INTRODUCTION AUX PROBABILITES
IX-X-XI-XII

Support du cours donné en 1^{re} année

par Marietta Manolessou
EISTI - Département Mathématiques

Année 2011-2012

Table des matières

Table des figures

Chapitre 9

Indépendance des variables aléatoires

1 Variables aléatoires indépendantes

Définition 1.1
Soient $X,\ Y$ deux variables aléatoires définies sur $(\Omega,\ \mathcal{A},\ \mathcal{P})$ donc $(X,\ Y) : (\Omega,\ \mathcal{A},\ \mathcal{P}) \to (\mathbb{R}\times\mathbb{R},\ \mathcal{B})$
On dit que $X,\ Y$ sont indépendantes si quelque soit le couple de boréliens $B_i,\ B_j$, on a :

$$P[\{X \in B_i\} \cap \{Y \in B_j\}] = P[\{X \in B_i\}]P[\{Y \in B_j\}]$$

Autrement : *la loi de probabilité P_{XY} du couple $(X,\ Y)$ est égale au produit des lois P_X et P_Y :*

$$P_{XY} = P_X\ P_Y$$

Définition 1.2 (Généralisation)
Soient $\{X_1,\ X_2, \dots X_n\}$ une famille de variables aléatoires définies sur $(\Omega,\ \mathcal{A},\ \mathcal{P})$ donc

$$\{X_1,\ X_2, \dots X_n\} : \quad (\Omega,\ \mathcal{A},\ \mathcal{P}) \to (\mathbb{R}^n,\ \mathcal{B}(\mathbb{R}^n))$$

On dit que ces variables aléatoires sont indépendantes si quelque soient les boréliens $B_1, \dots\ B_n$, on a :

$$P[\bigcap_{i=1}^{n}\{X_i \in B_i\}] = \prod_{i=1}^{n}[P[\{X_i \in B_i\}]$$

Théorème 1.1
Si deux variables aléatoires X_1 et X_2 sont indépendantes :
$(P[\{X_1 \in A\} \cap \{X_2 \in B\}] = P[\{X_1 \in A\}]\ P[\{X_2 \in B\}])$

$$\Rightarrow\ E(X_1X_2) = E(X_1)\ E(X_2)$$

Remarque 1.1
Attention : *La réciproque n'est pas toujours vraie !*

2 Variables aléatoires dépendantes

Définition 2.1

$$\begin{aligned} P[\{X_i \in A_i\} \cap \{X_j \in A_j\}] &= P[X_i \in A_i\}|\{X_j \in A_j\}] \quad P[\{X_j \in A_j\}] \\ &\Leftrightarrow\ P[\{X_iX_j\} \in A] \qquad A = “A_i \cap A_j'' \end{aligned}$$

1 Covariance de deux variables aléatoires $X_1,\ X_2$

Définition 2.2

$$cov[X_1,\ X_2] = E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})] = E[X_1 X_2] - \mu_{X_1}\mu_{X_2}$$

Remarque 2.1

Quand X_1 et X_2 sont indépendantes, d'après le théorème précédent :
$E[X_1,\ X_2] = E[X_1]\ E[X_2]$

$$\Rightarrow \quad cov[X_1,\ X_2] = 0$$

Théorème 2.1

Soient X_k, *avec* $k \in \{1,\ \dots\ p\}$, *variables aléatoires et,* $a_k,\ b_k$, *avec* $k \in \{1\ \dots\ p\},$ *une suite de nombres réels alors :*

$$\Rightarrow var\left[\sum_{k=1}^{p} a_k X_k + b_k\right] = \sum_{k=1}^{p} a_k^2 var\left[X_k\right] + \sum_{i \neq j} a_i a_j cov\left[X_i X_j\right]$$

Théorème 2.2

Soient X_k, *avec* $k \in \{1,\ \dots\ p\}$,, *variables aléatoires indépendantes et,* $a_k,\ b_k$, *avec* $k \in \{1\ \dots\ p\},$ *une suite de nombres réels alors :*

$$\begin{aligned} E\left[\sum_{k=1}^{p} a_k X_k + b_k\right] &= \sum_{k=1}^{p} a_k E\left[X_k\right] + b_k \\ var\left[\sum_{k=1}^{p} a_k X_k + b_k\right] &= \sum_{k=1}^{p} a_k^2 var\left[X_k\right] \end{aligned}$$

3 Echantillon

Définition 3.1 (échantillon)

On définit un échantillon par les données suivantes.

