Probabilités de base

Table des matières

I Espace de probabilité

Un espace de probabilité est un triplet $(\Omega, \mathcal{F}, \mathbb{P})$ où :

- Ω est un ensemble,

- $\mathcal{F}$ est une tribu (ou σ-algèbre) sur Ω,

- $\mathbb{P}$ est une (mesure de) probabilité sur $(\Omega, \mathcal{F})$.

Définition 1.1.1. Une tribu (ou σ-algèbre) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω (appelés “événements”) tels que

$$\begin{cases} i) & \phi \in \mathcal{F} \\ ii) & A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F} \\ iii) & (A_n)_{n=1}^{\infty} \subset \mathcal{F} \Longrightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F} \end{cases}$$

En particulier : $A, B \in \mathcal{F} \Longrightarrow A \cup B \in \mathcal{F}$.

De même, $A, B \in \mathcal{F} \Longrightarrow A \cap B \in \mathcal{F}$.

Exemple 1.1.2. - Soit $\Omega = \{1, ..., 6\}$. On peut définir plusieurs tribus sur Ω :

$\mathcal{F} = \mathcal{P}(\Omega) = \{\phi, \{1\}, \{2\}, ..., \{1, 2\}, ..., \Omega\}$ = tribu complète (la plus grande),

$\mathcal{F}_0 = \{\phi, \Omega\}$= tribu triviale (la plus petite) (ϕ= évén. impossible, Ω évén. arbitraire),

$\mathcal{F}_1 = \{\phi, \{1\}, \{2, ..., 6\}, \Omega\}$, $\mathcal{F}_2 = \{\phi, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$, etc.

- Soit $\Omega = [0; 1]$ et $I_1, ..., I_n$ une famille d'intervalles formant une partition de Ω. La famille de sous-ensembles définie par

$$\mathcal{G} = \{\phi, I_1, I_2, ..., I_1 \cup I_2, ..., I_1 \cup I_2 \cup I_3, ..., \Omega\}$$

est une tribu sur Ω.

Définition 1.1.3. Soit $\mathcal{A} = \{A_i,\ i \in I\}$ une famille de sous-ensembles de Ω. Alors la tribu engendrée par $\mathcal{A}$ est la plus petite tribu sur Ω qui contient tous les sous-ensembles $A_i,\ i \in I$. Elle est notée $\sigma(\mathcal{A})$.

Exemple 1.1.4. Reprenons l'exemple 1.1.2.

- Soit $\Omega = \{1, ..., 6\}$. Si $\mathcal{A}_1 = \{\{1\}\}$, alors $\sigma(\mathcal{A}_1) = \mathcal{F}_1$. Si $\mathcal{A}_2 = \{\{1, 3, 5\}\}$,alors $\sigma(\mathcal{A}_2) = \mathcal{F}_2$.

- Soit $\Omega = [0; 1]$. Si $\mathcal{A} = \{I_1, ..., I_n\}$, alors $\sigma(\mathcal{A}) = \mathcal{G}$.

Définition 1.1.5. Soit $\Omega = [0; 1]$. La tribu borélienne sur $[0, 1]$ est la tribu engendrée par la famille de sous-ensembles $\mathcal{A} = \{]a; b[: 0 \leq a < b \leq 1\}$ ={intervalles ouverts dans$[0, 1]$}. Elle est notée $\mathcal{B}([0, 1])$. Elle contient un très grand nombre de sous-ensembles de $[0, 1]$, mais pas tous.

Remarque 1.1.6. - Pour Ω ensemble fini, on choisit souvent $\mathcal{F} = \mathcal{P}(\Omega)$.

- Pour $\Omega \subset \mathbb{R}$ ou $\Omega \subset \mathbb{R}^n$, on choisit souvent $\mathcal{F} = \mathcal{B}(\Omega)$.

Définition 1.1.7. Une sous-tribu de $\mathcal{F}$ est une tribu $\mathcal{G}$ telle que si $A \in G$ alors $A \in \mathcal{F}$.

