Probabilités GM3

Ioana CIOTIR

2021 - 2022

Table des matières

Chapitre 1

Espaces probabilisés

1.1 Paradoxe de Bertrand

Soit un cercle de rayon 20 cm et un autre cercle concentrique de rayon 10 cm. On dessine une corde dans le grand cercle (**au hasard**). On veut calculer la probabilité qu'elle intersecte le petit cercle.

Solution 1 *Chaque corde est uniquement déterminée par la position de son centre (à une rotation prêt).*

On calcule

$$\mathbb{P} = \frac{\textit{Aire du petit cercle}}{\textit{Aire du grand cercle}} = \frac{100\pi}{400\pi} = \frac{1}{4}.$$

Solution 2 *Si on suppose les cordes verticales (à une rotation prêt) alors les cas favorables sont quand le milieu de la corde se trouve sur le diametre horizontal du petit cercle*

$$\mathbb{P} = \frac{\textit{Diametre du petit cercle}}{\textit{Diametre du grand cercle}} = \frac{20}{40} = \frac{1}{2}.$$

Solution 3 *On suppose que toutes les cordes ont (à une rotation prêt) une extremité dans le point le plus à gauche du grand cercle. On note l'angle formé par la corde et l'horizontale par θ.*

Le nombre de cas possible est entre $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ donc π.

Le nombre de cas favorable est entre $\left(-\frac{\pi}{6}, \frac{\pi}{6}\right)$ donc $\frac{\pi}{3}$.

$$\mathbb{P} = \frac{\frac{\pi}{3}}{\pi} = \frac{1}{3}.$$

Il faut donc correctement définir et comprendre le mot "hasard" et il faut bien choisir l'espece de probabilité.

1.2 Rappel de théorie de la mesure

Soit Ω un ensemble et $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω.

Définition 4 *On appelle $\mathcal{A} \subset \mathcal{P}(\Omega)$ une **tribu** (ou **une** $\sigma-$**algèbre**) si elle vérifie :*
- $\Omega \in \mathcal{A}$
- *Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$*
- *Pour toute famille dénombrable $\{A_n\}_{n\geq 1} \subset \mathcal{A}$ on a $\bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$.*

On apelle **ensemble mesurable** tout élément de la $\sigma-$algèbre $\mathcal{A}$.

Un espace Ω muni d'une $\sigma-$algèbre s'appelle **espace mesurable**. (Not. $(\Omega, \mathcal{A})$).

Exemple 5 *Les ensembles suivants sont des tribue (des $\sigma-$algèbres)*
- $\mathcal{A}_1 = \{\emptyset, \Omega\}$
- $\mathcal{A}_2 = \mathcal{P}(\Omega)$.

Proposition 6 *Soit Ω muni d'une $\sigma-$algèbre $\mathcal{A}$. On a les propriétés suivantes*
- *Si $A, B \in \mathcal{A}$ alors $A \cap B \in \mathcal{A}$*
- $\emptyset \in \mathcal{A}$.

Définition 7 *Soit $G \subset \mathcal{P}(\Omega)$. On dit que $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une **tribu ($\sigma-$algèbre) engendrée par** G si elle est la plus petite tribu ($\sigma-$algèbre) qui contient tous les éléménts de G.*

Définition 8 *Soit Ω un espace topologique (muni d'une famille d'ouverts). On appelle **tribu borélienne** la tribu engendrée par tous les ouverts de Ω. (Not. $\mathcal{B}(\Omega)$).*

Définition 9 *Une application $f : (\Omega, \mathcal{A}) \to \left(\widetilde{\Omega}, \widetilde{\mathcal{A}}\right)$ est dite **mesurable** si*

$$\forall \widetilde{A} \in \widetilde{\mathcal{A}} \text{ on a } f^{-1}\left(\widetilde{A}\right) \in \mathcal{A}.$$

Définition 10 *Une* ***mesure*** μ *sur* $(\Omega, \mathcal{A})$ *est une application*

$$\mu : \mathcal{A} \to [0, +\infty]$$

telle que

- $\mu(\emptyset) = 0$
- *si* $\{A_n\}_{n\geq 1}$ *est une famille dénombrable d'ensembles de* A *, deux à deux disjoints, alors*

$$\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Exemple 11 *La mesure de Lebesgue sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *qui généralise la notion de longueur des intervalles*

$$\lambda([a, b]) = b - a \quad \textit{et} \quad \lambda(x + A) = \lambda(A).$$

Proposition 12 *Soit une mesure* μ *sur* $(\Omega, \mathcal{A})$. *Alors on a les propriétés suivantes*

i) $\mu(A^c) = \mu(\Omega) - \mu(A)$ *(si* $\mu(\Omega) < \infty$*)*
ii) $\mu(A \backslash B) = \mu(A) - \mu(A \cap B)$
iii) $A \subset B \Longrightarrow \mu(A) \leq \mu(B)$
iv) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Le triplet $(\Omega, \mathcal{A}, \mu)$ est appelé un espace mesuré (espace mesurable muni d'une mesure).

1.3 Théorie des probabilités

Définition 13 *On appelle* ***mesure de probabilité*** *ou probabilité une mesure note* $\mathbb{P}$ *pour qui la masse totale est 1 (*$\mathbb{P}(\Omega) = 1$*).*

Un espace mesuré dont la mesure est une proabilité est appelé *espace de probabilité*. Not. $(\Omega, \mathcal{A}, \mathbb{P})$.

Un ensemble $A \in \mathcal{A}$ est appelé évènement.

Exemple 14 *(Espace de probabilité discrète)*

Soit $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$ *et une famille de valeurs positives* $\{p_1, p_2, ..., p_N\}$ *telles que* $\sum\limits_{i=1}^{N} p_i = 1$.

On peut definir une probabilité $\mathbb{P}$ *donnée par*

$$\mathbb{P}(\omega_i) = p_i, \quad \forall i \in \{1, 2, ..., N\}.$$

Pour un ensemble $A = \{\omega_1, \omega_3, \omega_7\}$ *on a* $\mathbb{P}(A) = p_1 + p_3 + p_7$.

Exemple 15 *(Espace de probabilité avec densité)*

Soit $\Omega = \mathbb{R}$ *et* $\mathcal{B}(\mathbb{R})$ *la tribu borélienne.*

On appelle dénsité une fonction $f : \mathbb{R} \to \mathbb{R}$ *telle que* $f \geq 0$, f *est intégrable et* $\int_{\mathbb{R}} f(x)\, dx = 1$.

On peut definir une probabilité $\mathbb{P}$ *donnée par*

$$\mathbb{P}(A) = \int\limits_A f(x)\, dx, \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

Remarque 16 *Les deux cas précédents sont des cas particuliers. On ne peut pas trouver une dénsité pour toute mesure de probabilité.*

Chapitre 2

Variables aléatoires

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

Définition 17 *On appelle variable aléatoire (notée v.a.) toute application mesurable*

$$X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Plus précisément X est une variable aléatoire si et seulement si elle vérifie

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

Définition 18 *Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. On définit*

$$\mathcal{U}(X) = \left\{ X^{-1}(B) : \ \forall B \in \mathcal{B}(\mathbb{R}) \right\}$$

une $\sigma-$algèbre qui est engendré par X.

Remarque 19 *On voit que :*

1) C'est la plus petite $\sigma-$algèbre par rapport à qui la variable aléatoire est mesurable.

2) $\mathcal{U}(X)$ contient toute les informations sur X.

3) Si on considère $Y = \Phi(X)$ pour Φ une fonction C^2, alors Y est aussi une fonction $\mathcal{U}(X)-$mesurable. Réciproquement, si Y est $\mathcal{U}(X)-$mesurable, alors il existe une fonction Φ telle qui $Y = \Phi(X)$.

Définition 20 *On appelle $X : [0, T] \times \Omega \to \mathbb{R}$ un processus stochastique si*

$$\begin{array}{ccll} t & \longmapsto & X(t, \omega) & \text{est continue} \\ \omega & \longmapsto & X(t, \omega) & \text{est une variable aléatoire.} \end{array}$$

On note par $\{X_t, \quad t \in [0, T]\}$.

2.1 Loi d'une variable aléatoire

Définition 21 *Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. On appelle loi de la variable aléatoire X une mesure $\mathbb{P}_X$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie par*

$$\begin{aligned} \mathbb{P}_X &: \mathcal{B}(\mathbb{R}) \to [0,1] \\ \mathbb{P}_X(A) &= \mathbb{P}(X \in A) \\ &= \mathbb{P}\left(X^{-1}(A)\right) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}), \quad \forall A \in \mathcal{B}(\mathbb{R}). \end{aligned}$$

Remarque 22 $\mathbb{P}_X$ *est une mesure de probabilité définie sur* $\mathcal{B}(\mathbb{R})$.

Loi d'une variable aléatoire discrète

On rappelle que une variable aléatoire discrète une variable qui a des valeurs dans un ensemble au plus dénombrable.

Soit X une variable aléatoire discrète.

Pour chaque $a \in \mathbb{R}$, on considère la mesure de Dirac en a qui est définie par

$$\delta_a(A) = \begin{cases} 1, & \text{si } a \in A \\ 0, & \text{sinon.} \end{cases}$$

On a la loi de X définie par

$$\mathbb{P}_X(A) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x)\, \delta_x(A).$$

Exemple 23 *Soit $X : \Omega \to \{x_1, x_2, x_3\}$ et $A = \{x_1, x_3\}$. On a*

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}(X = x_1)\,\delta_{x_1}(A) + \mathbb{P}(X = x_2)\,\delta_{x_2}(A) + \mathbb{P}(X = x_3)\,\delta_{x_3}(A) \\ &= \mathbb{P}(X = x_1) + \mathbb{P}(X = x_3). \end{aligned}$$

On voit que les ensembles $\{X = x_1\}$ et $\{X = x_3\}$ sont disjointe.

Si $X : \Omega \to \{x_1, x_2, x_3, ..., x_n\}$ et $\mathbb{P}(X = x_i) = p_i, \quad \forall i = \overline{1, n}$. Alors

$$\mathbb{P}_X(A) = \sum_{i=1}^{n} p_i \delta_{x_i}(A).$$

Intégrale de Lebesgue

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $f : \Omega \to \mathbb{R}$ une fonction mesurable, donc une fonction qui vérifie $f^{-1}(B) \in \mathcal{F}$, $\forall B \in \mathcal{B}(\mathbb{R})$.

On définit

$$\int_\Omega f(\omega)\, d\mathbb{P}(\omega)$$

- *Etape 1*
  Soit
  $$f(x) = \sum_{i=1}^{k} c_i I_{A_i}(x)$$
  une fonction simple (escalier) où $\{c_i\}_{i=1}^{k} \subset \mathbb{R}$ et $\{A_i\}_{i=1}^{k} \subset \mathcal{F}$.
  Pour une fonction simple et positive, on peut définir
  $$\int_\Omega f(\omega)\, d\mathbb{P}(\omega) = \sum_{i=1}^{k} c_i \mathbb{P}(A_i).$$
- *Etape 2*
  Soit f une fonction positive mesurable. On a toujours une suite de fonctions simples positives telles que
  $$f_n(\omega) \to f(\omega), \quad \forall \omega \in \Omega.$$
  On peut définire l'intégrale par
  $$\int_\Omega f(\omega)\, d\mathbb{P}(\omega) = \lim_{n\to\infty} \int_\Omega f_n(x)\, d\mathbb{P}(\omega).$$
  L'integrale est bien définie car
  1) On peut toujours construir une telle suite $\{f_n\}$
  2) Si on a deux suites $\{f_n\}$ et $\{g_n\}$ qui vérifient
  $$f_n(\omega) \to f(\omega), \quad \forall \omega \in \Omega.$$
  $$g_n(\omega) \to f(\omega), \quad \forall \omega \in \Omega.$$
  alors
  $$\lim_{n\to\infty} \int_\Omega f_n(x)\, d\mathbb{P}(\omega) = \lim_{n\to\infty} \int_\Omega g_n(x)\, d\mathbb{P}(\omega).$$
- *Etape 3*
  Soit f une fonction mesurable. On peut l'écrire comme
  $$f = f^+ - f^-$$
  où
  $$f^+(x) = \begin{cases} f(x), & \text{si } f(x) \geq 0 \\ 0, & \text{sinon} \end{cases} \qquad f^-(x) = \begin{cases} -f(x), & \text{si } f(x) \leq 0 \\ 0, & \text{sinon} \end{cases}.$$
  Comme f^+ et f^- sont positives et mesurables, on peut définir
  $$\int_\Omega f(\omega)\, d\mathbb{P}(\omega) = \int_\Omega f^+(\omega)\, d\mathbb{P}(\omega) - \int_\Omega f^-(\omega)\, d\mathbb{P}(\omega).$$

On dit qu'une fonction f est **Lebesgue intégrable** si et seulement si

$$\int_{\Omega} f(\omega)\, d\mathbb{P}(\omega) < \infty.$$

Proposition 24 *L'intégrale de Lebesgue a les propriétés suivantes*

1) $\int_A f(\omega)\, d\mathbb{P}(\omega) = \int_{\Omega} f(\omega)\, 1_A(\omega)\, d\mathbb{P}(\omega)$.