1. Le n-uplet de variables aléatoires $(X_1,\ \dots,\ X_n)$ indépendantes qui suivent la même loi de probabilités avec paramètres $(\mu,\ \sigma^2)$.
2. X une variable aléatoire abstraite appelée *variable aléatoire parente,* (qui suit la même loi) avec paramètres $(\mu,\ \sigma^2)$
3. $\bar{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ *la moyenne empirique de l'échantillon,* qui vérifie les propriétés :

$$E[\bar{X}] = \mu \qquad var[\bar{X}] = \frac{\sigma^2}{n}.$$

Chapitre 10

Inégalités de Tchebycheff

1

Théorème 1.1 *Soient* $(\Omega,\ \mathcal{A},\ P)$ *et* X *variable aléatoire sur* $(\Omega,\ \mathcal{A})$
Si $h : \mathbb{R} \to [0,\ +\infty[$ *tel que* $h(X)$ *variable aléatoire et* $E[h(X)]$ *existe*

$$\Rightarrow \forall k > 0\ P[\{h(X)\} \geq k] \leq \frac{E[h]}{k}$$

Théorème 1.2
Soient $(\Omega,\ \mathcal{A},\ P)$ *et* X *variable aléatoire sur* $(\Omega,\ \mathcal{A})$ *avec* $(\mu_X,\ \sigma_X^2)$ *données*

$$\Rightarrow$$

1. $\forall\, \delta > 0\ P[\{|X - \mu_X| \geq \delta\sigma_X\}] \leq \dfrac{1}{\delta^2}$
2. $\forall\, \delta > 0\ P[\{|X - \mu_X| < \delta\sigma_X\}] \geq 1 - \dfrac{1}{\delta^2}$
3. (a) $\forall\, \epsilon > 0\ P[\{|X - \mu_X| \geq \epsilon\}] \leq \dfrac{\sigma_X^2}{\epsilon^2}$

   (b) $\forall\, \epsilon > 0\ P[\{|X - \mu_X| < \epsilon\}] \geq 1 - \dfrac{\sigma_X^2}{\epsilon^2}$

Chapitre 11

Convergence d'une suite de variables aléatoires

1

1 Convergence en probabilité

Définition 1.1 (Convergence en probabilité)
Soit $(\Omega,\ \mathcal{A},\ P)$. Une suite de variables aléatoires $\{X_1,\ \ldots,\ X_n,\ \ldots\}$ sur $(\Omega,\ \mathcal{A})$ converge en probabilité vers X (variable aléatoire : $\Omega \to \mathbb{R}$)
Si :

$$\forall\ \epsilon > 0 \qquad \lim_{n\to+\infty} P[\{|X_n - X| \geq \epsilon\}] = 0$$

notations équivalentes :

$$X_n \xrightarrow{P} X \qquad \Leftrightarrow \lim_{n\to+\infty} X_n = X(P)$$

Théorème 1.1 *Soit* $(\Omega,\ \mathcal{A},\ P)$. *Soient* X, $\{X_1,\ \ldots\}$ *variables aléatoires définies sur* $(\Omega,\ \mathcal{A})$, *dont le second moment par rapport à l'origine existe et il est tel que :* $\lim\limits_{n\to+\infty} E[(X_n - X)^2] = 0$

$$\Rightarrow$$

–

$$\lim_{n\to+\infty} X_n = X(P)$$

– *Cette limite en probabilité est unique.*

Pour la démonstration de ce théorème on utilise les inégalités de Kolmogorov et Tchebycheff.