On note $\mathcal{G} \subset \mathcal{F}$.

Exemple 1.1.8. Reprenons l'exemple 1.1.2. On a $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}$, $\mathcal{F}_0 \subset \mathcal{F}_2 \subset \mathcal{F}$, mais pas $\mathcal{F}_1 \subset \mathcal{F}_2$, ni $\mathcal{F}_2 \subset \mathcal{F}_1$.

Remarque importante :

Il est toujours vrai que $A \in \mathcal{G}$ et $\mathcal{G} \subset \mathcal{F} \Longrightarrow A \in \mathcal{F}$.

Mais il est faux de dire que $A \subset B$ et $B \in \mathcal{F} \Longrightarrow A \in \mathcal{F}$. Contre-exemple :

$\{1\} \subset \{1,3,5\}$, $\{1,3,5\} \in \mathcal{F}_2$, mais $\{1\} \notin \mathcal{F}_2$.

Définition 1.1.9. Soit $\mathcal{F}$ une tribu sur Ω. Une (mesure de) probabilité sur $(\Omega, \mathcal{F})$ est une application $\mathbb{P} : \mathcal{F} \to [0;1]$ telle que

$$\begin{cases} i) & \mathbb{P}(\phi) = 0 \text{ et } P(\Omega) = 1, \\ ii) & (A_n)_{n=1}^{\infty} \subset \mathcal{F} \text{ disjoints (i.e. } A_n \cap A_m = \phi,\ \forall n \neq m) \Longrightarrow \mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n). \end{cases}$$

En particulier : $A, B \in \mathcal{F}$ et $A \cap B = \phi \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. De plus :

$$\begin{cases} i) \ Si \ (A_n)_{n=1}^{\infty} \subset \mathcal{F}, A_n \subset A_{n+1} et \ \cup_{n=1}^{\infty} A_n = A, \ alors \ \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A), \\ ii) \ Si \ (A_n)_{n=1}^{\infty} \subset \mathcal{F}, A_n \supset A_{n+1} et \ \cap_{n=1}^{\infty} A_n = A, \ alors \ \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A). \end{cases}$$

Pour d'autres propriétés, cf. exercices.

Exemple 1.1.10. Soient $\Omega = \{1, ..., 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$. On définit :

- $\mathbb{P}_1(\{i\}) = \frac{1}{6}\ \forall i$ (mesure de probabilité associée à un dé équilibré).

Dans ce cas, on voit p. ex. que $\mathbb{P}_1(\{1,3,5\}) = \mathbb{P}_1(\{1\}) + \mathbb{P}_1(\{3\}) + \mathbb{P}_1(\{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.

- $\mathbb{P}_2(\{i\}) = 0\ \forall i \leq 5,\ \mathbb{P}_2(\{6\}) = 1$ (mesure de probabilité associée à un dé pipé).

Définition 1.1.11. Soient $\Omega = [0,1]$ et $\mathcal{F} = \mathcal{B}([0,1])$. On appelle mesure de Lebesgue sur $[0,1]$ la mesure de probabilité définie par

$$\mathbb{P}(]a;b[) = b - a, \qquad \forall 0 \leq a < b \leq 1.$$

$\mathbb{P}$ n'est définie a priori que sur les intervalles, mais est uniquement extensible à tout ensemble borélien $B \in \mathcal{B}([0,1])$. Elle est notée $\mathbb{P}(B) = |B|$, $B \in \mathcal{B}([0,1])$.

En utilisant la propriété (ii) ci-dessus, on déduit que pour tout $x \in [0,1]$:

$$|\{x\}| = \lim_{n\to\infty} \left| \left] x - \frac{1}{n}, x + \frac{1}{n} \right[\right| = \lim_{n\to\infty} \frac{2}{n} = 0.$$

Généralisation à n dimensions : Soit $\Omega = [0,1]^n$.