2) *Si* $f = g$ *presque partout, alors* $\int_{\Omega} f(\omega)\, d\mathbb{P}(\omega) = \int_{\Omega} g(\omega)\, d\mathbb{P}(\omega)$.

3) *Si* $f \geq 0$ *alors* $\int_{\Omega} f(\omega)\, d\mathbb{P}(\omega) \geq 0$.

4) $\int_{\Omega} (f(\omega) + g(\omega))\, d\mathbb{P}(\omega) = \int_{\Omega} f(\omega)\, d\mathbb{P}(\omega) + \int_{\Omega} g(\omega)\, d\mathbb{P}(\omega)$

5) $\int_A f(\omega)\, d\mathbb{P}(\omega) + \int_B f(\omega)\, d\mathbb{P}(\omega) = \int_{A\cup B} f(\omega)\, d\mathbb{P}(\omega)$, *pour* $\forall A, B \in \mathcal{F}$ *tels que* $A \cap B = \emptyset$

Loi d'une variable aléatoire continue

Définition 25 *Soit* $(\Omega, \mathcal{F})$ *un espace mesurable muni de deux mesures* μ *et* σ. *On dit que* μ *est absolument continue par rapport à* σ *si, pour tout* $A \in \mathcal{F}$ *on a*

$$\sigma(A) = 0 \Longrightarrow \mu(A) = 0.$$

On note $\mu \ll \sigma$.

Théorème 26 *Si on a deux mesures* μ *et* σ *telles que* $\mu \ll \sigma$, *alors il existe une fonction* $f : \Omega \to \mathbb{R}$ *mesurable, telle que pour tout* $A \in \mathcal{F}$ *on a que*

$$\mu(A) = \int_A f(x)\, d\sigma(x).$$

On note

$$d\mu = f d\sigma.$$

Remarque 27 *La fonction* f *s'appelle la densité de* μ *par rapport à* σ *et on a le lien suivant entre les intégrales*

$$\int_A g(x)\, d\mu(x) = \int_A g(x)\, f(x)\, d\sigma(x).$$

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction densité (positive, intégrable et avec $\int_{\mathbb{R}} f(x)\, dx = 1$).

Alors on peut définir une mesure de probabilité sur $\mathcal{B}(\mathbb{R})$ définie par

$$\mathbb{P}(A) = \int_A f(x)\, dx, \quad \forall A \subset \mathcal{B}(\mathbb{R}).$$

Définition 28 *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité et* X *une variable aléatoire telle que*

$$X : \Omega \to \mathbb{R}.$$

On dit que X *est une varianle aléatoire "à dénsité" s'il existe une fonction* $f : \mathbb{R} \to \mathbb{R}_+$ *qui vérifie les propriétés d'une densité (intégrable, et telle que* $\int_{\mathbb{R}} f(x)\, dx = 1$*) telle que la loi de* X *peut être écrite comme suit*

$$\mathbb{P}_X(A) = \int_A f(x)\, dx, \quad \forall A \subset \mathcal{B}(\mathbb{R}).$$

Remarque 29 *Il y a des variables aléatoires qui ne sont pas "à densité". Une variable aléatoire est "à densité" seulement si sa loi est absolument continue par rapport à la mesure de Lebesgue.*

2.2 Fonction de répartition d'une variable aléatoire

Définition 30 *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité et* $X : \Omega \to \mathbb{R}$ *une variable aléatoire. On appelle fonction de répartition de* X *la fonction*

$$F_X : \mathbb{R} \to [0, 1]$$

définie par

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x]).$$

Proposition 31 ***(Propriétés de la fonction de répartition)***

1) F_X *est une fonction croissante*

2) $\lim_{x \to -\infty} F_X(x) = 0$ *et* $\lim_{x \to \infty} F_X(x) = 1$

3) F_X *est continue à droite et admet une limite à gauche*

$$\begin{aligned} F_X(a^+) &= \lim_{x \to a,\ x \geq a} F_X(x) = F_X(a) \\ F_X(a^-) &= \lim_{x \to a,\ x \leq a} F_X(x). \end{aligned}$$

4) $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$

Démonstration.

4)

$$
\begin{aligned}
(X \leq b) &= ((X \leq b) \cap (X > a)) \cup ((X \leq b) \cap (X \leq a)) \\
&= ((X \leq b) \cap (X > a)) \cup (X \leq a)
\end{aligned}
$$

$$
\begin{aligned}
F_X(b) &= \mathbb{P}((X \leq b) \cap (X > a)) + \mathbb{P}(X \leq a) \\
&= \mathbb{P}((X \leq b) \cap (X > a)) + F_X(a)
\end{aligned}
$$

Donc

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a).$$

■

Proposition 32 *Si X est une variable aléatoire de densité f, sa fonction de répartition F_X vérifie*

1) $F_X(x) = \int_{-\infty}^{x} f(t)\,dt, \quad \forall x \in \mathbb{R}$

2) F_X *est continue sur* $\mathbb{R}$

3) Si f est continue, alors F_X est dérivable et $F_X' = f$.

Démonstration.

1) On a par la définition de la fonction de répartition que

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(X^{-1}(]-\infty, x])\right) = \mathbb{P}_X(]-\infty, x]).$$

Comme X est une variable aléatoire de densité f, on a

$$\mathbb{P}_X(]-\infty, x]) = \int_{]-\infty, x]} f(x)\,dx = \int_{-\infty}^{x} f(x)\,dx.$$

2) On fixe un point a quelconque. Comme F_X est continue à droite, car fonction de répartition, il reste à montrer la continuité à gauche.
Soit $(a_n)_n$ une suite croissante qui converge vers a. Il faut vérifier que

$$\lim_{n \to \infty} F_X(a_n) = F_X(a).$$

En effet

$$F_X(a) - F_X(a_n) = \int_{a_n}^{a} f(x)\,dx = \int_{\mathbb{R}} f(x)\,1_{[a_n,a]}(x)\,dx.$$

Comme $\left|f(x)\,1_{[a_n,a]}(x)\right| \leq |f(x)|$ qui est intégrable et $f(x)\,1_{[a_n,a]}(x) \to 0$ pour $\forall x \in \mathbb{R}$, on peut appliquer le théorème de convergence dominée de Lebesgue et on obtient que

$$\int_{\mathbb{R}} f(x)\,1_{[a_n,a]}(x)\,dx \to 0$$

et donc
$$F_X(a) - F_X(a_n) \to 0$$
ce qui implique la continuité.

■

Remarque 33 *Si X est une variable aléatoire de densité f, on a, pour tout interval $[a,b]$, que*

$$\begin{aligned} F_X(b) &= \mathbb{P}(X \in]-\infty, b]) \\ &= \mathbb{P}(X \in]-\infty, a] \cup]a, b]) \\ &= F_X(a) + \mathbb{P}(X \in]a, b]). \end{aligned}$$

On obtient donc

$$\mathbb{P}(X \in]a,b]) = F_X(b) - F_X(a) = \int_a^b f(t)\,dt.$$

2.3 Espérance probabiliste

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

Rappel
Si on a $X : \Omega \to \{x_1, x_2, ..., x_n\}$ une variable aléatoire discrète alors on a

$$\begin{aligned} \mathbb{E}(X) &= \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) \\ &= \sum_{i=1}^{n} x_i \mathbb{P}_X(\{x_i\}). \end{aligned}$$

On observe que c'est la même formule que l'integrale de Lebesgue d'une fonction simple.

Définition 34 *Si X est une variable aleatoire, on appelle espérance de X l'intégrale de X par rapport à la mesure $\mathbb{P}$. Plus précisement*

$$\mathbb{E}(X) = \int_\Omega X(\omega)\,d\mathbb{P}(\omega).$$

Exemple 35 *Soit $A \in \mathcal{F}$ et sa fonction indicatrice 1_A (qui est une variable aléatoire). On a alors*

$$\mathbb{E}(X) = \mathbb{E}(1_A) = \int_\Omega 1_A(\omega)\,d\mathbb{P}(\omega) = \mathbb{P}(A).$$

Remarque 36 *Pour une variable aléatoire à densité on a*

$$\begin{aligned} \mathbb{P}_X(A) &= \int_A f(x)\,dx \\ d\mathbb{P}_X(A) &= f(x)\,dx. \end{aligned}$$

Par un changement de variable on a

$$\begin{aligned} \mathbb{E}(X) &= \int_\Omega \underbrace{X(\omega)}_{=x} d\mathbb{P}(\omega) \\ &= \int_\mathbb{R} x d\mathbb{P}\left(X^{-1}(x)\right) \\ &= \int_\mathbb{R} x d\mathbb{P}_X(x). \end{aligned}$$

On obtient que

$$\mathbb{E}(X) = \int_\mathbb{R} xf(x)\,dx.$$

Définition 37 *Une variable aléatoire X est dite **intégrable** si $|X|$ est d'espérance finie.*

*Une variable aléatoire X est dite **centrée** si $\mathbb{E}(X) = 0$.*

Proposition 38 ***(Propriétés de l'espérance)***

1) $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$

2) Si $X \geq 0$ *et* $\mathbb{E}(X) = 0$ *alors* $\mathbb{P}(X=0) = 1$

Changement de variable

Soit X une variable aléatoire et $h : \mathbb{R} \to \mathbb{R}$ une fonction mesurable et $Y = h(X)$.

