

Table des matières

1	Variables aléatoires réelles	1
1.1	Rappels fondamentaux sur les variables aléatoires	1
1.2	Caractéristiques numériques d'une variable aléatoire	5
1.2.1	Espérance mathématique d'une variable aléatoire	6
1.2.2	Variance d'une variable aléatoire	8
1.3	Transformation d'une variable aléatoire	9
2	Couples de variables aléatoires	13
2.1	Définition et exemple d'un couples de variables aléatoires	13
2.2	Couples de variables aléatoires discrètes et indépendance	13
2.2.1	Lois conjointes, marginales et conditionnelles.	14
2.2.2	Indépendance de deux variables aléatoires discrètes	17
2.3	Couples de variables aléatoires continues et indépendance	17
2.3.1	Indépendance de deux variables aléatoires continues	20
2.4	Transformation d'un couple de variables aléatoires	21
2.4.1	Transformation d'un couples de variables aléatoires continue	21
2.4.2	Transformation d'un couples de variables aléatoires discrètes	24
2.4.3	Covariance et Corrélation linéaire de deux variables aléatoires	26
2.5	Principales lois de probabilité	28
3	Fonctions caractéristiques et génératrices	29
3.1	Fonctions caractéristiques	29
3.1.1	Fonctions caractéristiques des lois de probabilité usuelles	31
3.1.2	Propriétés des fonctions caractéristiques	32
3.1.3	Fonction caractéristique d'un couple de variables aléatoires	34
3.2	Fonctions génératrices des moments	35
3.2.1	Fonction génératrice d'un couple de variables aléatoires.	37
4	Modes de convergence des suites de variables aléatoires	39
4.1	Modes de convergence usuels	39
4.1.1	Convergence en loi	39
4.1.2	Convergence en probabilité	40
4.1.3	La convergence presque sûre	41
4.1.4	Convergence dans L^p	42
4.1.5	La convergence presque complète	42
4.2	Liens entre les modes de convergence usuels	43

4.2.1	Propriétés des convergence en probabilité, presque sure, dans L^p et presque complète	47
-------	--	----

Chapitre 1

Variables aléatoires réelles

Une variable aléatoire est utilisée pour modéliser des phénomènes aléatoires ou un mécanisme non déterministe. C'est une variable qui peut prendre différentes valeurs avec une probabilité définie pour chacune des occurrences, au contraire d'une variable certaine qui ne prend qu'une seule valeur définie, avec une probabilité égale à 1. La variable aléatoire est une fonction définie sur l'ensemble des résultats possibles d'une expérience aléatoire, telle qu'il soit possible de déterminer la probabilité pour qu'elle prenne une valeur donnée ou qu'elle prenne une valeur dans un intervalle donné.

Dans tout le chapitre, $(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace de probabilité et E est un ensemble muni de la tribu ζ .

1.1 Rappels fondamentaux sur les variables aléatoires

Définition 1.1.1 (la variable aléatoire)

Une variable aléatoire X (v.a. en abrégé) est une application mesurable de $(\Omega, \mathcal{A}, \mathbb{P})$ dans un espace mesurable (E, ζ) , c'est à dire

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (E, \zeta),$$

vérifiant la condition

$$\forall B \in \zeta, X^{-1}(B) \in \mathcal{A}.$$

- X est dit **variable aléatoire réelle** (ou à valeurs réelles) si $(E, \zeta) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
- X est dit **vecteur aléatoire** si $(E, \zeta) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ et on le note $X = (X_1, \dots, X_n)$ où $X_i, i = \overline{1, n} : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une **variable aléatoire** appelée **i .ème marginale** de X . En particulier, X est dit **couple de variables aléatoires** si $(E, \zeta) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.
- X est dit **suite de variables aléatoires** si $(E, \zeta) = (\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$.

L'ensemble des valeurs possibles de la v.a. X est noté $X(\Omega)$. Si cet ensemble est fini ou dénombrable, la v.a. X est dite **discrète** et s'il est nondénombrable, la v.a. X est dite **continue**.

Remarque 1.1.1 Les variables aléatoires sont toujours notées par des lettres majuscules : X, Y, Z, \dots et les valeurs possibles des variables aléatoires sont notées par des lettres minuscules : x, y, z, \dots

Exemple 1.1.1 1. 1- L'application $X_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{si } \omega \notin A \end{cases}$, pour tout $A \in \Omega$ est dite \mathcal{A} -mesurable si $A \in \mathcal{A}$, et dans ce cas X_A est une variables aléatoire.

2. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ l'espace de probabilité du dé équilibré : $\Omega = \{1, 2, 3, 4, 5, 6\}$ et $\mathbb{P}(\{w_i\}) = \frac{1}{6}$ pour tout $i = 1, \dots, 6$.

On considère les applications X_1 et X_2 définies par

$$\begin{aligned} X_1(\omega) &= \omega, \forall \omega \in \Omega. \\ X_2(\omega) &= \mathbf{1}_{\{1,3,5\}}(\omega), \forall \omega \in \Omega. \end{aligned}$$

Alors les ensembles des valeurs possibles sont respectivement $X_1(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $X_2(\Omega) = \{1, 0\}$.

Soient les deux tribus $\mathcal{A}_1 = \mathcal{P}(\Omega)$ et $\mathcal{A}_2 = \{\emptyset, \{1, 2, 3\}, \{2, 4, 6\}, \Omega\}$ et on vérifie si X_1 et X_2 sont des v.a.

- Pour $\mathcal{A}_1 = \mathcal{P}(\Omega)$ on remarque que X_1 et X_2 sont \mathcal{A}_1 -mesurables, donc ce sont des v.a.
- Pour $\mathcal{A}_2 = \{\emptyset, \{1, 2, 3\}, \{2, 4, 6\}, \Omega\}$ on remarque que X_2 est \mathcal{A}_2 -mesurable mais X_1 n'est pas

\mathcal{A}_2 -mesurable.

En effet : on a $X_2(\Omega) = \{0, 1\}$, $\mathcal{P}(X_2(\Omega)) = \{\emptyset, \{1\}, \{0\}, \{0, 1\}\}$

$$X_2^{-1}(\{1\}) = \{\omega \in \Omega, X_2(\omega) = 1\} = \{1, 3, 5\} \in \mathcal{A}_2.$$

$$X_2^{-1}(\{0\}) = \{\omega \in \Omega, X_2(\omega) = 0\} = \{2, 4, 6\} \in \mathcal{A}_2.$$

$$X_2^{-1}(\{0, 1\}) = \{\omega \in \Omega, X_2(\omega) = 0 \vee X_2(\omega) = 1\} = \Omega \in \mathcal{A}_2.$$

$$X_2^{-1}(\emptyset) = \emptyset \in \mathcal{A}_2.$$

$$\text{Tandis que } X_1^{-1}(\{1\}) = \{\omega \in \Omega, X_1(\omega) = 1\} = \{1\} \notin \mathcal{A}_2.$$

Définition 1.1.2 (Tribu engendrée par une variable aléatoire)

La tribu engendrée par la variable aléatoire X , notée $\sigma(X)$, est la plus petite tribu rendant l'application X mesurable. Elle est définie comme suit

$$\sigma(X) = \{X^{-1}(B), B \in \zeta\}.$$

Exemple 1.1.2 La tribu engendrée par la v.a. 1_A , pour tout $A \in \Omega$, est

$$\sigma(1_A) = \{A, A^c, \emptyset, \Omega\}.$$

Définition 1.1.3 (Variable aléatoire simple et élémentaire)

1. On appelle variable aléatoire étagée ou simple, à valeurs dans \mathbb{R} , une variable aléatoire définie sur Ω , de la forme :

$$X(\omega) = \sum_{i=1}^k x_i 1_{A_i}(\omega),$$

où les $(A_i)_{i=1, \dots, k} \in \mathcal{A}$ sont disjoints (les $(A_i)_{i=1, \dots, k}$ forment une partition de Ω) et où $a_i \in \mathbb{R}$.

2. La variable aléatoire X est dite élémentaire si

$$X(\omega) = \sum_{i=1}^{\infty} x_i 1_{A_i}(\omega),$$

où $(x_i)_{i \geq 1}$ sont des nombres réels et les $(A_i)_{i \geq 1}$ forment une partition infinie de Ω .

Proposition 1.1.1 (*Approximation*)

Toute variable aléatoire X est limite simple de variables aléatoires étagées. Si de plus X est réelle positive, la limite peut être choisie croissante.

Preuve. TD ■

Définition 1.1.4 (*Loi de probabilité d'une variable aléatoire*)

1- Soit X une variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ dans (E, ζ) . La loi de probabilité (dite aussi distribution) de X est la probabilité, notée \mathbb{P}_X , définie sur (E, ζ) par :

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega ; X(\omega) \in B\}), B \in \zeta.$$

\mathbb{P}_X est appelée **loi image** de \mathbb{P} par l'application X .

2- Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire, la loi de X , notée $\mathbb{P}_{X_1, \dots, X_n}$, et appelée **loi jointe** des variables aléatoires X_1, \dots, X_n est définie par

$$\mathbb{P}_{X_1, \dots, X_n}(B) = \mathbb{P}(\{\omega \in \Omega ; X_1(\omega) \in B_1; X_2(\omega) \in B_2, \dots, X_n(\omega) \in B_n\}), B \in \mathcal{B}(\mathbb{R}^n).$$

Les lois des variables aléatoires marginales $(X_i)_{i=1, \dots, n}$ s'appellent les **lois marginales**.

