© Théo Jalabert

Cours de Probabilités Avancées

Pierre R.

I Vecteurs Aléatoires

1 Fonction de répartition, mesurabilité, généralités

a. Mesurabilité

Définition I.1 *Une famille $X = (X_1, ..., X_n)$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^n$ par :*

$$X(\omega) = (X_1(\omega), X_2(\omega), ..., X_n(\omega)), \quad \omega \in \Omega$$

est une variable aléatoire de dimension n, où un vecteur aléatoire mesurable si l'image réciproque de l'intervalle de dimension n :

$$I = \{(x_1, ..., x_n); -\infty < x_i \leq a_i, a_i \in \mathbb{R}, i = 1, ..., n\}$$

est aussi dans $\mathcal{A}$.
Autrement dit :

$$X^{-1}(I) = \{\omega : X_1(\omega) \leq a_1, ..., X_n(\omega) \leq a_n\} \in \mathcal{A}, \quad \forall\ a_i \in \mathbb{R}$$

ou encore $\forall\ B \in \mathcal{B}(\mathbb{R}^n),\ X^{-1}(B) \in \mathcal{A}$.

Remarque I.1 *Si X est un vecteur aléatoire de $\mathbb{R}^n$ définie sur (Ω, d), on peut écrire $X = (X_1, \ldots, X_n)$ et $X_1, \ldots, X_n$ sont des v.a.r. ; Réciproquement, si $X_1, \ldots, X_n$ sont des v.a.r. sur (Ω, d) alors $(X_1, \ldots, X_n)$ est un vecteur aléatoire de $\mathbb{R}^n$*

b. Fonctions de répartition

Définition I.2 *On appelle fonction de répartition (fdr) ou fonction de répartition jointe ou cumulative distribution function (cdf) d'une v.a. $X = (X_1, ..., X_n)$ vectorielle la fonction à plusieurs variables :*

$$F_X(x)(x_1, ..., x_n) = \mathbb{P}\left((X_1 \leq x_1) \cap ... \cap (X_n \leq x_n)\right), \quad \forall\ (x_1, ..., x_n) \in \mathbb{R}^n.$$

Propriétés I.1 *Les propriétés sont similaires à celles des fonctions de répartition dans le cas univarié :*

- *$F(x_1, ..., x_n) \longrightarrow 0$ quand n'importe lequel des $x_i \longrightarrow -\infty$.*
- *$F(x_1, ..., x_n) \longrightarrow 1$ quand tous les $x_i \longrightarrow \infty$.*
- *F est croissante, cadlag par rapport à toutes ses coordonnées.*

© Théo Jalabert

Remarque I.2 *Dans le cas uni dimensionnel, cela suffit à caractériser les fonctions de répartition, mais ce n'est plus le cas en dimension* ≥ 2.

Exemple I.1 ***Cas bi-dimensionnel*** *: Si* $F(x'y) = \mathbb{P}(X \leq x, Y \leq y)$*, on montre que :*

1. $F(x, y)$ *croissante, cadlag par rapport à* x *et* y.
2. — $\lim_{x\to\infty, y\to\infty} F(x, y) = 1$,
   — $\lim_{y\to-\infty} F(x, y) = 0, \quad \forall\ x$,
   — $\lim_{x\to-\infty} F(x, y) = 0, \quad \forall\ y$.

Mais (1) et (2) ne sont pas des conditions suffisantes pour faire de toutes fonctions $F(.,.)$ *une fonction de répartition.*

Par exemple, prenons :

$$F(x, y) = \begin{cases} 0 & si\ x < 0\ ou\ y < 0\ ou\ x + y < 1 \\ 1 & sinon \end{cases}$$

F *satisfait bien les 2 propriétés mais ce n'est pas une fonction de répartition :*

$$\begin{aligned} \mathbb{P}(1/3 < X \leq 1; 1/3 < Y \leq 1) &= F(1, 1) + F(1/3, 1/3) - F(1, 1/3) - F(1/3, 1) \\ &= 1 + 0 - 1 - 1 = -1 < 0 \ \textit{impossible !} \end{aligned}$$

Plus généralement il faut :

$$\begin{aligned} \mathbb{P}(x_1 < X \leq x_2; y_1 < Y \leq y_2) &= \mathbb{P}(\underbrace{x_1 < X \leq x_2; Y \leq y_2}_{A}) - \mathbb{P}(\underbrace{x_1 < X \leq x_2; Y \leq y_1}_{B}) \\ & \quad car\ B \subset A \Rightarrow \mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(B) \\ &= \mathbb{P}(X \leq x_2; Y \leq y_2) - \mathbb{P}(X \leq x_1, Y \leq y_2) - \mathbb{P}(X \leq x_2, Y \leq y_1) \\ &= F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) \\ &\geq 0 \end{aligned}$$

pour que ce soit bien une fonction de répartition.

Théorème I.1 *Une fonction* F *à 2 variables est une fonction de répartition d'une variable aléatoire de dimension 2 si et seulement si elle satisfait les conditions suivantes :*

1. F *est croissante et cadlag par rapport à ses 2 arguments,*
2. $F(-\infty, y) = F(x, -\infty) = 0$ *et* $F(\infty, \infty) = 1$,
3. $\forall\ (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ *avec* $x_1 < x_2$ *et* $y_1 < y_2$,

$$F(x_2, y_2) - F(x_2, y_1) + F(x_1, y_1) - F(x_1, y_2) \geq 0 \ (et\ \leq 1)$$

Théorème I.2 (Généralisation en dimension n) *Une fonction* $F(x_1, ..., x_n)$ *est une fonction de répartition jointe d'une v.a. de dimension* n *si et seulement si elle vérifie :*

1. F *est croissante et cadlag par rapport à* $x_1, ..., x_n$,

© Théo Jalabert

2. $F(-\infty, x_2, ..., x_n) = F(x_1, -\infty, x_3, ...x_n) = F(x_1, x_2, ..., -\infty) = 0$ *et* $F(\infty, \infty, ..., \infty) = 1$,
3. $\forall\ (x_1, ..., x_n) \in \mathbb{R}^n$ *et* $\forall\ \varepsilon_i > 0,\quad (i = 1, 2, ..., n)$, *on a :*

$$
\begin{aligned}
&F(x_1+\varepsilon_1, ..., x_n+\varepsilon_n) - \sum_{i=1}^{n} F(x_1+\varepsilon_1, ..., x_{i-1}+\varepsilon_{i-1}, \underline{x_i}, x_{i+1}+\varepsilon_{i+1}, ..., x_n+\varepsilon_n) \\
&+ \sum_{i,j=1;i<j}^{n} F(x_1+\varepsilon_1, ..., x_{i-1}+\varepsilon_{i-1}, \underline{x_i}, x_{i+1}+\varepsilon_{i+1}, \\
&\qquad ..., x_{j-1}+\varepsilon_{j-1}, \underline{x_j}, x_{j+1}+\varepsilon_{j+1}, ..., x_n+\varepsilon_n)) \\
&+... \\
&+(-1)^n F(x_1, ..., x_n)
\end{aligned}
$$

c. densités

v.a. bivariée discrète

Définition I.3 *Une v.a. bivariée (X, Y) est dite discrète si elle prend ps ses valeurs dans un ensemble dénombrable A de (couples) valeurs. Toute paire (x_i, y_j) qui a une probabilité strictement positive p_{ij} est appelée point de saut de la fonction de répartition, et p_{ij} est appelé le saut en (x_i, y_j). A est le support de la distribution de (X, Y).*

Remarque I.3 *Clairement, on a $\sum_{i,j} p_{ij} = 1$. La fonction de répartition est donc*

$$F(x, y) = \sum_{B} p_{ij}$$

où $B = \{(i, j); x_i \leq x, y_j \leq y\}$.

Définition I.4 *Si (X, Y) est une v.a. discrète prenant comme valeurs (x_i, y_j), les $p_{ij} = \mathbb{P}(X = x_i, Y = y_j)$ forment les masses de probabilités jointes de (X, Y) ou la distribution de masse ou la joint probability mass function (PMF), sorte de densité discrète.*

Théorème I.3 *Toute famille $(p_{ij})_{i,j\in\mathbb{N}} > 0$ tq $\sum_{i,j=1}^{\infty} p_{ij} = 1$ est une famille de masses de proba d'une v.a.*

v.a. bivariée continue

Définition I.5 *Une variable aléatoire bivariée (X, Y) est dite continue si il existe une fonction positive $f(.,.)$ telle que $\forall\ (x, y) \in \mathbb{R}^2$, on a*

$$F(x, y) = \int_{-\infty}^{x} \left[\int_{-\infty}^{y} f(u, v) dv \right] du$$

où F est la fonction de répartition de (X, Y), f est la fonction de densité de probabilité (jointe) de (X, Y) ou (en anglais) probability density function (PDF).

© Théo Jalabert

Théorème I.4 *Si f est positive et satisfait*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1,$$

alors f est une fonction de densité d'une v.a.

Cas général

Définition I.6 *La fonction de répartition marginale de $X = (X_1, ..., X_n)$ relativement à X_i est définie par*

$$F_{X_i}(x_i) = F_X(\infty, ..., x_i, ..., \infty)$$

C'est la fonction de répartition marginale de X_i. C'est une fonction de répartition univariée. On peut généraliser cette défintion l̀a marginale par rapport à plusieurs coordonnées.

Définition I.7 *Une fonction de densité de probabilité existe si la fonction de répartition est différentiable.*

— *Pour une v.a. scalaire, la densité f est la dérivée de la fonction de répartition :*

$$f_X(x) = F'_X(x)$$

— *Pour una v.a. bivariée,*

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

— *Ainsi, la densité jointe d'une v.a. vectorielle (ou d'un ensemble de v.a. réelles) est la dérivée de la fonction de répartition en dérivant pa rrapport à toutes les variables :*

$$f_X(x_1, ..., x_n) = \frac{\partial^n F_X(x_1, ..., x_n)}{\partial x_1 ... \partial x_n}$$

Propriétés I.2 *Un résultat plus utile est que l'intégrale d'une fonction de densité de probabilité jointe par rapport à certains arguments fournit la densité marginale des variables par rapport auxquelles elle n'a pas été intégrée :*

$$f_{X_i,X_j}(x_i,x_j) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(u_1, ..., x_i, ..., x_j, ..., u_n)du_1...du_{i-1}du_{i+1}..du_{j-1}du_{j+1}...du_n$$

et le même résultat pour les fonctions de répartitions :

$$F_{X_i,X_j}(x_i,x_j) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{x_i} ... \int_{-\infty}^{x_j} ... \int_{-\infty}^{\infty} f(u_1, ..., u_n)du_1du_n$$

© Théo Jalabert

d. Indépendance

Définition I.8 *Soit X une v.a. vectorielle de dimension n, partitionnée en $X = [X_1, X_2]$ avec $X_1 \in \mathbb{R}^{n_1}$, $X_2 \in \mathbb{R}^{n_2}$ et $n_1 + n_2 = n$. Alors X_1 et X_2 sont dit indépendants si la fonction de répartition jointe du vecteur X est le produit des fonctions de répartition de X_1 et X_2 :*

$$F_X(x_1, x_2) = \underbrace{F_X(x_1, \infty_2)}_{\textit{marginale de } x_1} \underbrace{F(\infty_2, x_2)}_{\textit{marginale de } x_2} \quad \textit{(Attention ! Vectoriel !}$$

soit :

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbb{P}(X_1 \leq x_1)\mathbb{P}(X_2 \leq x_2)$$

ce qui implique aussi :

$$f_X(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

Définition I.9 *(pour plus tard) On définit la densité de probabilité conditionnelle de Y sachant $X = x$ (donné) par*

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Exercice I.1 *La v.a. (X, Y) admet la densité jointe suivante :*

$$f_{X,Y}(x,y) = \begin{cases} 2xy + \frac{3}{2}y^2 & 0 < x < 1, 0 < y < 1 \\ 0 & \textit{sinon} \end{cases}$$

1. *Vérifier que $f_{X,Y}$ est une densité.*
2. *Trouver les densités marginales $f_X(x)$ et $f_Y(y)$.*
3. *Trouver les densités conditionnelles $f_{X|Y=y}(x)$ et $f_{Y|X=x}(y)$.*
4. *Calculer $\mathbb{P}((x,y) \in [0, 1/2]^2)$.*
5. *Trouver $\mathbb{P}(X < Y)$*

2 Espaces L^p, Moments

Rappels en dimension 1

On appelle moment d'ordre k d'une v.a. réelle

$$m_k = \mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx \text{ si } X \text{ admet une densité}$$

ou

$$m_k = \mathbb{E}(X^k) = \sum_{i=1}^{\infty} x_i^k \mathbb{P}(X = x_i) \text{ si } X \text{ discrète}$$

On appelle moment centré d'ordre k la quantité

$$\overline{m_k} = \mathbb{E}\left[(X - \mathbb{E}(X))^k\right]$$

Si X admet un moment d'ordre 1, X admet une espérance et est donc dans l'espace L^1. Si elle admet un moment d'ordre 2, X admet une variance et appartient donc à l'espace L^2

© Théo Jalabert

a. Rappel sur les espaces L^p

Définition I.10 — *On appelle $L^1(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel des v.a. réelles définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ admettant une espérance.*
— *On appelle $\widetilde{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel quotient de $L^1(\Omega, \mathcal{A}, \mathbb{P})$ par la relation d'équivalence $X \approx Y \Leftrightarrow X = Y$ p.s..*

Propriétés I.3 *L'espace $\widetilde{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ muni de la norme $\|X\| = \int |X| d\mathbb{P}$ est un espace normé complet.*

Définition I.11 — *On appelle $L^2(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel des v.a. réelles définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ admettant des moments d'ordre 2.*
— *On appelle $\widetilde{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel quotient de $L^2(\Omega, \mathcal{A}, \mathbb{P})$ par la relation d'équivalence $X \approx Y \Leftrightarrow X = Y$ p.s..*

Propriétés I.4 — $L^1(\Omega, \mathcal{A}, \mathbb{P}) \subset L^2(\Omega, \mathcal{A}, \mathbb{P})$
— *$\widetilde{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ muni du produit scalaire $< X, Y >= \mathbb{E}(XY)$ est un espace euclidien complet.*
— *Inégalité de Cuachy-Schwartz : Si X et Y sont dans L^2 alors $XY \in L^1$ et*

$$|\mathbb{E}(XY)| \leq \mathbb{E}(X^2)^{1/2}\mathbb{E}(Y^2)^{1/2}$$

Définition I.12 — *Si X et $Y \in L^2$, on appelle covariance de X et de Y*

$$cov(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)$$

— *On appelle Coefficient de corrélation linéaire la quantité*

$$\rho(X, Y) = \frac{cov(X, Y)}{\mathbb{V}(X)^{1/2}\mathbb{V}(Y)^{1/2}}$$

Propriétés I.5
1. $\rho \in [-1, 1]$.
2. *Si X et Y sont indépendants, alors $cov(X, Y) = 0$ (réciproque fausse) et $\rho = 0$.*
3. *Si $\rho = 1$ (resp. -1), il existe une relation linéaire entre X et Y de pente positive (resp. négative).*

Définition I.13 — *On appelle $L^p(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel des v.a. réelles définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ admettant des moments d'ordre p, c'est à dire tq $\int |X|^p d\mathbb{P} < \infty$. On note alors*

$$\|X\|_p = \left(\int_\Omega |X|^p d\mathbb{P}\right)^{1/p}$$

— *On appelle $L^\infty(\Omega, \mathcal{A}, \mathbb{P})$ l'espace vectoriel des v.a. réelles définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ tq*

$$\sup\{x|\mathbb{P}(|X| > x) > 0\} < \infty$$

et alors

$$\|X\|_\infty = \sup\{x|\mathbb{P}(|X| > x) > 0\} = \inf\{x|\mathbb{P}(|X| > x) = 0\}$$

On dit alors que X est essentiellement (ou $\mathbb{P}$.p.s.) bornée.