2 Convergence en moyenne quadratique

Définition 1.2 (Convergence en moyenne quadratique)
Soient $(\Omega,\ \mathcal{A},\ P)$ et $X,\ (X_1,\ \ldots,\ X_n,\ \ldots)$ variables aléatoires définies sur $(\Omega,\ \mathcal{A})$ et dont le second moment μ_2' existe
On dit que la suite $X_1,\ \ldots,\ X_n$ *converge en moyenne quadratique* si

$$\lim_{n\to+\infty} E\left[(X_n - X)^2\right] = 0 \Leftrightarrow \lim_{n\to+\infty} X_n = X(mq) \Leftrightarrow X_n \xrightarrow{mq} X$$

Proposition 1.1
Si :

$$\lim_{n\to+\infty} X_n = X(mq) \Rightarrow \lim_{n\to+\infty} X_n = X(P)$$
$$\Leftrightarrow$$

Convergence en moyenne quadratique $\Rightarrow$ Convergence en probabilité.

3 Convergence presque sûre

Définition 1.3 (Convergence presque sûre)
Soient $(\Omega,\ \mathcal{A},\ P)$ et $X,\ (X_1,\ \ldots,\ X_n,\ \ldots)$ variables aléatoires définies sur $(\Omega,\ \mathcal{A})$. On dit que la suite $(X_1,\ \ldots,\ X_n,\ \ldots)$ converge *presque sûrement vers la variable aléatoire X* si

$$P[\{\lim_{n\to+\infty} X_n = X\}] = 1 \Leftrightarrow X_n \underset{p.s}{\to} X \Leftrightarrow \lim_{n\to+\infty} X_n = X(p.s).$$

Remarque : Si $X = C$ *(avec $C \in \mathbb{R}$ constante)* $\Leftrightarrow$ *X var. al. dégénérée en C* alors on a convergence presque sûre vers C

Théorème 1.2

1- $\lim_{n\to+\infty} X_n = X(ps)$ *si et seulement si*

$$\forall\ \epsilon > 0, \quad \exists\ N > 0 : \quad \lim_{n\geq N} P[\{\sup |X_n - X| < \epsilon\}] = 1$$

2- *Si*

$$\lim_{n\to+\infty} X_n = X(p.s) \quad \Rightarrow \quad \lim_{n\to+\infty} X_n = X(P)$$

2 Lois des grands nombres

1 Loi forte (resp.Loi faible) des grands nombres

Définition 2.1
Soit $(\Omega,\ \mathcal{A},\ P)$, et $(X_1,\ \ldots,\ X_n,\ \ldots)$ une suite de variables aléatoires sur$(\Omega,\ \mathcal{A})$ pas forcément indépendantes et tel que $\forall\ i\ \ E[X_i]$ existe.
Si la variable aléatoire transformée définie par :

$$Y_n = \frac{\displaystyle\sum_{i=1}^{n} [X_i - E(X_i)]}{n} = \bar{X} - \frac{\sum E[X_i]}{n}$$

converge vers 0 suivant un mode de convergence, alors la suite obéit à la loi des grands nombres :

* Si $\lim_{n\to+\infty} Y_n = 0(P) \Rightarrow$ la suite obéit à la loi faible des grands nombres.

* Si $\lim_{n\to+\infty} Y_n = 0(p.s) \Leftrightarrow$ la suite obéit à la loi forte des grands nombres.

Théorème 2.1 (Application à un échantillon)
Soient $(\Omega,\ \mathcal{A},\ P)$, *et* $X_1,\ X_2,\ \ldots\ X_n$ *une suite de variables aléatoires indépendantes sur* $(\Omega,\ \mathcal{A})$, *suivant la même loi de probabilité et telles que :*

$$\forall\ i \in \{1,\ 2, \ldots\ \ n\}\ \ E[X_i] = \mu,\ \ var[X_i] = \sigma^2$$

existent, alors la suite $\{X_1,\ \ldots,\ X_n\}$ *obéit à la loi faible et aussi à la loi forte des grands nombres.*

Pour la preuve de ce théorème on utilise la généralisation de l'inégalité de Tchebycheff $\Leftrightarrow$ inégalité de Kolmogorov.