- Tribu borélienne : $\mathcal{B}(\Omega) = \sigma(\mathcal{A})$, où $\mathcal{A} = \{]a_1, b_1[\times]a_2, b_2[\times ... \times]a_n, b_n[,\ 0 \leq a_i < b_i \leq 1\}$.

$\mathcal{A}$ est la famille des "rectangles" dans Ω.

- Mesure de Lebesgue : $\mathbb{P}(]a_1, b_1[\times]a_2, b_2[\times ... \times]a_n, b_n[) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$.

Comme dans le cas uni-dimensionnel, $\mathbb{P}$ n'est définie a priori que sur certains ensembles (les rectangles), mais est uniquement extensible à tout $B \in \mathcal{B}(\Omega)$ (p.ex. B = disque, ovale...).

Notation : $\mathbb{P}(B) = |B|$, pour $B \in \mathcal{B}(\Omega)$.

II Variable aléatoire

Définition 1.2.1. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une variable aléatoire (souvent abrégé v.a. par la suite) est une application $X : \Omega \to \mathbb{R}$ telle que

$$\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Proposition 1.2.2. X est une v.a. ssi $\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{F}$, $\forall t \in \mathbb{R}$.

Remarque 1.2.3. - X est aussi dite une fonction (ou v.a.) $\mathcal{F}$-mesurable.

- Si $\mathcal{F} = \mathcal{P}(\Omega)$, alors X est toujours $\mathcal{F}$-mesurable.

- Si $\Omega = \mathbb{R}$ et $\mathcal{F} = \mathcal{B}(\mathbb{R})$, alors X est dite une fonction borélienne.

Définition 1.2.4. Pour $A \subset \Omega$, on pose $1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$

On vérifie que la v.a. 1_A est $\mathcal{F}$-mesurable ssi $A \in \mathcal{F}$.

Exemple 1.2.5. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ l'espace de probabilité du dé équilibré (cf. exemple 1.1.10).

$X_1(\omega) = \omega : \ \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = i\}) = \mathbb{P}(\{i\}) = \frac{1}{6}$.

$X_2(\omega) = 1_{\{1,3,5\}}(\omega) : \ \mathbb{P}(\{\omega \in \Omega : X_2(\omega) = 1\}) = \mathbb{P}(\{1,3,5\}) = \frac{1}{2}$.

Soit $\mathcal{F} = \mathcal{P}(\Omega)$. X_1 et X_2 sont toutes deux $\mathcal{F}$-mesurables.

Soit $\mathcal{F}_2 = \{\phi, \{1,3,5\}, \{2,4,6\}, \Omega\}$. Seule X_2 est $\mathcal{F}_2$-mesurable ; X_1 ne l'est pas. En effet :

$\{\omega \in \Omega : X_2(\omega) = 1\} = \{1,3,5\} \in \mathcal{F}_2$ et $\{\omega \in \Omega : X_2(\omega) = 0\} = \{2,4,6\} \in \mathcal{F}_2$

tandis que $\{\omega \in \Omega : X_1(\omega) = 1\} = \{1\} \notin \mathcal{F}_2$.

Définition 1.2.6. La tribu engendrée par une famille de v.a. $\{X_i, \ i \in I\}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est définie par

$$\sigma(X_i, \ i \in I) = \sigma(\{X_i \in B\}, \ i \in I, \ B \in \mathcal{B}(\mathbb{R})) = \sigma(\{X_i \leq t\}, \ i \in I, \ t \in \mathbb{R}).$$

Exemple 1.2.7. Reprenons l'exemple précédent : $\sigma(X_1) = \mathcal{F} = \mathcal{P}(\Omega), \ \sigma(X_2) = \mathcal{F}_2 = \mathcal{P}(\Omega)$.

Proposition 1.2.8. Si $g : \mathbb{R} \to \mathbb{R}$ est borélienne et $X : \Omega \to \mathbb{R}$ est une v.a., alors $g(X)$ est une v.a.