© Théo Jalabert

b. Moments

Définition I.14 *Soit $X \in L^1(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$ un vecteur aléatoire. Le moment d'ordre 1, $\bar{X}$, de X est simplement le vecteur des espérances des composantes du vecteur X :*

$$\bar{X} = \mathbb{E}(X) = \mathbb{E}\left[(X_1, ..., X_n)\right] = (\mathbb{E}(X_1), ..., \mathbb{E}(X_n))$$

c'est-à-dire $\bar{X}$ est le vecteur dont les composantes sont

$$\bar{X}_i = \mathbb{E}(X_i).$$

Le vecteur aléatoire $X - \bar{X}$ est appelé vecteur aléatoire centré *associé à X.*

Proposition I.1 *Soit $X \in L^1(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$ et $Y \in L^1(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$ La linéarité de l'espérance d'une variable aléatoire est bien entendu valable pour un vecteur aléatoire :*

$$\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y) \forall \lambda, \mu \in \mathbb{R}$$

Mais cela peut être étendu à la multiplication matricielle : Soit $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ (matrice $n \times p$), et $B \in \mathbb{R}^p$ (vecteur de dimension p), alors $AX + B \in L^1(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^p)$ et

$$\mathbb{E}(AX + B) = A\mathbb{E}(X) + B$$

Définition I.15 *La matrice de variances-covariances d'un vecteur aléatoire $X \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$ est défini par :*

$$\Gamma_X = V(X) = \mathbb{E}\left((X - \bar{X}).{}^t(X - \bar{X})\right)$$

où $\bar{X}$ désigne le moment d'ordre 1 de X.
Autrement dit, l'élément (i, j) de la matrice de variances-covariances s'écrit :

$$V_{i,j}(X) = \mathbb{E}(\bar{X}_i \bar{X}_j) = \mathbb{E}((X_i - \mathbb{E}(X_i)(X_j - \mathbb{E}(X_j))) = Cov(X_i, X_j)$$

Les éléments diagonaux sont les variances des composantes de X, et les éléments non diagonaux sont les covariances entre les composantes de X.

Proposition I.2 Propriétés de la variance et de la covariance
Soit $X \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R})$, on note σ_X l'écart-type de X et $Var(X)$ sa variance. Alors on a :

1. $$\sigma_X^2 = V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}\left((X - \mathbb{E}(X))^2\right).$$
   $$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)$$
2. *La variance est invariante par translation, et l'écart-type est positivement homogène*
   $$\forall a, b \in \mathbb{R}, \;\; V(aX + B) = \sigma_{aX+B}^2 = a^2 V(X)$$
3. $\forall Y \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R})$,
   $$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$$

Proposition I.3 Propriétés de la matrice de variances-covariances
Pour tout $X \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$*, on a :*

1. *La matrice* $V(X)$ *est une matrice carrée* $n \times n$ *symétrique, positive*
2. *La matrice* $V(X)$ *est non inversible si le vecteur* X *est* $\mathbb{P}$*-p.s. dans un hyperplan (c'est-à-dire qu'il existe une relation de liaison (dépendance) entre ses coordonnées)*
3. *Pour* $A \in \mathbb{R}^n$*, la variance de la variable* $Y = {}^tA.X = \sum_{i=1}^n a_i X_i$ *est égale à*
$$V(Y) = V({}^tA.X) = {}^tA.V(X).A$$
4. *Soit* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ *(matrice* $n \times p$*), et* $B \in \mathbb{R}^p$ *(vecteur de dimension* p*), alors* $AX + B \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^p)$ *et*
$$V(A.X + B) = A.V(X).{}^tA.$$

c. Régression linéaire

On rappelle ici la définition du coefficient de corrélation linéaire :

Définition I.16 *Soit* X *et* $Y \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R})$ *de variance non nulle. On appelle* coefficient de corrélation linéaire *de* X *et* Y *le réel :*
$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Proposition I.4 Propriétés du coefficient de corrélation linéaire

1. $|\rho_{X,Y}| \leq 1$ *(vient de l'inégalité de Cauchy-Schwartz).*
2. *Pour que* $|\rho_{X,Y}| = 1$*, il faut et il suffit que* $\exists a, b, c \in \mathbb{R}$ *non tous nuls tels que*
$$\mathbb{P}(aX + bY + c = 0) = 1$$

Le coefficient de corrélation linéaire "mesure" la liaison/corrélation entre des variables aléatoires.

Exercice I.2 *Soit* X *et* Y *deux variables aléatoires de* $L^2(\Omega, \mathcal{A}, \mathbb{P})$*.*

1. *Déterminer la constante* a *qui minimise la quantité* $\mathbb{E}((Y-a)^2)$*.*
2. *Déterminer le couple* a *et* b *qui minise la quantité* $\mathbb{E}((Y - aX - b)^2)$*. (On cherche la meilleure approximation de* Y *comme fonction affine de* X *au sens des moindres carrés.)*
3. *Généralisation. Chercher les constantes* a_1*,* $a_2..a_n$*,* b *qui minisent* $\mathbb{E}(Y - a_1X_1 - ... - a_nX_n - b)^2)$*. On appelle régression linéaire affine de* Y *par rapport aux variables* $X_1, ..., X_n$ *la solution de ce problème.*

© Théo Jalabert

4. *Soeint $Y, Z, X_1, ..., X_n$ des variables de L^2. On appelle coefficient de corrélation partiel de Y et Z contre $X_1, ..., X_n$ le coefficient de corrélation linéaire des variables Y et Z privées de leurs régressions linéaires affines (notées $\widetilde{Y}$ et $\widetilde{Z}$) par rapport à $(X_1, ..., X_n)$. Montrer que ce coefficient est égale à :*

$$\alpha \frac{\mathbb{V}(Z-\widetilde{Z})^{1/2}}{\mathbb{V}(Y-\widetilde{Y})^{1/2}}$$

où α désigne le coeff. de Z dans la régression linéaire affine de Y par rapport aux variables $X_1, ..., X_n, Z$.

3 Transformations de variables aléatoires

a. Cas général

Rappel : cas univarié Soit $Y = g(X)$ où g est une fonction connue. On veut calculer F_Y et f_y à parti de F_X et f_X.

Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$ et $g^{-1}(B) \subset \mathbb{R}$ l'ensemble pour lequel $g(g^{-1}(B)) = B$. Alors :

$$\mathbb{P}(Y \in B) = \mathbb{P}(g(X) \in B) = \mathbb{P}(X \in g^{-1}(B))$$

Car $X \in g^{-1}(B)$ si et seulement si $g(X) = Y \in g(g^{-1}(B)) = B$.

Donc, pour trouver $F_Y(y)$, on prend $B_y =]-\infty, y]$ et on obtient :

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \in B_y) = \mathbb{P}(X \in g^{-1}(B_y))$$

Cas bivarié On veut calculer $\mathbb{P}(Y \in B)$ avec $Y \in \mathbb{R}^d$ fonction de $X \in \mathbb{R}^d$ telle que :

$$Y = \begin{pmatrix} Y_1 \\ ... \\ Y_d \end{pmatrix} = \begin{pmatrix} g_1(X_1, X_2) \\ ... \\ g_d(X_1, X_2) \end{pmatrix} = g(X)$$

Soient $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^d$ une fonction connue, $B \in \mathbb{R}^d$ et $g^{-1}(B) \subset \mathbb{R}^2$ l'ensemble pour lequel $g(g^{-1}(B))$. Alors :

$$\mathbb{P}(Y \in B) = \mathbb{P}(g(X) \in B) = \mathbb{P}(X \in g^{-1}(B))$$

b. Transformation de densité conjointe continue

Théorème I.5 *Si pour toute fonction borélienne bornée*

$$\mathbb{E}(h(g(X))) = \int_{\mathbb{R}} h(y) f_Y(y) dy$$

alors f_Y est la densité de $Y = g(X)$.

Cela s'applique également pour des fonctions $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

© Théo Jalabert

Théorème I.6 Cas bivarié
Soient $X = (X_1, X_2) \in \mathbb{R}^2$ *un vecteur aléatoire de loi continue,* $Y = (Y_1, Y_2) = g(X)$ *défini par* $Y_1 = g_1(X_1, X_2)$ *et* $Y_2 = g_2(X_1, X_2)$*, où*

1. *Le système d'équations* $\begin{cases} y_1 = g_1(x_1, x_2) \\ y_1 = g_1(x_1, x_2) \end{cases}$ *peut être résolu pour tout* (y_1, y_2)*, donnant les solutions* $x_1 = h_1(y_1, y_2)$*,* $x_2 = h_2(y_1, y_2)$ *(*$h = g^{-1}$ *image inverse).*
2. g_1 *et* g_2 *sont continuement différentiables, et ont pour matrice jacobienne :*

$$J(x_1, x_2) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix}$$

Le jacobien est le déterminant de la matrice jacobienne.

Alors

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)\ |J(x_1, x_2)|^{-1}\big|_{(x_1,x_2)=h(y_1,h_2)}$$

Théorème I.7 Cas multivarié
Le théorème ci-dessus s'tend aux cas des vecteurs aléatoires de $\mathbb{R}^n$*, admettant une densité. Soit* $Y = g(X) \in \mathbb{R}^n$*, où* $X \in \mathbb{R}^n$ *est une variable aléatoire continue,* $g : (X_1, ..., X_n) \implies (Y_1, ..., Y_n) = (g_1(X_1, ..., X_n), ..., Y_n = g_n(X_1, ..., X_n))$*. Si la transformation invesre* $h = g^{-1}$ *existe, et si* g *a pour jacobien*

$$J(x_1, ..., x_n) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & ... & \frac{\partial g_1}{\partial x_n} \\ ... & ... & ... \\ \frac{\partial g_2}{\partial x_1} & ... & \frac{\partial g_2}{\partial x_n} \end{pmatrix}$$

On trouve que :

$$f_Y(y) = f_X(g^{-1}(y))\ |detJ(y)|^{-1}\big|_{x=h(y)}$$

ou encore de manière équivalente

$$f_{Y_1,...,Y_n}(y_1, ..., y_n) = f_{X_1,...,X_n}(x_1, ..., x_n)\ |detJ(y)|^{-1}\big|_{(x_1,...,x_2)=h(y_1,...,y_n)}$$

Exercice I.3 *Soit le couple* (X, Y) *de densité*

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 + y^2)\right), \quad (x, y) \in \mathbb{R}^2$$

Effectuer le changement de variable $(X, Y) \longrightarrow (R, \Theta)$ *avec :*

$$X = R\cos\Theta, \qquad Y = R\sin\Theta.$$

et trouvez la loi de (R, Θ)*.*

© Théo Jalabert

c. Sommes de variables aléatoires indépendantes

Théorème I.8 dim 2
Soient X et Y des variables aléatoires indépendantes, alors la fonction de densité de leur somme $S = X + Y$ est le produit de convolution $f_X \star f_Y$ des fonctions de densité f_X, f_Y :

$$f_S(s) = f_X \star f_Y(s) = \begin{cases} \int_{-\infty}^{+\infty} f_X(x) f_Y(s-x)dx & \text{si } X \text{ et } Y \text{ sont continues,} \\ \sum_{x \in Supp(X,Y)} f_X(x) f_Y(s-x) & \text{si } X \text{ et } Y \text{ sont discrètes.} \end{cases}$$

Remarque I.4 *Dans le cas discret, les f sont les masses de probabilités.*

Remarque I.5 *La convolution peut être vu comme une généralisation de l'idée de moyenne mobile.*

Théorème I.9 dim n
Soient $X_1, ..., X_n$ des vecteurs aléatoires indépendants, alors la densité de $S = X_1 + ... + X_n$ est le produit de convolution des densités des X_i :

$$f_S(s) = f_{X_1} \star ... \star f_{X_n}(s)$$

d. Exercices

Exercice I.4

1. *Soient X_1 et X_2 deux variables aléatoires indépendantes représentant les résultats de 2 lancers de dés équilibrés.*
   Quelle est la loi de $X_1 + X_2$?
2. *Soient X_1, X_2 deux variables aléatoires indépendantes de loi exponentielle de paramètre λ. Quelle est la loi de $X_1 + X_2$? de $X_1 - X_2$? du couple $(X_1 + X_2, X_1 - X_2)$?*