Définition 1.1.5 (*Egalité de deux variables aléatoires*)

Soient X et Y deux variables aléatoires, on dit que :

1. X et Y sont égales en loi ou en distribution, et on note $X \stackrel{L}{=} Y$ ssi

$$\forall A \in \mathcal{A}, \mathbb{P}(X \in A) = \mathbb{P}(Y \in A).$$

2. X et Y sont égales presque sûrement, et on note $X \stackrel{p.s.}{=} Y$ ssi

$$\mathbb{P}(\{\omega \in \Omega, X(\omega) = Y(\omega)\}) = 1.$$

Exemple 1.1.3 On lance une pièce de monnaie équilibrée et on définit les v.a. suivantes :

$$X = \begin{cases} 1, & \text{si le résultat est face} \\ 0, & \text{si le résultat est pile} \end{cases}, Y = \begin{cases} 1, & \text{si le résultat est pile} \\ 0, & \text{si le résultat est face} \end{cases}.$$

Les v.a X et Y sont-elles égales en loi ? sont-elles égales presque sûrement ?

Solution 1.1.1 On a $\Omega = \{P, F\}$ et $\mathbb{P}(\{P\}) = \mathbb{P}(\{F\}) = \frac{1}{2}$. On dit que $X \stackrel{L}{=} Y$ ssi la loi de probabilité \mathbb{P}_X , de X , est identique à celle \mathbb{P}_Y , de Y . Il remarque que

$$\mathbb{P}_X(1) = \mathbb{P}(X = 1) = \frac{1}{2}, \mathbb{P}_X(0) = \mathbb{P}(X = 0) = \frac{1}{2},$$

et

$$\mathbb{P}_Y(1) = \mathbb{P}(Y = 1) = \frac{1}{2}, \mathbb{P}_Y(0) = \mathbb{P}(Y = 0) = \frac{1}{2}.$$

c'est à dire X et Y ont la même loi de probabilité, alors $X \stackrel{L}{=} Y$.

Pour l'égalité presque sûre il est clair que X et Y prennent des valeurs différentes pour chaque élément de Ω . Autrement dit

$$\mathbb{P}(\{\omega \in \Omega, X(\omega) = Y(\omega)\}) = 0.$$

Donc X et Y ne sont pas égales presque sûrement.

Remarque 1.1.2 Si deux variables aléatoire sont égales presque sûrement alors elle sont égales en loi, mais la réciproque est fausse.

Définition 1.1.6 (Fonction de répartition)

La fonction de répartition F_X de la variable aléatoire réelle X , est définie de \mathbb{R} dans $[0, 1]$ par

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

Proposition 1.1.2 La fonction de répartition F_X possède les propriétés suivantes :

1. F_X est croissante ;
2. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$;
3. F_X est continue à droite et admet une limite à gauche $\lim_{x \rightarrow x_0, x \leq x_0} F_X(x) = \mathbb{P}(X < x_0)$;
4. Si x n'est pas un atome (valeur possible) de X , alors F_X est continue à gauche (donc continue) en x . L'ensemble des points de discontinuité de F_X est l'ensemble des atomes de X .

Preuve.

1. Soit $x \leq y \in \mathbb{R}$, on a $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\} \subset \{\omega \in \Omega : X(\omega) \leq y\} = \{X \leq y\}$, donc

$$F_X(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq y) = F_X(y).$$

D'où la fonction F_X est croissante.

2. Pour démontrer cette propriété, on prend d'abord l'ensemble $A_n =]-\infty, -n]$.

$$\lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow +\infty} F_X(-n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq -n) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(A_n),$$

et comme $(A_n)_{n \geq 1}$ est une suite décroissant des événements de $\mathcal{B}(\mathbb{R})$, avec $\lim_{n \rightarrow +\infty} A_n = \bigcap_{n \geq 1} A_n = \emptyset$, alors

$$\lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(A_n) = \mathbb{P}_X\left(\bigcap_{n \geq 1} A_n\right) = \mathbb{P}_X(\emptyset) = 0.$$

Puis on prend l'ensemble $B_n =]-\infty, n]$

$$\begin{aligned} \lim_{x \rightarrow +\infty} F_X(x) &= \lim_{n \rightarrow +\infty} F_X(n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq n) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(B_n) \\ &= \mathbb{P}_X\left(\bigcup_n B_n\right) = \mathbb{P}_X(\mathbb{R}) = 1, \end{aligned}$$

car $(B_n)_{n \geq 1}$ est une suite croissante des événements de $\mathcal{B}(\mathbb{R})$, avec $\bigcup_n B_n = \lim_{n \rightarrow +\infty} B_n = \mathbb{R}$.

3. Pour démontrer la continuité à droite, on prend la suite $(B_n)_{n \geq 1}$ des événements de la forme $B_n =]-\infty, x - \frac{1}{n}[$ avec $\bigcup_{n \geq 1} B_n =]-\infty, x[$ et $\mathbb{P}_X\left(\bigcup_{n \geq 1} B_n\right) = \mathbb{P}(X < x)$,

$$\lim_{x \rightarrow x^-} F_X(y) = \lim_{n \rightarrow +\infty} F_X\left(x - \frac{1}{n}\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(X \leq x - \frac{1}{n}\right) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(B_n) = \mathbb{P}_X\left(\bigcup_n B_n\right) = \mathbb{P}(X < x).$$

4. On constate que si $\mathbb{P}(X = x) = 0$ alors $\mathbb{P}(X < x) = \mathbb{P}(X \leq x)$ est la continuité à gauche manquante vient.
Réciproquement si x est une atome alors $\mathbb{P}(X < x) \neq \mathbb{P}(X \leq x)$ et $\lim_{y \nearrow x} F_X(y) < F_X(x)$ il y a une discontinuité de F_X en x .

■

Remarque 1.1.3 1. Si X est une variable aléatoire discrète, alors

$$*\mathbb{P}_X(x) = \mathbb{P}(X = x) \geq 0 \quad \text{et} \quad \sum_{x \in X(\Omega)} \mathbb{P}_X(x) = 1.$$

$$*\sigma(X) = \sigma(\{X = x\}, x \in X(\Omega))$$

et dans ce cas la fonction de répartition \mathbb{F}_X vérifie :

$$1 - \mathbb{F}_X(y) = \sum_{x \in X(\Omega), x \leq y} \mathbb{P}_X(x), \quad \text{pour } y \in \mathbb{R}.$$

$$2 - \mathbb{P}(a \leq X \leq b) = \mathbb{F}_X(b) - \mathbb{F}_X(a).$$

2. Si X est continue (à densité, i.e. admet une densité de probabilité, une fonction f_X , vérifiant

$$\begin{cases} 1 - \forall x \in \mathbb{R}, f_X(x) \geq 0. \\ 2 - \int_{\mathbb{R}} f_X(x) dx = 1. \\ 3 - f_X \text{ est continue.} \end{cases}$$

alorssi $x < 0$

$$(a) \quad 1 - \forall A \in \mathcal{B}(\mathbb{R}), \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \int_{\mathbb{A}} f_X(x) dx.$$

$$(b) \quad 2 - F_X(y) = \int_{-\infty}^y f_X(x) dx, \quad \text{pour tout } y \in \mathbb{R}.$$

$$(c) \quad 3 - \mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b) \\ = \int_a^b f_X(x) dx = F_X(b) - F_X(a).$$

1.2 Caractéristiques numériques d'une variable aléatoire

L'espace des variables aléatoires, définies sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} , dont la valeur absolue est de puissance p -ième intégrable, est noté $L^p(\Omega, \mathcal{A}, \mathbb{P})$, pour $p \geq 1$, est défini comme suit :

- 1) Si $p \in [1, +\infty[$, $L^p(\Omega, \mathcal{A}, \mathbb{P}) = \left\{ X \text{ est une v.a.} : X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R}; \mathbb{E}|X|^p = \int_{\Omega} |X|^p d\mathbb{P} < +\infty \right\}$ est un espace vectoriel normé avec pour norme

$$\|X\|_p = (\mathbb{E}|X|^p)^{\frac{1}{p}} \quad (\text{norme } L^p).$$

- 2) Si $p = +\infty$, alors $L^p(\Omega, \mathcal{A}, \mathbb{P}) = \{X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R} : \exists c; \mathbb{P}(|X| > c) = 0\}$ c'est l'ensemble des v.a. presque sûrement bornée. Sa norme est définie par

$$\|X\|_{\infty} = \inf \{c > 0 : \mathbb{P}(|X| > c) = 0\} \quad (\text{Supremum essentiel}).$$

$\|X\|_{\infty}$ est le plus petit nombre réel positif vérifiant $\|X\|_p \leq \|X\|_{\infty}$.

1.2.1 Espérance mathématique d'une variable aléatoire

Définition 1.2.1 (moment d'ordre k d'une variable aléatoire)

On appelle moment d'ordre k ($k \in \mathbb{N}^*$) d'une variable aléatoire X , lorsqu'il existe, le réel $\mathbb{E}(X^k)$.

