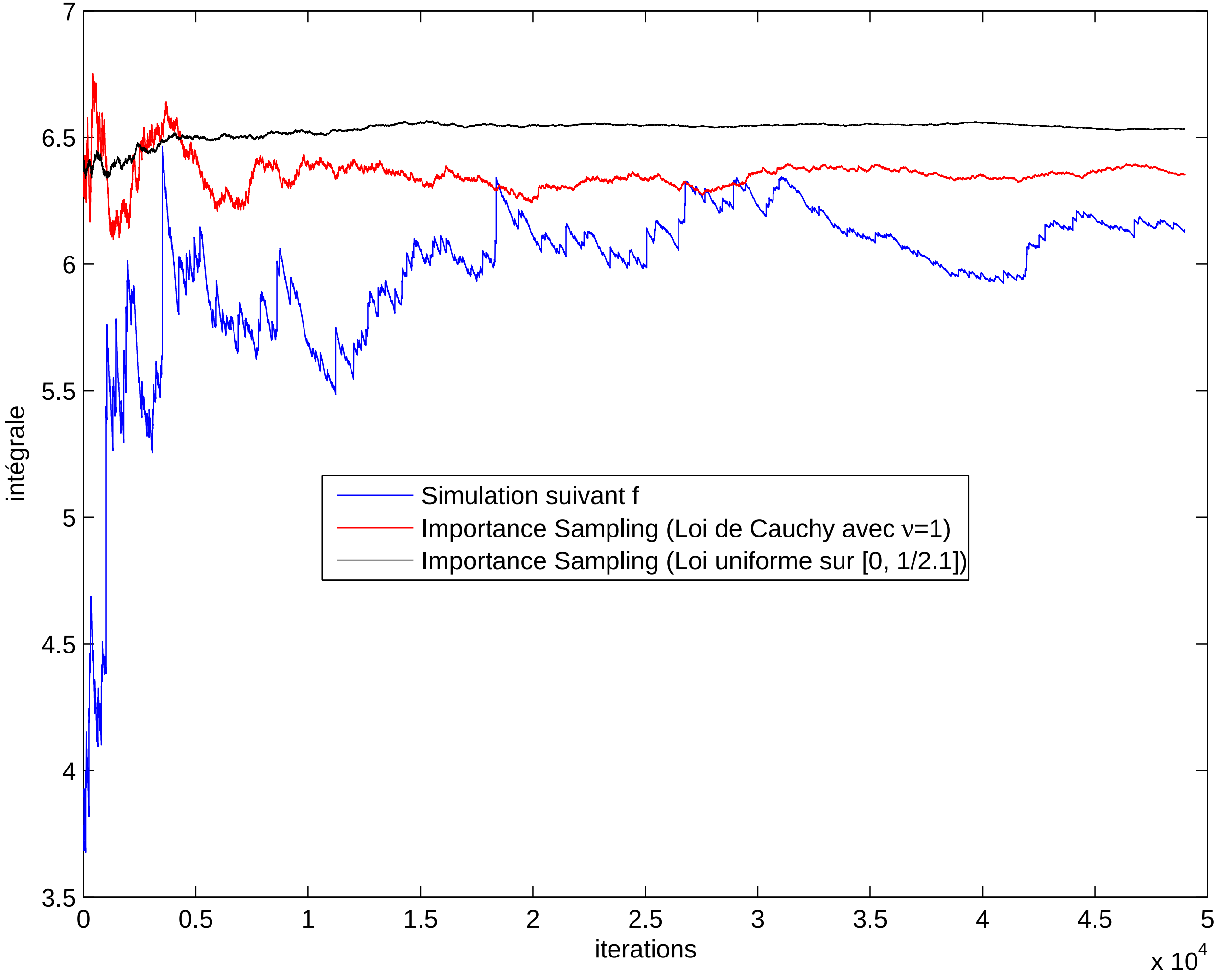
$$I = \int_a^{\infty} u^5 f(u) du = \int_{\mathbb{R}} u^5 \mathbb{I}_{]a,+\infty[} f(u) du \simeq 6.54$$

  - Simulation suivant f
  - Échantillonnage d'importance avec loi de Cauchy
  - Changement de variables $u = 1/v$ permet d'obtenir

$$I = \int_0^{\frac{1}{a}} a \frac{1}{av^7} f\left(\frac{1}{v}\right) dv \simeq \frac{1}{a} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{v_j^7} f\left(\frac{1}{v_j}\right),$$

  où V suit une loi uniforme sur $\left]0, \frac{1}{a}\right[$.

Échantillonnage d'importance



Vraie valeur $I \simeq 6.54$ et $\nu = 12, a = 2.1$.

Théorème de la limite centrale

- Théorème de la limite centrale
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E\left[X_k\right] = m < +\infty$ et de variance $\sigma^2 < +\infty$, alors la va centrée réduite $Y_n = \frac{\sum_{k=1}^{n} X_k - nm}{\sqrt{n\sigma^2}}$ converge en loi vers la loi normale $\mathcal{N}(0,1)$.
- Preuve

$$\varphi_{Y_n}(t) = E\left[e^{itY_n}\right] = e^{-\frac{itm\sqrt{n}}{\sigma}} \prod_{k=1}^{n} E\left[e^{i\frac{t}{\sigma\sqrt{n}}X_k}\right]$$

Mais

$$E\left[e^{i\frac{t}{\sigma\sqrt{n}}X_k}\right] = \varphi\left(\frac{t}{\sigma\sqrt{n}}\right)$$

Donc

$$\ln\left[\varphi_{Y_n}(t)\right] = -\frac{itm\sqrt{n}}{\sigma} + n\ln\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)$$

Preuve

En utilisant le développement de Taylor de φ

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + t^2\lambda(t).$$

On en déduit

$$\ln\left[\varphi_{Y_n}(t)\right] = -\frac{t^2}{2} + \frac{t^2}{n}\lambda\left(\frac{t}{\sigma\sqrt{n}}\right)$$

$$\lim_{n\to+\infty}\varphi_{Y_n}(t) = e^{-\frac{t^2}{2}} \qquad \forall t \;\Leftrightarrow\; Y_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}(0,1)$$

Que faut-il savoir ?

- Convergence en loi ?
- Convergence en moyenne quadratique ?
- $\frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers ? Conditions ?
- $\frac{1}{n}\sum_{k=1}^{n} X_k$ converge en moyenne quadratique vers ? Conditions ?
- $Y_n = \frac{\sum_{k=1}^{n} X_k - ?}{?}$ converge en loi vers ?