Exercice I.5

1. *Calculer la densité jointe de $X_1 + X_2$ et $X_1 - X_2$ lorsque X_1, X_2 suivent deux lois normales centrées réduites indépendantes.*
2. *Calculer la densité jointe de $X_1 + X_2$ et $\frac{X_1}{X_1+X_2}$ lorsque X_1, X_2 suivent deux exponentielles de paramètre λ indépendantes.*
3. *Calculer la densité de $\frac{X_2}{X_1}$ si X_1, X_2 suivent deux lois normales centrées réduites indépendantes.*

Exercice I.6 *On considère la fonction suivante*

$$f_{X,Y}(x,y) = e^{-y}\mathbf{1}_{\{0<x<y\}}.$$

1. *Vérifier que $f_{X,Y}$ définit une densité de probabilité sur $\mathbb{R}^2$.*
2. *Calculer les densités marginales f_X et f_Y de X et Y. Les variables X et Y sont-elles indépendantes ?*
3. *Calculer $E(X), E(Y), cov(X,Y)$.*
4. *Déterminer la loi jointe $f_{Z,T}$ du couple (Z,T) défini par :*

$$\begin{cases} Z &= X + Y \\ T &= Y - X \end{cases}$$

© Théo Jalabert

5. *En déduire les densités marginales de Z et T.*

Exercice I.7 *Soit (X,Y) un vecteur aléatoire bidimensionnel de densité*

$$f(x,y) = \begin{cases} \frac{1}{4}(1+xy) & si\ -1 \leq x \leq 1\ et\ -1 \leq y \leq 1 \\ 0 & sinon \end{cases}$$

et soit (U,V) un autre vecteur aléatoire bidimensionnel de densité

$$g(u,v) = \begin{cases} \frac{1}{4} & si\ -1 \leq u \leq 1\ et\ -1 \leq v \leq 1 \\ 0 & sinon \end{cases}$$

1. *Vérifier que f et g sont bien des densités de probabilités sur $\mathbb{R}^2$.*
2. *Trouver les densités marginales de (X,Y) et (U,V) notées respectivement f_X, f_Y, g_U et g_V.*
3. *Vérifier que malgré les égalités (en loi) $X \sim U$ et $Y \sim V$, on n'a pas $(X,Y) \sim (U,V)$!*

II Fonction Caractéristique

1 Transformée de Fourier d'une mesure - Fonction caractéristique

Définition II.1 *Soit X une variable aléatoire* **complexe** *définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que X est intégrable par rapport à la mesure de probabilité $\mathbb{P}$ si $\mathbb{E}(|X) < +\infty$. On définit alors*

$$\mathbb{E}_{\mathbb{P}}(X) = \mathbb{E}_{\mathbb{P}}(Re(X)) + i\mathbb{E}_{\mathbb{P}}(Im(X))$$

où $Re(X)$ et $Im(X)$ désignent des parties réelles et imaginaires de X.

Remarque II.1 *On rappelle que X est une variable aléatoire complexe ssi $Re(X)$ et $Im(X)$ sont des variables aléatoires.*
Et X est intégrable ssi $Re(X)$ et $Im(X)$ sont intégrables, et :

$$|X| = \sqrt{Re(X)^2 + Im(X)^2}.$$

Définition II.2 *Soit $\mathbb{P}$ une mesure de probabilité définie sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On appelle* **transformée de Fourier de la mesure** $\mathbb{P}$ *la fonction ϕ définie sur $\mathbb{R}$, à valeurs dans $\mathbb{C}$, par :*

$$\phi(t) := \int_{\mathbb{R}} e^{itx} d\mathbb{P}(x)$$

Définition II.3 **Fonction caractéristique : cas univarié**
Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A})$. On définit la **fonction caractéristique de** X *comme étant la transformée de Fourier de la mesure image de $\mathbb{P}$ par X, $\mathbb{P}_X$, et on note*

$$\phi_X(t) := \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X(x) = \mathbb{E}_{\mathbb{P}}\left(e^{itX}\right)$$

© Théo Jalabert

Définition II.4 **Fonction caractéristique : cas multivarié**
Soit $(X_1, ..., X_n)$ *un vecteur aléatoire défini sur un espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$*. La fonction caractéristique de* X *est la fonction définie par :*

$$\forall t = (t_1, ..., t_n) \in \mathbb{R}^n, \quad \phi_X(t) = \mathbb{E}\left(e^{i<t,X>}\right) = \mathbb{E}_{\mathbb{P}}\left(e^{i\sum_{k=1}^{n} t_k X_k}\right).$$

© Théo Jalabert

Proposition II.1 *1.* $\phi_X(0) = 1$

2. $|\phi_X(t)| \leq 1$

3. ϕ_X *est uniformément continue*

4. $\phi_X(-t) = \overline{\phi_X(t)}$

5.

$$\forall a, b \in \mathbb{R}, \;\; \phi_{aX+b}(t) = e^{itb}\phi_X(at) \text{ en dimension 1,}$$

$$\forall A \in \mathcal{M}_{k,n}(\mathbb{R}), \forall B \in \mathbb{R}^k, \phi_{A.X+B}(t) = e^{i<t,B>}\phi_X(A^{\star}.t) \text{ en dimension } n.$$

Proposition II.2 **Lien entre le cas multivarié et le cas univarié**
Soit $X = (X_1, ..., X_n)$ *un vecteur aléatoire à valeurs dans* $\mathbb{R}^n$ *défini sur un espace de probabilités* $(\Omega, \mathcal{A}, \mathbb{P})$*, alors*

$$\phi_X(t) = \phi_{{}^tt.X}(1)$$

2 Fonction caractéristique : utilisation

a. Injectivité

Théorème II.1 *Injectivité de la transformée de Fourier d'une mesure de probabilités*
Soient $\mathbb{P}$ *et* $\mathbb{Q}$ *deux probabilités sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*. Si*

$$\int_{\mathbb{R}} e^{itx} d\mathbb{P}(x) = \int_{\mathbb{R}} e^{itx} d\mathbb{Q}(x), \;\; \forall t \in \mathbb{R}$$

Alors $\mathbb{P} = \mathbb{Q}$.

Corollaire II.1 *Soit* X *et* Y *deux variables aléatoires définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$*. L'égalité entre les fonctions caractéristiques de* X *et* Y *implique que* X *et* Y *ont la même loi :*

$$\phi_X(t) = \phi_Y(t), \;\; \forall t \in \mathbb{R} \Longrightarrow \mathbb{P}_X = \mathbb{P}_Y.$$

Consequence *: La fonction caractéristique caractérise la loi d'une variable aléatoire.*

Remarque II.2 *L'injectivité de la transformée de Fourier reste vraie dans le cas vectoriel :*

Proposition II.3 *Soient* $\mathbb{P}$ *et* $\mathbb{Q}$ *deux probabilités sur* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$*. Si*

$$\int_{\mathbb{R}} e^{i<t,x>} d\mathbb{P}(x) = \int_{\mathbb{R}} e^{i<t,x>} d\mathbb{Q}(x), \;\; \forall t \in \mathbb{R}^n$$

Alors $\mathbb{P} = \mathbb{Q}$.

Proposition II.4 **Fonctions caractéristiques de lois symétriques**
Soit X *une variable aléatoire réelle sur* $(\Omega, \mathcal{A}, \mathbb{P})$*. On a :*

$$\mathbb{P}_X = \mathbb{P}_{-X} \Longleftrightarrow \phi_X \text{ est une fonction réelle.}$$

© Théo Jalabert

b. Inversion

Théorème II.2 **Formule d'inversion avec fonction de r épartition**
Soit $\mathbb{P}$ une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ayant pour transformée de Fourier $\phi(t) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}(x)$.
On note $F(x) = \mathbb{P}(]-\infty, x[)$ sa fonction de répartition.

1. *Soient a et $b \in \mathbb{R}$ $(a < b)$. Alors :*

$$\frac{F(b)+F(b^+)}{2} - \frac{F(a)+F(a^+)}{2} = \frac{1}{2\pi} \lim_{c \longrightarrow +\infty} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \phi(t) dt$$

2. *Si F est continue en a et b, alors :*

$$F(b) - F(a) = \frac{1}{2\pi} \lim_{c \longrightarrow +\infty} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \phi(t) dt$$

Remarque II.3 *Ce théorème est vrai en dimension d.*
$a \leq b$ signifie $a_i \leq b_i \forall i \in [1, ..., d]$, et on obtient alors comme formule :

$$F(b) - F(a) = \frac{1}{(2\pi)^d} \lim_{c_1, ..., c_i, ..., c_d \longrightarrow +\infty} \int_{-c_1}^{c_1} ... \int_{-c_i}^{c_i} ... \int_{-c_d}^{c_d} \frac{e^{-i<t,a>} - e^{-i<t,b>}}{i < t, \mathbb{1} >} \phi(t) dt$$

Théorème II.3 *Soit $\mathbb{P}$ une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, de transformée de Fourier $\phi(t)$. Si ϕ est intégrable par rapport à la mesure de Lebesgue sur $\mathbb{R}$ (ie $\int_{\mathbb{R}} |\phi(t)| dt < +\infty$), alors $\mathbb{P}$ possède une densité f par rapport à la mesure de Lebesgue, donnée par la formule :*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{itx} \phi(t) dt, \ \ \forall x \in \mathbb{R}$$

Ce qui s'adapte pour une mesure sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$:

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i<t,x>} \phi(t) dt \ \ \forall x \in \mathbb{R}^d$$

Cette formule d'inversion permet de retrouver la loi d'une va lorsque l'on connaît sa transformée de Fourier (sa fonction caractéristique).
Elle permet également d'obtenir des méthodes numériques d'implémentation et de simulation de variables aléatoires, en simulant leur fonction caractéristiques, puis en utilisant cette formule pour inverser et retrouver la loi (méthode de Fast Fourier Transform).

c. Fonction caractéristique et indépendance

Proposition II.5 *La transformée de Fourier de la mesure produit de 2 mesures et le produit de leurs transformée de leurs transformées de Fourier :*

$$\widehat{\mu_1 \otimes \mu_2}(t_1, t_2) = \widehat{\mu_1}(t_1) \widehat{\mu_2}(t_2), \ \ \forall (t_1, t_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$$

Cela permet d'obtenir un critère d'indépendance de variables aléatoires en termes de fonctions caractéristiques :

© Théo Jalabert

Théorème II.4 **Critère d'indépendance**
Soit $X = (X_1, X_2)$ *une variable aléatoires à valeurs dans* $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. *Alors* X_1 *et* X_2 *sont indépendantes ssi*

$$\forall (t_1, t_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, \ \ \phi_{(X_1,X_2)}(t_1, t_2) = \phi_{X_1}(t_1).\phi_{X_2}(t_2)$$

Définition II.5 *La fonction caractéristique d'une marginale s'obtient facilement : Soit* $X = (X_1, X_2)$ *une variable aléatoire dans* $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. *Alors*

$$\begin{array}{rcl} \forall t_1 \in \mathbb{R}^{d_1}, \quad \phi_{X_1}(t_1) & = & \phi_X(t_1, 0) \\ \forall t_2 \in \mathbb{R}^{d_2}, \quad \phi_{X_2}(t_2) & = & \phi_X(0, t_2). \end{array}$$

Corollaire II.2 **Critère d'indépendance (bis)**
X_1 *et* X_2 *deux variables aléatoires à valeurs respectivement dans* $\mathbb{R}^{d_1}$ *et* $\mathbb{R}^{d_2}$ *sont indépendantes ssi, en notant* $X = (X_1, X_2)$:

$$\forall t = (t_1, t_2) \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}, \ \ \phi_X(t) = \phi_X(t_1, 0).\phi_X(0, t_2)$$

C'est juste une autre manière d'écrire le premier critère d'indépendance.

Proposition II.6 *La transformée de Fourier du produit de convolution de 2 mesures* μ_1 *et* μ_2 *sur* $\mathbb{R}^d$ *est le produit de leurs transformées de Fourier :*

$$\widehat{\mu_1 \star \mu_2} = \widehat{\mu_1}.\widehat{\mu_2}$$

Théorème II.5 **Fonction caractéristique de la somme de 2 vai**
Soient X *et* Y *deux variables aléatoires indépendantes à valeurs dans* $\mathbb{R}^d$. *Alors la fonction caractéristique de leur somme est*

$$\forall t \in \mathbb{R}^d, \ \ \phi_{X+Y}(t) = \phi_X(t).\phi_Y(t)$$

C'est évidemment la même chose pour toute somme finie de variables aléatoires indépendantes :

$$\phi_{X_1+\ldots+X_n}(t) = \phi_{X_1}(t)\ldots\phi_{X_n}(t), \ \ \forall t \in \mathbb{R}^d$$

Exercice II.1 *Soient* X_1, X_2 *deux variables aléatoires réelles indépendantes, de même loi de Laplace (densité* $\frac{1}{2}e^{-|x|}$*). On définit* $X = (X_1, X_2)$*, ainsi que* Y_1 *et* Y_2 *par* $Y_1 = X_1 - X_2$, $Y_2 = X_1 + X_2$.