Définition 1.2.2 L'espérance mathématique (moment d'ordre 1) de la variable aléatoire intégrable X , notée $\mathbb{E}(X)$, est définie comme l'intégrale de X par rapport à la probabilité \mathbb{P} comme suit :

$$\mathbb{E}(X) = \int_{\mathcal{X}} X d\mathbb{P}(x).$$

$\mathbb{E}(X)$ joue le rôle dévolu à la moyenne en statistiques : elle correspond à la valeur moyenne espérée par un observateur lors d'une réalisation de la variable aléatoire X .

Exemple 1.2.1 1. Si $X = 1_{A_i}$, alors $\mathbb{E}(X) = \mathbb{E}(1_{A_i}) = \int 1_{A_i} d\mathbb{P}(x) = \int_{A_i} d\mathbb{P}(x) = \mathbb{P}(A_i)$.

2. Si $X = \sum_{i=1}^n a_i 1_{A_i}$, alors $\mathbb{E}(X) = \sum_{i=1}^n a_i \mathbb{P}(A_i)$.

3. Si $X \geq 0$ ($\forall \omega \in \Omega, X(\omega) \geq 0$), on pose

$$X_n(\omega) = \sum_{k=0}^{\infty} \frac{k}{2^n} 1_{\{\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\}},$$

où $(X_n)_{n \geq 0}$ est une suite croissante de v.a. étagées qui converge vers X . On définit $\mathbb{E}(X)$ par

$$\mathbb{E}(X) = \lim_{n \rightarrow +\infty} \mathbb{E}(X_n) = \lim_{n \rightarrow +\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \mathbb{P}\left(\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\right).$$

4. Si X est une v.a. quelconque telle que $\mathbb{E}(X) < \infty$, alors

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

$$\text{et } |X(\omega)| = X^+(\omega) + X^-(\omega).$$

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

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

Espérance d'une variable aléatoire discrète

Définition 1.2.3 Soit X une v.a. discrète avec $X(\Omega) = \{x_i, i \in I\}$ dénombrable. On dit que X est intégrable ssi

$$\mathbb{E}|X| = \sum_{i \in I} |x_i| \mathbb{P}(X = x_i) < +\infty.$$

Dans ce cas l'espérance de X est

$$\mathbb{E}(X) = \sum_{i \in I} x_i \mathbb{P}(X = x_i).$$

- Si $h : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable, alors $h(X)$ est une variable aléatoire discrète, elle est dite intégrable ssi

$$\sum_{i \in I} |h(x_i)| \mathbb{P}(X = x_i) < +\infty.$$

L'espérance de $h(X)$ est définie par

$$\mathbb{E}(h(X)) = \sum_{i \in I} h(x_i) \mathbb{P}(X = x_i).$$

Exemple 1.2.2 Soit X une v.a. de loi de probabilité donné par le tableau suivant :

x_i	-1	0	2	5
$\mathbb{P}(X = x_i)$	0,01	0,2	0,43	0,36

1) Calculer F_X et $E(X)$

2) On définit une nouvelle variable aléatoire $Y = 2X + 1$.

a- Déterminer la loi de probabilité de Y

b- Calculer F_Y et $\mathbb{E}(Y)$.

Exemple 1.2.3 Soit $X \sim \mathbf{P}(\lambda)$ $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda), k \in \mathbb{N}$.

1) Calculer $\mathbb{E}(X)$.

On montre d'abord que X est intégrable

$$\begin{aligned} E(|X|) &= \sum |k| P(X = k) = \sum_{\infty} k \frac{\lambda^k}{k!} \exp(-\lambda) \text{ (car } k \text{ est positif)} \\ &= \lambda < \infty \end{aligned}$$

alors X est intégrable et son espérance est $E(X) = \lambda$.

Espérance d'une variable aléatoire continue

Définition 1.2.4 Soit X une variable aléatoire continue de densité de probabilité f_X . Si X est intégrable, c'est à dire

$$\mathbb{E}|X| = \int_{\mathbb{R}} |x| f_X(x) dx < +\infty.$$

Alors l'espérance de X est donnée par la formule suivante :

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

- Si $h : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable, alors $h(X)$ est une v.a. continue, elle est intégrable ssi

$$\int_{\mathbb{R}} |h(x)| f(x) dx < +\infty.$$

L'espérance de $h(X)$ est alors

$$\mathbb{E}(h(x)) = \int_{\mathbb{R}} h(x) f(x) dx.$$

Propriétés de l'espérance mathématique

Soit X et Y deux variables aléatoires intégrables, $(a, b) \in \mathbb{R}$. Alors :

- 1) *La positivité* : si $X \geq 0$ p.s., alors $\mathbb{E}(X) \geq 0$ p.s. ;
- 2) *La monotonie* : si $0 \leq X \leq Y$ p.s., alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$ p.s. ;
- 3) *Inégalité de Markov* : Si X admet un moment d'ordre 1, alors pour tout $a > 0$,

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

- 4) *La linéarité* : $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

5) *Inégalité de Jensen* : pour φ une fonction réelle convexe ($\varphi : \mathbb{R} \rightarrow \mathbb{R}$), avec $\varphi(X)$ est intégrable. Alors

$$\varphi[\mathbb{E}(X)] \leq \mathbb{E}[\varphi(X)].$$

- 6) *Inégalité de Cauchy-Schwartz* : Si X et Y admettent des moments d'ordre 2, alors

$$\mathbb{E}|X.Y| \leq \sqrt{\mathbb{E}(X^2).\mathbb{E}(Y^2)}.$$

Preuve. (TD) ■

1.2.2 Variance d'une variable aléatoire

Définition 1.2.5 (*Variance d'une variable aléatoire*)

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

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2].$$

La variance mesure la déviation moyenne autour de la moyenne espérée $\mathbb{E}(X)$ et l'écart type est donné par

$$\sigma_X = \sqrt{\text{Var}(X)}.$$

utilisé souvent pour mesurer la dispersion d'une variable aléatoire X .

Remarque 1.2.1 1. 1- Pour une v.a. discrète X , avec $X(\Omega) = \{x_i, i \in I\}$, on a

$$\text{Var}(X) = \sum_{i \in I} (x_i - \mathbb{E}(X))^2 \mathbb{P}_X(x_i).$$

2. 2- Pour une v.a. continue X à densité f_X , on a

$$\text{Var}(X) = \int_{\mathbb{R}} (x - \mathbb{E}(X))^2 f_X(x) dx.$$

Propriétés de la variance

Soient X et Y deux variables aléatoires de $L^2(\Omega, \mathcal{A}, \mathbb{P})$ et $(a, b) \in \mathbb{R}^2$.

- 1) $\text{Var}(X) \geq 0$;
- 2) $\text{Var}(aX) = a^2 \text{Var}(X)$;
- 3) $\text{Var}(X + b) = \text{Var}(X)$;
- 4) $\text{Var}(X) = 0 \Leftrightarrow X = \mathbb{E}(X)$;
- 5) $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$
- 6) $\text{Var}(X - Y) \neq \text{Var}(X) + \text{Var}(Y)$
- 7) *Inégalité de Tchebychev* : si $\text{Var}(X) < +\infty$, alors $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}$.
- 8) *Formule de koenig* : $\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$.

1.3 Transformation d'une variable aléatoire

Une transformation d'une variable aléatoire X , appelée aussi fonction d'une variable aléatoire X , est la composée d'une fonction h et X . Autrement dit, c'est l'application $h(X)(\omega) = hoX(\omega)$ définie sur l'espace (Ω, \mathcal{A}) à valeurs dans l'ensemble d'arrivée de h .

Théorème 1.3.1 *Soient X une variable aléatoire et h une fonction mesurable (continue) à valeurs réelles, alors $h(X)$ est une variable aléatoire.*

Preuve. Pour montrer que $h(X)$ est une variable aléatoire on utilise la définition, c'est à dire on montre que $h(X)$ est une application mesurable. D'abord, on a

$$X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) \text{ et } h : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

donc la composée des deux fonctions h et X , notée $h(X) = hoX$ est une application définie de (Ω, \mathcal{A}) à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ comme suit

$$\begin{array}{ccc} hoX : & (\Omega, \mathcal{A}) & \xrightarrow{X} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \xrightarrow{h} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega \mapsto & & hoX(\omega). \end{array}$$

Pour montrer la mesurabilité de $h(X)$ on montre que pour tout $B \in \mathcal{B}(\mathbb{R})$, $[h(X)]^{-1}(B) \in \mathcal{A}$. Soit $B \in \mathcal{B}(\mathbb{R})$, on a par définition

$$[h(X)]^{-1}(B) = [h \circ X]^{-1}(B) = X^{-1}[h^{-1}(B)],$$

Comme h est une fonction mesurable, alors $A = h^{-1}(B) \in \mathcal{B}(\mathbb{R})$ et X est une v.a. (une application mesurable), par conséquent $X^{-1}[h^{-1}(B)] = X^{-1}[A] \in \mathcal{A}$.