2 Convergence en loi

Définition 2.2 (Convergence en loi)
Soient les variables aléatoires $X,\ \{X_1,\ \ldots\ X_n\}$ et les fonctions de répartition correspondantes, $F_X,\ \{F_{X_1},\ \ldots\ F_{X_n}\}$. On dit que la suite $\{X_1,\ X_2,\ \ldots,\ X_n\}$ *converge en loi vers* X si

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x)$$

en tout point de continuité x de F_X.

$$\Leftrightarrow$$

$$X_n \underset{\mathcal{L}}{\to} X \qquad \Leftrightarrow \qquad \lim_{n\to+\infty} X_n = X(\mathcal{L})$$

Exemple 2.1

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \quad e^{-\frac{1}{2}t^2} dt$$

$$X_n : \mathcal{N}\left(\frac{1}{n},\ 1 + \frac{1}{n}\right)$$

$$\lim_{n\to+\infty} F_{X_n}(x) = \lim_{n\to+\infty} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\left(1+\frac{1}{n}\right)} e^{-\frac{1}{2}\frac{\left(t-\frac{1}{n}\right)^2}{1+\frac{1}{n}}} dt = F_X(x)$$

On pose $z = \dfrac{t - \frac{1}{n}}{\sqrt{1+\frac{1}{n}}}$

Théorème 2.2 (de Lévy-Cramer-Duguc)

1.
$$X_n \underset{\mathcal{L}}{\to} X \Rightarrow \Phi_{X_n}(\omega) \to \Phi_X(\omega)$$

uniformément dans tout intervalle fini.

2.
$$\begin{array}{ll} \textit{Si } \Phi_{X_n}(\omega) & \to \Phi(\omega) \textit{ avec } \Re\Phi(\omega) \textit{ continue à } \omega = 0 \\ & \Rightarrow \Phi(\omega) \textit{ fonction caractéristique d'une variable aléatoire} X \\ & \textit{et} \quad X_n \underset{\mathcal{L}}{\to} X \end{array}$$

Remarque 2.1 *Tableau de la hiérarchie des convergences*

Conv. en moyenne quadr. $\searrow$

Conv. en probabilité $\rightarrow$ Conv. en loi

Convergence presque sûre $\nearrow$

3 Th. "Limite Centrale"

Théorème 2.3 (Théorème de la limite centrale)

Soient $(\Omega,\ \mathcal{A},\ P)$, *et* $X_1,\ X_2,\ \dots\ X_n$ *une suite de variables aléatoires indépendantes sur* $(\Omega,\ \mathcal{A})$, *suivant la même loi de probabilité et telles que :*

$$\forall\ i \in \{1,\ 2, \dots\ \ n\}\ \ E[X_i] = \mu,\ \ var[X_i] = \sigma^2$$

existent,

$$\Rightarrow$$

la suite $Y_1,\ Y_2,\ \dots, Y_n,\ \dots$ *où*

$$Y_n = \sqrt{n}\ \frac{\bar{X} - \mu}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)$$

converge en loi vers une variable aléatoire $Y : \mathcal{N}(0,\ 1)$.

4 Th. de De Moivre- Laplace

Théorème 2.4 (Théorème de De Moivre-Laplace)

$$\mathcal{B}(n,\ p) \rightarrow \mathcal{N}(0,\ 1)$$

Convergence en loi de la loi binomiale vers la loi $\mathcal{N}(0,\ 1)$.

Soient $(\Omega,\ \mathcal{A},\ P)$, *et* $X_1,\ X_2,\ \dots\ X_n$ *une suite de variables aléatoires indépendantes sur* $(\Omega,\ \mathcal{A})$, *suivant la même loi de probabilité binomiale* $\mathcal{B}(n,\ p)$.

$$\Rightarrow$$

La suite $Y_n = \dfrac{X_n - np}{\sqrt{npq}}$ *converge en loi vers la loi Normale centrée réduite :*

$$Y_n \underset{n \to \infty}{\overset{\mathcal{L}}{\rightarrow}} Y : \mathcal{N}(0,\ 1)$$

Preuve :