1. *Calculer la fonction caractéristique de* X_1 *et* X_2*, puis celle de* X.
2. *Calculer les fonctions caractéristiques de* Y *puis des marginales* Y_1 *et* Y_2.
3. *Montrer que* Y_1 *et* Y_2 *ont la même loi.*
4. *Montrer que* $Cov(Y_1, Y_2) = 0$
5. *Montrer que* Y_1 *et* Y_2 *ne sont pas indépendantes.*

d. Fonction caractéristique et moments

Théorème II.6 *Soit X une variable aléatoire réelle, et ϕ_X sa fonction caractéristique.*

- *Si X admet un moment d'ordre n, alors ϕ_X est de classe C^n, et $\forall 1 \leq k \leq n$, on a :*

$$\forall t \in \mathbb{R}, \quad \phi_X^{(k)}(t) = i^k \int_{\mathbb{R}} x^k e^{itx} d\mathbb{P}_X(x)$$

et en particulier

$$(\star) \qquad \phi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$$

- *Inversement, si ϕ_X est k fois dérivable en 0, $k \geq 2$, alors X admet des moments jusqu'à l'ordre $2\left\lfloor \frac{k}{2} \right\rfloor$, donnés par la formule $(\star)$.*

3 Exemples de fonctions caractéristiques de lois usuelles

Loi discrètes	*probabilités*	$\phi_X(t)$
Bernouilli $b(p)$	$p\delta_1 + (1-p)\delta_0$	$(1-p) + pe^{it}$
Binomiale $B(n,p)$	$C_n^k p^k q^{(n-k)}$	$\left((1-p) + pe^{it}\right)^n$
Poisson $\mathcal{P}(\lambda)$	$e^{-\lambda}\frac{\lambda^k}{k!}$	$\exp\left(\lambda e^{it} - 1\right)$

Lois continues	*densités*	$\phi_X(t)$
Uniforme	$\mathbb{1}_{[0,1]}(x)$	$\frac{e^{it}-1}{it}$
Normale $\mathcal{N}(\mu,\sigma)$	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{i\mu t - \sigma^2\frac{t^2}{2}}$
Exponentielle $\mathcal{E}(\theta)$	$\theta e^{-\theta x}$	$\frac{\theta}{\theta - it}$
Gamma $\Gamma(\theta,p)$	$\frac{\theta^p}{\Gamma(p)}e^{-\theta x}x^{p-1}$	$\left(\frac{\theta}{\theta-it}\right)^p$
Cauchy	$\frac{1}{\pi(x^2+1)}$	$e^{-\lvert t\rvert}$
Laplace	$\frac{1}{2}e^{-\lvert x\rvert}$	$\frac{1}{1+t^2}$

Remarque II.4 *On retrouve facilement grâce aux fonctions caractéristiques quelques propriétés bien connues des lois :*

- *Une somme de lois exponentielles indépendantes = une loi Gamma*
- *Une somme de lois de bernouilli indépendantes = une loi Binomiale*
- *La loi de Cauchy n'a aucun moment d'aucun ordre ($e^{-|x|}$ n'est pas dérivable en 0).*
- *Si on somme 2 lois normales indépendantes, de paramètres (m, σ^2) et (l, ρ^2), on obtient une loi normale de paramètres $(m + l, \sigma^2 + \rho^2)$.*

Exercice II.2 *Etude de la loi Lognormale : on dit que X suit une loi log-normale lorsque $ln(X)$ suit une loi normale. C'est-à-dire X est l'exponentielle d'une loi normale : $X = e^Y$ où Y suit une loi normale.*

© Théo Jalabert

1) Soit φ une app mesurable positive $Y \sim \mathcal{N}(m,\sigma)$

$$\mathbb{E}[\varphi(e^Y)] = \int_{\mathbb{R}} \varphi(e^y) f_Y(y)\, dy$$

$$= \int_{\mathbb{R}} \varphi(e^y) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-m)^2}{2\sigma^2}} dy$$

Changement de variable :

$u = e^y \Rightarrow y = \ln(u)$

$du = e^y dy \Rightarrow dy = \frac{du}{u}$

$$\Rightarrow \mathbb{E}[\varphi(e^Y)] = \int_{\mathbb{R}_+^*} \varphi(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\ln(u)-m)^2}{2\sigma^2}} \frac{du}{u}$$

$$\Rightarrow f_X(x) = \frac{1}{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-m)^2}{2\sigma^2}} \mathbb{1}_{\mathbb{R}_+^*}(x)$$

2) $$\mathbb{E}(X) = \int_0^\infty x \frac{1}{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-m)^2}{2\sigma^2}} dx$$

$$= \int_0^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-m}{\sigma}\right)^2} dx$$

$u = \frac{\ln(x)-m}{\sigma}$

Rappel : $\int_{-\infty}^{\infty} e^{-(ax^2+bx+c)} = \sqrt{\frac{\pi}{a}}\, e^{\frac{b^2-4ac}{4a}}$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \sigma e^{u\sigma+m} du$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} e^{u\sigma+m} du$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{1}{2}u^2 - \sigma u - m\right)} du$$

$$= e^{\frac{\sigma^2+4\times\frac{1}{2}\times m}{4\times\frac{1}{2}}} = e^{\frac{\sigma^2+2m}{2}} = e^{\frac{\sigma^2}{2}+m}$$

et $$V(X) = e^{2m+\sigma^2}\left(e^{\sigma^2}-1\right)$$

© Théo Jalabert

1. *Donner la fonction de densité d'une loi lognormale de paramètres* (m, σ).
2. *Calculer l'espérance et la variance de* X

Exercice II.3 **Fonctions génératrices**

1. *Calculer la fonction génératrice d'une variable aléatoire suivant une loi de Bernouilli, retrouver l'espérance et la variance de cette loi.*
2. *Mêmes questions pour une loi binomiale.*
3. *Mêmes questions pour une loi exponentielle.*

Exercice II.4 **Loi Binomiale Négative**

1. *La loi de Pascal, ou loi géométrique, est la loi du temps d'atteinte* T_1 *du premier succès dans une suite de tirages indépendants de Bernouilli de paramètre* p.
   (a) *Quelle est la loi de* T_1 *?*
   (b) *Quelle est sa fonction génératrice ?*
   (c) *En déduire sa moyenne ? sa variance ?*
2. *On note à présent* T_k *le* k^{ime} *succès d'une suite de tirages indépendants de Bernouilli de paramètre* p. *On admettra que l'intuition qui voit* T_k *comme la somme de* k *vai* $I_1, ..., I_k$ *de même loi* T_1 *est vraie.*
   (a) *Quel est l'ensemble des valeurs possibles de* T_k *?*
   (b) *Quelle est la loi de* T_k *?*
   (c) *Calculer la fonction génératrice de* T_k.
   (d) *En déduire la moyenne et la variance de* T_k.
   (e) *Quel est l'ensemble de valeurs possibles prises par* $T_k - k$ *? Donnez la loi de* $T_k - k$ *Quelle est la fonction génératrice de* $T'_k = T_k - k$ *?*
3. *Généralisation : si on remplace dans l'étude précédente des caractéristiques de* T'_k, $k \in \mathbb{N}$ *par* $\lambda > 0$, *on n'a plus l'interprétation en termes de temps d'atteinte du k-ième succès, mais on obtient ce que l'on appelle la* **Loi Binomiale-Négative** $BN(\lambda, p)$.
   Déduire des résultats précédent l'expression la loi de X, *sa fonction génératrice ainsi que sa moyenne et sa variance, quand* X *suit une loi* $BN(\lambda, p)$.

Exercice II.5 1. *Quelle est la transformée de Laplace d'une va suivant une loi de Laplace ? Pour quelles valeurs de* t *est-elle définie ?*

2. *Même question pour une loi uniforme ?*
3. *Même question pour une loi exponentielle ? Gamma ?*

Exercice II.6 **Loi hypoexponentielle**

© Théo Jalabert

II-3

1) $X \sim b(p)$ $\quad \mathbb{P}(X=x) = p\,\delta_1(x) + (1-p)\,\delta_0(x)$

$$\phi_X(t) = p e^{it\times 1} + (1-p) e^{it \times 0}$$
$$= (1-p) + p e^{it}$$

$$\phi_X^{(k)}(0) = i^k \, \mathbb{E}(X^k)$$

$$\mathbb{E}(X) = -i\,\phi_X'(0)$$
$$= p$$

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

avec $\mathbb{E}(X^2) = -\phi_X''(0) = p$

$$\Rightarrow V(X) = p - p^2 = p(1-p)$$

$X \sim B(n,p)$ $\quad \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$$\Rightarrow \phi_X(t) = \sum_{k=1}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\Rightarrow \phi_X(t) = \left(1 - p + e^{it} p\right)^n$$

$X \sim \mathcal{E}(\theta)$ $\quad f(x) = \theta e^{-\theta x}$

$T_1 \sim \mathcal{G}(p)$ $\quad \mathbb{P}(T_1 = k) = pq^{k-1}$ © Théo Jalabert

$$\phi_{T_1}(t) = \sum_{k=1}^{\infty} e^{itk} pq^{k-1}$$

Rappel : $\phi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$

$$= \frac{p}{q} \sum_{k=1}^{\infty} (qe^{it})^k$$

$$= \frac{p}{q} \frac{qe^{it}}{1-qe^{it}}$$

$$= \frac{pe^{it}}{1-qe^{it}}$$

$\Rightarrow \mathbb{E}(T_1) = -i\,\phi_X'(0)$

$= -i$

Définition II.6 *Soit $\lambda_1, ..., \lambda_n$, n réels positifs différents. Une variable aléatoire X continue à valeurs dans $\mathbb{R}^+$ suit une loi hypoexponentielle d'ordre n et de paramètres $\lambda_1, ..., \lambda_n$ si sa fonction de densité est donnée par*

$$f_X(t) = \sum_{i=1}^{n} a_i \lambda_i e^{\lambda_i t} \mathbb{1}_{\mathbb{R}^+}(t) \text{ où } a_i = \prod_{j=1, j\neq i}^{n} \frac{\lambda_j}{\lambda_j - \lambda_i}, \ i = 1, ..., n$$

Soit $n \geq 2$, et $X_1, ..., X_n$, n variables aléatoires indépendantes distribuées selon des lois exponentielles de taux respectifs $\lambda_1, ..., \lambda_n$ tels que $\forall i \neq, \lambda_i \neq \lambda_j$.

1. *Montrer que $Z = X_1 + ... + X_n$ suit une loi hypoexponentielle d'ordre n et de paramètres $\lambda_1, ..., \lambda_n$.*
2. *Déterminer la transformée de Laplace des X_i. En déduire l'espérance et la variance des X_i.*
3. *Déterminer la transformée de Laplace de Z, son espérance et sa variance.*

Exercice II.7 **Calcul de fonction caractéristiques**

1. *Déterminer les fonctions caractéristiques des lois de probabilités suivantes :*
   (a) *binomiale de paramètres n et p,*
   (b) *poisson de paramètre λ,*
   (c) *exponentielle de paramètre λ,*
   (d) *normale centrée réduite, puis $N(m, \sigma^2)$.*
2. *Soit X et Y 2 variables aléatoires indépendantes de fonctions caractéristiques ϕ_X et ϕ_Y. Déterminer ϕ_{aX+b} où $a, b \in \mathbb{R}$ et ϕ_{X+Y} en fonction de ϕ_X et ϕ_Y.*
3. *En utilisant les questions précédentes, déterminer les lois des v.a. suivantes :*
   (a) *$X+Y$ où X et Y sont indépendantes de lois respectives $B(m,p)$ et $B(n,p)$,*
   (b) *$X+Y$ où X et Y sont indépendantes de lois respectives $P(\lambda)$ et $P(\mu)$,*
   (c) *$\sigma X + m$ où $X \sim \mathcal{N}(0,1)$,*
   (d) *$X - Y + 2Z$ où X, Y et Z sont indépendantes de loi $\mathcal{N}(0,1)$.*

III Vecteurs gaussiens

1 Définitions

a. Variable aléatoire gaussienne

Définition III.1 *Une variable aléatoire X réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ est dite* normale *(ou* gaussienne*) si sa loi image $\mathbb{P}_X$ a une densité par rapport à la mesure de Lebesgue de la forme :*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{\sigma^2}}, \ \text{avec } \mu \in \mathbb{R}, \ \sigma \in \mathbb{R}_+^*$$

On note $\mathbb{P}_X = \mathcal{N}(\mu, \sigma^2)$.
Si $\mu = 1$ et $\sigma = 0$, la variable aléatoire est dite gaussienne centrée réduite.

Proposition III.1 Théorème de changement de variable

$$X \sim \mathcal{N}(0,1) \Longrightarrow aX + b \sim \mathcal{N}(\mu, \sigma^2)$$

Proposition III.2 Moment d'une va gaussienne centrée réduite
Soit $X \sim \mathcal{N}(0,1)$, *et* $r \in \mathbb{N}$. *Alors*
- *Si* r *est impair,* $\mathbb{E}(X^r) = 0$
- *Si* r *est pair,*

$$\mathbb{E}(X^r) = 2^{\frac{r}{2}} \frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi}} = \frac{1}{\sqrt{\pi}}(r-1)(r-3)...3 \times 1$$

Proposition III.3 Fonction caractéristique d'une va gaussienne

$$\begin{aligned} X \sim \mathcal{N}(0,1) &\implies \phi_X(t) = e^{-\frac{t^2}{2}} \\ X \sim \mathcal{N}(\mu, \sigma^2) &\implies \phi_X(t) = e^{it\mu} e^{-\frac{\sigma^2 t^2}{2}} \end{aligned} \tag{1}$$

Proposition III.4 *Stabilité du caractère gaussien par combinaison linéaire*
Soit $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ *et* $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ *indépendantes, alors toute combinaison linéaire de* X *et* Y *est une variable aléatoire gaussienne :*

$$\alpha X + \beta Y \sim \mathcal{N}\left(\alpha\mu_X + \beta\mu_Y \,;\, \alpha^2\sigma_X^2 + \beta^2\sigma_Y^2\right)$$

b. Vecteurs gaussiens

Définition III.2 **Caractérisation par la loi**
$X = (X_1, ..., X_n)$ *vecteur aléatoire sur* $(\Omega, \mathcal{A}, \mathbb{P})$. X *est un* **vecteur gaussien** *(non dégénéré) s'il admet une densité par rapport à la mesure de Lebesgue sur* $\mathbb{R}^n$ *de la forme :*

$$\frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} e^{-\frac{{}^t(x-\mu)\Sigma^{-1}(x-\mu)}{2}}$$

où $\mu \in \mathbb{R}^n$ *et* Σ *matrice* $n \times n$ *symétrique définie positive.*

Définition III.3 **Caractérisation par stabilité par combinaison linéaire**
Soit $X = (X_1, ..., X_n)$ *un vecteur aléatoire sur* $(\Omega, \mathcal{A}, \mathbb{P})$.
X *est un vecteur gaussien ssi* $\forall \alpha \in \mathbb{R}^n$, ${}^t\alpha X = \sum_{i=1}^n \alpha_i X_i$ *est une variable aléatoire gaussienne.*