D'où $h(X)$ est une variable aléatoire. ■

Proposition 1.3.1 *Soient X_1, \dots, X_n des variables aléatoires. Les fonctions suivantes sont des variables aléatoires aussi*

- 1- $\max_n(X_1, \dots, X_n), \min_n(X_1, \dots, X_n), \sup_n(X_1, \dots, X_n)$ et $\inf_n(X_1, \dots, X_n)$.
- 2- $\limsup_{n \rightarrow \infty} X_n$ et $\liminf_{n \rightarrow \infty} X_n$.
- 3- La somme $X_1 + \dots + X_n$, le produit $X_1 \dots X_n$
- 4- L'inverse $\frac{1}{X_i}, X_i \in \{X_1, \dots, X_n\}$.
- 5- La division $\frac{X_i}{X_j}, X_i, X_j \in \{X_1, \dots, X_n\}$ pour tout $i \neq j$.

Transformation d'une variable aléatoire continue (à densité)

On a une variable aléatoire continue X et une fonction h qui est une **bijection**, et on définit une autre variable aléatoire $Y = h(X)$ dont le but est de déterminer la densité de Y .

Méthode 1 : En utilisant la fonction de répartition F_X de X , on suivra les étapes suivantes :

1- On détermine $Y(\Omega)$. Si ce n'est pas immédiat, il suffit d'étudier la fonction h avec pour ensemble de départ $X(\Omega)$. Un tableau de variation donne l'ensemble d'arrivée, donc $Y(\Omega)$.

2- On sait désormais que pour les t à gauche de $Y(\Omega)$, on a $F_Y(t) = 0$ et pour les t à droite $F_Y(t) = 1$.

3- On prend $t \in Y(\Omega)$ et on se ramène à la fonction de répartition de X à l'aide de la méthode suivante :

3-a- Si h est croissante :

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}(h(X) \leq t) \\ &= \mathbb{P}(X \leq h^{-1}(t)) \text{ (car } h \text{ est strictement croissante sur } Y(\Omega)) \\ &= F_X(h^{-1}(t)). \end{aligned}$$

3-b- Si h est décroissante :

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}(h(X) \leq t) \\ &= \mathbb{P}(X \geq h^{-1}(t)) \text{ (car } h \text{ est strictement décroissante sur } Y(\Omega)) \\ &= 1 - \mathbb{P}(X < h^{-1}(t)) \\ &= 1 - \mathbb{P}(X \leq h^{-1}(t)) \text{ (car } X \text{ est continue donc } \mathbb{P}(X = h^{-1}(t)) = 0) \\ &= 1 - F_X(h^{-1}(t)). \end{aligned}$$

4- Une fois qu'on a obtenu l'expression de F_Y sur $Y(\Omega)$, il suffit de dériver pour obtenir la densité f_Y .

Exemple 1.3.1 Soit X une variable aléatoire continue telle que $X(\Omega) = \mathbb{R}$ et de fonction de répartition F_X de classe C^1 sur \mathbb{R} . On pose $Y = \ln(1 + |X|)$ dont le but est de déterminer la densité de probabilité de Y .

D'abord on a $Y(\Omega) = \mathbb{R}^+$ car

Puis on remarque que la fonction $h(x) = \ln(1 + |x|)$ est croissante et que pour tout $t < 0$ on a $F_Y(t) = 0$.

- soit $t \in \mathbb{R}^+$,

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}(\ln(1 + |X|) \leq t) \\ &= \mathbb{P}(1 + |X| \leq e^t) \text{ (car la fonction } x \mapsto e^x \text{ est une bijection croissante)} \\ &= \mathbb{P}(|X| \leq e^t - 1) \\ &= \mathbb{P}(1 - e^t \leq X \leq e^t - 1) \text{ (car } t \geq 0) \\ &= F_X(e^t - 1) - F_X(1 - e^t). \end{aligned}$$

Comme $F_X \in C^1$, alors par dérivation des composées des fonctions continues $F_X, e^t - 1$ et $1 - e^t$, on obtient

$$f_Y(t) = \begin{cases} 0, & \text{si } t < 0 \\ e^t (f_X(e^t - 1) + f_X(1 - e^t)) & \end{cases}.$$

Méthode 2 : On peut combiner les deux cas précédents afin d'obtenir la densité de probabilités de $Y = h(X)$ par la formule suivantes

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right|.$$

Exemple 1.3.2 Soit X une variable aléatoire qui suit la loi uniforme sur l'intervalle $[a, b]$, c'est à dire sa fonction de densité est de donnée par

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b] \\ 0, & \text{sinon} \end{cases}$$

On définit une nouvelle variable aléatoire $Y = h(X) = 2 + X$, puis on cherche l'expression de la densité de probabilité de Y .

On

$$y = h(x) = 2 + x \Rightarrow x = y - 2 = h^{-1}(y).$$

la densité f_Y de la v.a. Y est donnée par la formule

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right| = f_X(y - 2) \left| \frac{d}{dy} (y - 2) \right| \\ &= f_X(y - 2) |1| = f_X(y - 2) \\ &= \begin{cases} \frac{1}{b-a}, & \text{si } y - 2 \in [a, b] \\ 0, & \text{sinon} \end{cases} \\ &= \begin{cases} \frac{1}{b-a}, & \text{si } y \in [2 + a, 2 + b] \\ 0, & \text{sinon} \end{cases}. \end{aligned}$$

Transformation d'une variable aléatoire discrète.

Supposons qu'on a une variable aléatoire discrète X d'ensemble de valeurs possibles est $X(\Omega)$ et la loi de probabilité de X étant connue. On définit une nouvelle variable aléatoire $Y = g(X)$, où g est une fonction continue sur $X(\Omega)$. Pour déterminer la loi de probabilité de Y , il suffit de :

- 1- Déterminer $Y(\Omega)$.
- 2- Déterminer $\mathbb{P}(Y = y_j) = \mathbb{P}(g(X) = y_j), \forall y_j \in Y(\Omega)$.

Exemple 1.3.3 On lance trois fois une pièce de monnaie équilibrée et on note X la v.a. représentant le nombre de faces obtenues.

- 1- Déterminer la loi de probabilité de la v.a. X et sa fonction de répartition
- 2- Calculer l'espérance mathématique de X et sa variance.
- 3- Soit la v.a. $Y = X^2 - 1$. Déterminer la loi de probabilité de la v.a. Y et donner sa fonction de répartition.
- 4- Calculer l'espérance mathématique de Y et sa variance par deux méthodes différentes.

Solution 1.3.1 1- On détermine la loi de probabilité de X et sa fonction de répartition F_X .
D'abord l'ensemble des résultats possibles de cette expérience aléatoire est

$$\Omega = \{(P, P, P), (F, F, F), (P, P, F), (P, F, P), (P, F, F), (F, P, P), (F, F, P), (F, P, F)\}$$

on a les valeurs possible de X sont $X(\Omega) = \{0, 1, 2, 3\}$, puis on calcule les probabilités suivantes :

$$\begin{aligned} \mathbb{P}(X = 0) &= \mathbb{P}(\{(P, P, P)\}) = \frac{1}{8}, \mathbb{P}(X = 1) = \mathbb{P}(\{(P, P, F), (P, F, P), (F, P, P)\}) = \frac{3}{8}, \\ \mathbb{P}(X = 2) &= \mathbb{P}(\{(P, F, F), (F, F, P), (F, P, F)\}) = \frac{3}{8}, \mathbb{P}(X = 3) = \mathbb{P}(\{(F, F, F)\}) = \frac{1}{8}. \end{aligned}$$

$$\text{On a pour tout } x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \leq x) = \sum_{x_i \in X(\Omega), x_i \leq x} \mathbb{P}(X = x_i) = \begin{cases} 0, & \text{si } x < 0 \\ \frac{1}{8}, & \text{si } 0 \leq x < 1 \\ \frac{4}{8}, & \text{si } 1 \leq x < 2 \\ \frac{7}{8}, & \text{si } 2 \leq x < 3 \\ 1, & \text{si } x \geq 3 \end{cases}$$

2- Calcul de l'espérance mathématique de X et la variance.

$$\mathbb{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i) = \frac{12}{8}.$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) \text{ et } \mathbb{E}(X^2) = \sum_{x_i \in X(\Omega)} x_i^2 \mathbb{P}(X = x_i) = \frac{24}{8}, \text{ donc } \text{Var}(X) = \frac{5}{8}.$$

3- On détermine la loi de probabilité de la v.a. Y et on donne sa fonction de répartition

On a $Y(\Omega) = \{-1, 0, 3, 8\}$, $\mathbb{P}(Y = -1) = \mathbb{P}(X = 0) = \frac{1}{8}$, $\mathbb{P}(Y = 0) = \mathbb{P}(X = 1) = \frac{3}{8}$, $\mathbb{P}(Y = 3) = \mathbb{P}(X = 2) = \frac{3}{8}$,
 $\mathbb{P}(Y = 8) = \mathbb{P}(X = 3) = \frac{1}{8}$.

On a pour tout $y \in \mathbb{R}$, $F_Y(y) = \mathbb{P}(Y \leq y) = \sum_{y_i \in Y(\Omega), y_i \leq y} \mathbb{P}(Y = y_i) = \begin{cases} 0, & \text{si } y < -1 \\ \frac{1}{8}, & \text{si } -1 \leq y < 0 \\ \frac{4}{8}, & \text{si } 0 \leq y < 3 \\ \frac{7}{8}, & \text{si } 3 \leq y < 8 \\ 1, & \text{si } y \geq 8 \end{cases}$

4- On calcule l'espérance mathématique de Y et la variance par deux méthodes différentes.