La fonction caractéristique de X_n :

$$\Phi_{X_n}(\omega) = \left[pe^{i\omega} + 1 - p\right]^n$$

$$\Rightarrow \Phi_{Y_n} = \left[pe^{\dfrac{i\omega}{\sqrt{npq}}} + 1 - p \right]^n \times e^{\dfrac{-i\omega np}{\sqrt{npq}}}$$

$$\Rightarrow \ln \Phi_{Y_n} = n \ln \left[p e^{\dfrac{i\omega}{\sqrt{npq}}} + 1 - p \right] - \frac{i\omega np}{\sqrt{npq}}$$

Développement limité de l'exponentiel à l'ordre 2 :

$$\Rightarrow \ln \Phi_{Y_n} \underset{n\to\infty}{\simeq} n \ln \left[1 + p \left(\frac{i\omega}{\sqrt{npq}} - \frac{\omega^2}{2npq} \right) \right] - \frac{i\omega np}{\sqrt{npq}}$$

Développement limité du $\ln$ à l'ordre 2 :

$$\ln(1+x) \underset{x\to 0}{\simeq} x - \frac{x^2}{2}$$

$$\Rightarrow \quad \Phi_{Y_n}(\omega) \underset{n\to\infty}{\simeq} -\frac{\omega^2}{2q} + \frac{1}{2}\frac{p\omega^2}{q}$$

avec $q = 1 - p$

$\Rightarrow \ln \Phi_{Y_n} \underset{n\to\infty}{\simeq} -\dfrac{\omega^2}{2}$ et

$$\Phi_{Y_n} \underset{n\to\infty}{\simeq} e^{-\dfrac{\omega^2}{2}}$$

On reconnaît la fonction caractéristique de la loi normale $\mathcal{N}(0,\ 1)$.

Chapitre 12

Vecteurs aléatoires

1 Vecteurs aléatoires de dimension 2

1 Définition

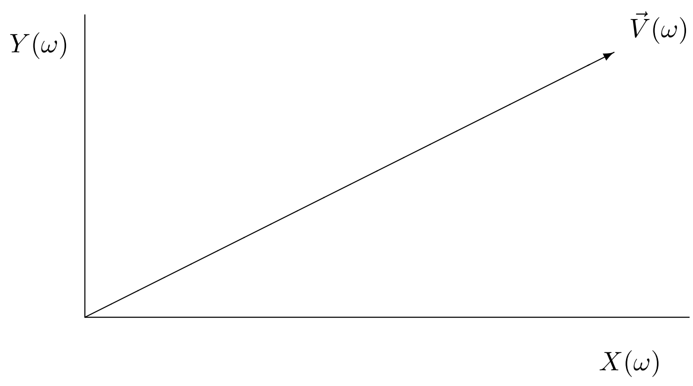
Soit un espace de probabilité $(\Omega,\ \mathcal{A},\ P)$ et soient deux variables aléatoires (X, Y) définies sur cet espace. Le vecteur aléatoire $V(X, Y)$ est défini par :

$$\begin{array}{lll} X\ :\ (\Omega, \mathcal{A},\ P) & \rightarrow \mathbb{R} & \\ Y\ :\ (\Omega, \mathcal{A},\ P) & \rightarrow \mathbb{R} & \\ V(XY) : (\Omega, \mathcal{A},\ P) & \rightarrow & \mathbb{R}^2 \\ \quad \forall \omega\ \in \Omega & \mapsto & V(X(\omega), Y(\omega)) \end{array}$$

Exemple 1.1

Sur la figure on représente le vecteur aléatoire $\vec{V}(\omega) \in \mathbb{R}^2$. Et pour l' événement : $A(D) \equiv \vec{V} \in D, \quad D \subset \mathbb{R}^2$, on associe l'ensemble des pts ω t.q.

$$A(x, y) = \{\omega | X(\omega) \leq x\ ;\ Y(\omega) \leq y\}$$



2 Mesure de Probabilité - Fonction de répartition conjointe

$$F(x, y) \equiv P[\{X \leq x\} \bigcap \{Y \leq y\}]$$

*** Cas continu :**

La mesure de probabilité définie par :

$$P(x, y) = \frac{\partial^2 F}{\partial x \partial y}$$

si elle existe on l'appelle **fonction de densité conjointe** (notée "$f(x, y)$)") d'un vecteur aléatoire continu et inversement :

$$\begin{array}{rl} F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} & P(x',y')\,dx'dy' \\ & \Updownarrow \\ & (f(x',y')) \end{array}$$