Démonstration. Soit $B \in \mathcal{B}(\mathbb{R})$. On a

$$\{\omega \in \Omega : g(X(\omega)) \in B\} = \left\{\omega \in \Omega : X(\omega) \in g^{-1}(B)\right\}.$$

Or $g^{-1}(B) = \{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$, car g est borélienne. Comme X est une v.a., on en déduit que $\{\omega \in \Omega : X(\omega) \in g^{-1}(B)\} \in \mathcal{F}$, et donc finalement que $g(X)$ est une v.a.

Proposition 1.2.9. Toute fonction continue est borélienne (et pratiquement toute fonction discontinue l'est aussi !).

III Loi d'une variable aléatoire

Définition 1.3.1. La loi d'une v.a. X est l'application $\mu_X : \mathcal{B}(\mathbb{R}) \to [0,1]$ définie par

$$\mu_X(B) = \mathbb{P}(\{X \in B\}), \qquad B \in \mathcal{B}(\mathbb{R}).$$

NB : $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ forme un nouvel espace de probabilité !

Exemple 1.3.2.-. Soit $\Omega = \{1, ..., 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{i\}) = \frac{1}{6}$, $\forall i$.

$X_1(\omega) = \omega$, $\mu_X(\{i\}) = \mathbb{P}(\{X = i\}) = \mathbb{P}(\{i\}) = \frac{1}{6}$.

-.Soit $\Omega = \{1, ..., 6\} \times \{1, ..., 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{(i,j)\}) = \frac{1}{36}$, $\forall i, j \in \Omega$.

$X(\omega) = X(\omega_1, \omega_2) = \omega_1 + \omega_2$. On a alors, p.ex :

$\mu_X(\{7\}) = \mathbb{P}(\{X = 7\}) = P(\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}) = 6 \times \frac{1}{36} = \frac{1}{6}$.

Fonction de répartition d'une variable aléatoire

Définition 1.3.3. La fonction de répartition d'une v.a. X est l'application $F_X : \mathbb{R} \to [0;1]$ définie par

$$F_X(t) = \mathbb{P}(\{X \le t\}) = \mu_X(]-\infty, t]), \quad t \in \mathbb{R}.$$

Proposition 1.3.4. La donnée de F_X équivaut à celle de μ_X.

Cette dernière proposition est à rapprocher de la proposition 1.2.2.

Deux types particuliers de variables aléatoires

A) Variable aléatoire discrète :

X prend ses valeurs dans un ensemble D dénombrable ($X(\omega) \in D; \forall \omega \in \Omega$). Dans ce cas, on a : $\sigma(X) = \sigma(\{X = x\}, x \in D)$, $p(x) = \mathbb{P}(\{X = x\}) \ge 0$ et $\sum_{x \in D} p(x) = \mathbb{P}(\{x \in D\}) = 1$.

De plus, $F_X(t) = \sum_{x \in D : x \le t} p(x)$.

B) Variable aléatoire continue :

$\mathbb{P}(X \in B) = 0$ si $|B| = 0$ (en part. $\mathbb{P}(X = x) = 0 \quad \forall x$). Sous cette condition, le théorème de Radon-Nikodym assure l'existence d'une fonction borélienne $f_X : \mathbb{R} \to \mathbb{R}$ (appelée densité) telle que

$$f_X(x) \ge 0, \ \forall x \in \mathbb{R}, \quad \int_{\mathbb{R}} f_X(x)dx = 1 \quad \text{et} \quad \mathbb{P}(\{X \in B\}) = \int_B f_X(x)dx.$$

De plus, $F_X(t) = \int_{-\infty}^t f_X(x)dx$ et $F'_X(t) = f_X(t)$.

Exemple 1.3.5.

A) Loi binomiale $\mathcal{B}(n; p)$, $n \ge 1$, $p \in [0,1]$:

$$p(k) = \mathbb{P}\{(X = k\}) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \text{pour } 0 \le k \le n,$$

où $\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!}$.