Exemple III.1 *Soit* $X \sim \mathcal{N}(0,1)$, ε *une variable aléatoire indépendante* X *de loi*

$$\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2}.$$

Soit $Y = \varepsilon X$. *Alors*

$\boxed{Y \sim \mathcal{N}(0,1)}$ *mais* (X, Y) *n'est pas un vecteur gaussien.*

Remarque III.1 *Mais si les va sont indépendantes, on a la réciproque.*

© Théo Jalabert

Proposition III.5 *Si X_1, X_2 sont des va gaussiennes* **indépendantes***, (X_1, X_2) est un couple gaussien, tout comme tout couple $(aX_1+bX_2, cX_1+dX_2), \forall a, b, c, d \in \mathbb{R}$.*

Remarque III.2 *Les va constantes sont considérées comme gaussiennes alors qu'elles ne satisfont pas la définition III.2 ou III.3. C'est ce qu'on appelle des variables gaussiennes* **dégénérées***, correspondant à $\sigma = 0$.*

Définition III.4 *Caractérisation par la fonction caractéristique*
$X = (X_1, ..., X_n)$ un vecteur aléatoire est dit **gaussien** *ssi il existe une matrice symétrique positive $\Sigma \in \mathcal{M}_n(\mathbb{R})$ (semi-définie) (qui est la matrice de variance-covariance de X) et μ un vecteur de $\mathbb{R}^n$ (égal à l'espérance de X) tels que la fonction caractéristique de X s'écrive :*

$$\boxed{\phi_X(t) = e^{i<t,\mu> - \frac{{}^t t \Sigma t}{2}}}$$

Proposition III.6 *Quand la matrice de variance-covariance de X est* **inversible***, les 3 définitions précédentes sont équivalentes.*
Dans ce cas, on parle de vecteur gaussien **non dégénéré***.*

c. Indépendance et vecteurs gaussiens

Théorème III.1 *Soit $X = (X_1, ..., X_n)$ un vecteur aléatoire gaussien non dégénéré. Les composantes de X sont indépendantes si et seulement si elles sont non corrélées : $Cov(X_j, X_k) = 0$, $\forall j \neq k$. Autrement dit, ssi la matrice de variance-covariance est diagonale.*

Corollaire III.1 *Si (X_1, X_2) est un vecteur gaussien, alors $X_1 \perp X_2$ ssi $Cov(X_1, X_2) = 0$.*

Remarque III.3 *Il est* **indispensable** *de savoir au préalable que le vecteur X est gaussien.*
C'est le **SEUL** *cas où $Cov(X, Y) = 0$ entraîne $X \perp Y$!*

Exemple III.2 *(même avec des va gaussiennes) Reprenons l'exemple précédent, nous avons montré que X et $Y = \varepsilon X$ suivaient la même loi normale. La covariance est nulle mais elles ne sont cependant pas indépendantes.*

Proposition III.7 *$Z = (X, Y)$, si $X \perp Y$ alors Z gaussien $\Leftrightarrow$ X et Y gaussien.*

2 Transformation de vecteurs gaussiens

a. Loi du χ^2

Si $X_1, ..., X_n$ sont des v.a. gaussiennes centrées réduites indépendantes, alors $U_n = X_1^2 + ... + X_n^2$ suit une loi de gamma de paramètres $(n/2, 1/2)$, appelée loi du χ^2 à un degrés de liberté (ou loi de Pearson). Sa densité est donnée par :

$$f(u) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} \exp\left(-\frac{u}{2}\right) u^{\frac{n}{2}-1} \quad u > 0$$

Ses moments sont obtenus par :

$$\mathbb{E}(U_n^r) = 2^r \frac{\Gamma(n/2 + r)}{\Gamma(n/2)}$$

avec en particulier $\mathbb{E}(U_n) = n$ *et* $\mathbb{V}(U_n) = 2n$ *(en utilisant* $\Gamma(n+1) = n\Gamma(n)$*.*

b. Autres lois dérivées

— *Loi de Student : Y et Z sont indépendants,* $Y \sim \mathcal{N}(0,1)$ *et* $Z \sim \chi^2(n)$*, alors :*

$$T_n = \frac{Y}{\sqrt{2/n}}$$

suit une loi de Student à n degrés de liberté. Sa densité vaut

$$f(u) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{u^2}{n}\right)^{-\frac{n+1}{2}}, \quad \forall\, u \in \mathbb{R}$$

— *Loi de Fisher-Snedecor : X et Y sont indépendantes et suivent des lois du* χ^2 *à n et m degrés de libretés respectivement, alors :*

$$F = \frac{X/n}{Y/m}$$

suit une loi de Fisher à n et m degrés de liberté notée $F(n,m)$ *et*

$$B = \frac{X}{Y}$$

suit une loi de Beta $\beta(n/2, m/2)$ *de deuxième espèce. Les densités d'une* $\beta(p,q)$ *et* $F(n,m)$ *sont*

$$f_B(x) = \frac{1}{B(p,q)} \frac{x^{p-1}}{(1+x)^{p+q}} \quad \forall\, x > 0$$

avec

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 x^{p-1}(1-x)^{q-1} dx$$

et

$$f_F(x) = \frac{1}{B\left(\frac{n}{2}, \frac{m}{2}\right)} n^{n/2} m^{m/2} \frac{x^{\frac{n}{2}-1}}{(m+nx)^{\frac{n+m}{2}}} \quad \forall\, x > 0$$

c. Statistique gaussienne

***Théorème III.2* Théorème de Cochran**

Soit $(X_1, ..., X_n)$ *un* $n-$*échantillon gaussien, de loi* $\mathcal{N}(m, \sigma^2)$*,* $\sigma > 0$*. Soit* $\bar{X}_n = \frac{X_1 + ... + X_n}{n}$ *la moyenne empirique de l'échantillon, et* $\bar{V}_n = \frac{1}{n}\left[(X_1 - \bar{X}_n)^2 + ... + (X_n - \bar{X}_n)^2\right]$ *sa variance empirique. Alors :*

1. $\bar{X}_n \sim \mathcal{N}(m, \frac{\sigma^2}{n})$ *;*
2. $\left(\frac{n}{\sigma^2}\bar{V}_n\right) \sim \chi^2_{n-1}$ *;*
3. $\bar{X}_n$ *et* $\bar{V}_n$ *sont indépendantes ;*
4. $\frac{\bar{X}_n - m}{\sqrt{\bar{V}_n/(n-1)}} \sim T_{n-1}$

© Théo Jalabert

3 Exercices

Questions 1 *Supposons que l'indice Dow Jones, noté D, soit une v.a. distribuée selon la loi $\mathcal{N}(\nu, \tau^2)$. Soit S le Swiss Market Index. La densité conditionnelle de S sachant que $D = d$ est une loi normale d'espérance d et de variance σ^2.*

1. *Quelle est la distribution du couple (D, S) ?*
2. *Quelle est la distribution conditionnelle de D sachant que $S = s$?*
3. *A partir de données historiques, on connait les valeurs de ν, τ et σ. Ayant observé une valeur s pour le Swiss Market index, calculer le meilleur prédicateur pour l'indice Dow Jones.*

IV Convergence de suites de variables aléatoires

1 Convergence presque sûre et en probabilité

Définition IV.1 — *On dit qu'une suite de vecteurs aléatoires X_n converge presque sûrement vers X ($X_n \overset{p.s.}{\to} X$)si*

$$\mathbb{P}\left(\omega \ : \ X_n(\omega) \to X(\omega)\right) = 1$$

Autrement dit, $X_n \overset{p.s.}{\to} X$ si l'ensemble des états ω pour lesquels X_n ne converge pas est de mesure nulle.

— *On dit qu'une suite de vecteurs aléatoires X_n converge en probabilité vers X ($X_n \overset{\mathbb{P}}{\to} X$) si*

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(|X_n - X| > \varepsilon\right) = 0$$

Remarque IV.1 *On note dans la suite $|x|$ la norme euclidienne de $x \in \mathbb{R}^k$ i.e. $(\sum x_j^2)^{1/2}$ mais le choix de la norme est indifférent.*

Remarque IV.2 *On peut aussi dire, quitte à enlever un ensemble de mesure nulle (celui pour lequel $X_n(\omega)$ ne converge pas vers $X(\omega)$), que $X_n \to X$ p.s. si et seulement si X_n converge ponctuellement vers X en tant que suite de fonctions de Ω dans $\mathbb{R}$. Il s'ensuit que si ϕ est une fonction continue $\mathbb{R}$ alors $\phi(X_n)$ converge vers $\phi(X)$ p.s.*

Lemme IV.1 (Borel-Cantelli Version 2) *Soit (X_n) une suite de v.a. et X une v.a. sur $(\Omega, \mathcal{A}, \mathbb{P})$ on a :*

$$\text{Si } \forall\ \varepsilon > 0 \quad \sum_{n=0}^{\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) < \infty \Rightarrow X_n \overset{p.s.}{\to} X$$

Si les X_n sont indépendants alors :

$$X_n \overset{p.s.}{\to} 0 \Leftrightarrow \sum_{n=0}^{\infty} \mathbb{P}(|X_n| \geq \varepsilon) < \infty \quad \forall\ \varepsilon > 0$$

Théorème IV.1 (CNS de cv p.s.) *Soit X une v.a. réelle et X_n une suite de v.a.r. définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ une condition nécessaire et suffisante pour que $X_n \stackrel{p.s.}{\to} X$ est*

$$\forall\ \varepsilon > 0 \quad \mathbb{P}\left(\left\{\omega \in \Omega,\ \sup_{m \geq n} |X_m(\omega) - X(\omega)| > \varepsilon\right\}\right) \to 0$$

ou

$$\mathbb{P}\left(\cup_{k \geq n}(|X_k - X| > \varepsilon)\right) \to 0$$

Théorème IV.2 (CS de cv p.s.) *Si $\forall\ \varepsilon > 0$ on a :*

$$\sum_{n=0}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$$

alors on a

$$X_n \stackrel{p.s.}{\to} X$$

Théorème IV.3 *Soit $(X_n)_n$ une suite de v.a. dans $\mathbb{R}^k$ alors*

$$X_n \stackrel{p.s.}{\to} X \Rightarrow X_n \stackrel{\mathbb{P}}{\to} X$$

Remarque IV.3 **Attention !** *La réciproque est fausse !*

Théorème IV.4

1. *Soit X_n une suite de v.a. de $\mathbb{R}^k$. Si $X_n \stackrel{\mathbb{P}}{\to} X$ alors on peut extraire une sous-suite $(X_{n_k})_{k \in \mathbb{N}}$ qui converge presque sûrement.*
2. *En fait on a même, $X_n \stackrel{\mathbb{P}}{\to} X \Leftrightarrow$ toute sous-suite de (X_n) contient une sous-suite qui converge p.s. (vers X).*
3. *Pour que X_n converge en proba vers X il faut et il suffit qu'elle soit de Cauchy pour la convergence en proba i.e. pour tout $\varepsilon > 0$, on a :*

$$\lim_{n,m} \mathbb{P}(|X_n - X_m| > \varepsilon) = 0$$

Théorème IV.5 (Invariance de la cv en proba par transformation continue) *$(X_n)_n$, X des v.a. sur $\mathbb{R}^k$ et f une fonction mesurable continue de $\mathbb{R}^k \to \mathbb{R}^l$, alors, si $X_n \stackrel{\mathbb{P}}{\to} X$ alors $f(X_n) \stackrel{\mathbb{P}}{\to} f(X)$.*

2 Petit Catalogue d'inégalités

Théorème IV.6 (Inégalité de Markov)

$$\forall \alpha > 0, \quad \mathbb{P}[|X| \geq \alpha] \leq \frac{1}{\alpha^k} \mathbb{E}[|X|^k]$$

Théorème IV.7 (Bienaymé-Tchebichev) *Supposons que X possède une espérance et une variance i.e. $X \in L^2$ alors*

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(|X - m| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sigma^2$$

© Théo Jalabert

Théorème IV.8 (Jensen) *φ est une fonction convexe sur un intervalle I. X est une v.a. à valeurs dans I telles que X et $\varphi(X)$ soient intégrables. Alors*

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$$

Théorème IV.9 (Hölder) *Soient p et q tels que $1/p + 1/q = 1$ alors :*

$$\mathbb{E}(|XY|) \leq \mathbb{E}^{1/p}(|X|^p) + \mathbb{E}^{1/q}(|X|^q)$$

Théorème IV.10 (Schwartz)

$$cov(X,Y) \leq \sqrt{\mathbb{V}(X)\mathbb{V}(Y)}$$

Théorème IV.11 (Lyapounov) *$0 < \alpha < \beta$ alors $\mathbb{E}^{1/\alpha}(|X|^\alpha) \leq \mathbb{E}^{1/\beta}(|X|^\beta)$*

3 Convergence dans L^p

Définition IV.2 (Convergence L^p) *Soit $(X_n)_n$ et X des v.a. de $\mathbb{R}^k$. on dit que X_n converge en moyenne d'ordre p (ou dans L^p) vers X si $|X|^p$ est intégrable et si*

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|^p) = 0$$

Remarque IV.4 .

Si $p = 2$, on parle de convergence en moyenne quadratique.