On voit que

$$\begin{aligned} \mathbb{E}(Y) &= \int_\Omega Y(\omega)\,d\mathbb{P}(\omega) \\ &= \int_\Omega h(X(\omega))\,d\mathbb{P}(\omega) \\ &= \int_\mathbb{R} h(x)\,d\mathbb{P}\left(X^{-1}(x)\right) \\ &= \int_\mathbb{R} h(x)\,d\mathbb{P}_X(x) \\ &= \int_\mathbb{R} h(x) f(x)\,dx. \end{aligned}$$

Proposition 39 *(Inégalité de Markov) Si X est une variable aléatoire positive*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Définition 40 *Soit $(X_n)_n$ une suite de variables aléatoires. On dit que $X_n \to X$ presque sûrement (noté p.s.) si $\mathbb{P}(X_n \to X) = 1$, ce qui signifie $X_n(\omega) \to X(\omega)$ pour $\forall \omega \in A$ tel que $\mathbb{P}(A) = 1$.*

Définition 41 *Soit $(X_n)_n$ une suite de variables aléatoires. On dit que $X_n \to X$ en espérance si $\mathbb{E}(X_n) \to \mathbb{E}(X)$ pour $n \to \infty$.*

Théorème 42 ***(de la convergence monotone)*** *Soit $(X_n)_n$ une suite croissante de variables aléatoires telle que $X_n \geq 0$ et $X_n \to X$ presque sûrement. Alors*

$$\mathbb{E}(X_n) \to \mathbb{E}(X), \quad \text{pour } n \to \infty.$$

Théorème 43 ***(de la convergence dominée)*** *Soit $(X_n)_n$ une suite de variables aléatoires telle que $X_n \to X$ presque sûrement. S'il existe une variable aléatoire Y telle que $\mathbb{E}(Y) < \infty$ (intégrable) et $|X_n| < Y$ p.s., alors $\mathbb{E}(X_n) \to \mathbb{E}(X)$, pour $n \to \infty$.*

2.4 Moments des variables aléatoires

Définition 44 *Une variable aléatoire $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ a un moment d'ordre $p \geq 1$ si et seulement si*

$$\mathbb{E}(|X|^p) = \int_\Omega |X|^p \, d\mathbb{P} < \infty.$$

On peut construire l'espace

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) = \left\{ \widehat{X} \text{ les classes d'équivalence de } X : \Omega \to \mathbb{R};\ \mathbb{E}(|X|^p) < \infty \right\}$$

muni de la norme

$$\|X\|_p = (\mathbb{E}(|X|^p))^{1/p}.$$

Remarque 45 *On peut montrer que c'est un espace de Banach.*

Proposition 46 *On a les propriétés suivantes*

- (Inégalité de Hölder) Soient X et Y deux variables aléatoires qui vérifient $\mathbb{E}(|X|^p) < \infty$ et $\mathbb{E}(|Y|^q) < \infty$ pour $p, q \in \mathbb{R}$ telles que $p, q > 1$ et $\frac{1}{p} + \frac{1}{q} = 1$. Montrer que
$$\mathbb{E}(XY) \leq \mathbb{E}(|X|^p)^{1/p} \, \mathbb{E}(|Y|^q)^{1/q}.$$
On retrouve l'inégalité de Cauchy-Schwartz pour $p = q = 2$.

- (Inégalité de Minkowski) Soient X et Y deux variables aleatoires $\mathbb{E}(|X|^p) < \infty$ et $\mathbb{E}(|Y|^p) < \infty$, $1 \leq p < \infty$. Montrer que
$$\mathbb{E}(|X+Y|^p)^{1/p} \leq \mathbb{E}(|X|^p)^{1/p} + \mathbb{E}(|Y|^p)^{1/p}.$$

Proposition 47 *Si une variable aléatoire est bornée, elle admet des moments de tous les ordres.*

Définition 48 *Soit X une variable aléatoire. On appelle*

- *moment d'ordre n*
$$\mu_n = \mathbb{E}(X^n), \quad \forall n \in \mathbb{N}$$
- *transformé de Fourier*
$$\Phi_X(t) = \mathbb{E}\left(e^{itX}\right), \quad \forall t > 0$$
- *fonction génératrice de moments*
$$M_X(t) = \mathbb{E}\left(e^{tX}\right), \quad \forall t > 0.$$

Proposition 49 *Soit X une variable aléatoire positive et $t \geq 0$. Alors*
$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n \mu_n}{n!}$$
où μ_n est le moment d'ordre n.

Démonstration. On sait que
$$e^{tx} = \sum_{n=0}^{\infty} \frac{t^n x^n}{n!}.$$

On applique l'espérance
$$\begin{aligned} M_X(t) &= \mathbb{E}\left(e^{tX}\right) \\ &= \mathbb{E}\left(\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right) \\ &= \sum_{n=0}^{\infty} \frac{t^n \mathbb{E}(X^n)}{n!} \\ &= \sum_{n=0}^{\infty} \frac{t^n \mu_n}{n!}. \end{aligned}$$

■

Proposition 50 *Soit X une variable aléatoire positive telle que $M_X(t) < \infty$ pour t tel que $|t| < \varepsilon$, où $\varepsilon > 0$. Alors $M_X(t)$ est infiniment dérivable sur $(-\varepsilon, \varepsilon)$ et la dérivé d'ordre r est*

$$M_X^{(r)}(t) = \mathbb{E}\left(e^{tX} X^r\right).$$

Démonstration. On calcule

$$\begin{aligned} M_X^{(r)}(t) &= \frac{d^r}{dt^r}\left(\sum_{n=0}^{\infty} \frac{t^n \mu_n}{n!}\right) \\ &= \sum_{n=0}^{\infty} \frac{\mu_n}{n!} \frac{d^r}{dt^r}\left(t^n\right) \\ &= \sum_{n=r}^{\infty} \frac{\mu_n}{n!} n(n-1)\ldots(n-r+1)\, t^{n-r} \\ &= \sum_{n=r}^{\infty} \frac{\mu_n}{(n-r)!} t^{n-r}. \end{aligned}$$

On note $k = n - r$ et on obtient

$$\begin{aligned} & \sum_{k=0}^{\infty} \frac{\mu_{k+r}}{k!} t^k \\ = & \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{E}\left(X^{k+r}\right) \\ = & \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} X^{k+r}\right) \\ = & \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{t^k X^k}{k!} X^r\right) \\ = & \mathbb{E}\left(e^{tX} X^r\right). \end{aligned}$$

■

Remarque 51 *On voit que pour $t = 0$ on a $M_X^{(r)}(0) = \mathbb{E}(X^r)$.*

Proposition 52 *(Inégalité de Tchebychev) Soit X une variable aléatoire et $p \in [0, \infty)$. Alors*

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}(|X|^p).$$

Démonstration. On voit que

$$\begin{aligned}
\mathbb{E}\left(|X|^p\right) &= \int_\Omega |X|^p \, d\mathbb{P} \\
&\geq \int_{\{|X|\geq\lambda\}} |X|^p \, d\mathbb{P} \\
&\geq \lambda^p \int_{\{|X|\geq\lambda\}} d\mathbb{P} \\
&= \lambda^p \mathbb{P}\left\{|X| \geq \lambda\right\}.
\end{aligned}$$

■

2.5 Variance

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $X : \Omega \to \mathbb{R}$ une variable aléatoire.

Définition 53 *Si $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, on définit la variance de X par*

$$V(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \int_\Omega |X - \mathbb{E}(X)|^2 \, d\mathbb{P}.$$

Si on prend

$$h(x) = |x - \mathbb{E}(X)|^2$$

et par la formule de changement de variables, on a

$$\begin{aligned}
V(X) &= \int_\Omega h(X(\omega)) \, d\mathbb{P}(\omega) \\
&= \int_\mathbb{R} h(x) \, d\mathbb{P}\left(X^{-1}(x)\right) \\
&= \int_\mathbb{R} |x - \mathbb{E}(X)|^2 \, d\mathbb{P}_X(x).
\end{aligned}$$

Si X est une variable aléatoire de densité f, alors on a

$$V(X) = \int_\mathbb{R} |x - \mathbb{E}(X)|^2 f(x) \, dx.$$

Proposition 54 *On a les propriétés suivantes*

1) $V(X) \geq 0$

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

3) $V(aX) = a^2 V(X)$

4) $V(b) = 0$ *si* $b = const.$

5) $V(X + b) = V(X)$

6) $V(X) = 0$ *ssi* $X = const.\ p.s.$

Démonstration.
1) $\mathbb{E}\left((X - \mathbb{E}(X))^2\right) \geq 0$

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

3) $V(aX) = \mathbb{E}\left((aX - \mathbb{E}(aX))^2\right) = a^2\mathbb{E}\left((X - \mathbb{E}(X))^2\right) = a^2V(X)$

4) $V(b) = \mathbb{E}\left((b - \mathbb{E}(b))^2\right) = \mathbb{E}\left((b-b)^2\right) = 0$

5) $V(X + b) = \mathbb{E}\left((X + b - \mathbb{E}(X + b))^2\right) = \mathbb{E}\left((X + b - \mathbb{E}(X) + b)^2\right) = V(X)$

6) Si $X = c$ (const) alors on a par 4) que $V(X) = 0$.
Si $V(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = 0$ alors la variable $(X - \mathbb{E}(X))^2$ est positive d'espérance nulle et donc nulle presque sûrement.
On obtient $X = \mathbb{E}(X)$.

■

Covariance

Définition 55 *Si* X *et* Y *sont deux variables aléatoire avec des variances finies, alors on peut définir la covariance de* X *et* Y *par*

$$\begin{aligned} cov(X, Y) &= \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

Proposition 56 *On a les propriétés suivantes*
1) $(X, Y) \longmapsto cov(X, Y)$ *est une application bilinéaire*
2) Si X *et* Y *sont des variables aléatoires centrés alors* $cov(X, Y) = \mathbb{E}[XY]$
3) Si $X = Y$ *alors* $cov(X, X) = V(X)$

Proposition 57 *Si X et Y sont deux variables aléatoire avec des variances finies, alors*

$$V(X+Y) = V(X) + V(Y) + 2cov(X,Y).$$

Démonstration. On développe

$$\begin{aligned} V(X+Y) &= \mathbb{E}\left[(X+Y)^2\right] - (\mathbb{E}(X+Y))^2 \\ &= \mathbb{E}\left[X^2+Y^2+2XY\right] - (\mathbb{E}(X)+\mathbb{E}(Y))^2 \\ &= \mathbb{E}\left[X^2\right] + \mathbb{E}\left[Y^2\right] + 2\mathbb{E}[XY] \\ &\quad - (\mathbb{E}(X))^2 - (\mathbb{E}(Y))^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\ &= V(X) + V(Y) + 2cov(X,Y). \end{aligned}$$

■

Proposition 58 *On a*

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

Démonstration. On applique l'inégalité de Holder

$$\begin{aligned} |cov(X,Y)| &= |\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]| \\ &\leq \sqrt{\mathbb{E}\left((X-\mathbb{E}(X))^2\right)\mathbb{E}\left((Y-\mathbb{E}(Y))^2\right)} \\ &\leq \sqrt{V(X)V(Y)}. \end{aligned}$$

■

Corrélation

Définition 59 *Soit X, Y deux variables aléatoires avec des variances finies. On définit le coefficient de corrélation est*

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{V(X)V(Y)}}.$$

Proposition 60 *Si $\rho(X,Y) = \pm 1$ alors il y a un lien linéaire entre X et Y de la forme $Y = aX + b$, pour $a, b \in \mathbb{R}$. En plus on peut montrer que*

$$a = \frac{cov(X,Y)}{V(X)}, \quad b = \mathbb{E}(Y) - a\mathbb{E}(X).$$

Proposition 61 *Si X est une variable aléatoire pour qui $V(X)$ existe, on a*

$$\mathbb{P}(|X-\mathbb{E}(X)|\geq t)\leq\frac{V(X)}{t^2}.$$

Démonstration. On a

$$\begin{aligned}\mathbb{P}(|X-\mathbb{E}(X)|\geq t) &= \mathbb{P}\left(|X-\mathbb{E}(X)|^2\geq t^2\right)\\ &\leq \frac{\mathbb{E}\left[(X-\mathbb{E}(X))^2\right]}{t^2}\\ &\leq \frac{V(X)}{t^2}.\end{aligned}$$

■

Chapitre 3

Vecteurs aléatoires

3.1 Rappels d'intégration

Définition 62 *Soit* $(X, \mathcal{A}, \mu)$ *et* $(Y, \mathcal{B}, \nu)$ *deux espaces mesurés.*
On définit l'espace produit

$$X \times Y = \{(x, y) \,|x \in X,\ y \in Y\}$$

$$\begin{array}{rcl} \mathcal{A} \otimes \mathcal{B} & = & \textit{tribu produit sur } X \times Y \\ & & \textit{engendrée par les produits } A \times B \end{array}$$

$$\begin{array}{rcl} \mu \otimes \nu & = & \textit{mesure produit sur } (X \times Y, \mathcal{A} \otimes \mathcal{B}) \\ (\mu \otimes \nu)(A \times B) & = & \mu(A)\,\nu(B)\,. \end{array}$$

Exemple 63 $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ *où*

$$\mathcal{B}\left(\mathbb{R}^d\right) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R})}_{d\ fois}$$

et

$$\lambda_d = \underbrace{\lambda \otimes ... \otimes \lambda}_{d\ fois}$$

où

$$\begin{array}{rcl} \lambda & : & \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+ \\ \lambda([a, b]) & = & b - a. \end{array}$$

On obtient

$$\lambda_d([a_1, b_1] \times ... \times [a_d, b_d]) = (b_1 - a_1)(b_2 - a_2) ... (b_d - a_d)\,.$$

Théorème 64 *(Fubini-Tonelli) Soit* $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow \mathbb{R}_+$ *et* μ *et* ν *sont deux mesures sur* $(X, \mathcal{A})$ *et* $(Y, \mathcal{B})$. *Alors*

1)

$$x \longmapsto \int_Y f(x,y)\, d\nu(y) \ \text{ et } y \longmapsto \int_X f(x,y)\, d\mu(x) \ \text{ sont mesurables.}$$

2)

$$\begin{aligned} \int_{X \times Y} f(x,y)\, d(\mu \otimes \nu)(x,y) &= \int_X \left(\int_Y f(x,y)\, d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x,y)\, d\mu(x) \right) d\nu(y) . \end{aligned}$$

Changement de variable

Soit $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurables et $\varphi : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ une fonction telle que $\nu = \mu\varphi^{-1}$.