Méthode 1 : On utilise les définitions de $\mathbb{E}(Y)$ et $\text{Var}(Y)$.

$$\mathbb{E}(Y) = \sum_{y_i \in Y(\Omega)} y_i \mathbb{P}(Y = y_i) = 2.$$

$$\text{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) \text{ et } \mathbb{E}(Y^2) = \sum_{y_i \in Y(\Omega)} y_i^2 \mathbb{P}(Y = y_i) = \frac{92}{8}, \text{ donc } \text{Var}(Y) = \frac{15}{2}.$$

Méthode 2 : On utilise les propriétés de $\mathbb{E}(Y)$ et $\text{Var}(Y)$.

$$\mathbb{E}(Y) = \mathbb{E}(X^2 - 1) = \mathbb{E}(X^2) - 1 = \frac{24}{8} - 1 = 2.$$

$$\text{Var}(Y) = \text{Var}(X^2 - 1) = \text{Var}(X^2) = \mathbb{E}(X^4) - \mathbb{E}^2(X^2), \mathbb{E}(X^4) = \sum_{x_i \in X(\Omega)} x_i^4 \mathbb{P}(X = x_i) = \frac{132}{8}, \text{ donc } \text{Var}(Y) = \frac{15}{2}.$$

Chapitre 2

Couples de variables aléatoires

2.1 Définition et exemple d'un couples de variables aléatoires

Définition 2.1.1 On appelle couple de variables aléatoires réelles, tout couple (X, Y) où X et Y sont deux variables aleatoires réelles définies sur le même espace Ω .

Exemple 2.1.1 On lance deux dés équilibrés à six faces, l'un bland, l'autre rouge. On appelle X (resp. Y)

le numéro obtenu avec le dé bland (resp. rouge).

Comme X et Y sont des variables aléatoires, alors (X, Y) est un couple de variables aléatoires.

Définition 2.1.2 Soit (X, Y) un couple de variables aléatoires. On appelle loi **conjointe** (loi **jointe**) du couple (X, Y) , et on la note $\mathbb{P}_{(X,Y)}$, la probabilité image de \mathbb{P} par l'application (X, Y) , définie comme suit

$$\mathbb{P}_{(X,Y)}(B) = \mathbb{P}(\{\omega \in \Omega ; X(\omega) \in B_1, Y(\omega) \in B_2\}), B = (B_1, B_2) \in \mathcal{B}(\mathbb{R}^2).$$

Définition 2.1.3 La **fonction de répartition jointe** du couple de variables aléatoires $Z = (X, Y)$, norée $F_{(X,Y)}$ ou F_Z , est définie pour tout $z = (x, y) \in \mathbb{R}^2$ comme suit

$$\begin{aligned} F_Z(z) &= F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y) \\ &= \mathbb{P}(\{X \leq x\} \cap \{Y \leq y\}) . \end{aligned}$$

2.2 Couples de variables aléatoires discrètes et indépendance

Soient X et Y deux variables aléatoire discrètes définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans deux ensembles **finis** ou **dénombrables** $X(\Omega)$ et $Y(\Omega)$:

1. prenant soit **un nombre fini** de valeurs, "n" pour la variable X et "m" pour la variable Y .
2. soit **un ensemble dénombrable** de valeurs, notées x_i pour X et y_j pour Y .

Définition 2.2.1 L'application Z définie de Ω dans $X(\Omega) \times Y(\Omega)$ par :

$$\forall \omega \in \Omega, Z(\omega) = (X(\omega), Y(\omega)),$$

est une variable aléatoire discrète sur Ω .

On notera pour tout valeur possible z de la v.a. Z , $z = (x_i, y_j) \in X(\Omega) \times Y(\Omega)$:

$$Z(x_i, y_j) = (X = x_i, Y = y_j) = (\{X = x_i\} \cap \{Y = y_j\}).$$

Remarque 2.2.1 L'ensemble des valeurs possibles de la variable aléatoire Z est noté $Z(\Omega)$ tel que $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$.

Exemple 2.2.1 On lance simultanément deux dés, l'un blanc, l'autre rouge et on note X le plus grand des deux chiffres obtenus et Y le plus petit (si les numéros sont égaux, X et Y prennent la valeur commune). Les applications X et Y sont deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ où $\Omega = \{(i, j), 1 \leq i, j \leq 6\}$ avec $\text{card}(\Omega) = 36$, \mathbb{P} est la probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$, c'est à dire

$$\forall A \in \mathcal{P}(\Omega); \mathbb{P}(A) = \frac{\text{le nombre de cas favorables}}{\text{le nombre de cas possibles}}.$$

$X(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $Y(\Omega) = \{1, 2, 3, 4, 5, 6\}$
 l'ensemble des valeurs possibles de la v.a. $Z = (X, Y)$ est

$$Z(\Omega) = \left\{ \begin{array}{l} (1, 1), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1), (2, 2), (3, 2), (3, 3), (4, 2), (5, 2), (6, 2), (4, 3), \\ (4, 4), (5, 3), (5, 4), (5, 5), (6, 3), (6, 4), (6, 5), (6, 6) \end{array} \right\}$$

2.2.1 Lois conjointes, marginales et conditionnelles.

Loi conjointe d'un couple de variables aléatoires discrètes

Soient X et Y deux variables aléatoires discrètes à valeurs $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ respectivement, où I et J sont deux ensembles finis ou dénombrables.

Définition 2.2.2 La loi de la variable aléatoire $Z = (X, Y)$, appelée aussi loi de probabilité **simultanée** ou loi **conjointe** des deux variable aléatoires X et Y , est définie par l'ensemble des nombres réels $(p_{i,j})_{i \in I, j \in J}$, définis par

$$p_{i,j} = \mathbb{P}(X = x_i, Y = y_j),$$

vérifiant les deux conditions suivantes :

$$\left\{ \begin{array}{l} 1 - \forall (i, j) \in I \times J, p_{i,j} \in [0, 1]; \\ 2 - \sum_{i \in I} \sum_{j \in J} p_{i,j} = \sum_{j \in J} \sum_{i \in I} p_{i,j} = 1. \end{array} \right.$$

Exemple 2.2.2 Déterminons la loi jointe des deux v.a. X et Y de l'exemple précédent. On note par R le numéro obtenu du dé rouge et B le numéro obtenu du dé blanc. Puisque $X(\Omega)$ et $Y(\Omega)$ sont finis, on peut représenter la loi de Z par un tableau de contingence (à double entrée) comme suit

$X \setminus Y$	1	2	3	4	5	6
1	$\frac{1}{36}$	0	0	0	0	0
2	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0
3	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0
4	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0
5	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0
6	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

En effet,

1. si $x < y$, alors $\mathbb{P}_{(X,Y)}(x, y) = \mathbb{P}(X = x, Y = y) = 0$.

2. si $x > y$ alors

$$\begin{aligned}\mathbb{P}_{(X,Y)}(x,y) &= \mathbb{P}(X=x, Y=y) = \mathbb{P}(\{R=x, B=y\} \cup \{B=x, R=y\}) \\ &= \mathbb{P}(\{R=x, B=y\}) + \mathbb{P}(\{B=x, R=y\}) = \frac{2}{36},\end{aligned}$$

car les deux événements $\{R=x, B=y\}$ et $\{B=x, R=y\}$ sont disjoints.

3. si $x = y$ alors

$$\mathbb{P}_{(X,Y)}(x,y) = \mathbb{P}(X=x, Y=y) = \frac{1}{36}.$$

On remarque que $\forall (x,y) \in X(\Omega) \times Y(\Omega), \mathbb{P}_{(X,Y)}(x,y) \in [0,1]$ et

$$\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} \mathbb{P}(X=x, Y=y) = \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} \mathbb{P}(X=x, Y=y) = 1.$$

Donc on a défini la loi de probabilité conjointe du couple (X,Y) .

Définition 2.2.3 La **fonction de répartition jointe** du couple de variables aléatoires discrètes $Z = (X,Y)$ est définie pour tout $z = (x,y) \in \mathbb{R}^2$ comme suit

$$\begin{aligned}F_Z(z) &= F_{(X,Y)}(x,y) = \mathbb{P}(X \leq x, Y \leq y) \\ &= \sum_{x_i \in X(\Omega), x_i \leq x} \sum_{y_j \in Y(\Omega), y_j \leq y} \mathbb{P}(X=x_i, Y=y_j)\end{aligned}$$

Lois marginales d'un couple de variables aléatoires discrètes

Définition 2.2.4 Soit (X,Y) un couple de variables aléatoires discrètes. La loi de X est appelée **première loi**

marginale du couple et celle de Y est appelée **deuxième loi marginale** du couple.