- *Si $p = 1$, on parle de convergence en moyenne.*
- *Si $r < p$, on a par Hölder $\|X\|_r \leq \|X\|_p$ Donc si $\|X_n - X\|_p \to 0$, on a $\|X_n - X\|_r \to 0$ quelque soit $r < p$. Donc la convergence $L^p \Rightarrow$ convergence L^r si $r < p$. En particulier si on converge dans L^2, on converge dans L^1.*

Proposition IV.1 (Convergence $L^1 \Rightarrow$ convergence en proba.) *$(X_n)_n$ une suite de v.a. de $\mathbb{R}^k$ alors $X_n \overset{L^1}{\to} X$ en moyenne $\Rightarrow X_n \overset{\mathbb{P}}{\to} X$*

Proposition IV.2 *$(X_n)_n$ est une suite de v.a. de $\mathbb{R}^k$. On suppose que :*

- *Tous les X_n sont p.s. majorés par K (constante), i.e. $|X_n| \leq K$ p.s.*
- *$X_n \overset{\mathbb{P}}{\to} X$*

Alors $X_n \overset{L^1}{\to} X$

Théorème IV.12 *En fait : $X_n \overset{L^1}{\to} X$ si et seulement si les 2 conditions suivantes sont vérifiées :*

1. *$X_n \overset{\mathbb{P}}{\to} X$*
2. *la suite X_n est uniformément intégrable, i.e.*

$$\forall \varepsilon > 0,\ \exists K > 0, \forall n \in \mathbb{N}, \mathbb{E}\left(|X_n| \mathbb{1}_{|X_n|>K}\right) < \varepsilon$$

© Théo Jalabert

4 Convergence en loi

a. Convergence étroite de mesure de probas

Définition IV.3 *On dit qu'une suite de mesures positives bornées $\mathbb{P}_1, ..., \mathbb{P}_n, ...$ sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ converge étroitement vers une mesure positive bornée P, et on note $\mathbb{P}_n \overset{e}{\to} \mathbb{P}$ si pour toute fonction réelle f définie sur $\mathbb{R}^k$, <u>continue et bornée</u>, on a :*

$$\int_{\mathbb{R}^k} f(x)\mathbb{P}_n(x) \to \int_{\mathbb{R}^k} f(x)d\mathbb{P}(x) \quad n \to \infty$$

Théorème IV.13 *Soient $\mathbb{P}_1, \mathbb{P}_2, ...$ des mesures de probas sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ de fonctions de répartition F_1, F_2,... Alors $\mathbb{P}_n \overset{e}{\to} \mathbb{P}$ si et seulement si $F_n(x) \to F(x)$ en tout point x de continuité de F.*

Remarque IV.5 *Si $F_n \to F$ pour tout point x de continuité de F, on dit que F_n converge complétement vers F et on note $\mathcal{F}_n \overset{c}{\to} F$*

Propriétés IV.1 *$(\mathbb{P}_n)$ une suite de mesures de probas sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ de densités f_n par rapport à la mesure de Lebesgue sur $\mathbb{R}^k$. Si f_n converge presque partout vers une densité f alors $\mathbb{P}_n \overset{e}{\to} \mathbb{P}$ où $\mathbb{P}$ est la mesure de densité de f.*

Théorème IV.14 *Soit (ϕ_n) une suite de transformées de Fourier de mesures de proba $\mathbb{P}_n)$ sur $(\mathbb{R}^k, (\mathbb{R}^k))$. Si ϕ_n converge en tout point de $\mathbb{R}^k$ vers ϕ, alors ϕ est la transformée de Fourier d'une mesure de proba $\mathbb{P}$ sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ et on a $\mathbb{P}_n \overset{e}{\to} \mathbb{P}$.*

b. Convergence en loi

Définition IV.4 *Soit $(X_n)_n$ une suite de vecteurs aléatoires de $\mathbb{R}^k$. On dit que X_n converge en loi vers le vecteur aléatoire X et on note $X_n \xrightarrow{\mathcal{L}} X$ si $\mathbb{P}_{X_n} \overset{e}{\to} \mathbb{P}_X$*

Propriétés IV.2 (Récapitulatif) *Il y a équivalence entre*
- *Convergence en loi de v.a.*
- *Convergence étroite de mesures de proba.*
- *Convergence complète de fonction de répartition.*

Propriétés IV.3 *Soir $(X_n)_n$ une suite de v.a. de $\mathbb{R}^k$ alors $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si quelque soit $\lambda = (\lambda_1, ..., \lambda_k) \in \mathbb{R}^k$, on a ${}^t\lambda.X_n \xrightarrow{\mathcal{L}} {}^t\lambda.X$.*

Remarque IV.6 *En prenant $\lambda_i = 1$ et $\lambda_j = 0$ pour tout $j \neq i$, on obtient que la convergence en loi d'un vecteur aléatoire implique la convergence en loi des composantes de X_n (notée X_n^i vers les composantes de X (notées X^i).*

Propriétés IV.4 *Si $X_n \xrightarrow{\mathcal{L}} X$ (vecteurs aléatoires de $\mathbb{R}^k$) et si h est une fonction continue de $\mathbb{R}^k$ dans $\mathbb{R}^n$ alors :*

$$h(X_n) \xrightarrow{\mathcal{L}} h(X)$$

© Théo Jalabert

Théorème IV.15 (Paul Levy) *Soit $(X_n)_n$ une suite de vecteurs aléatoires de $\mathbb{R}^d$ de fonctions caractéristiques ϕ_{X_n}. Alors :*

1. *Si $X_n \xrightarrow{\mathcal{L}} X$, alors $\phi_{X_n} \to \phi_X$ la fonction caractéristique de X (convergence simple de fonction et même convergence uniforme sur tout compact de $\mathbb{R}^d$).*
2. *Inversement, si (ϕ_{X_n}) convergence simplement vers une fonction ϕ continue en 0, alors ϕ est la transformée de Fourier d'une proba μ sur $\mathbb{R}^d$ et $X_n \xrightarrow{\mathcal{L}} \mu$. De plus, il existe un v.a. définie sur un espace probabilisé $\Omega, \mathcal{A}, \mathbb{P})$ à valeur dans $\mathbb{R}^d$ telle que $X_n \xrightarrow{\mathcal{L}} X$.*

Théorème IV.16 (cv en proba ⇒ cv en loi) *$X, X_1, X_2, ...$ une suite de v.a. de $\mathbb{R}^k$ alors*

$$X_n \xrightarrow{\mathbb{P}} X \Longrightarrow X_n \xrightarrow{\mathcal{L}} X$$

Théorème IV.17 (Portmanteau) *Les trois conditions suivantes sont équivalentes :*

1. *$X_n \xrightarrow{\mathcal{L}} X$.*
2. *Pour toute fonction réelle f continue et bornée, $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$.*
3. *Pour tout borélien de A de $\mathbb{R}$ tq $\mathbb{P}(X_n \in \partial A = \bar{A}\backslash\mathring{A}) = 0$, $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$.*

Théorème IV.18 *Si a est une constante alors $X_n \xrightarrow{\mathcal{L}} a \Longrightarrow X_n \xrightarrow{\mathbb{P}} a$.*

5 Liens entre convergence : Cas gaussien

Proposition IV.3 *Soit (X_n) une suite de v.a. centrées de loi Normale. Alors les propriétés suivantes sont équivalentes :*

1. *$\sum X_n$ converge p.s.*
2. *$\sum X_n$ converge en Proba*
3. *$\sum X_n$ converge dans L^2*
4. *$\sum X_n$ converge en loi*

6 Théorème Limite

a. Loi des grands nombres

Proposition IV.4 *Si X est une v.a. admettant un moment d'ordre 2 alors la fonction caractéristique admet un DL d'ordre 2 en 0, donné par, quelque soit t dans $\mathbb{R}$*

$$\varphi_X(t) = 1 + it\mathbb{E}(X) - \frac{t^2}{2}\mathbb{E}(X^2) + o(t^2)$$

© Théo Jalabert

Théorème IV.19 (Loi faible des grands nombres ou Théorème de Khintchine) *Soit* X_n *une suite de v.a. définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$ *indépendantes, de même loi, intégrables et d'espérance* $\mathbb{E}(X_i) = \mu$, *alors*

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{S_n}{n} \xrightarrow{\mathbb{P}} \mu$$

Théorème IV.20 (Loi forte des grands nombres (Kolmogorov)) X_n *une suite de v.a. de* $\mathbb{R}^d$ *définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$ *indépendantes, de même loi, intégrables d'espérance* μ. *Alors*

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{S_n}{n} \xrightarrow{p.s.} \mu$$

Théorème IV.21 (Loi faible des grands nombres dans L^2***)*** X_n *une suite de v.a. dans* $\mathbb{R}^d$ *sur* $(\Omega, \mathcal{A}, \mathbb{P})$, *non corrélées, de carré intégrable, de même espérance* μ *et même variance* σ^2. *Alors*

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{S_n}{n} \xrightarrow{L^p} \mu$$

Théorème IV.22 (Troisième loi faible des grands nombres) *Soit* $(X_n)_n$ *une suite de v.a. définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$ *admettant un moment d'ordre 2, deux à deux non corrélées. On suppose la convergence des suites :*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_i) \to m \quad \text{et} \quad \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(X_i) \to 0$$

alors la suite de v.a. $\overline{X}$ *converge en proba vers* m.

Définition IV.5 *Soit* $X_1, ..., X_n$ *des v.a. de même loi que* X. *La fonction* F_n *de* $\mathbb{R}, \Omega$ *dans* $[0, 1]$ *définie par :*

$$\forall\ (x, \omega) \in \mathbb{R} \times \Omega \quad F_n(x, \omega) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(X_j \leq x)}(\omega)$$

est appelée fonction de répartition empirique (associé à X*) basée sur l'échantillon* $X_1, ..., X_n$

Théorème IV.23 (Glivenko-Cantelli) *Pour* $\mathbb{P}$*-presque tout* ω, *la suite des fonctions de répartitions* $F_n(., \omega)$ *converge uniformément vers* F, *autrement dit, on a :*

$$\mathbb{P} - p.s. \quad \lim_n \sup |F_n(x, .) - F(x, .)| = 0.$$

b. Théorème Central Limite

Théorème IV.24 *TCL unidimensionnel $(X_n)_n$ une suite de v.a.r. indépendantes de même loi, de carré intégrable (i.e. $\mathbb{E}(X_n^2) < \infty$), d'espérance μ et de variance σ^2. On note par*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Alors

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma} = \sqrt{n}\frac{\overline{X}-\mu}{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

Théorème IV.25 *TCL multi (X_n) suite de v.a. de $\mathbb{R}^d$ indépendants, de même loi, de carré intégrable. On note μ et Γ le vecteur moyenne et la matrice de variance. Alors*

$$\sqrt{n}\left(\overline{X}-\mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Gamma)$$

c. Approximations de lois

Théorème IV.26 (Limite central Poissonien) *Soit S_n une v.a. de loi $\mathcal{B}(n,p_n)$. Si $\lim_n np_n = \lambda > 0$, S_n converge en loi vers une v.a. de Poisson de paramètre λ.*

Théorème IV.27 (Approximation de la loi Binomiale par la Normale) *Soit (X_n) une suite de v.a. réelle avec $X_n \sim \mathcal{B}(n,p)$. Alors $(q = 1-p)$*

$$\frac{X_n - np}{\sqrt{npq}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

Théorème IV.28 *Soit (X_n) une suite de vecteurs aléatoires dans $\mathbb{R}^k$. X_n suit une loi multinomiale $\mathcal{M}(n,p_1,...,p_m)$ $(p_i > 0, \sum p_i = 1)$, p le vecteur de dimension m et p le vecteur de composantes p_i. Soit $q_i = 1-p_i$, alors :*

$$\frac{X_n - np}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Gamma)$$

où Γ est la mtrice de termes diagonaux p_iq_i et de termes croisés $-p_iq_j$

Théorème IV.29 (Convergence de la loi de Poisson) *Si X_n une suite de v.a. de Poisson de paramètre λ_n avec $\lambda_n \to \infty$ alors*

$$\frac{X_n - \lambda}{\sqrt{\lambda}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

V Théorème de Radon-Nikodym

1 Rappels sur les mesures de probabilités

Définition V.1 Absolue continuité et équivalence de mesures
Soit P et Q deux mesures positives définies sur $(\Omega,\mathcal{A})$.

© Théo Jalabert

1. *On dit que* $\mathbb{Q}$ *est* **absoument continue** *par rapport à* $\mathbb{P}$ *(ou aussi que* $\mathbb{P}$ *domine* $\mathbb{Q}$*) et on note* $\mathbb{Q} \ll \mathbb{P}$ *si*

$$\forall A \in \mathcal{A},\ \mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0$$

2. *On dit que* $\mathbb{Q}$ *et* $\mathbb{P}$ *sont* **équivalentes** *et on note* $\mathbb{Q} \sim \mathbb{P}$ *si* $\mathbb{Q} \ll \mathbb{P}$ $\mathbb{P} \ll \mathbb{Q}$.

Remarques :
- *Si* $\mathbb{Q} \ll \mathbb{P}$*, les ensembles négligeables pour* $\mathbb{P}$ *sont négligeables pour* $\mathbb{Q}$.
- *Si* $\mathbb{Q} \sim \mathbb{P}$*, les 2 mesures ont les mêmes ensembles négligeables.*
- *Deux mesures sont dites* **étrangères** $(\mathbb{P} \perp \mathbb{Q})$ *si* $\exists A \in \mathcal{A}$ *tel que* $\mathbb{P}(A) = 0$ *et* $\mathbb{Q}(A) = 1$ $(\mathbb{Q}(^{A}) = 0)$.

2 Théorème de Radon-Nikodym

Proposition V.1 *Soit X une variable aléatoire réelle* positive*, définie sur un espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$*, d'espérance* $\mathbb{E}^{\mathbb{P}}(X) = 1$.
L'application $\mathbb{Q}$ *définie sur* $\mathcal{A}$ *par*

$$\boxed{\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}(X \mathbb{1}_A)}$$

est une mesure de probabilité absolument continue par rapport à la mesure de probabilité $\mathbb{P}$.

Définition V.2 *Soit* $(\Omega, \mathcal{A})$ *un espace probabilisable,* $\mathbb{P}$ *une mesure* σ*-finie et X une variable aléatoire réelle positive, telle que* $\int_\Omega X(\omega) d\mathbb{P}(\omega) = 1$.
Soit $\mathbb{Q}$ *la probabilité définie sur* $(\Omega, \mathcal{A})$ *par*

$$\forall A \in \mathcal{A}, \mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}(X \mathbb{1}_A).$$

On appelle X **densité de la probabilité** $\mathbb{Q}$ *par rapport à la probabilité* $\mathbb{P}$*, et on note* $\frac{d\mathbb{Q}}{d\mathbb{P}} = X$.

Théorème V.1 **(Théorème de Radon-Nikodym)**
Soit $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probablisé et* $\mathbb{Q}$ *une mesure de probabilité sur* $(\Omega, \mathcal{A})$ *absolument continu par rapport à* $\mathbb{P}$.
alors il existe une classe (pour l'égalité $\mathbb{P}$*-p.s.) unique de variables aléatoires X réelles positives d'espérance 1 sous* $\mathbb{P}$ *telle que*

$$\forall A \in \mathcal{A}, \mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[X \mathbb{1}_A].$$

On note $X = \frac{d\mathbb{Q}}{d\mathbb{P}}$*, et on l'appelle* **densité de probabilité** *de* $\mathbb{Q}$ *par rapport à* $\mathbb{P}$*, ou* **dérivée de Radon-Nikodym** *de* $\mathbb{Q}$ *par rapport à* $\mathbb{P}$.