Théorème 65 *(Transfert) Si* $h : (Y, \mathcal{B}) \rightarrow \mathbb{R}$ *est une fonction mesurable alors*

$$\int_X h(\varphi(x))\, d\mu(x) = \int_Y h(y)\, d\nu(y) .$$

Intégrale de Riemann

1) Sur $\mathbb{R}$

$$\int_I f(x)\, dx = \int_{\varphi(I)} f\left(\varphi^{-1}(y)\right) \left| \left(\varphi^{-1}\right)'(y) \right| dy.$$

2) Sur $\mathbb{R}^d$

$$\int_V f(x)\, d\lambda(x) = \int_{\varphi(V)} f\left(\varphi^{-1}(y)\right) \left| J_{\varphi^{-1}}(y) \right| d\lambda(y) .$$

Coordonnées polaires et sphériques

1) Soit $(x,y) \in \mathbb{R}^2$ et $(r,\theta) \in [0,\infty) \times [0, 2\pi)$ telles que

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} .$$

Alors

$$\varphi^{-1}(r,\theta) = (r\cos\theta, r\sin\theta) = (x(r,\theta), y(r,\theta))$$

et donc

$$J_{\varphi^{-1}}(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

Donc

$$\iint_{\mathbb{R}^2} f(x,y)\, dxdy = \int_0^\infty \int_0^{2\pi} f(r\cos\theta, r\sin\theta)\, r \ drd\theta.$$

2) Soit $(x, y, z) \in \mathbb{R}^3$ et $(r, \varphi, \theta) \in [0, \infty) \times [0, 2\pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ telles que

$$\begin{cases} x = r\cos\theta\cos\varphi \\ y = r\cos\theta\sin\varphi \\ z = r\sin\theta \end{cases}.$$

Alors

$$\begin{aligned} \varphi^{-1}(r, \theta, \varphi) &= (r\cos\theta\cos\varphi, r\cos\theta\sin\varphi, r\sin\theta) \\ &= (x(r, \theta, \varphi), y(r, \theta, \varphi), z(r, \theta, \varphi)) \end{aligned}$$

et donc

$$J_{\varphi^{-1}}(r, \theta, \varphi) = \begin{vmatrix} \cos\theta\cos\varphi & -r\sin\theta\cos\varphi & -r\cos\theta\sin\varphi \\ \cos\theta\sin\varphi & -r\sin\theta\sin\varphi & -r\sin\theta\cos\varphi \\ \sin\theta & r\cos\theta & 0 \end{vmatrix} = r^2\cos\theta.$$

Donc

$$\begin{aligned} & \iiint_{\mathbb{R}^3} f(x, y, z)\, dxdydz \\ = \quad & \int_0^{\infty}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(r\cos\theta\cos\varphi, r\cos\theta\sin\varphi, r\sin\theta)\, r^2\cos\theta d\theta d\varphi dr. \end{aligned}$$

3.2 Vecteurs aléatoires

Définition 66 *On appelle vecteur aléatoire une application $X : \Omega \to \mathbb{R}^d$ qui est mesurable. Si $X = (X_1, X_2, ..., X_d)$, on appelle les applications X_i marginales de X.*

Définition 67 *La loi d'un vecteur aléatoire est une mesure de probabilité sur $\mathbb{R}^d$ définie par*

$$\begin{aligned} & \mathbb{P}_X(A_1 \times A_2 \times ... \times A_d) \\ = \quad & \mathbb{P}(X_1 \in A_1, X_2 \in A_2, ..., X_d \in A_d), \quad \forall A_i \in \mathcal{B}(\mathbb{R}). \end{aligned}$$

Définition 68 *Un vecteur aléatoire X est* ***discret*** *si $X(\Omega)$ est discret dans $\mathbb{R}^d$.*

Un vecteur aléatoire X est ***de loi à densité*** *si la loi est absolument continue par rapport à la mesure de Lebesgue dans $\mathbb{R}^d$. ($\mathbb{P}_X \ll \lambda_{\mathbb{R}^d}$). Alors*

$$\mathbb{P}_X(A) = \int_A f(x_1, x_2, ..., x_d)\, dx_1 dx_2 ... dx_d, \quad \forall A \in \mathcal{B}\left(\mathbb{R}^d\right)$$

où f est une densité i.e., $f \geq 0$ et $\int_{\mathbb{R}^d} f(x_1, x_2, ..., x_d)\, dx_1 dx_2 ... dx_d = 1$.

Proposition 69 *Si (X, Y) est un vecteur aléatoire de loi $\mathbb{P}_{(X,Y)} = \mu$ alors on peut obtenir les lois marginales $\mathbb{P}_X$ et $\mathbb{P}_Y$ de X et Y par*

$$\mathbb{P}_X(A) = \mu(A \times \mathbb{R}) \text{ et } \mathbb{P}_Y(B) = \mu(\mathbb{R} \times \mathbb{B}).$$

Démonstration. On voit facilement que

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}(\{X \in A\}) \\ &= \mathbb{P}(\{(x, y) \in A \times \mathbb{R}\}) \\ &= \mu(A \times \mathbb{R}). \end{aligned}$$

■

Exemple 70 ***(Cas discret)***

$$(X, Y)(\Omega) = \{(x_1, y_2), ..., (x_i, y_i), ...\}$$

$$\begin{aligned} X(\Omega) &= \{x_1, x_2, ..., x_i, ...\} \\ Y(\Omega) &= \{y_1, y_2, ..., y_i, ...\} \end{aligned}$$

$$\mathbb{P}_X(x_i) = \mathbb{P}(X = x_i) = \sum_{y_j \in Y(\Omega)} \mathbb{P}(X = x_i; Y = y_j).$$

Exemple 71 ***(Cas à densité)***
Soit (X, Y) un couple de variables aléatoires de loi $\mathbb{P}_{(X,Y)}$ de densité f. Alors les lois marginales $\mathbb{P}_X$ et $\mathbb{P}_Y$ sont des lois de densités donnés par

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f(x, y)\, dy \\ f_Y(y) &= \int_{\mathbb{R}} f(x, y)\, dx. \end{aligned}$$

Démonstration. On voit que

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}(X \in A;\ Y \in \mathbb{R}) \\ &= \mathbb{P}((X, Y) \in A \times \mathbb{R}) \\ &= \mathbb{P}_{(X,Y)}(A \times \mathbb{R}) \\ &= \int_{A \times \mathbb{R}} f(x, y)\, dxdy \\ &= \int_A \left(\int_{\mathbb{R}} f(x, y)\, dy \right) dx \\ &= \int_A f_X(x)\, dx, \end{aligned}$$

donc f_X est la densité de loi $\mathbb{P}_X$. ■

Généralisation

Soit $X = (X_1, X_2, ..., X_d)$ de densité f. Alors la loi de la variable marginale X_i est de densité

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{d-1}} f(x_1, x_2, ..., x_d)\, dx_1 ... dx_{i-1} dx_{i+1} ... dx_d.$$

En effet

$$\begin{aligned}
\mathbb{P}(X_i \in A) &= \mathbb{P}(X_1 \in \mathbb{R}, ..., X_i \in A, ..., X_d \in \mathbb{R}) \\
&= \int_{A \times \mathbb{R}^{d-1}} f(x_1, x_2, ..., x_i, ..., x_d)\, dx_1 ... dx_i ... dx_d \\
&= \int_A f_{X_i}(x_i)\, dx_i.
\end{aligned}$$

3.3 Fonction d'un vecteur aléatoire

Soit $Z = (X, Y)$ un vecteur aléatoire et $h : \mathbb{R}^2 \to \mathbb{R}$.

Esperance

$$\begin{aligned}
\mathbb{E}(h(X, Y)) &= \int_\Omega h(X(\omega), Y(\omega))\, d\mathbb{P}(\omega) \\
&= \iint_{\mathbb{R}^2} h(x, y)\, d\mathbb{P}_{(X,Y)}(x, y).
\end{aligned}$$

Loi

$$\begin{aligned}
\mathbb{E}(g(h(X, Y))) &= \iint_{\mathbb{R}^2} g(h(x, y))\, d\mathbb{P}_{(X,Y)}(x, y) \\
&= \iint_{\mathbb{R}^2} g(h(x, y))\, f_{(X,Y)}(x, y)\, dxdy.
\end{aligned}$$

On fait un changement de variable $u = h(x, y)$ on obtient

$$\mathbb{E}(g(h(X, Y))) = \iint_{\mathbb{R}^2} g(u)\, f_{(X,Y)}(x(u, v), y(u, v))\, |J|\, dudv.$$

On intégre par rapport à v.

Chapitre 4

Indépendance

Définition 72 *Deux événements A et B d'un espace de probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$ *sont indépendantes si*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)\,.$$

Remarque 73 *Il ne faut pas confondre avec incompatibilité, i.e.* $A \cap B = \emptyset$.

Définition 74 *Une famille d'évènements $A_1,\ A_2,\ ..., A_n$ sont mutuellement indépendantes si pour toute sous-famille $A_{i_1},\ A_{i_2}, ..., A_{i_k}$ on a*

$$\mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}\right) = \mathbb{P}\left(A_{i_1}\right) ... \mathbb{P}\left(A_{i_k}\right).$$

Définition 75 *Deux tribus ($\sigma-$algébre) $\mathcal{F}$ et $\mathcal{G}$ sur un même espace de probabilité Ω sont indépendantes si pour tout $A \in \mathcal{F}$ et $B \in \mathcal{G}$ on a*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)\,.$$

Indépendance de variables aleatoires

Définition 76 *Deux variabiles aléatoires X et Y sont indépendantes si les tribus engendrées par X et Y sont indépendantes.*

Proposition 77 *Deux variables aléatoires X et Y sont indépendantes si pour tout $A, B \in \mathcal{B}(\mathbb{R})$ les évènements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendantes, i.e.*

$$\mathbb{P}(X \in A; Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$$

ou

$$\mathbb{P}\left(X^{-1}(A) \cap Y^{-1}(B)\right) = \mathbb{P}\left(X^{-1}(A)\right)\mathbb{P}\left(Y^{-1}(B)\right).$$

Définition 78 *Une famille de variables aléatoires* $X_1,\ X_2, ..., X_n$ *sont dites mutuellement indépendantes si pour toute famille d'évènements* $A_{i_1}, A_{i_2}, ..., A_{i_k} \in \mathcal{B}(\mathbb{R})$ *les évènements* $\{X_{i_1} \in A_{i_1}\}, ..., \{X_{i_k} \in A_{i_k}\}$ *sont indépendantes.*

Définition 79 *Une suite de variables aléatoires* $(X_i)_{i\in\mathbb{N}}$ *est dite indépendante si toute famille finie d'éléments de* $(X_i)_{i\in\mathbb{N}}$ *a la propriété précédente.*