1- **La loi marginale de X** est définie pour tout $x \in X(\Omega)$ par

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

2- **La loi marginale de Y** est définie pour tout $y \in Y(\Omega)$ par

$$\mathbb{P}_Y(y) = \mathbb{P}(Y=y) = \sum_{x_i \in X(\Omega)} \mathbb{P}(X=x_i, Y=y).$$

Exemple 2.2.3 Cherchons les lois marginales des variables aléatoires X et Y définies dans l'exemple précédent en utilisant le tableau de contingence

$X \setminus Y$	1	2	3	4	5	6	$\mathbb{P}_X(x)$
1	$\frac{1}{36}$	0	0	0	0	0	$\mathbb{P}_X(1) = \frac{1}{36}$
2	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	$\mathbb{P}_X(2) = \frac{3}{36}$
3	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	$\mathbb{P}_X(3) = \frac{5}{36}$
4	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	$\mathbb{P}_X(4) = \frac{7}{36}$
5	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	$\mathbb{P}_X(5) = \frac{9}{36}$
6	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\mathbb{P}_X(6) = \frac{11}{36}$
$\mathbb{P}_Y(y)$	$\mathbb{P}_Y(1) = \frac{11}{36}$	$\mathbb{P}_Y(2) = \frac{9}{36}$	$\mathbb{P}_Y(3) = \frac{7}{36}$	$\mathbb{P}_Y(4) = \frac{5}{36}$	$\mathbb{P}_Y(5) = \frac{3}{36}$	$\mathbb{P}_Y(6) = \frac{1}{36}$	1

Remarque 2.2.2 On peut obtenir les lois marginales à partir de la loi conjointe, mais la réciproque est fausse, autrement dit les lois marginales ne suffit pas pour connaître la loi conjointe.

Lois conditionnelles de variables aléatoires discrètes

Définition 2.2.5 Soit (X, Y) un couple de variables aléatoires.

1- Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}_Y(y) \neq 0$, on appelle loi de X **conditionnellement à l'événement** $\{Y = y\}$ (**loi conditionnelle** de X sachant que $\{Y = y\}$), et on la note $\mathbb{P}_{X|Y}$, la loi de probabilité définie, pour tout $x \in X(\Omega)$, par

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

2- Pour tout $x \in X(\Omega)$ tel que $\mathbb{P}_X(x) \neq 0$, on appelle loi de Y **conditionnellement à l'événement** $\{X = x\}$ (**loi conditionnelle** de Y sachant que $\{X = x\}$), et on la note $\mathbb{P}_{Y|X}$, la loi de probabilité définie, pour tout $y \in Y(\Omega)$, par

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

Exemple 2.2.4 (suite de l'exemple précédent)

1- La loi conditionnelle de $X|Y = 1$ est calculer comme suit

$$\begin{aligned} \mathbb{P}_{X|Y}(1|1) &= \frac{\mathbb{P}(X = 1, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{1}{36}}{\frac{11}{36}} = \frac{1}{11}, \mathbb{P}_{X|Y}(2|1) = \frac{\mathbb{P}(X = 2, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}, \\ \mathbb{P}_{X|Y}(3|1) &= \frac{\mathbb{P}(X = 3, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}, \mathbb{P}_{X|Y}(4|1) = \frac{\mathbb{P}(X = 4, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}, \\ \mathbb{P}_{X|Y}(5|1) &= \frac{\mathbb{P}(X = 5, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}, \mathbb{P}_{X|Y}(6|1) = \frac{\mathbb{P}(X = 6, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}. \end{aligned}$$

On remarque que

$$\mathbb{P}_{X|Y}(1|1) + \mathbb{P}_{X|Y}(2|1) + \mathbb{P}_{X|Y}(3|1) + \mathbb{P}_{X|Y}(4|1) + \mathbb{P}_{X|Y}(5|1) + \mathbb{P}_{X|Y}(6|1) = 1$$

et pour tout $x \in X(\Omega)$; $\mathbb{P}_{X|Y}(x|y) \in [0, 1]$.

Donc on a déterminé la loi de probabilité conditionnelle de $X|Y = 1$.

2- La loi conditionnelle de $Y|X = 1$ est calculer comme suit

$$\begin{aligned} \mathbb{P}_{Y|X}(1|1) &= \frac{\mathbb{P}(X = 1, Y = 1)}{\mathbb{P}(X = 1)} = \frac{\frac{1}{36}}{\frac{1}{36}} = 1, \mathbb{P}_{Y|X}(2|1) = \frac{\mathbb{P}(X = 1, Y = 2)}{\mathbb{P}(X = 1)} = 0, \\ \mathbb{P}_{Y|X}(3|1) &= \frac{\mathbb{P}(X = 1, Y = 3)}{\mathbb{P}(X = 1)} = 0, \mathbb{P}_{Y|X}(4|1) = \frac{\mathbb{P}(X = 1, Y = 4)}{\mathbb{P}(X = 1)} = 0, \\ \mathbb{P}_{Y|X}(5|1) &= \frac{\mathbb{P}(X = 1, Y = 5)}{\mathbb{P}(X = 1)} = 0, \mathbb{P}_{Y|X}(6|1) = \frac{\mathbb{P}(X = 1, Y = 6)}{\mathbb{P}(X = 1)} = 0. \end{aligned}$$

Théorème 2.2.1 (Théorème de transfert)

Soient (X, Y) un couple de variables aléatoires discrètes et h une fonction définie sur $X(\Omega) \times Y(\Omega)$ à valeurs réelles. $h(X, Y)$ admet une espérance mathématique si et seulement si

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |h(x, y)| \mathbb{P}(X = x, Y = y) < \infty.$$

Dans ce cas

$$\mathbb{E}[h(X, Y)] = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} h(x, y) \mathbb{P}(X = x, Y = y).$$

2.2.2 Indépendance de deux variables aléatoires discrètes

Définition 2.2.6 Deux variables aléatoires discrètes X et Y sont dites indépendantes si tous les couples d'événements $\{X = x\}$, $\{Y = y\}$ sont indépendants. Autrement dit,

$$\forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}(\{X = x\}, \{Y = y\}) = \mathbb{P}(\{X = x\}) \cdot \mathbb{P}(\{Y = y\}).$$

Remarque 2.2.3 1- Deux variables aléatoires X et Y sont indépendantes si et seulement si toutes les lois conditionnelles de X sachant $Y = y$ sont identiques à la loi de X . (De même pour Y)

$$\mathbb{P}_{X|Y}(x|y) = \mathbb{P}(X = x) \text{ pour tout } x \in X(\Omega).$$

$$\mathbb{P}_{Y|X}(y|x) = \mathbb{P}(Y = y) \text{ pour tout } y \in Y(\Omega).$$

2- Dans le cas où X et Y sont des variables aléatoires indépendantes, on peut obtenir la loi conjointe du

couple (X, Y) à partir des deux lois marginales.

3- Deux variables aléatoires X et Y sont indépendantes si et seulement si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), F_{(X,Y)}(x, y) = F_X(x) \cdot F_Y(y)$$

2.3 Couples de variables aléatoires continues et indépendance

Définition 2.3.1 (Densité de probabilité conjointe)

Une fonction f de \mathbb{R}^2 dans \mathbb{R} est dite densité de probabilité conjointe si

$$\begin{aligned} & 1 - \forall (x, y) \in \mathbb{R}^2, f(x, y) \geq 0 \\ & 2 - \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy = 1. \end{aligned}$$

Nous allons maintenant voir comment associer une loi de probabilité à une fonction de densité conjointe f .

Définition 2.3.2 Un couple de variables aléatoires (X, Y) est dit (conjointement) continu si sa fonction de répartition $F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y)$ peut s'écrire sous la forme

$$F_{(X,Y)}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(u, v) du dv,$$

pour tout $(x, y) \in \mathbb{R}^2$, où $f_{(X,Y)}$ est une fonction de densité conjointe. Et donc

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

Remarque 2.3.1 La fonction de densité $f_{(X,Y)}$ caractérise la loi du couple aléatoire (X, Y) en représentant la distribution de la masse de probabilité dans le plan.

Exemple 2.3.1 1- Soit f une fonction à valeurs réelles définie par

$$f(x, y) = \begin{cases} \frac{3}{5}(xy + x^2), & \text{si } x \in [0, 1] \text{ et } y \in [0, 2] \\ 0, & \text{sinon} \end{cases}$$

- Montrer que f est une densité de probabilité.

En effet, $f(x, y)$ est une fonction positive pour tout $x \in [0, 1]$ et $y \in [0, 2]$, donc $\forall (x, y) \in \mathbb{R}^2, f(x, y) \geq 0$.

Pour l'intégrale on a

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy = \int_0^2 \int_0^1 \frac{3}{5}(xy + x^2) dx dy = 1.$$

donc f est une densité de probabilité

2- Soit g une fonction à valeurs réelles définie par

$$g(x, y) = \begin{cases} a(x + y^2), & \text{si } x \in [0, 1] \text{ et } y \in [0, 1] \\ 0, & \text{sinon} \end{cases}$$

- Calculer la valeur de a pour g soit une densité de probabilité.