Proposition V.2 **(Calcul d'espérance sous une nouvelle probabilité)**
Soit $(\Omega, \mathcal{A})$ *un espace probabilisable. Soit* $\mathbb{Q}$ *une mesure de probabilité sur* $(\Omega, \mathcal{A})$*, absolument continue par rapport à* $\mathbb{P}$ *mesure de probabilité sur* $(\Omega, \mathcal{A})$.
Soit $\frac{d\mathbb{Q}}{d\mathbb{P}}$ *la densité de* $\mathbb{Q}$ *par rapport à* $\mathbb{P}$.

© Théo Jalabert

Soit Y la variable aléatoire définie sur $(\Omega, \mathcal{A})$, intégrable par rapport à $\mathbb{Q}$.
Alors $Y\frac{d\mathbb{Q}}{d\mathbb{P}}$ est intégrable par rapport à $\mathbb{P}$ et

$$\mathbb{E}^{\mathbb{P}}\left[Y\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \mathbb{E}^{\mathbb{Q}}[Y]$$

Soit

$$\int_\Omega Y(\omega)\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)d\mathbb{P}(\omega) = \int_\Omega Y(\omega)d\mathbb{Q}(\omega)$$

Proposition V.3

1. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable, $\mathbb{P}$ et $\mathbb{Q}$ deux probabilités équivalentes„ $\mathbb{P} \sim \mathbb{Q}$. Alors*

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \frac{1}{\frac{d\mathbb{Q}}{d\mathbb{P}}},\ \mathbb{P}\text{-}p.s.$$

2. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable, $\mathbb{P}, \mathbb{Q}, \mathbb{R}$ trois probabilités avec $\mathbb{Q} \ll \mathbb{P}$ et $\mathbb{R} \ll \mathbb{Q}$. Alors $\mathbb{R} \ll \mathbb{P}$ et*

$$\frac{d\mathbb{R}}{d\mathbb{P}} = \frac{d\mathbb{R}}{d\mathbb{Q}} \times \frac{d\mathbb{Q}}{d\mathbb{P}},\ \mathbb{P}\text{-}p.s.$$

3 Probabilités équivalentes dans les modèles gaussiens

Soit U une variable aléatoire gaussienne de moyenne $m_{\mathbb{P}}$ et de variance $\sigma^2_{\mathbb{P}} > 0$. Posons

$$Y = \exp\left(\lambda(U - m_{\mathbb{P}}) - \frac{\lambda^2}{2}\sigma^2_{\mathbb{P}}\right).$$

Alors Y est une variable aléatoire strictement positive, d'espérance 1, qui définit une nouvelle probabilité $\mathbb{Q}$ sous laquelle U est également une variable aléatoire gaussienne.

$$m_{\mathbb{Q}} = m_{\mathbb{P}} + \lambda\sigma^2_{\mathbb{P}},\ \sigma^2_{\mathbb{Q}} = \sigma^2_{\mathbb{P}}$$

La variable aléatoire $U + \lambda\sigma^2_{\mathbb{P}}$ sous $\mathbb{P}$ est gaussienne, de variance $\sigma^2_{\mathbb{P}}$ et de moyenne $m_{\mathbb{Q}} = m_{\mathbb{P}} + \lambda\sigma^2_{\mathbb{P}}$. Elle a la même loi que U sous la probabilité $\mathbb{Q}$.

$$\mathbb{E}_{\mathbb{P}}\left(f\left(U + \lambda\sigma^2_{\mathbb{P}}\right)\right) = \mathbb{E}_{\mathbb{Q}}(f(U)) = \mathbb{E}_{\mathbb{P}}\left(\exp\left(\lambda(U - m_{\mathbb{P}}) - \frac{\lambda^2}{2}\sigma^2_{\mathbb{P}}\right) f(U)\right)$$

VI Conditionnement

1 Probabilités conditionnelles

a. Rappels sur les probabilités conditionnelles

Définition VI.1 *Soit $(\Omega, \mathcal{A}, \mathbb{P})$un espace probabilisé, et B un évènement* non négligeable *($\mathbb{P}(B) \neq 0$). On appelle* **probabilité conditionnelle de A sachant B** *le réel noté $\mathbb{P}(A|B)$:*

$$\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}} \tag{2}$$

Proposition VI.1 *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$*un espace probabilisé.*

1. *Si A et B sont deux évènements indépendants non négligeables, alors*
$$\mathbb{P}(A|B) = \mathbb{P}(A) \text{ et } \mathbb{P}(B|A) = \mathbb{P}(B).$$
2. *Si A et B sont deux évènements non négligeables, alors*
$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B).\mathbb{P}(B) = \mathbb{P}(B|A).\mathbb{P}(A).$$
3. *Soient* A_i $i \in \mathbb{N}$ *une suite d'évènements mutuellement disjoints, alors*
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B).$$

Exercice VI.1 *Soit* (X, Y) *un vecteur aléatoire défini sur un espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$. *On suppose que la loi de* (X, Y) *admet une densité jointe par rapport à la mesure de Lebesgue sur* $\mathbb{R}^2$ *donnée par*
$$f(x, y) = (x + y)\mathbb{1}_{[0,1]}(x, y)$$

1. *Vérifier que f définit bien une densité de probabilité.*
2. *Calculer* $\mathbb{P}(2X \leq 1|X + Y \leq 1)$.

Exercice VI.2 *Soit X une variable aléatoire suivant une loi exponentielle de paramètre* λ. *Soit* x, y *deux réels strictement positifs. Montrer que*
$$\mathbb{P}(X > x + y|X > x) = \mathbb{P}(X > y).$$

Proposition VI.2 *(Chain Rule)*
Soient $A_1, A_2, ..., A_n$ *une suite d'évènements dont l'intersection* $\bigcap_{i=1}^{n} A_i$ *est non négligeable. Alors*
$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2).....\mathbb{P}(A_n|A_1 \cap ... \cap A_{n-1})$$

Preuve : *par récurrence. Poser* $B_n = A_1 \cap ... \cap A_n$.
le cas $n = 2$ *est facile. On suppose ensuite que l'hypothèse de récurrence est vraie au rang* n, *on l'applique à* B_n *et* A_{n+1} *et on applique le cas* $n = 2$.

Exercice VI.3 *On tire consécutivement 3 balles d'une urne contenant 5 boules blanches et 5 boules noires, sans remise. Quelle est la probabilité que toutes les boules tirées soient noires ?*

Proposition VI.3 *(Formule des probabilités totales)*
Soit $(\Omega, \mathcal{A}, \mathbb{P})$*un espace probabilisé et* $(A_i)_{i \in I \subset \mathbb{N}}$ *une partition de* Ω. *Alors pour tout évènement B :*
$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Proposition VI.4 *(Formule de Bayes)*
Soit $(\Omega, \mathcal{A}, \mathbb{P})$*un espace probabilisé, et* $(A_i)_{i\in\mathbb{N}}$ *une partition de* Ω*, formée d'évènements non négligeables. Alors pour tout évènement* B*,* $\forall i \in \mathbb{N}$ *:*

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i\in\mathbb{N}} \mathbb{P}(B|A_i)\mathbb{P}(A_i)}$$

Proposition VI.5 *Soient* A_1 *et* A_2 *deux évènements non négligeables disjoints,* B *un évènement quelconque. Alors*

$$\mathbb{P}(B|A_1 \cup A_2) = \mathbb{P}(B|A_1)\frac{\mathbb{P}(A_1)}{\mathbb{P}(A_1) + \mathbb{P}(A_2)} + \mathbb{P}(B|A_2)\frac{\mathbb{P}(A_2)}{\mathbb{P}(A_1) + \mathbb{P}(A_2)}$$

Preuve : *Exercice... Il suffit d'appliquer la définition de la probabilité conditionnelle, et utiliser le fait que* A_1 *et* A_2 *sont disjoints.*

b. Loi d'une v.a. conditionnelle à un événement

Définition VI.2 *Soit* X *une variable aléatoire définie sur un espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$*, à valeurs dans un espace probabilisable* $(\Omega', \mathcal{A}')$*. Soit* B *un évènement non négligeable de* $\mathcal{A}$ *(*$\mathbb{P}(B) \neq 0$*). On définit la* **loi conditionnelle de** X **sachant** B*, notée* $\mathbb{P}_{X|B}$ *la* probabilité image de $\mathbb{P}(.|B)$ par X*, définie sur* $(\Omega', \mathcal{A}')$*, et qui à tout évènement* $A' \in \mathcal{A}'$ *associe :*

$$\boxed{\mathbb{P}_{X|B}(A') = \mathbb{P}(X^{-1}(A')|B) = \mathbb{P}(X \in A'|B)}$$

Définition VI.3 (Fonction de répartition et densité conditionnelle)

— *On peut alors définir* **la fonction de répartition de la loi conditionnelle de** X **sachant** B *comme :*

$$\forall x \in \mathbb{R},\ F_{X|B}(x) = \mathbb{P}(X < x|B) = \mathbb{P}_{X|B}(]-\infty, x[).$$

C'est la fonction de répartition de la probabilité image $\mathbb{P}_{X|B}$ *de* $\mathbb{P}(.|B)$ *par* X*.*

— *Si la loi de* X *sachant* B*,* $\mathbb{P}_{X|B}$ *admet une densité (par rapport à la mesure de Lebesgue), on l'appelle* **densité conditionnelle de** X **sachant** B*, on la note* $f_{X|B}$ *et on a :*

$$\forall A \in \mathcal{B}\mathbb{R},\ \mathbb{P}_{X|B}(A) = \mathbb{P}(X \in A|B) = \int_A f_{X|B}(x)dx.$$

© Théo Jalabert

Définition VI.4 (Espérance d'une variable aléatoire conditionnelle à un évènement)
On définit à présent l'espérance de X conditionnelle à B comme l'espérance de X sous la probabilité $\mathbb{P}(.|B)$.
Par le théorème de tansfert,

$$\mathbb{E}(X|B) = \mathbb{E}^{\mathbb{P}(.|B)}(X) = \mathbb{E}^{\mathbb{P}_{X|B}}(x) = \int_{\mathbb{R}} x dF_{X|B}(x)$$

Dans le cas d'existence d'une densité conditionnelle $f_{X|B}$, l'espérance de X sachant B est

$$\mathbb{E}(X|B) = \int_{\mathbb{R}} x f_{X|B}(x) dx.$$

2 Espérance conditionnelle

a. Espérance conditionnelle sachant une sous-tribu

Définition VI.5 *Soit $(\Omega, \mathcal{A}, \mathbb{P})$un espace probabilisé, et $\mathcal{F}$ une sous-tribu de $\mathcal{A}$. Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{A})$, $\mathbb{P}$-intégrable.*
On définit **l'espérance conditionnelle de X sachant** $\mathcal{F}$*, et on note* $Y = \mathbb{E}(X|\mathcal{F})$ *comme la* variable aléatoire Y *sur* $(\Omega, \mathcal{F})$*, telle que*

1. *Y est* mesurable *par rapport à $\mathcal{F}$, càd*

$$\forall B \in \mathcal{B}(\mathbb{R}),\ Y^{-1}(B) \in \mathcal{F}$$

2. *$\forall A \in \mathcal{F}$, on a*

$$\boxed{\int_A X d\mathbb{P} = \int_A Y d\mathbb{P}}$$

autrement dit :

$$\boxed{\mathbb{E}^{\mathbb{P}}(X 1\!\!1_A) = \mathbb{E}^{\mathbb{P}}(Y 1\!\!1_A)}$$

Proposition VI.6 **Conséquences directes de la définition**

1. *Si $\mathcal{F} = \mathcal{A}$, alors $\mathbb{E}(X|\mathcal{A}) = X$.*
   Ou encore, si X est $\mathcal{F}$-mesurable, $\mathbb{E}(X|\mathcal{F}) = X$.
2. *Si $\mathbb{E}(X|\mathcal{F})$ existe, alors*

$$\mathbb{E}\left(\mathbb{E}(X|\mathcal{F})\right) = \mathbb{E}(X)$$

3. *Si X est constante,*

$$\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X) = X$$

4. *Si $\mathcal{F}$ est la tribu triviale $\{\Omega, \emptyset\}$, alors*

$$\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X)$$

Proposition VI.7 (Existence et Unicité de l'espérance conditionnelle) *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ et $\mathcal{F}$ une sous-tribu de $\mathcal{A}$. Alors*

© Théo Jalabert

1. *l'espérance conditionnelle de X sachant $\mathcal{F}$, $\mathbb{E}(X|\mathcal{F})$ existe,*
2. *Si Z et Z' sont deux espérances conditionnelles de X sachant $\mathcal{F}$ alors $Z = Z'$ p.s.*

De même, si Z' est une espérance conditionnelle de $X|\mathcal{F}$ et $Z' = Z$ p.s. alors Z' est une espérance conditionnelle de $X|\mathcal{F}$.