Indépendance de vecteurs aleatoires

Proposition 80 *Un vecteur aléatoire* $X = (X_1, X_2, ..., X_n)$ *est à composantes indépendantes si et seulement si sa loi* $\mathbb{P}_X$ *est une loi produit de ses lois marginales*

$$\mathbb{P}_X = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \otimes ... \otimes \mathbb{P}_{X_n}.$$

Démonstration. Soit

$$B = B_1 \times B_2 \times ... \times B_n \in \mathcal{B}(\mathbb{R}^n).$$

Alors

$$\begin{aligned}
\mathbb{P}_X(B) &= \mathbb{P}((X_1, X_2, ..., X_n) \in B_1 \times B_2 \times ... \times B_n) \\
&= \mathbb{P}(X_1 \in B_1, X_2 \in B_2, ..., X_n \in B_n) \\
&\overset{ind\acute{e}p}{=} \mathbb{P}(X_1 \in B_1)\,\mathbb{P}(X_2 \in B_2)\,...\mathbb{P}(X_n \in B_n) \\
&= \mathbb{P}_{X_1}(B_1)\,\mathbb{P}_{X_2}(B_2)\,...\mathbb{P}_{X_n}(B_n) \\
&= \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \otimes ... \otimes \mathbb{P}_{X_n}(B_1 \times B_2 \times ... \times B_n) \\
&= \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \otimes ... \otimes \mathbb{P}_{X_n}(B).
\end{aligned}$$

■

Corollaire 81 *Les variables aléatoires* $X_1, X_2, ...X_n$ *sont indépendantes si et seulement si*

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

Applications

1) Cas discret

Les variables aléatoires X et Y sont indépendantes si et seulement si $\forall x \in X(\Omega)$ et $\forall y \in Y(\Omega)$ on a

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y).$$

2) Cas à densité

Les variables aléatoires X et Y de densités f et g sont indépendantes si et seulement si (X,Y) a comme densité

$$\begin{aligned} f\otimes g &: \mathbb{R}^2\to\mathbb{R} \\ (f\otimes g)(x,y) &= f(x)\,g(y)\,. \end{aligned}$$

Proposition 82 *Soient $\{X_i\}_{1\leq i\leq n}$ des variables aléatoires indépendates et $h_i:\mathbb{R}\to\mathbb{R}$, $1\leq i\leq n$, des fonctions mesurables telles que $h_i(X_i)$ sont intégrables.*

Alors le produit $\prod_{i=1}^n h_i(X_i)$ est intégrable et

$$\mathbb{E}\left(\prod_{i=1}^n h_i(X_i)\right)=\prod_{i=1}^n \mathbb{E}\left(h_i(X_i)\right).$$

Démonstration. On voit que

$$\begin{aligned} \mathbb{E}\left(\prod_{i=1}^n h_i(X_i)\right) &= \int_\Omega \prod_{i=1}^n h_i(X_i)\,d\mathbb{P} \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n h_i(x_i)\,d\mathbb{P}_{(X_1,X_2,\dots X_n)}(x_1,x_2,\dots,x_n) \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n h_i(x_i)\,d\mathbb{P}_{X_1}(x_1)\,d\mathbb{P}_{X_2}(x_2)\dots d\mathbb{P}_{X_n}(x_n) \\ &= \int_{\mathbb{R}} h_1(x_1)\,d\mathbb{P}_{X_1}(x_1)\int_{\mathbb{R}} h_2(x_2)\,d\mathbb{P}_{X_2}(x_2)\dots\int_{\mathbb{R}} h_n(x_n)\,d\mathbb{P}_{X_n}(x_n) \\ &= \mathbb{E}(h_1(x_1))\,\mathbb{E}(h_2(x_2))\dots\mathbb{E}(h_n(x_n))\,. \end{aligned}$$

■

Corollaire 83 *Soient $X_1,X_2,\dots,X_n$ des variables aléatoires indépendantes et integrables. Alors*

$$\mathbb{E}(X_1X_2\dots X_n)=\mathbb{E}(X_1)\,\mathbb{E}(X_2)\dots\mathbb{E}(X_n)\,.$$

Covariance et indépendance

Proposition 84 *Soient X et Y deux v.a. indépendantes avec variances finies. Alors $cov(X,Y)=0$.*

Démonstration.

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

■

Corollaire 85 *Si X et Y sont deux variables aléatoires indépendantes avec variances finies alors on a*

$$V\left(X+Y\right)=V\left(X\right)+V\left(Y\right).$$

Proposition 86 *Si X et Y sont deux variables aléatoires indépendantes alors pour toute fonction f et g on a que $f\left(X\right)$ et $g\left(Y\right)$ sont aussi indépendantes.*

Chapitre 5

Espérance conditionnelle

On rappelle que la proabilité conditionnelle est définie par

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}(1_A 1_B)}{\mathbb{P}(B)}$$

car $\mathbb{P}(A) = \mathbb{E}(1_A)$ et $1_A 1_B = 1_{A\cap B}$.

Pour une variable aléatoire $X : \Omega \to \mathbb{R}$ on définit

$$\mathbb{E}(X|B) = \frac{\mathbb{E}(X \cdot 1_B)}{\mathbb{P}(B)} = \frac{1}{\mathbb{P}(B)} \int_\Omega X 1_B d\mathbb{P} = \int_\Omega X \frac{1_B}{\mathbb{P}(B)} d\mathbb{P}.$$

Comme $\frac{1_B}{\mathbb{P}(B)}$ est une dénsité on peut noter $\frac{1_B}{\mathbb{P}(B)} d\mathbb{P} = d\mathbb{P}_B$ où

$$\mathbb{P}_B(A) = \int_A 1_B \frac{1}{\mathbb{P}(B)} d\mathbb{P} = \int_\Omega 1_{A\cap B} \frac{1}{\mathbb{P}(B)} d\mathbb{P} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Soit X une variable aléatoire integrable sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et $\mathcal{B}$ une $\sigma-$algébre (tribu) telle que $\mathcal{B} \subset \mathcal{F}$.

Définition 87 *On définit l'espérance conditionnelle de X sachant $\mathcal{B}$ une variable aléatoire integrable Z qui a les proprietés suivantes*

i) Z est mesurable par rapport à $\mathcal{B}$

ii) pour tout $A \in \mathcal{B}$ on a

$$\mathbb{E}(X \cdot 1_A) = \mathbb{E}(Z \cdot 1_A)$$

ou

ii)' pour tout variable aléatoire Y qui est borné et $\mathcal{B}-$mesurable on a

$$\mathbb{E}(XY) = \mathbb{E}(ZY).$$

On note $Z = \mathbb{E}(X|\mathcal{B})$.

Remarque 88 *La variable aléatoire Z est unique.*

Démonstration. Soit Z' une variable aléatoire qui vérifie les deux conditions. On a

$$\mathbb{E}((Z - Z')\,1_A) = 0 \text{ et } \mathbb{E}((Z - Z')\,1_{\overline{A}}) = 0.$$

Soit

$$A = \{\omega \in \Omega \mid Z'(\omega) < Z(\omega)\}$$

$$\begin{array}{rcl} \mathbb{E}((Z - Z')\,1_A) & = & 0 \\ (Z - Z')\,1_A & \gneqq & 0 \end{array}$$

on a $Z = Z'$ p.s. A.

De la même façon, pour

$$\overline{A} = \{\omega \in \Omega \mid Z'(\omega) \geq Z(\omega)\}$$

$$\begin{array}{rcl} \mathbb{E}((Z' - Z)\,1_{\overline{A}}) & = & 0 \\ (Z' - Z)\,1_{\overline{A}} & \gneqq & 0 \end{array}$$

on a $Z = Z'$ p.s. $\overline{A}$. ∎

Proposition 89 (Propriétés spécifiques)

1) On a

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

2) Si X est indépendant de $\mathcal{B}$, alors

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

3) Si X est $\mathcal{B}$ - mesurable, alors

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

4) Si Y est une variable aléatoire borné et $\mathcal{B}$ - mesurable, alors

$$\mathbb{E}(YX|\mathcal{B}) = Y\mathbb{E}(X|\mathcal{B})$$

5) Soit $\mathcal{G} \subset \mathcal{B}$ deux $\sigma-$algébre. Alors

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

Démonstration.

1) On choisi $A = \Omega$ et on applique $ii)$

$$\left.\begin{array}{r} \mathbb{E}(X1_\Omega) = \mathbb{E}(X) \\ \mathbb{E}(Z1_\Omega) = \mathbb{E}(Z) \end{array}\right\} \Rightarrow \mathbb{E}(X) = \mathbb{E}(Z)$$

donc $\mathbb{E}(\mathbb{E}(X|\mathcal{B})) = \mathbb{E}(X)$.

2) On a

$$\begin{aligned} \mathbb{E}(X1_A) &= \mathbb{E}(X)\mathbb{E}(1_A) \\ &= \mathbb{E}(\mathbb{E}(X)1_A) \\ &= \mathbb{E}(\mathbb{E}(X|\mathcal{B})1_A), \quad \forall A \in \mathcal{B}. \end{aligned}$$

Donc

$$\mathbb{E}(X) = \mathbb{E}(X|\mathcal{B}) \text{ p.s.}$$

3) Comme X est $\mathcal{B}-$mesurable, alors c'est évident que

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

4) On a

$$\left.\begin{array}{r} \mathbb{E}(Y\mathbb{E}(X|\mathcal{B})1_A) = \mathbb{E}(XY1_A) \\ \mathbb{E}(\mathbb{E}(XY|\mathcal{B})1_A) = \mathbb{E}(XY1_A) \end{array}\right\} \Rightarrow \mathbb{E}(YX|\mathcal{B}) = Y\mathbb{E}(X|\mathcal{B}).$$

5) C'est évident en utilisant 3).

■

Proposition 90 (Propriétés similaires aux celles de l'éspérance)

1) Linéarité

$$\mathbb{E}((aX+b)|\mathcal{B}) = a\mathbb{E}((X)|\mathcal{B}) + b.$$

$$\mathbb{E}((X_1+X_2)|\mathcal{B}) = \mathbb{E}((X_1)|\mathcal{B}) + \mathbb{E}((X_2)|\mathcal{B}).$$

2) Monotonie

$$X_1 \leq X_2 \textit{ p.s. alors } \mathbb{E}(X_1|\mathcal{B}) \leq \mathbb{E}(X_2|\mathcal{B}).$$

3) Si X et X_n sont des variables aléatoire réelles dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ alors $X_n \nearrow X$ implique $\mathbb{E}(X_n|\mathcal{B}) \nearrow \mathbb{E}(X|\mathcal{B})$.

4) Soit f une fonction continue et convexe et X une variable aléatoire réelle telle que X et $f(X)$ sont intégrables, alors

$$f(\mathbb{E}(X|\mathcal{B})) \leq \mathbb{E}(f(X)|\mathcal{B}).$$

5) Si on a $X_n \to X$ p.s. et $|X_n| \leq Z \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. Alors

$$\lim_{n\to\infty} \mathbb{E}(X_n|\mathcal{B}) = \mathbb{E}(X|\mathcal{B}).$$

Chapitre 6

Somme de deux variables aléatoires

6.1 Convolution de mesures et de fonctions

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité qui a une structure d'espace vectoriel.pour Ω.

Définition 91 *Pour deux mesures de probailité μ et ν sur Ω on définit la convolution*

$$(\mu * \nu)(A) = \int_{\Omega} \mu(A - x)\, d\nu(x)$$

où $A - x = \{a - x \,|a \in A\}$.

On sait que :
- $\mu * \nu = \nu * \mu$
- $\mu * \nu$ est une mesure de probabilité
- $\mu * \delta_0 = \delta_0 * \mu = \mu$ où $\delta_0(A) = \begin{cases} 1, \text{si } 0 \in A \\ 0, \ \text{si } 0 \notin A \end{cases}$ (la mesure de Dirac).