Pour que g soit une densité de probabilité il faut que a soit **positif** et que $\int_{\mathbb{R}} \int_{\mathbb{R}} g(x, y) dx dy = 1$.
on a

$$\int_{\mathbb{R}} \int_{\mathbb{R}} g(x, y) dx dy = 1 \Leftrightarrow \int_0^1 \int_0^1 a(x + y^2) dx dy = 1 \Leftrightarrow a = \frac{6}{5}.$$

Théorème 2.3.1 (Théorème de transfert)

Soient (X, Y) un couple de variables aléatoires réelles, admettant une densité conjointe $f_{(X, Y)}$. Une fonction continue h définie de \mathbb{R}^2 dans \mathbb{R} admet une espérance si et seulement si.

$$\int_{\mathbb{R}^2} |h(x, y)| f_{(X, Y)}(x, y) dx dy < \infty.$$

Dans ce cas

$$\mathbb{E}[h(X, Y)] = \int_{\mathbb{R}^2} h(x, y) f_{(X, Y)}(x, y) dx dy.$$

Densités marginales et conditionnelles

Soit (X, Y) un couple de variables aléatoires de densité conjointe $f_{(X, Y)}$ et de fonction de répartition conjointe $F_{(X, Y)}$.

Définition 2.3.3 (Densités marginales)

1- On appelle première densité marginale de $f_{(X, Y)}$ la fonction f_X représentant la densité marginale de X , définie pour tout $x \in \mathbb{R}$, par

$$f_X(x) = \int_{\mathbb{R}} f_{(X, Y)}(x, y) dy.$$

2- On appelle seconde densité marginale de $f_{(X, Y)}$ la fonction f_Y représentant la densité marginale de Y , définie pour tout $y \in \mathbb{R}$, par

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

Définition 2.3.4 (Fonctions de répartitions marginales)

1- La fonction de répartition marginale de la variable aléatoire X est définie comme suit

$$F_X(x) = \mathbb{P}(X \leq x) = \lim_{y \rightarrow \infty} F_{(X,Y)}(x, y).$$

2- La fonction de répartition marginale de la variable aléatoire Y est définie comme suit

$$F_Y(y) = \mathbb{P}(Y \leq y) = \lim_{x \rightarrow \infty} F_{(X,Y)}(x, y).$$

Définition 2.3.5 (Densités conditionnelles)

1- La densité conditionnelle de X sachant Y , notée $f_{X|Y}$, est définie par

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

2- La densité conditionnelle de Y sachant X , notée $f_{Y|X}$, est définie par

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

Exemple 2.3.2 Soit (X, Y) un couple de variables aléatoires de densité conjointe $f_{(X,Y)}$ définie par

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{3}{5}(xy + x^2), & \text{si } 0 < x \leq 1 \text{ et } 0 < y \leq 2 \\ 0, & \text{si non} \end{cases}$$

1- Montrer que $f_{(X,Y)}$ est une densité de probabilité.

2- Calculer $\mathbb{E}(XY)$ et $\mathbb{E}(X + Y)$.

3- Calculer les densités marginales de X et Y .

4- Calculer les densités conditionnelles $f_{X|Y}$ et $f_{Y|X}$.

5- Calculer la fonction de répartition $F_{(X,Y)}$.

Solution 2.3.1 1- On montre que $f_{(X,Y)}$ est une densité de probabilité.

$$f_{(X,Y)} \text{ est une densité de probabilité} \Leftrightarrow \begin{cases} 1 - \forall (x, y) \in \mathbb{R}; f_{(X,Y)} \geq 0; \\ 2 - f_{(X,Y)} \text{ est continue}; \\ 3 - \int_{\mathbb{R}^2} f_{(X,Y)}(x, y) dx dy = 1. \end{cases}$$

On a $x \mapsto \frac{3}{5}(xy + x^2)$ est une fonction continue est positive sur $0 < x \leq 1$ et $0 < y \leq 2$, alors $f_{(X,Y)}$ est continue est positive. Pour l'intégrale

$$\int_{\mathbb{R}^2} f_{(X,Y)}(x, y) dx dy = \int_0^2 \int_0^1 \frac{3}{5}(xy + x^2) dx dy = 1.$$

2- Calculons $\mathbb{E}(XY)$ et $\mathbb{E}(X + Y)$ en utilisant le théorème de transfert

$$\mathbb{E}(XY) = \int_{\mathbb{R}^2} xy f_{(X,Y)}(x, y) dx dy = \int_0^2 \int_0^1 \frac{3}{5} xy(xy + x^2) dx dy = \frac{5}{6}.$$

$$\mathbb{E}(X + Y) = \int_{\mathbb{R}^2} (x + y) f_{(X,Y)}(x, y) dx dy = \int_0^2 \int_0^1 \frac{3}{5} (x + y)(xy + x^2) dx dy = \frac{19}{10}.$$

3- On calcule les densités marginales de X et Y .

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y) dy = \int_0^2 \frac{3}{5}(xy + x^2) dy = \frac{6}{5}(x + x^2) \cdot 1_{\{x \in [0,1]\}}.$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y) dx = \int_0^1 \frac{3}{5}(xy + x^2) dx = \frac{3}{5} \left(\frac{y}{2} + \frac{1}{3} \right) \cdot 1_{\{y \in [0,2]\}}.$$

4- On calcule les densités conditionnelles $f_{X|Y}$ et $f_{Y|X}$.

$$f_{X|Y}(x|y) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{\frac{3}{5}(xy + x^2)}{\frac{3}{5} \left(\frac{y}{2} + \frac{1}{3} \right)} = \frac{(xy + x^2)}{2 \left(\frac{y}{2} + \frac{1}{3} \right)} \cdot 1_{\{x \in [0,1], y \in [0,2]\}}$$

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{\frac{3}{5}(xy + x^2)}{\frac{6}{5}(x + x^2)} = \frac{(xy + x^2)}{\left(\frac{y}{2} + \frac{1}{3} \right)} \cdot 1_{\{x \in [0,1], y \in [0,2]\}}.$$

5- On calcule la fonction de répartition $F_{(X,Y)}$.

$$F_{(X,Y)}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(u,v) dudv$$

$$= \begin{cases} 0, & \text{si } x < 0 \text{ et } y < 0 \\ \frac{3}{5} \left(\frac{1}{4}x^2y^2 + \frac{1}{3}yx^3 \right), & \text{si } 0 \leq x < 1 \text{ et } 0 \leq y < 2 \\ 1, & \text{si } x > 1 \text{ et } y > 2. \end{cases}$$

2.3.1 Indépendance de deux variables aléatoires continues

Définition 2.3.6 Deux variables aléatoires absolument continues X et Y sont dites indépendantes ssi

$$f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y).$$

Remarque 2.3.2 1- Deux variables aléatoires X et Y sont indépendantes si et seulement si la densité conditionnelles de $f_{X|Y}$ identiques à f_X . (De même pour Y)

$$f_{X|Y}(x|y) = f_X(x) \text{ pour tout } x \in X(\Omega)$$

$$f_{Y|X}(y|x) = f_Y(y) \text{ pour tout } y \in Y(\Omega).$$

2- Pour X et Y sont des variables aléatoires indépendantes, on peut obtenir la loi conjointe du couple (X,Y) à partir des deux lois marginales.

3- Deux variables aléatoires X et Y sont indépendantes si et seulement si

$$\forall (x,y) \in \mathbb{R}^2, F_{(X,Y)}(x,y) = F_X(x) \cdot F_Y(y)$$

Exemple 2.3.3 Soit (X,Y) un couple de variables aléatoires de densité conjointe $f_{(X,Y)}$ définie par

$$f_{(X,Y)}(x,y) = \begin{cases} 2ye^{-x}, & \text{si } x \geq 0, 0 < y < 1 \\ 0, & \text{si non} \end{cases}$$

1- Les densités marginales de X et Y respectivement sont

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy = \int_0^1 2ye^{-x} dy = e^{-x} \cdot 1_{\{x \geq 0\}}.$$

$$f_Y(y) = \int_{\mathbb{R}} f(x, y) dx = \int_0^{\infty} 2ye^{-x} dx = 2y \cdot 1_{\{0 < y < 1\}}.$$

On remarque que

$$f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y),$$

donc les variables aléatoires X et Y sont indépendantes.

2- Les densités conditionnelles sont

$$f_{X|Y}(x|y) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)} = \frac{2ye^{-x}}{2y} = e^{-x} \cdot 1_{\{x \geq 0\}},$$

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x, y)}{f_X(x)} = \frac{2ye^{-x}}{e^{-x}} = 2y \cdot 1_{\{0 < y < 1\}},$$

On observe que $f_{X|Y}(x|y) = f_X(x)$ et $f_{Y|X}(y|x) = f_Y(y)$ qui est un résultat immédiat de l'indépendance de X et Y .

3- La fonction de répartition conjointe est

$$F_{(X,Y)}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(u, v) dudv = \begin{cases} (1 - e^{-x})y^2, & \text{si } x \geq 0, 0 < y < 1 \\ 1 - e^{-x}, & \text{si } x \geq 0, y \geq 1 \\ 0, & \text{sinon} \end{cases}$$

et les fonctions de répartitions marginales sont

$$F_X(x) = \lim_{y \rightarrow +\infty} F_{(X,Y)}(x, y) = \begin{cases} 0, & \text{si } x < 0 \\ 1 - e^{-x}, & \text{si } x \geq 0 \end{cases}$$

$$F_Y(y) = \lim_{x \rightarrow +\infty} F_{(X,Y)}(x, y) = \begin{cases} 0, & \text{si } y < 0 \\ y^2, & \text{si } 0 \leq y < 1 \\ 1, & \text{si } 1 \leq y \end{cases}$$

on obtient aussi $F_{(X,Y)}(x, y) = F_X(x) \cdot F_Y(y)$.