Propriétés VI.1 *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, X, $Y \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, $\mathcal{F}$ sous tribu de $\mathcal{A}$. Alors on a les propriétés suivantes :*

1. **Linéarité** *: Quelque soit $a, b \in \mathbb{R}$, $\mathbb{E}(aX + bY|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + b\mathbb{E}(Y|\mathcal{F})$ p.s.*
2. **Positivité** *: Si $X \geq 0$ alors $\mathbb{E}(X|\mathcal{F}) \geq 0$ $\mathbb{P}$-p.s.*
3. **Croissance** *: Si $X \leq Y$ $\mathbb{P}$-p.s. alors $\mathbb{E}(X|\mathcal{F}) \leq \mathbb{E}(Y|\mathcal{F})$ $\mathbb{P}$-p.s.*
4. *On a*
$$(\mathbb{E}(X|\mathcal{F}))^+ \leq \mathbb{E}(X^+|\mathcal{F})$$
et pareil avec -
$$(\mathbb{E}(X|\mathcal{F}))^- \leq \mathbb{E}(X^-|\mathcal{F})$$

Proposition VI.8 *Soit X une v.a. $\mathbb{P}$ intégrable, $\{B_i, i \in \mathbb{N}\}$ un système complet d'événements non négligeables. Soit $\mathcal{F}$ la sous-tribu engendrée par les $B_i, i \in \mathbb{N}$. Alors*
$$\mathbb{E}(X|\mathcal{F}) = \sum_{i \in \mathbb{N}} \frac{\mathbb{E}(X \mathbb{1}_{B_i})}{\mathbb{P}(B_i)} \mathbb{1}_{B_i}$$

Proposition VI.9 *X et Y deux v.a. sur $(\Omega, \mathcal{A}, \mathbb{P})$, $\mathcal{F}$ sous -tribu de $\mathcal{A}$*

1. *Si XY et Y sont $\mathbb{P}$ intégrables et si X est $\mathcal{F}$ mesurable alors*
$$\mathbb{E}(XY|\mathcal{F}) = X\mathbb{E}(Y|\mathcal{F}) \quad \mathbb{P}.p.s.$$
2. *Esprances itérées : si $\mathcal{F} \subset \mathcal{G}$,*
$$\mathbb{E}\left(\mathbb{E}(X|\mathcal{G})|\mathcal{F}\right) = \mathbb{E}(X|\mathcal{F})$$
$$\text{et } \mathbb{E}\left(\mathbb{E}(X|\mathcal{F})|\mathcal{G}\right) = \mathbb{E}(X|\mathcal{F})$$
3. *Inégalité de Jensen : Si f est une fonction réelle convexe alors :*
$$f\left(\mathbb{E}(X|\mathcal{F})\right) \leq \mathbb{E}\left(f(X)|\mathcal{F}\right)$$

Théorème VI.1 (Beppo-Levi conditionnel) *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $\mathcal{F}$ une sous -tribu de $\mathcal{A}$ et X_n une suite croissante de v.a. positives, convergeant vers X alors la suite de v.a. $\mathbb{E}(X_n|\mathcal{F})$ converge vers $\mathbb{E}(X|\mathcal{F})$.*

Théorème VI.2 (Convergence dominé conditionnel) *Soit X_n une suite de v.a. positives de Z une v.a. intégrables tq quelque soit n $|X_n| \leq Z$. On suppose que $X_n \to X$ p.s. Alors $\mathbb{E}(X_n|\mathcal{F})$ converge p.s. vers $\mathbb{E}(X|\mathcal{F})$.*

Définition VI.6 *Si $\mathcal{F}$ est engendrée par une variable aléatoire discrète Y, on note $\mathbb{E}(X|Y) = \mathbb{E}(X|\mathcal{F})$.*

Exemple VI.1 *Soit X une v.a. sur $(\Omega, \mathcal{A}, \mathbb{P})$ suivant une loi de poisson de paramètre $\lambda > 0$ et soit $Y = 2\lfloor X/Z \rfloor$. Calculer $\mathbb{E}(X|Y)$ et $\mathbb{E}(Y|X)$.*

© Théo Jalabert

b. Espérance conditionnelle par rapport à une variable aléatoire

Définition VI.7 *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé,* $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ *et* $Y_1, ..., Y_n$ *des v.a. définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$*. On définit* $\mathbb{E}(X|Y_1, ..., Y_n)$ *(l'espérance conditionnelle de* X *sachant* $Y_1, ..., Y_n$*) comme* $\mathbb{E}(X|\sigma(Y_1, ..., Y_n))$*.*

Proposition VI.10 (Lemme de Doob) *Si* $(\Omega, \mathcal{A}, \mathbb{P})$ *est un espace probabilisable et si* $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ *et* Y *est une v.a. alors* $\mathbb{E}(X|Y)$ *est une fonction mesurable de* Y*.*

Remarque VI.1 $\mathbb{E}(X|Y)$ *est donc la meilleure approximation de* X *qui soit* Y *mesurable.*

Proposition VI.11 *Si* X *et* Y *sont deux v.a. indep. avec* $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ *alors*

$$\mathbb{E}(X|Y) = \mathbb{E}(X) \quad \mathbb{P} - p.s.$$

Propriétés VI.2 (X, Y) *couple de v.a.r.* Ψ *une fonction mesurable sur* $\mathbb{R}^2$ *et* h *une fonction mesurable sur* $\mathbb{R}$*. Alors si* $\Psi(X, Y)$ *et* $\Psi(X, Y)h(Y)$ *sont intégrables alors :*

1. $\mathbb{E}\left(\mathbb{E}(\Psi(X, Y)|Y)\right) = \mathbb{E}\left(\Psi(X, Y)\right)$.
2. $\mathbb{E}\left(h(Y)\Psi(X, Y)|Y\right) = h(Y)\mathbb{E}\left(\Psi(X, Y)|Y\right)$

c. Variance conditionnelle

Définition VI.8 $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé,* X *une v.a.r. de carré intégrable,* $\mathcal{F}$ *une sous-tribu de* $\mathcal{A}$*. On définit la variance condtionnelle de* X *sachant* $\mathcal{F}$ *et on note*

$$\mathbb{V}(X|\mathcal{F}) = \mathbb{E}\left((X - \mathbb{E}(X|\mathcal{F}))^2|\mathcal{F}\right)$$

De manière identique, on définit pour Y *v.a.*

$$\mathbb{V}(X|Y) = \mathbb{E}\left((X - \mathbb{E}(X|Y))^2|Y\right)$$

Proposition VI.12 $\mathbb{V}(X|\mathcal{F})$ *est bien définie, p.s. finie et positive et*

$$\mathbb{V}(X|\mathcal{F}) = \mathbb{E}\left(X^2|\mathcal{F}\right) - \mathbb{E}(X|\mathcal{F})^2$$

Proposition VI.13 (Formule de décomposition de la variance) *Soit* (X, Y) *un couple de v.a.r. définies sur* $(\Omega, \mathcal{A}, \mathbb{P})$*,* X *de carré intégrable. Alors :*

$$\mathbb{V}(X) = \mathbb{E}\left(\mathbb{V}(X|Y)\right) + \mathbb{V}\left(\mathbb{E}(X|Y)\right)$$

d. Lois de probabilités conditionnelles

Propriétés VI.3 (de la proba si il y a densité) *Soit X une v.a.r. de densité $f_{X,Y}(x,y)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. On note $f_X(x)$ et $f_Y(y)$ les densités marginales. Notons*

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Alors, on a :

$$\forall\ C \in \mathcal{B},\ \mathbb{P}(X \in C|Y = y) = \int_C f_{X|Y}(x|y)dx \qquad p.s.$$

Propriétés VI.4 *Soit (X,Y) un couple de v.a.r. de densité $f_{X,Y}(x,y)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$, X étant $\mathbb{P}$-intégrable. On note $f_X(x)$ et $f_Y(y)$ les densités marginales de X et Y. On note*

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Soit g une fonction mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$ tq $g(X,Y)$ soit $\mathbb{P}$-intégrable. Alors :

$$\mathbb{E}(g(X,Y)|Y = y) = \int_{\mathbb{R}} g(X,Y) f_{X|Y}(x|y)dx$$

et en particulier

$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x f_{X|Y}(x|y)dx.$$

Propriétés VI.5 (Cas particulier : discret) *Soit (X,Y) un couple de v.a. discrètes, prenant les valeurs (x_i, y_j), $i = 1, 2...$ et $j = 1, 2...$ défini sur $(\Omega, \mathcal{A}, \mathbb{P})$ espace probabilisé. On suppose que quelque soit i et j $\mathbb{P}(X = x_i, Y = y_j) > 0$ alors :*

$$\mathbb{P}(X = x_i|Y = y_j) = \frac{\mathbb{P}(X = x_i, Y = Y_j)}{\mathbb{P}(Y = y_j)}$$

Questions 2 *Soit U_1 et U_2 deux lois uniformes entre 0 et 1. On pose $X = \min(U_1, U_2)$ et $Y = \max(U_1, U_2)$. Calculer $\mathbb{E}(X|Y)$ et $\mathbb{E}(Y|X)$.*

e. Lois normales conditionnelles

Proposition VI.14 *Soit $X = (X_1, X_2)$ un vecteur al{eatoire défini sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^n$ (X_1 dans $\mathbb{R}^p$ et X_2 dans $\mathbb{R}^{n-p}$). On suppose que*

$$\mathbb{P}_{X_1} \sim \mathcal{N}(\mu_1, Q_1)$$

avec $\mu_1 \in \mathbb{R}^p$ et $Q_1 \in \mathcal{M}(p,p)$ et

$$\mathbb{P}_{X_2|X_1=x_1} \sim \mathcal{N}(\mu_2 + M(x_1 - \mu_1), Q_2)$$

© Théo Jalabert

avec $\mu_2 \in \mathbb{R}^{n-p}$ *et* $Q_2 \in \mathcal{M}(n-p, n-p)$*. Alors la loi jointe* $\mathbb{P}_X = \mathbb{P}_{X_1,X_2}$ *est une loi gaussienne*

$$\mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, Q\right)$$

avec

$$Q = \begin{pmatrix} Q_1 & Q_1^t M \\ MQ_1 & Q_2 + MQ_1^t M \end{pmatrix}$$

Proposition VI.15 *Soit* X *un vecteur aléatoire gaussien,* $X = (X_1, X_2) \in \mathbb{R}^p \times \mathbb{R}^{n-p}$*. Soit* $\mu = (\mu_1, \mu_2)'$ *son espérance et*

$$\Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}$$

sa matrice de variance-covariance. On suppose que Γ_{11} *est régulière, alors*

$$\mathbb{P}_{X_2|X_1=x_1} \sim \mathcal{N}\left(\mu_2 + \Gamma_{21}\Gamma_{11}^{-1}(x_1 - \mu_1), \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}\right)$$

Proposition VI.16 *Soit* $(X, Y_1, ..., Y_n)$ *un vecteur gaussien centré. Alors l'espérance conditionnelle* $\mathbb{E}(X|Y_1, ..., Y_n)$ *est p.s. égale à la projection orthogonale dans* $\mathcal{L}^2(\mathbb{P})$ *de* X *sur l'espace vectoriel engendré par* $\{Y_1, ..., Y_n\}$*. Donc l'espérance conditionnelle est la régression linéaire.*

Remarque VI.2 — *On peut vérifier, sous l'hypothèse de normalité que si* $Z = \mathbb{E}(X|Y)$ *alors* $\mathbb{E}((X-Z)Y) = 0$ *i.e.* $X - Z$ *est orthogonal à* Y*.*

$$\begin{aligned} \mathbb{E}((X - \mathbb{E}(X|Y)).Y) &= \mathbb{E}\left(\mathbb{E}((X - \mathbb{E}(X|Y)).Y|Y)\right) = \mathbb{E}\left(Y.\mathbb{E}(X - \mathbb{E}(X|Y)|Y)\right) \\ &= \mathbb{E}\left(Y.(\mathbb{E}(X|Y) - \mathbb{E}(X|Y))\right) \end{aligned}$$

— *en pratique, pour calculer l'espérance conditionnelle d'une v.a. gaussienne* X*, on écrit* $Z = \mathbb{E}(X|Y_1, ..., Y_n)$ *sous la forme d'une combinaison lin{eaire des* Y_i *et les coefficients sont déterminés par les conditions d'orthogonalité :*

$$\begin{cases} Z = \sum_{i=1}^n \alpha_i Y_i \\ \forall\ j, \quad \mathbb{E}\left((X - Z).Y_j\right) = 0 \end{cases}$$

les valeurs des espérances $\mathbb{E}(XY_i)$ *et* $\mathbb{E}(Y_iY_j)$ *sont données par les termes de la matrice de covariance. (puisque les v.a. sont centrées).*

— *Si le vecteur est non centré, on le centre* $\overline{X} = X - \mathbb{E}(X)$*,* $\overline{Y_i} = Y_i - \mathbb{E}(Y_i)$*. Si* $(X, Y_1, ..., Y_n =$ *est gaussin alors* $(\overline{X}, \overline{Y_1}, ..., \overline{y_n})$ *aussi et de matrice de variance covariance identique. De plus,*

$$\mathbb{E}(X|Y_1, ..., Y_n) = \mathbb{E}(X) + \mathbb{E}(\overline{X}|\overline{Y_1}, ..., \overline{Y_n})$$

f. Martingales

Définition VI.9 *Une famille de v.a. est un processus. Si elle est indexée par* $\mathbb{N}$*, il est à temps discret, si il est indexé par* $\mathbb{R}_+$*, il est à temps continu.*

© Théo Jalabert

Définition VI.10 *On appelle Filtration une suite croissante de sous tribus de* $\mathcal{A}$ $(\mathcal{F}_n)_n$. $(X_n)_n$ *est un processus adapté si* X_n *est* $\mathcal{F}_n$ *mesurable quelque soit* n *et prévisible si* X_n *est* $\mathcal{F}_{n-1}$ *mesurable.*

Définition VI.11 *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé muni d'une filtration* $(\mathcal{F}_n)_n$ *et* $(X_n)_n$ *un processus adapté. Si quelque soit* n, X_n *est* $\mathbb{P}$ *int{egrable et si*

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$$

on dit que X_n *est une martingale.*
Si $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n$ *(resp* $\leq$*), c'est une sous-martingale (resp. sur-martingale).*

Théorème VI.3 (Jensen) (X_n) *une (sous) martingale, soit* ϕ *convexe (croissante) tq* $\phi(X_n)$ *soit intégrable alors* $(\phi(X_n))_n$ *est une sous martingale.*

Théorème VI.4 *Si* (X_n) *et* (Y_n) *sont deux martingales alors* $(X_n \vee Y_n)$ *est une sous martingale et* $(X_n \wedge Y_n)$ *est une sur-martingales.*

Théorème VI.5 $(X_n)_n$ *une sous martingale. Alors*

$$\forall\ \alpha > 0 \quad \mathbb{P}\left(\max_{0 \leq i \leq n} X_i > \alpha\right) \leq \frac{\mathbb{E}(|X_n|)}{\alpha}$$