Définition 92 *Pour deux fonctions réelles f et g, la convolution est une fonction donné par*

$$(f * g)(x) = \int_{\mathbb{R}} f(x - y)\, g(y)\, dy.$$

Proposition 93 *Soient μ et ν des mesures de densité f et g par rapport à une mesure λ, i.e. $d\mu = fd\lambda$ et $d\nu = gd\lambda$. Alors on a*

$$(\mu * \nu)(A) = \int_A (f * g)\, d\lambda.$$

Démonstration. On voit que

$$\begin{aligned}
(\mu * \nu)(A) &= \int\limits_{\Omega} \mu(A - x)\, d\nu(x) \\
&= \int\limits_{\Omega}\int\limits_{\Omega} 1_{(A-x)}(y)\, d\mu(y)\, d\nu(x) \\
&= \int\limits_{\Omega}\int\limits_{\Omega} 1_A(x+y)\, d\mu(y)\, d\nu(x) \\
&= \int\limits_{\{x+y\in A\}}\int f(x)\, g(y)\, dxdy \\
&\overset{t=x \text{ et } s=x+y}{=} \int_A \left(\int_{\mathbb{R}} f(t)\, g(s-t)\, dt \right) ds \\
&= \int_A (f * g)(s)\, ds.
\end{aligned}$$

■

6.2 Loi d'une somme de variables aléatoires indépendantes à densité

Proposition 94 *Soient X et Y deux variables aléatoires indépendantes et de loi $\mathbb{P}_X$ et $\mathbb{P}_Y$. Alors la loi de $X+Y$ est $\mathbb{P}_X * \mathbb{P}_Y$.*

Proposition 95 *Soient X et Y deux variables aléatoires indépendantes et de densités f et g. La loi de $X+Y$ est donné par*

$$\mathbb{P}_{X+Y}(A) = \mathbb{P}(X+Y \in A) = \int_A (f * g)(x)\, dx$$

pour tout $A \in \mathcal{B}(\mathbb{R})$. En fait, $X+Y$ a la densité $f * g$.

Démonstration. On sait par la proposition précédente que

$$\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$$

et par la Proposition 93

$$(\mathbb{P}_X * \mathbb{P}_Y)(A) = \int_A (f * g)(x)\, dx.$$

On obtient que

$$\mathbb{P}_{X+Y}(A) = \int_A (f * g)(x)\, dx.$$

■

Chapitre 7

Convergences de variables aléatoires

7.1 Convergence presque sûre

Définition 96 *Une suite de variables aléatoires $(X_n)_n$ converge presque sûrement vers X si on a*

$$\mathbb{P}\left(\left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

On note

$$X_n \to X \quad \text{p.s.}$$

7.2 Convergence en norme L^p

On rappelle que pour tout $p > 1$ on définit les espaces

$$L^p(\Omega) = \left\{\widehat{X} : \Omega \to \mathbb{R} \,\middle|\, \|X\|_p = (\mathbb{E}|X|^p)^{1/p} < \infty\right\}.$$

On sait que c'est une espace de Banach pour la norme $\|.\|_p$. On remarque que pour $p = 2$ on a un espace de Hilbert avec

$$\langle X, Y\rangle = \mathbb{E}(XY).$$

Définition 97 *Une suite de variables aléatoires $(X_n)_n$ converge en norme L^p vers X si on a*

$$\lim_{n\to\infty} \|X_n - X\|_p = 0$$

ou

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

On note

$$X_n \xrightarrow{L^p} X.$$

7.3 Convergence en loi

Une suite de variables aléatoires $(X_n)_n$ converge en loi vers une variable aléatoire X si les fonctions de répartition F_n des variables X_n convergent vers la fonction de répartition F de la variable X en tout point . Donc

$$F_n(x) \longrightarrow F(x), \quad \forall x \in \mathbb{R}.$$

On note

$$X_n \xrightarrow{\mathcal{L}} X.$$

Formes équivalentes :

- On dit que la suite de variables aléatoires $(X_n)_n$ converge en loi vers une variable aléatoire X si

$$\lim_{n\to\infty} \mathbb{E}(f(X_n)) = \mathbb{E}(f(X)),$$

pour toute fonction continue et bornée f.
- On dit que la suite de variables aléatoires $(X_n)_n$ converge en loi vers une variable aléatoire X si

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x)\, d\mathbb{P}_{X_n}(x) = \int_{\mathbb{R}} f(x)\, d\mathbb{P}_X(x),$$

pour toute fonction continue et bornée f.

7.4 Convergence en probabilité

Une suite de variables aléatoires $(X_n)_n$ converge en probabilité vers X si pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathbb{P}(|X - X_n| \geq \varepsilon) = 0.$$

On note

$$X_n \xrightarrow{\mathbb{P}} X.$$

7.5 Relatoins entre les types de convergence

Proposition 98 *La* ***convergence en probabilité*** *implique la* ***convergence en loi****. La réciproque est en général fausse.*

Proposition 99 *Si une suite de variables aléatoires $(X_n)_n$ converge en loi vers la constante C, alors $(X_n)_n$ converge en probabilité vers C.*

Démonstration. On sait que

$$\mathbb{P}\left(|X_n - C| \geq \varepsilon\right) = \mathbb{E}\left(1_{\{|X_n - C| \geq \varepsilon\}}\right).$$

De l'autre côté on observe que

$$|X_n - C| \geq \varepsilon \Leftrightarrow X_n \in (-\infty, C - \varepsilon) \cup (C + \varepsilon, +\infty)$$

et donc

$$1_{\{|X_n - C| \geq \varepsilon\}} = 1_{\{(-\infty, C-\varepsilon) \cup (C+\varepsilon, +\infty)\}}\left(X_n\right) \stackrel{Not}{=} h\left(X_n\right).$$

On définit aussi une fonction

$$g(x) = \begin{cases} 1, & x \leq C - \varepsilon \text{ ou } x \geq C + \varepsilon \\ afine, & x \in (C - \varepsilon, C + \varepsilon) \backslash C \\ 0, & x = C \end{cases}$$

On remarque que g est continue, bornée et que $h \leq g$.
On revient à la premiere relation

$$\begin{aligned} \mathbb{P}\left(|X_n - C| \geq \varepsilon\right) &= \mathbb{E}\left(1_{\{|X_n - C| \geq \varepsilon\}}\right) \\ &= \mathbb{E}\left(1_{\{(-\infty, C-\varepsilon) \cup (C+\varepsilon, +\infty)\}}\left(X_n\right)\right) \\ &= \mathbb{E}\left(h\left(X_n\right)\right) \leq \mathbb{E}\left(g\left(X_n\right)\right) \\ &\leq \int_\Omega g\left(X_n\right) d\mathbb{P} \rightarrow \int_\Omega g\left(C\right) d\mathbb{P} = g\left(C\right) = 0. \end{aligned}$$

Donc

$$\mathbb{P}\left(|X_n - C| \geq \varepsilon\right) \underset{n \to \infty}{\longrightarrow} 0$$

ce qui est la convergence en probabilité. ■

Proposition 100 *Si X_n et X sont des variables aléatoires dans $L^p(\Omega)$, alors la convergence $X_n \xrightarrow{L^p} X$ entraîne la convergence en probabilité $X_n \xrightarrow{\mathbb{P}} X$.*

Démonstration. On applique l'inégalité de Tchevychev

$$\begin{aligned} \mathbb{P}\left(|X_n - X| \geq \varepsilon\right) &= \mathbb{P}\left(|X_n - X|^p \geq \varepsilon^p\right) \\ &\leq \frac{\mathbb{E}\left(|X_n - X|^p\right)}{\varepsilon^p} \\ &= \frac{\|X_n - X\|_p^p}{\varepsilon^p} \rightarrow 0 \text{ pour } n \rightarrow \infty. \end{aligned}$$

■

Proposition 101 *Si une suite de variables aléatoires $\left(X_n\right)_n$ converge en presque sûrement vers une variable aléatoire X alors elle converge aussi en probabilité.*

Proposition 102 *Si la suite $\left(X_n\right)_n$ converge en probabilité vers X, alors elle admet une sous-suite $\left(X_{n_k}\right)_k$ qui converge presque sûrement vers X.*

Schéma

$$\begin{array}{ccc} \textit{Convergence p.s.} & \overset{\textit{Lebesgue}}{\rightsquigarrow} & \textit{Convergence } L^p \\ \searrow & & \swarrow \\ & \textit{Convergence en } \mathbb{P} & \\ & \downarrow & \\ & \textit{Convergence en loi} & \end{array}$$

Chapitre 8

Lois des grandes nombres

La loi des grands nombres représente la convergence de la moyenne arithmétique vers la moyenne probabiliste (espérance).

8.1 Version faible de la LGN

Théorème 103 *Soit $(X_n)_n$ une suite de variables aléatoires indépendantes et de même loi (i.i.d.) avec un moment d'ordre deux. Alors*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mathbb{E}\left(X_1\right), \quad \textit{pour } n \to \infty.$$

Démonstration. On observe que la variable aléatoire limite est la constante $\mathbb{E}(X_1)$ qui est égale à $\mathbb{E}(X_i)$ pour tout i car les variables aléatoires X_i ont la même loi, donc la même espérance.

On doit montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}\left(X_1\right)\right| \geq \varepsilon\right) = 0. \tag{8.1}$$

Posons

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Par la linéarité de l'espérance on a

$$\mathbb{E}\left(M_n\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left(X_i\right) = \mathbb{E}\left(X_1\right).$$

D'autre part, par l'indépendance des variables aléatoires X_i, on a

$$\begin{aligned} V(M_n) &= V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{V(X_1)}{n}. \end{aligned}$$

On rappelle l'inégalité de Tchebychev

$$\begin{aligned} \mathbb{P}(|X| > \varepsilon) &\leq \frac{\mathbb{E}(|X|^p)}{\varepsilon^p} \\ \mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) &\leq \frac{\mathbb{E}\left((X - \mathbb{E}(X))^2\right)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}. \end{aligned}$$

On revient à l'équation (8.1) et en remplaçant, on obtient que

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1)\right| \geq \varepsilon\right) \\ = & \lim_{n\to\infty} \mathbb{P}(|M_n - \mathbb{E}(M_n)| \geq \varepsilon) \\ \leq & \lim_{n\to\infty} \frac{\mathbb{E}\left((M_n - \mathbb{E}(M_n))^2\right)}{\varepsilon^2} \\ = & \lim_{n\to\infty} \frac{V(M_n)}{\varepsilon^2} \\ = & \lim_{n\to\infty} \frac{V(X_1)}{\varepsilon^2 n} = 0. \end{aligned}$$

■

Exemple 104 *On souhaite estimer le paramètre p inconnu d'une loi de Bernoulli en observant un grand nombre de fois le phénomène aléatoire de loi $\mathcal{B}(p)$, c'est à dire en observant les réalisations d'une suite de variables aléatoires indépendantes X_i, $1 \leq i \leq n$ de loi $\mathcal{B}(p)$.*

Notons

$$X_i = \begin{cases} 0, & \text{avec probabilité } 1-p \\ 1, & \text{avec probabilité } p \end{cases}.$$

et

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i \underset{n\to\infty}{\longrightarrow} p.$$

On a

$$\mathbb{P}(|M_n - p| > \varepsilon) \leq \frac{V(X_1)}{n\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$$

(car on majore $p(1-p)$ par $\frac{1}{4}$).

On obtient

$$|M_n - p| > \varepsilon \Leftrightarrow |p - M_n| > \varepsilon$$

$$\begin{aligned}
\mathbb{P}(|p - M_n| > \varepsilon) &= 1 - \mathbb{P}(|p - M_n| < \varepsilon) \\
&= 1 - \mathbb{P}(-\varepsilon < p - M_n < \varepsilon).
\end{aligned}$$

Finalement

$$\begin{aligned}
\mathbb{P}(M_n - \varepsilon < p < M_n + \varepsilon) &= 1 - \mathbb{P}(|p - M_n| > \varepsilon) \\
&= 1 - \mathbb{P}(|M_n - p| > \varepsilon) \\
&\geq 1 - \frac{1}{4n\varepsilon^2}.
\end{aligned}$$

et donc

$$\mathbb{P}(M_n - \varepsilon < p < M_n + \varepsilon) \geq \underbrace{1 - \frac{1}{4n\varepsilon^2}}_{\overset{Not}{=} 1-\alpha}.$$

On note par $1-\alpha$ la probabilité d'erreur et par $M_n - \varepsilon < p < M_n + \varepsilon$ l'intervalle de confiance.