*** Cas discret - fonction de masse conjointe :**

(i)

$$P[x_i, y_j] \equiv P[\{X = x_i\} \bigcap \{Y = y_j\}]$$

et $\forall\ D \in \mathbb{R}^2$ fonction de répartition

$$P[D] = \sum_{(x_i, y_j) \in D} P[x_i, y_j]$$

3 Fonction de répartition marginale

*** Cas continu :**

$$F_X(x) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(x', y') dx' dy'$$

et fonction de densité marginale $\Leftrightarrow$ (dériv. par rapport à x)

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

*** Cas discret**

Fonction de masse marginale

$$Px_i = \sum_{y_j} P(x_i y_j)$$

Fonction de répartition marginale

$$\begin{array}{rl} F_X(x_i) = & F[x_i, +\infty] \\ & \quad\quad\downarrow \\ & y_j \in]-\infty, +\infty[\end{array}$$

4 Espérance mathématique (cas général : $h(X, Y)$)

***Cas continu :**

$$E[h[X, Y]] = \iint h(x, y)\ f(x, y) dx dy$$

*** Cas discret :**

$$E[h[X,Y]] = \sum_{x_i, y_j} h(x_i, y_j) P(x_i, y_j)$$

Propriétés

(a) Si $h(x,y) = x + y$

$$E[h(X,Y)] = E[X] + E[Y]$$

(b) Si $h(x,y) = xy \quad \Rightarrow$ en général

$$E[(X \cdot Y] \neq E[X]E[Y]$$

5 Indépendance

Théorème 1.1 (Indépendance) *Les propriétés suivantes sont équivalentes :*

i) Deux variables aléatoires $X,\ Y$ sont indépendantes

ii) la fonction de répartition conjointe se factorise en produit de fonctions de répartition marginales :

$$\Leftrightarrow \quad F(x,y) = F_X(x) F_Y(y).$$

iii) La fonction de densité (resp. fonction de masse) conjointe se factorise en produit de fonctions de densité (resp. de masse) marginales :

$$f(x,y) = f_X(x) f_Y(y)$$

$$P(x,y) = P_X(x) P_Y(y)$$

Corollaire

Pour deux variables aléatoires (X,Y) indépendantes on a le résultat suivant :

$$E[(X \cdot Y] = E[X]E[Y]$$

mais : **Remarque importante**

Cette factorisation peut se produire même si $X,\ Y$ ne sont pas indépendantes (liées).

6 Covariance (définition)

$$\text{Cov}(X,Y) \equiv E[(X - \mu_X)(Y - \mu_Y)]$$

ou $\Rightarrow$

$$\text{Cov}[X,Y] = E[XY] - \mu_X \mu_Y$$

(i) Corollaire

Cov$(X,Y) = 0$
si $X,\ Y$ sont non corrélées

(ii) "Inégalité de Schwartz"

$$\text{Cov}^2(X,Y) \leq \sigma_X^2 \sigma_Y^2$$

$\Leftrightarrow$

$$\text{Cov}(X,Y) \leq \sigma_X \sigma_Y$$

e) Coefficient de corrélation (de $X,\ Y$) déf.

$$C \equiv \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

en général

$$|C| \leq 1$$

* Si $C = 0$ $\Leftrightarrow$ var. al. $X,\ Y$ non corrélées.
* Si $|C| = 1$ (C peutêtre aussi < 0) $\Leftrightarrow$ $X,\ Y$ **complètement corrélées**.