B) Loi gaussienne $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$:

$$\text{densité :} \qquad f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}.$$

Terminologie : - Si X suit une loi gaussienne (p.ex.), on écrit $X \sim \mathcal{N}(\mu, \sigma^2)$.

- Si X et Y suivent une même loi, on dit que X et Y sont identiquement distribuées (i.d.) et on note $X \sim Y$.

IV Espérance d'une variable aléatoire

Construction de l'espérance (= intégrale de Lebesgue !)

On procède en trois étapes :

1. Soit $X(\omega) = \sum_{i=0}^{\infty} x_i 1_{A_i}(\omega)$, $x_i \geq 0$, $A_i \in \mathcal{F}$. On définit l'espérance de telles v.a. (dites simples) comme suit :

$$\mathbb{E}(X) = \sum_{i=0}^{\infty} x_i \mathbb{P}(A_i) \in [0,\ +\infty].$$

Attention ! $\mathbb{E}(X)$ peut prendre la "valeur" $+\infty$.

Exemples : - Si $X = 1_A$, $\mathbb{P}(A) = p$, alors $\mathbb{E}(X) = \mathbb{P}(A) = p$.

- Si $X = c1_\Omega = cte$ sur Ω, alors $\mathbb{E}(X) = c$.

2. Soit X une v.a. $\mathcal{F}$-mesurable telle que $X(\omega) \geq 0$,$\forall \omega \in \Omega$. On pose

$$X_n(\omega) = \sum_{i=1}^{\infty} \frac{i}{2^n} 1_{\left\{\frac{i}{2^n} \leq X < \frac{i+1}{2^n}\right\}}(\omega).$$

Alors (X_n) est une suite croissante de v.a. qui tend vers X. On définit

$$\mathbb{E}(X) = \lim_{n\to\infty} \mathbb{E}(X_n) = \lim_{n\to\infty} \sum_{i=1}^{\infty} \frac{i}{2^n} \mathbb{P}\left(\left\{\frac{i}{2^n} \leq X < \frac{i+1}{2^n}\right\}\right) \in [0,\ +\infty].$$

3. Soit X une v.a. $\mathcal{F}$-mesurable quelconque. On pose

$$X(\omega) = X^+(\omega) - X^-(\omega) \text{ avec } \begin{cases} X^+(\omega) = \max(0, X(\omega)) \geq 0, \\ X^-(\omega) = \max(0, -X(\omega)) \geq 0. \end{cases}$$

On a alors $|X(\omega)| = X^+(\omega) + X^-(\omega) \geq 0$.

- Si $\mathbb{E}(|X|) < \infty$, alors on définit $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$.

- Si $\mathbb{E}(|X|) = \infty$, alors on dit que $\mathbb{E}(X)$ n'est pas définie.

Terminologie : - Si $\mathbb{E}(X) = 0$, alors on dit que X est une v.a. centrée.

- Si $\mathbb{E}(|X|) < \infty$, alors on dit que X est une v.a. intégrable.

- Si $\mathbb{E}(X^2) < \infty$ alors on dit que X est une v.a. de carré intégrable.

- On dit que X est une v.a. bornée s'il existe une cte $K > 0$ telle que $|X(\omega)| \leq K, \forall \omega \in \Omega$.

Remarque 1.4.1. X bornée $\Longrightarrow \mathbb{E}(X^2) < \infty \Longrightarrow \mathbb{E}(|X|) < \infty$.

Proposition 1.4.2. Soient X une v.a. et $g : \mathbb{R} \to \mathbb{R}$ une fonction borélienne telle que $\mathbb{E}(|g(X)|) < \infty$. Alors

A) Si X est une v.a. discréte (à valeurs dans D dénombrable), alors

$$\mathbb{E}(g(X)) = \sum_{x \in D} g(x) \mathbb{P}(\{X = x\}).$$

B) Si X est une v.a. continue (avec densité f_X), alors

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) f_X(x) dx.$$

Ceci s'applique en particulier si $g(x) = x$.