On a donc

$$\alpha = \frac{1}{4n\varepsilon^2} \text{ et } \varepsilon = \sqrt{\frac{1}{4n\alpha}}.$$

Sondage

- *Un institut de sondage interroge 1000 personnes. On note p la probabilité de dire oui et p non.*
  Dans le sondage cette proportion est $0,54$. On est interessé par un interval de confiance pour p au niveau $0,95$.
  On a

  $$\begin{aligned}
  M_{1000} &= 0,54 \\
  I &= [0,54 - \varepsilon\ ,\ 0,54 + \varepsilon] \\
  \mathbb{P}(p \in I) &\geq 1 - \frac{1}{4 \cdot 1000 \cdot \varepsilon^2} = 1 - \alpha.
  \end{aligned}$$

  Pour avoir un niveau de confience de $0,95$ on pose

  $$\begin{aligned}
  1 - \frac{1}{4 \cdot 1000 \cdot \varepsilon^2} &\geq 0,95 \\
  \frac{1}{4 \cdot 1000 \cdot \varepsilon^2} &\leq 0,05 \\
  \varepsilon &\geq \sqrt{\frac{1}{200}} \simeq 0,0707.
  \end{aligned}$$

- *On cherche n pour que $\varepsilon = 0,01$ et $1-\alpha = 0,95$.*
  Donc

$$\begin{aligned} 1 - \frac{1}{4 \cdot n \cdot (0,01)^2} &\geq 0,95 \\ \frac{1}{4 \cdot n \cdot (0,01)^2} &\leq 0,05 \\ n &\geq \frac{1}{4 \cdot 0,05 \cdot (0,01)^2} \\ &= 50.000 \end{aligned}$$

8.2 Version forte de la LGN

Théorème 105 *Soit $(X_n)_n$ une suite de variables aléatoires indépendantes et de même loi (i.i.d.) avec un moment d'ordre un (i.e. $\mathbb{E}(|X_1|) < \infty$, si l'espérance est bien définie). Alors*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} \mathbb{E}(X_1), \quad \text{pour } n \to \infty.$$

8.2.1 Estimateurs

Si X est une variable aléatoire, on appelle échantillon indépendant de la loi de X la suite $(X_i)_i$ qui est i.id. de même loi que X. On définit
- la moyenne empirique $M_n = \frac{1}{n}\sum_{k=1}^{n} X_k$
- la variance empirique $V_n = \frac{1}{n}\sum_{k=1}^{n} (X_k - M_n)^2$

Proposition 106 *Les estimateurs M_n et V_n sont consistants pour $\mathbb{E}(X)$ et $V(X)$ i.e.*

$$\begin{aligned} M_n &\xrightarrow{p.s.} \mathbb{E}(X), \quad \text{pour } n \to \infty \\ V_n &\xrightarrow{p.s.} V(X), \quad \text{pour } n \to \infty. \end{aligned}$$

Démonstration. On a

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} \mathbb{E}(X), \quad \text{pour } n \to \infty \text{ par LGN}$$

et donc

$$M_n \xrightarrow{p.s.} \mathbb{E}(X), \quad \text{pour } n \to \infty.$$

Pour la variance on a

$$\begin{aligned}
V_n &= \frac{1}{n}\left(\sum_{k=1}^{n}\left(X_k^2 - 2X_k M_n + M_n^2\right)\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^{n} X_k^2 - 2M_n \sum_{k=1}^{n} X_k + nM_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^{n} X_k^2 - 2nM_n^2 + nM_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^{n} X_k^2 - nM_n^2\right) \\
&\simeq \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = V(X)
\end{aligned}$$

■

8.2.2 Méthode de Monte-Carlo

Soit $f : B \longrightarrow \mathbb{R}^2$ et $(X_n)_n$ une suite de variables aléatoires i.id. uniformes sur B. On a

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow{p.s.} \frac{1}{\lambda(B)}\int_B f(x)\,dx, \quad n \to \infty.$$

En effet

$$\begin{aligned}
\mathbb{E}(f(X)) &= \int_{\mathbb{R}} f(x)\,d\mathbb{P}_X(x) \\
&= \int_{\mathbb{R}} f(x)\frac{1}{\lambda(B)}1_B(x)\,dx \\
&= \frac{1}{\lambda(B)}\int_B f(x)\,dx.
\end{aligned}$$

Donc on peut approcher l'intégrale en approchant $\mathbb{E}(f(X))$ par $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$.

8.3 Version L^1 de la LGN

Soit $(X_n)_n$ une suite de variables aléatoires indépendantes.

On note $S_n = \sum_{k=1}^{n} X_k$ la somme et on étudie la convergence de ces quantités quand $n \to \infty$.

Proposition 107 *Si* $(X_n)_n$ *est une suite de variables aléatoires indépendantes, alors on a*

$$S_n \text{ converge p.s.} \quad \Leftrightarrow \quad S_n \text{ converge en probabilité.}$$

Proposition 108 *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes centrées, de carrés intégrables. Alors on a*

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| > \alpha\right) \leq \frac{\mathbb{E}\left|S_n^2\right|}{\alpha^2}.$$

Cette inégalité est meilleure que celle de Tchebychev appliquée à S_n *qui donne*

$$\mathbb{P}\left(|S_n| > \alpha\right) \leq \frac{\mathbb{E}\left|S_n^2\right|}{\alpha^2}.$$

Proposition 109 *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes centrées, de* L^2*. Si* $\sum_n \mathbb{E}\left|X_n^2\right| < \infty$ *alors* S_n *converge p.s.*

Proposition 110 *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes centrées, de* L^2 *et* $a_n > 0$ *des réels qui cenvergent en croissant vers* $+\infty$*. Alors si* $\sum_{n=1}^{\infty} \frac{\mathbb{E}\left[X_n^2\right]}{a_n^2} < \infty$ *on a* $\frac{S_n}{a_n} \overset{p.s.,\ L^2}{\longrightarrow} 0.$

Théorème 111 *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes et de même loi (v.a.i.i.d.) dans* L^1*. Alors*

$$\frac{S_n}{n} \xrightarrow{L^1} \mathbb{E}\left[X_1\right].$$

Chapitre 9

Théorème limite centrale

9.1 Fonction caractéristique

Définition 112 *Soit X une variable aléatoire. On appelle fonction caractéristique une fonction de $\mathbb{R}$ dans $\mathbb{C}$*

$$\Phi_X(t) = \mathbb{E}\left(e^{itX}\right).$$

Remarque 113 *Si X est une variable aléatoire discrète avec $X(\Omega) = \{x_1, x_2, ..., x_n\}$, alors on a*

$$\Phi_X(t) = \sum_{k=1}^{n} e^{itx_k} \mathbb{P}(X = x_k).$$

Si X est une variable aléatoire à densité, alors on a

$$\begin{aligned} \Phi_X(t) &= \int_\Omega e^{itX} d\mathbb{P} \\ &= \int_\mathbb{R} e^{itx} d\mathbb{P}_X(x) \\ &= \int_\mathbb{R} e^{itx} f(x)\, dx \end{aligned}$$

Proposition 114 *La fonction caractéristique caractérise la loi. Donc si $\Phi_X = \Phi_Y$ alors les variables aléatoires X et Y on la même loi et réciproquement.*

Proposition 115 *Si on a deux variables aléatoires indépendantes X et Y alors*

$$\Phi_{X+Y}(t) = \Phi_X(t)\,\Phi_Y(t).$$

Démonstration. On voit facilement que

$$\Phi_{X+Y}(t) = \mathbb{E}\left(e^{it(X+Y)}\right) = \mathbb{E}\left(e^{itX}\right)\mathbb{E}\left(e^{itY}\right) = \Phi_X(t)\,\Phi_Y(t)\,.$$

■

Définition 116 *Soit (X,Y) un vecteur aléatoire. On définit la fonction caractéristique comme une fonction de $\mathbb{R}^2$ dans $\mathbb{C}$*

$$\begin{aligned}\Phi_{(X,Y)}(t,s) &= \mathbb{E}\left(e^{i\langle(t,s),(X,Y)\rangle}\right) \\ &= \mathbb{E}\left(e^{i(tX+sY)}\right).\end{aligned}$$

Proposition 117 *Deux variables aléatoires sont indépendantes si et seulement si*

$$\Phi_{(X,Y)}(t,s) = \Phi_X(t)\,\Phi_Y(s)\,.$$

Démonstration. On a

$$\begin{aligned}X \text{ et } Y \text{ indép} &\Leftrightarrow \mathbb{P}_{(X,Y)} = \mathbb{P}_X\mathbb{P}_Y \\ &\Leftrightarrow \int_{\mathbb{R}^2} e^{itx+isy}d\mathbb{P}_{(X,Y)} = \int_{\mathbb{R}^2} e^{itx+isy}d\mathbb{P}_X\mathbb{P}_Y \\ &\Leftrightarrow \Phi_{(X,Y)}(t,s) = \Phi_X(t)\,\Phi_Y(s)\,.\end{aligned}$$

■

Théorème 118 ***(Paul Lévy)*** *La convergence en loi des variables aléatoires est équivalente à la convergence simple de leur fonction caractéristique.*

Démonstration. $\Rightarrow$ On sait que

$$\mathbb{P}_{X_n} \to \mathbb{P}_X \text{ pour } n \to \infty$$

signifie

$$\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X)) \text{ pour } \forall f \in C_b(\mathbb{R}) \text{ quand } n \to \infty.$$

On choisi $f(x) = e^{itx}$ et on obtient

$$\mathbb{E}\left(e^{itX_n}\right) \to \mathbb{E}\left(e^{itX}\right) \text{ quand } n \to \infty$$

et donc

$$\Phi_{X_n}(t) \to \Phi_X(t)\,.$$

■

9.2 Variables et vecteurs gaussiens

Définition 119 *Une variable aléatoire X suit la loi normale standard $N(0,1)$ si elle admet la densité*

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Plus générale, une variable aléatoire X suit la loi normale $N\left(m, \sigma^2\right)$ si elle admet la densité

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t-m)^2/2\sigma^2}.$$

Proposition 120 *Si $X_0 \hookrightarrow N(0,1)$ alors $X = m + \sigma X_0 \hookrightarrow N\left(m, \sigma^2\right)$ et si $X \hookrightarrow N\left(m, \sigma^2\right)$ alors $X_0 = \dfrac{X-m}{\sigma} \hookrightarrow N(0,1)$.*

Proposition 121 *Une variable aléatoire X de loi $N\left(m, \sigma^2\right)$ a l'espérance $\mathbb{E}(X) = m$, la variance $V(X) = \sigma^2$ et la fonction caractéristique $\Phi_X(t) = \exp\left(imt - \sigma^2 t^2/2\right)$.*

Par contre, une variable aléatoire qui a l'espérance $\mathbb{E}(X) = m$ et la variance $V(X) = \sigma^2$ n'est pas forcement de loi normale.

Si une variable aléatoire a la fonction caractéristique $\Phi_X(t) = \exp\left(imt - \sigma^2 t^2/2\right)$ alors elle suit une loi normale $N\left(m, \sigma^2\right)$.