2 Vecteurs aléatoires à $n > 2)$ dimensions

Soit $(\Omega,\ \mathcal{A},\ P)$ et $(\mathbb{R},\ \mathcal{R})$

$$\vec{X}(\omega) = \{X_1(\omega), X_2(\omega) \ldots X_?(\omega)\}$$

avec

$$X_i\ :\ \Omega \to \mathbb{R} \text{ (ou } \mathbb{C})$$
$$\omega \to X_i(\omega)$$
$$(\forall\, i = 1, 2, \ldots, n)$$

et t.q.

$$\{\omega | X_i(\omega) \in A\} = \mathcal{A}$$

(ou $\forall\ A \in \mathcal{C}$) $\forall\ A \in \mathcal{R}$

Rappel :

$\mathcal{R} \Leftrightarrow$ Tribu de Borel de $\mathbb{R}$
$\mathcal{A} \Leftrightarrow$ Tribu de Ω
ou ($\mathcal{C}$ Tribu de $\mathbb{C}$)

1 Fonction de répartition

$\forall\ x \in \mathbb{R}$ (ou $x \in \mathbb{C}^n$)

$$F(x) = F(x_1, \ldots, x_n) = P[\{X_1 \leq x_1\} \bigcap \{X_2 \leq x_2\} \bigcap \cdots \bigcap \{X_n \leq x_n\}]$$

Remarque En général on peut décomposer :

$$\begin{array}{llll} F(x) = & F_d(x)+ & F_c(x)+ & F_s(x) \\ & \uparrow & \uparrow & \downarrow \\ & \text{(cas} & \text{(cas} & \text{(singulière} \\ & \text{discret)} & \text{continu)} & \text{car on ignore..)} \end{array}$$

(i) (Cas continu) fonction de densité

$$f(x) = \frac{\partial^n F}{\partial x_1 \partial x_2 \ldots \partial x_n}$$

(ii) (cas discret) fonction de masse

$$P(x^{(i)}) = P\left[\{\vec{X} = \{x_1^{(i)}, x_2^{(i)} \ldots x_n^{(i)}\}\right]$$

Attention

$x^{(i)}$ est un vecteur précis dans $\mathbb{R}^n$ ou dans $\mathbb{C}^n$.

2 Espérance Mathématique (cas général pour une fonction $h(X)$)

cas continu

$$E[\{h(X)\}] = \underset{\substack{\mathbb{R}^n \\ (\text{ou } \mathbb{C}^n)}}{\iint} \dots \int h(x) f(x) dx_1 \dots dx_n$$

cas discret

$$E[\{h(X)\}] = \sum_{\substack{x^{(i)} \in \mathbb{R}^n \\ \text{ou } x^{(i)} \in \mathbb{C}^n}} P(x^{(i)}) h(x^{(i)})$$

3 Matrice de covariance (Γ)

(i) Si $x \in \mathbb{R}^n$ $(\mu \in \mathbb{R}^n)$

$$\Gamma = E[XX^T] - \mu\mu^T$$

ou

$$\Gamma_{ij} = E[X_i X_j] - \mu_i \mu_j$$

éléments diagonaux

$$\Gamma_{ii} = E[X_i^2] - \mu_i^2 \equiv \sigma_{X_i}^2$$

$\Leftrightarrow$ variances des X_i
ou

$$\Gamma \equiv E[(X - \mu)(X - \mu)^T]$$

Remarque $E[X_i X_j]$ = Matrice de corrélation

(ii) Si $x \in \mathbb{C}^n$

$$\Gamma \overset{\text{def}}{=} E[XX^*] - \mu\mu^*$$

ou

$$\Gamma = E[(X - \mu)(X - \mu)^*]$$

et si X centré $(\mu = 0)$

$$\Gamma = E[XX^*]$$

4 Probabilités Conditionnelles par rapport à un événement

Cas scalaire

a) Soit $A \in \mathcal{R}$ (évènement) (ex. $A = [a, b]$)
Fonction de répartition conditionnelle

$$F[x|A] = \frac{P[\{X \leq x\} \bigcap A]}{P[A]}$$

Fonction de densité conditionnelle

$$f[x|A] = \frac{dF[x|A]}{dx}$$

Espérance Conditionnelle

$$E[h(X)|A] = \int f[x|A]h(x)dx$$

b) (def. analogue pour un vecteur de probabilité)
Fonction de répartition conditionnelle

$$F[X_1, X_2 \dots X_n|A] = \frac{P[\{X_1 \le x\} \bigcap \{X_2 \le x_2\} \bigcap \dots \bigcap \{X_n \le x_n\} \bigcap A]}{P[A]}$$