Variance et covariance de variables aléatoires

Définition 1.4.3. Soient X, Y deux v.a. de carré intégrable. On pose

$$\begin{aligned} Var(X) &= \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 \geq 0 \\ Cov(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

Terminologie : - Un événement $A \in \mathcal{F}$ est dit négligeable si $\mathbb{P}(A) = 0$.

- Un événement $A \in \mathcal{F}$ est dit presque sûr (souvent abréegé p.s.) si $\mathbb{P}(A) = 1$, i.e. si A^c est négligeable.

Exemple 1.4.4. Soit X une v.a. telle que $P(\{X = c\}) = 1$. Alors on dit que $X = c$ presque sûrement ("$X = c$ p.s.")

Proposition 1.4.5. Si $(A_n)_{n=1}^{\infty} \subset \mathcal{F}$ est une famille d'événements négligeables (i :e : $\mathbb{P}(A_n) = 0 \ \forall n$), alors $\cup_{n=1}^{\infty} A_n$ est négligeable.

Démonstration. $\mathbb{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n) = 0$.

Exemple 1.4.6. L'ensemble $A = [0; 1] \cap \mathbb{Q}$ est négligeable pour la mesure le Lebesgue, car $\mathbb{Q}$ est dénombrable et $|\{x\}| = 0$ pour tout $x \in [0, 1]$.

Propriétés de l'espérance

Soient X, Y deux v.a. intégrables.

- Linéarité : $\mathbb{E}(cX + Y) = c\mathbb{E}(X) + \mathbb{E}(Y),\ c \in \mathbb{R}$ et X, Y v.a. intégrables.

- Positivité : si $X \geq 0$ p.s., alors $\mathbb{E}(X) \geq 0$.

- Positivité stricte : si $X \geq 0$ p.s. et $\mathbb{E}(X) = 0$, alors $X = 0$ p.s.

- Monotonie : si $X \geq Y$ p.s., alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$.

Inégalité de Cauchy-Schwarz

Soient X, Y deux v.a. de carré intégrable. Alors

$$\begin{cases} i) & XY \text{ est intégrable}\,; \\ ii) & \left(\mathbb{E}\left(|XY|\right)\right)^2 \leq \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right). \end{cases}$$

En posant $Y \equiv 1$, on trouve que $\left(\mathbb{E}\left(|X|\right)\right)^2 \leq \mathbb{E}\left(X^2\right)$ (donc $\mathbb{E}\left(|X|\right) < \infty$ si $\mathbb{E}(X^2) < \infty$; cf. remarque 1.4.1).

Inégalité triangulaire

Soient X, Y deux v.a. intégrables. Alors

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

Inégalité de Jensen

Soient X une v.a. et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction borélienne et convexe telle que $\mathbb{E}\left(|\varphi\left(X\right)|\right) < \infty$.

Alors

$$\varphi\left(\mathbb{E}\left(X\right)\right) \leq \mathbb{E}\left(\varphi\left(X\right)\right).$$

En particulier, $|\mathbb{E}\left(X\right)| \leq \mathbb{E}\left(|X|\right)$.

Démonstration. Vu que φ est convexe, on a $\varphi\left(x\right) = \sup\limits_{a,b : ay+b \leq \varphi(y), \forall y \in \mathbb{R}} \left(ax+b\right)$ et donc :

$$\varphi\left(\mathbb{E}\left(X\right)\right) = \sup_{a,b:\ldots}\left(a\mathbb{E}\left(x\right)+b\right) = \sup_{a,b:\ldots}\left(\mathbb{E}\left(aX+b\right)\right) \leq \sup_{a,b:\ldots}\mathbb{E}\left(\varphi\left(X\right)\right) = \mathbb{E}\left(\varphi\left(X\right)\right).$$

Exemple 1.4.7. Si $X = a$ ou b avec prob.$\frac{1}{2}, \frac{1}{2}$ et φ est convexe, alors $\varphi\left(\mathbb{E}\left(X\right)\right) = \varphi\left(\frac{a+b}{2}\right) \leq \frac{\varphi(a)+\varphi(b)}{2} = \mathbb{E}\left(\varphi\left(X\right)\right)$.