Proposition 122 *Soient $X_1 \hookrightarrow N\left(m_1, \sigma_1^2\right)$ et $X_2 \hookrightarrow N\left(m_2, \sigma_2^2\right)$ indépendantes. Alors $X_1 + X_2 \hookrightarrow N\left(m_1 + m_2, \sigma_1^2 + \sigma_2^2\right)$.*

Définition 123 *Un vecteur aléatoire $X = (X_1, X_2, ..., X_n)$ est gaussien si et seulement si toutes les combinaisons linéaires de ces coordonnées $\langle a, X\rangle = a_1 X_1 + ... + a_n X_n$ suit une loi gaussienne dans $\mathbb{R}$, pour tout $a \in \mathbb{R}^n$.*

Remarque 124 *Si $X = (X_1, X_2, ..., X_n)$ est un vecteur aléatoire gaussien alors $\langle a, X\rangle$ suit une loi normale de paramètres*

$$\begin{aligned}
\mathbb{E}(\langle a, X\rangle) &= \mathbb{E}(a_1 X_1 + ... + a_n X_n) \\
&= \langle a, \mathbb{E}(X)\rangle
\end{aligned}$$

$$
\begin{aligned}
V(\langle a, X\rangle) &= V(a_1X_1 + ... + a_nX_n) \\
&= \sum_{i,j=1}^{n} a_i a_j Cov(X_i, X_j) \\
&= a^t Cov(X_i, X_j)\, a.
\end{aligned}
$$

Donc $\langle a, X\rangle$ *suit une loi* $N\left(\langle a, \mathbb{E}(X)\rangle, a^t Cov(X_i, X_j)\, a\right)$ *et sa fonction caractéristique est*

$$\Phi_{\langle a,X\rangle}(x) = \exp\left(ix\,\langle a, \mathbb{E}(X)\rangle - \frac{1}{2}\left(a^t Cov(X_i, X_j)\, a\right) x^2\right).$$

Proposition 125 *La fonction caractéristique d'un vecteur gaussien* $X = (X_1, X_2, ..., X_n)$ *est donné par*

$$\Phi_X(x) = \exp\left(i\,\langle x, \mathbb{E}(X)\rangle - \frac{1}{2}\left(x^t Cov(X_i, X_j)\, x\right)\right).$$

Proposition 126 *Soient* (X, Y) *un couple gaussien. Alors* X *et* Y *sont indépendantes si et seulement si* $Cov(X, Y) = 0$.

Le sense direct est vrai quelque soit la loi de X et Y. La réciproque non.

9.3 Théorème limite centrale

Théorème 127 *Soit* $(X_n)_n$ *une suite de variables aléatoires indépendantes et de même loi, d'espérance* m *et de variance* σ^2. *Soit* $S_n = X_1 + ... + X_n$ *la somme partielle. Alors quand* $n \to \infty$

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} N(0, 1).$$

Démonstration. Soit $Y_i = X_i - m$.

Alors les variables Y_i sont indépendantes de même loi avec

$$\mathbb{E}(Y_i) = 0 \text{ et } V(Y_i) = V(X_i).$$

Notons

$$S_n' = Y_1 + ... + Y_n = S_n - nm$$

et

$$Z_n = \frac{S_n - nm}{\sigma\sqrt{n}} = \frac{S_n'}{\sigma\sqrt{n}}.$$

On a

$$\begin{aligned}
\Phi_{Z_n}(t) &= \mathbb{E}\left(e^{itZ_n}\right) \\
&= \mathbb{E}\left(e^{it\frac{S'_n}{\sigma\sqrt{n}}}\right) \\
&= \Phi_{S'_n}\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&= \Phi_{Y_1}\left(\frac{t}{\sigma\sqrt{n}}\right)\Phi_{Y_2}\left(\frac{t}{\sigma\sqrt{n}}\right)...\Phi_{Y_n}\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&= \left(\Phi_{Y_1}\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n
\end{aligned}$$

et

$$\Phi_{Y_1}(x) = \Phi_{Y_1}(0) + x\Phi'_{Y_1}(0) + \frac{x^2}{2}\Phi''_{Y_1}(0) + x^2\varepsilon(x).$$

On calcule

$$\begin{aligned}
\Phi_{Y_1}(0) &= 1 \\
\Phi'_{Y_1}(0) &= i\mathbb{E}(Y_1) = 0 \\
\Phi''_{Y_1}(0) &= (i)^2\,\mathbb{E}\left(Y_1^2\right) = -\sigma^2.
\end{aligned}$$

Donc

$$\begin{aligned}
\Phi_{Z_n}(t) &= \left(1 - \frac{\sigma^2}{2}\frac{t^2}{n\sigma^2} + \frac{t^2}{n\sigma^2}\varepsilon\left(\frac{t}{\sqrt{n}\sigma}\right)\right)^n \\
&= \left(1 - \frac{t^2}{2n} + \frac{1}{n}\varepsilon\left(\frac{1}{\sqrt{n}}\right)\right)^n \\
&= e^{\ln\left(1-\frac{t^2}{2n}+\frac{1}{n}\varepsilon\left(\frac{1}{\sqrt{n}}\right)\right)^n} \\
&= e^{n\ \ln\left(1-\frac{t^2}{2n}+\frac{1}{n}\varepsilon\left(\frac{1}{\sqrt{n}}\right)\right)} \\
&\sim e^{n\ \left(1-\frac{t^2}{2n}+\frac{1}{n}\varepsilon\left(\frac{1}{\sqrt{n}}\right)\right)} \\
&= e^{-\frac{t^2}{2}+\varepsilon\left(\frac{1}{\sqrt{n}}\right)}.
\end{aligned}$$

On a

$$\Phi_{Z_n}(t) \simeq e^{-\frac{t^2}{2}+\varepsilon\left(\frac{1}{\sqrt{n}}\right)} \to e^{-\frac{t^2}{2}}.$$

Donc $Z_n \to N(0,1)$ converge p.s. ■

Exemple 128 *On lance une pièce et lorsqu'il obtient pile on gagne 100 euro et lorsqu'il obtient face, il perd 100 euro.*

Il faut estimer le nombre maximal de lancers à effectuer pour avoir plus de 95% de ne pas perdre plus de 2000 euro.

On note par n le nombre de lancers effectués.

La variable aléatoire S_n qui est égale au nombre de piles obtenus sur les n premiers lancers, suit une loi $B(n, 1/2)$ et le gain vaut :

$$\begin{aligned} G_n &= 100 \times S_n - 100 \times (n - S_n) \\ &= 200 S_n - 100n. \end{aligned}$$

On cherche alors n tel que

$$\mathbb{P}(G_n \geq -2000) \geq 0,95.$$

Or

$$\{G_n \geq -2000\} = \{S_n - n/2 \geq -10\}.$$

Comme S_n qui suit une loi binomiale peut être vue comme une somme

$$S_n = \epsilon_1 + ... + \epsilon_n$$

de variables aléatoires de loi $b(1/2)$ de Bernoulli.

On calcule

$$\begin{aligned} \mathbb{E}(S_n) &= \frac{n}{2} \\ V(S_n) &= \frac{n}{4}. \end{aligned}$$

D'après le TCL on a

$$S_n^* = \frac{S_n - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \hookrightarrow N(0,1).$$

On calcule

$$\mathbb{P}(G_n \geq -2000) \geq 0,95$$

$$\begin{aligned} &\Leftrightarrow \quad \mathbb{P}\left(\frac{S_n - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \geq \frac{-10}{\frac{\sqrt{n}}{2}}\right) \geq 0,95 \\ &\Leftrightarrow \quad \mathbb{P}\left(N(0,1) < \frac{20}{\sqrt{n}}\right) \geq 0,95. \end{aligned}$$

On obtient $n = 146$.

Chapitre 10

Temps d'arrêt

Définition 129 *Soit* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité. Une suite croissante* $(\mathcal{F}_n)_n$ *de sous-*$\sigma-$*algèbre de* $\mathcal{F}$ *est applé filtration. L'espace* $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathbb{P})$ *est donc un espace filtré.*

Définition 130 *Soit* $(X_n)_{n\in\mathbb{N}}$ *un processus stochastique. On dit que* $(X_n)_{n\in\mathbb{N}}$ *est* $\mathcal{F}_n-$ *adapté si pour tout* n *on a que* X_n *est* $\mathcal{F}_n-$ *mesurable.*

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathbb{P})$ un espace de probabilité filtré.

Définition 131 *Une variable aléatoire* $T : \Omega \to \mathbb{N}$ *est appelée temps d'arrêt si pour tout* $n \in \mathbb{N}$*, on a*

$$\begin{aligned} \{T = n\} &\in \mathcal{F}_n \Longleftrightarrow \{T \leq n\} \in \mathcal{F}_n \\ &\Longleftrightarrow \{T \geq n\} \in \mathcal{F}_n. \end{aligned}$$

Exemple 132

1) Le temps constant $T = k$*, pour* $k \in \mathbb{N}$ *est un temps d'arrêt.*

2) Soit $\{X_n\}$ *un processus stochastique discret adapté à la filtration* $\{\mathcal{F}_n, n \geq 0\}$ *.*
Pour $A \in \mathcal{B}(\mathbb{R})$ *on définit*

$$T_A = \begin{cases} \inf\{n \in \mathbb{N} \mid X_n \in A\}, & si\ A \neq \emptyset \\ \infty, \ sinon \end{cases}$$

C'est un temps d'arrêt appelé "temps d'entré dans A*".*
On voit que

$$\{T_A = n\} = \{X_0 \notin A, X_1 \notin A, ..., X_n \in A\} \in \mathcal{F}_n$$

$$X_0^{-1}(A^c) \cap X_1^{-1}(A^c) \cap ... \cap X_n^{-1}(A^c) \in \mathcal{F}_n$$

car $(\mathcal{F}_n)_n$ *est une filtration,* $A \in \mathcal{B}(\mathbb{R})$ *et* $\{X_n\}$ *est adapté.*

Proposition 133

1) Soit S et T deux temps d'arrêt. Alors $S+T$, $S \wedge T$ et $S \vee T$ sont aussi des temps d'arrêt.

2) Soit T deux temps d'arrêt. On définit

$$\mathcal{F}_T = \{A \mid A \cap \{T \leq n\} \in \mathcal{F}_n, \ \forall n \geq 0\}.$$

On montre que $\mathcal{F}_T$ est une tribu ($\sigma-$algèbre).

- $\emptyset, \Omega \in \mathcal{F}_T$
- $A^c \cap \{T \leq n\} = (A \cup \{T > n\})^c = ((A \cap \{T \leq n\}) \cup \{T > n\})^c \in \mathcal{F}_n$ *car*

$$A = (A \cap \{T \leq n\}) \cup (A \cap \{T > n\})$$

et

$$A \cap \{T > n\} \subset \{T > n\}.$$

On obtient donc $A^c \cap \{T \leq n\} \in \mathcal{F}_n \Longrightarrow A^c \in \mathcal{F}_T$.

- A *et* $B \in \mathcal{F}_T$

$$(A \cup B) \cap \{T \leq n\} = \underbrace{(A \cap \{T \leq n\})}_{\in \mathcal{F}_n} \cup \underbrace{(B \cap \{T \leq n\})}_{\in \mathcal{F}_n}$$

(Sur $\mathcal{F}_n$ on sait que c'est une $\sigma-$algèbre et on utilise ses proprietés).

3) Pour $\mathcal{F}_T$ définies avant on a que $S \leq T$ alors $\mathcal{F}_S \subset \mathcal{F}_T$.
On sait que $\{T \leq n\} \subset \{S \leq n\}$ car

$$\begin{aligned} \omega &\in \{T \leq n\} \Longrightarrow T(\omega) \leq n \\ &\Longrightarrow S(\omega) \leq T(\omega) \leq n \\ &\Longrightarrow \omega \in \{S \leq n\}. \end{aligned}$$

Donc

$$A \cap \{T \leq n\} = \underbrace{A \cap \{S \leq n\}}_{\in \mathcal{F}_n} \cap \underbrace{\{T \leq n\}}_{\in \mathcal{F}_n}$$

$$A \cap \{T \leq n\} \in \mathcal{F}_n \Longrightarrow A \in \mathcal{F}_T.$$

Bibliographie

[1] Jean-Christophe Breton "*Cours de probabilités*"

[2] Sébastien Martineau "*Cours de probabilités*"

[3] Bruno Saussereau "*Cours de théorie des probabilités avec exercices corrigés et devoirs*"