Espérance conditionnelle

$$E[h(X_1, \dots X_n)|A] = \iint f[x_1 \dots x_n|A]h(x)dx_1 \dots dx_n$$

5 Indépendance

Théorème 2.1 (indépendance (Généralisation à $n > 2$ dimensions)
Les propriétés suivantes sont équivalentes :

i) Les variables aléatoires $\{X_1, \dots, \quad X_n$ sont indépendantes,

ii) la fonction de répartition conjointe se factorise en produit de fonctions de répartition marginales.

iii) La fonction de densité (resp. fonction de masse) conjointe se factorise en produit de fonctions de densité (resp. de masse) marginales.

6 Distributions Conditionnelles entre variables et vecteurs aléatoires

Fonction de répartition conditionnelle

$$F[x|Y \le y] \stackrel{\text{def}}{=} P[\{X \le x\}|Y \le y\}] = \frac{F(x,y)}{F(+\infty, y)}$$

avec, $F(+\infty, y)$, fonction de répartition marginale par rapport à Y.

Fonction de masse conditionnelle

(cas discret)

Définition 2.1

$$P_{x_i|y_j} \stackrel{\textit{notation}}{\equiv} P[\{X = x_i\}|\{Y = y_j\}] \stackrel{\textit{def}}{\equiv} \frac{P[\{X = x_i\} \bigcap \{Y = y_j\}]}{P[Y = y_j]}$$

où

$$P[Y = y_j] = \sum_{x_i} P_{x_i y_j}$$

Fonction de densité conditionnelle (cas continu)

Cas vectoriel (sur $\mathbb{R}$)

Définition 2.2

$$f_X[x|y] \stackrel{def}{\equiv} \frac{f(x,y)}{f_Y(y)}$$

et aussi :

Proposition 2.1

$$i) \qquad \int f_X[x|y]dx_1 \dots dx_n = 1$$

$$ii) \qquad \int f_X[x|y]f_Y(y)dy_1 \dots dy_n = f_X(x)$$

Exercice 2.1 *Montrer la proposition 2.1*

7 Moments conditionnels (cas général-vectoriel)

Espérance Conditionnelle

Soit une fonction - var. aléatoire $h(X)$, avec,

X vect. de probabilité ; $X : \ (\Omega, \mathcal{A}, P) \to \mathbb{R}^n$
Y vect. de probabilité ; $Y : \ (\Omega, \mathcal{A}, P) \to \mathbb{R}^n$

Définition 2.3 *Espérance Conditionnelle*

$$E_X[h(X)|y] = \int h(x) f_X[x|y]dx$$

Propriété importante !

Proposition 2.2

$$E_Y[E_X[h(X)|y]] = E_X[h[X]]$$

Exercice 2.2 *Montrer la proposition 2.2*

Moments Conditionnels d'ordre p

Définition 2.4

$$E_X[X^p|y] = \int x^p \ f[x|y]dx$$

(Définition analogue pour le cas discret)

Variance conditionnelle

Définition 2.5

$$\begin{aligned} E_X[(X-\mu_X)^2|y] &= \int (X-\mu_x)^2 \ f[x|y]dx \\ &= E[X^2|y] + \mu_X^2 - 2\mu_X E_X[X|y] \end{aligned}$$

8 Transformations de vecteurs aléatoires (à densité)

Théorème 2.2 *Soit* $(\Omega, \mathcal{F}, P)$ *un espace de probabilité et soit*

$$X = \{X_1,\ X_2, ..., X_n\} : \ (\Omega, \mathcal{F}, P) \ \rightarrow \ \mathbb{R}^n$$

un vecteur aléatoire.

Soit $Y = \phi(X)$ *un vecteur aléatoire - transformation biunivoque du vecteur aléatoire* X.

$$\Rightarrow$$

la densité de probabilité du vecteur Y *est obtenue par la formule suivante :*

$$g(y) = \frac{f((\phi)^{-1}(y))}{|\det J|}$$

où J *est la matrice -**Jacobien** de la transformation.*