Inégalité de Chebychev (ou Markov)

Soient X une v.a. et $\psi : \mathbb{R} \to \mathbb{R}_+$ telle que ψ est borélienne et croissante sur $\mathbb{R}_+$, $\psi\left(a\right) > 0$ pour tout $a > 0$ et $\mathbb{E}\left(\psi\left(X\right)\right) < \infty$. Alors

$$\mathbb{P}\left(\{X \geq a\}\right) \leq \frac{\mathbb{E}\left(\psi\left(X\right)\right)}{\psi\left(a\right)}, \quad \forall a > 0.$$

Démonstration. Du fait que ψ est croissante sur $\mathbb{R}_+$, on a

$$\mathbb{E}\left(\psi\left(X\right)\right) \geq \mathbb{E}\left(\psi\left(X\right)1_{\{X \geq a\}}\right) \geq \mathbb{E}\left(\psi\left(a\right)1_{\{X \geq a\}}\right) = \psi\left(a\right)\mathbb{E}\left(1_{\{X \geq a\}}\right) = \psi\left(a\right)\mathbb{P}\left(\{X \geq a\}\right).$$

Comme $\psi\left(a\right) > 0$, ceci permet de conclure.

V Indépendance

V.1 Indépendance d'événements

Définition 1.5.1. Deux événements A et $B(\in \mathcal{F})$ sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Attention ! Ne pas confondre : A et B sont <u>disjoints</u> si $A \cap B = \phi(\Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B))$.

<u>Notation :</u> Si A est indépendant de B, on note $A \perp B$ (de même pour les tribus et les v.a. ; voir plus bas).

<u>Conséquence :</u>

$$\begin{aligned} \mathbb{P}(A^c \cap B) &= \mathbb{P}(B \backslash (A \cap B)) = \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = (1 - \mathbb{P}(A))\,\mathbb{P}(B) = \mathbb{P}(A^c)\mathbb{P}(B). \end{aligned}$$

De même, on a $\mathbb{P}(A \cap B^c) = \mathbb{P}(A)\mathbb{P}(B^c)$ et $\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c)$.

Définition 1.5.2. n événements $A_1, ..., A_n \in \mathcal{F}$ sont indépendants si

$$\mathbb{P}\left(A_1^* \cap ... \cap A_n^*\right) =_{i=1}^{n} \mathbb{P}\left(A_i^*\right), \quad \text{où } A_i^* = \text{soit } A_i, \quad \text{soit } A_i^c.$$

Proposition 1.5.3. n événements $A_1, ..., A_n \in \mathcal{F}$ sont indépendants si

$$\mathbb{P}\left({}_{i \in I} A_i\right) =_{i \in I} \mathbb{P}\left(A_i\right), \quad \forall I \subset \{1, ..., n\}.$$

Remarque 1.5.4. - Pour $n > 2$, la condition $\mathbb{P}(A_1 \cap ... A_n) = \mathbb{P}(A_1)...\mathbb{P}(A_n)$ <u>ne suffit pas !</u>

- L'indépendance de n événements telle que définie ci-dessus est <u>plus forte</u> que l'indépendance deux à deux $(A_i \perp A_j,\ \forall i \neq j)$.

V.2 Indépendance de tribus

Définition 1.5.5. Une famille $(\mathcal{F}_1 ... \mathcal{F}_n)$ de sous-tribus de $\mathcal{F}$ est indépendante si

$$\mathbb{P}(A_1 \cap ... \cap A_n) = \mathbb{P}(A_1)...\mathbb{P}(A_n),\ \forall A_1 \in \mathcal{F}_1, ..., A_n \in \mathcal{F}_n.$$

Proposition 1.5.6. $(\sigma(A_1), ..., \sigma(A_n))$ est une famille de sous-tribus indépendantes ssi les événements $(A_1, ..., A_n)$ sont indépendants.

V.3 Indépendance de variables aléatoires

Définition 1.5.7. Une famille $(X_1, ..., X_n)$ de v.a. ($\mathcal{F}$-mesurables) est indépendante si

$(\sigma(X_1), ..., \sigma(X_n))$ est indépendante.

Proposition 1.5.8. $(X_1, ..., X_n)$ est une famille de v.a. indépendantes

ssi les événements $\{X_1 \leq t_1\}, ..., \{X_n \leq t_n\}$ sont indépendants $\forall\ t_1, ..., t_n \in \mathbb{R}$.

En particulier : $X \perp Y$ ssi $\mathbb{P}(\{X \leq t; Y \leq s\}) = \mathbb{P}(\{X \leq t\})\mathbb{P}(\{Y \leq s\}), \forall\ t; s \in \mathbb{R}$.

Exemple 1.5.9. Soit $\Omega = \{1, ..., 6\} \times \{1, ..., 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{(i, j)\}) = \frac{1}{36}$. On pose

$$X(\omega) = X(\omega_1, \omega_2) = \omega_1, \quad X(\omega) = Y(\omega_1, \omega_2) = \omega_2.$$

Calculons $\mathbb{P}(\{X = i\}) = \mathbb{P}(\omega \in \Omega : \omega_1 = i) = \mathbb{P}(\{(i, 1), ..., (i, 6)\}) = \frac{1}{6} = \mathbb{P}(\{Y = j\})$.

D'autre part, $\mathbb{P}(\{X = i, Y = j\}) = \mathbb{P}(\{(i; j)\}) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = P(\{X = i\})\, P(\{Y = j\})$, donc X et Y sont indépendantes.

Exemple 1.5.10. Si $X(\omega) = c$, $\forall \omega \in \Omega$, alors $X \perp Y$, $\forall Y$ (une v.a. constante est indépendante de toute autre v.a.).

Exemple 1.5.11. Si $X \perp Y$ et g, h sont des fonctions boréliennes, alors $g(X) \perp h(Y)$; c'est vrai en particulier si $g = h$ (mais pas $X = Y$, bien sûr !). Cette propriété découle du fait que $g(X)$ est $\sigma(X)$-mesurable (resp. que $h(Y)$ est $\sigma(Y)$-mesurable) et de la définition d'indépendance pour les tribus.

Proposition 1.5.12. Soient $c \in \mathbb{R}$ et X, Y deux v.a. de carré intégrable. Si $X \perp Y$, alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y) \text{ et } Var(cX + Y) = c^2 Var(X) + Var(Y),$$

La première égalité dit que si deux v.a. sont indépendantes, alors elles sont décorrelées ; la réciproque n'est pas vraie.

Proposition 1.5.13. Si $X \perp Y$ et X, Y sont deux v.a. continues possédant une densité conjointe $f_{X,Y}$ (i.e. $\mathbb{P}((X, Y) \in B) = \int\int_B f_{X,Y}(x, y)\, dxdy,\ \forall B \in \mathcal{B}(\mathbb{R}^2)$), alors $f_{X,Y}(x, y) = f_X(x)\, f_Y(y),\ \forall x, y \in \mathbb{R}$.

VI Espérance conditionnelle

VI.1 Conditionnement par rapport à un événement $B \in \mathcal{F}$

Soit $A \in \mathcal{F}$: $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$, si $\mathbb{P}(B) \neq 0$.

Soit X une v.a. intégrable (i.e. $\mathbb{E}(|X|) < \infty$) : $\mathbb{E}(X \mid B) = \frac{\mathbb{E}(X 1_B)}{\mathbb{P}(B)}$, si $\mathbb{P}(B) \neq 0$.

VI.2 Conditionnement par rapport à une v.a. discrète Y (à valeurs dans D dénombrable)