2.4 Transformation d'un couple de variables aléatoires

Problème : On dispose d'un couple de variables aléatoires continue ou discret (X, Y) dont on connaît la loi conjointe et on voudrait connaître la loi de la variable aléatoire $h(X, Y)$, où h est une fonction continue définies de \mathbb{R} dans \mathbb{R} .

2.4.1 Transformation d'un couples de variables aléatoires continue

Considérons (X, Y) un couple de variables aléatoires admettant une fonction de densité jointe $f_{(X,Y)}$. Soient deux intervalles ouverts A et $B \subset \mathbb{R}^2$. Un difféomorphisme de A vers B est une **bi-jection** continûment différentiable ' $\varphi : A \rightarrow B$ dont la réciproque φ^{-1} est également continûment différentiable.

Théorème 2.4.1 (Théorème de changement de variables)

Soient (X, Y) un couple de variables aléatoires admettant la densité conjointe $f_{(X,Y)}$ et $A, B \subset \mathbb{R}^2$ des ouverts tels que

$$f_{(X,Y)}(x, y) = 0, \text{ si } (x, y) \notin A.$$

Soit h un difféomorphisme de A vers B .

Alors $Z = h(X, Y)$ est un couple de variables aléatoires (Z_1, Z_2) admettant la densité conjointe $f_{(Z_1, Z_2)}$ définie par

$$f_{(Z_1, Z_2)}(z_1, z_2) = f_{(X,Y)}(h^{-1}(z_1, z_2)) |Jac h^{-1}(z_1, z_2)|.$$

Remarque 2.4.1 pour une application différentiable h de B dans A

$$Jac h^{-1}(z_1, z_2) = \det \begin{pmatrix} \frac{\partial h_1^{-1}}{\partial z_1} & \frac{\partial h_1^{-1}}{\partial z_2} \\ \frac{\partial h_2^{-1}}{\partial z_1} & \frac{\partial h_2^{-1}}{\partial z_2} \end{pmatrix}.$$

Exemple 2.4.1 Soit (X, Y) un couple de variables aléatoires définies sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$, admettant une densité de probabilité par rapport à la mesure le Lebesgue sur \mathbb{R}^2 , donnée par

$$f_{(X,Y)}(x, y) = \frac{1}{4\pi} e^{-\frac{1}{2}(\frac{x^2}{2} + x^2 y + y^2)}.$$

1- Chercher les densités marginales f_X et f_Y .

2- En utilisant le théorème de changement de variable, chercher la densité de probabilité du couple aléatoire $(X, Y + \frac{X}{2})$.

Solution 2.4.1 1- Cherchons les densités marginales f_X et f_Y .

On a

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dy = \frac{1}{2\sqrt{2\pi}} e^{-\frac{x^2}{8}}.$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dx = \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}.$$

2- Cherchons la densité de probabilité du couple $(X, Y + \frac{X}{2})$

On introduit la fonction $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est bijective. En effet, pour tout $(u, v) \in \mathbb{R}^2$, on a

$$h(x, y) = (u, v) \Leftrightarrow \begin{cases} x = u \\ y + \frac{x}{2} = v \end{cases} \Leftrightarrow \begin{cases} x = u \\ y = v - \frac{u}{2} \end{cases}.$$

Donc en appliquant le théorème de changement de variable, on obtient

$$h^{-1}(u, v) = (h_1^{-1}(u, v), h_2^{-1}(u, v)) = \left(u, v - \frac{u}{2}\right),$$

$$\text{et } Jac h^{-1}(u, v) = \det \begin{pmatrix} \frac{\partial h_1^{-1}(u,v)}{\partial u} & \frac{\partial h_1^{-1}(u,v)}{\partial v} \\ \frac{\partial h_2^{-1}(u,v)}{\partial u} & \frac{\partial h_2^{-1}(u,v)}{\partial v} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} = 1.$$

D'où la densité demandée est

$$f_{(X, Y + \frac{X}{2})}\left(x, y + \frac{x}{2}\right) = f_{(X,Y)}\left(u, v - \frac{u}{2}\right) = \frac{1}{4\pi} e^{-\frac{1}{2}\left(\frac{u^2}{2} + u^2\left(v - \frac{u}{2}\right) + \left(v - \frac{u}{2}\right)^2\right)}.$$

Cas particulier des transformations des variables aléatoires

On s'intéresse à la loi de la somme de deux variables aléatoires continues X et Y .

Somme de variables aléatoires

Théorème 2.4.2 Soit (X, Y) un couple de variables aléatoires réelles, admettant une densité de probabilité conjointe $f_{(X,Y)}$. On définit la somme de ces deux variables aléatoire comme suit

$$Z = X + Y.$$

Alors la variable aléatoire Z admet la densité

$$f_Z(z) = \int_{\mathbb{R}} f_{(X,Y)}(z - y, y) dy.$$

Corollaire 2.4.1 Soient X et Y deux variables aléatoires réelles indépendantes, de densités marginales respectives f_X et f_Y

. Alors $Z = X + Y$ admet la densité de probabilité f_Z définie par

$$f_Z(z) = \int_{\mathbb{R}} f_X(z - y) f_Y(y) dy.$$

La densité f_Z est appelée produit de convolution de f_X et f_Y , noté $f_X * f_Y$.

Exemple 2.4.2 1- Somme de deux lois uniformes sur $[0, 1]$ indépendantes.

Soit X et Y deux variables aléatoires indépendantes telles que $X \sim U_{[0,1]}$ et $Y \sim U_{[0,1]}$, donc la loi de la variable aléatoire $Z = X + Y$ est donnée par

$$f_X * f_Y(z) = f_Z(z) = \int_{\mathbb{R}} f_X(z - y) f_Y(y) dy = f_Z(z) = \int_0^1 f_X(z - y) dy.$$

on a $f_X(z - y) = \begin{cases} 1, & \text{si } z - y \in [0, 1] \\ 0, & \text{sinon} \end{cases} = \begin{cases} 1, & \text{si } z - 1 \leq y \leq z \\ 0, & \text{sinon} \end{cases}$, donc

$$f_Z(z) = \int_{z-1}^z f_X(z - y) dy.$$

On en déduit :

$$f_Z(z) = \begin{cases} 0, & \text{si } z < 0 \\ \int_0^z dy = z, & \text{si } 0 < z < 1 \\ \int_{z-1}^1 dy = 2 - z, & \text{si } 1 < z < 2 \\ 0, & \text{si } z > 2. \end{cases}$$

Exemple 2.4.3 2- Somme de lois normales centrées réduites.

Soit X et Y deux variables aléatoires indépendantes définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que $X \sim \mathcal{N}(0, 1)$ et $Y \sim \mathcal{N}(0, 1)$. On cherche la densité de $Z = X + Y$.

Solution 2.4.2 On applique le corollaire précédent pour les densités de probabilité marginales suivantes

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

On obtient

$$\begin{aligned} f_Z(z) &= \int_{\mathbb{R}} f_X(z-y) f_Y(y) dy = \int_{-\infty}^{+\infty} dy \\ &= \lambda_1^2 e^{-\lambda_1 z} \int_0^{+\infty} e^{-(\lambda_1 + \lambda_2)y} dy \\ &= \lambda_1 \lambda_2 e^{-\lambda_1 z} \cdot 1_{\{z \geq 0\}} \end{aligned}$$

2.4.2 Transformation d'un couples de variables aléatoires discrètes

Soient X et Y deux variables aléatoires discrètes à valeurs dans $X(\Omega)$ et $Y(\Omega)$, et h une fonction définie de $X(\Omega) \times Y(\Omega)$ à valeurs dans \mathbb{R} , alors $h(X, Y)$ est une variable aléatoire discrète définie sur (Ω, \mathcal{A}) .

Théorème 2.4.3 Soit (X, Y) un couple de variables aléatoires discrètes, admettant une loi de probabilité conjointe $\mathbb{P}_{(X,Y)}$. La loi de la variable aléatoire discrète $Z = h(X, Y)$ est définie pour tout $z \in Z(\Omega)$ par

$$\mathbb{P}(Z = z) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ h(x,y)=z}} \mathbb{P}(X = x, Y = y).$$

Exemple 2.4.4 Soit (X, Y) un couple de variables aléatoires discrètes dont la loi de probabilité conjointe $\mathbb{P}_{(X,Y)}$ est donnée par le tableau de contingence suivant

$X \mid Y$	1	2	3	4
1	0.08	0.04	0.16	0.12
2	0.04	0.02	0.08	0.06
3	0.08	0.04	0.16	0.12

- 1- Déterminer les lois marginales du couple et préciser si X et Y sont indépendantes.
- 2- Déterminer la loi de probabilité de la variable aléatoire $Z = XY$.

Solution 2.4.3 1- On détermine les lois marginales \mathbb{P}_X et \mathbb{P}_Y .

$X \mid Y$	1	2	3	4	$\mathbb{P}_X(x)$
1	0.08	0.04	0.16	0.12	0.40
2	0.04	0.02	0.08	0.06	0.20
3	0.08	0.04	0.16	0.12	0.40
$\mathbb{P}_Y(y)$	0.2	0.1	0.4	0.3	1

2- La loi de $Z = XY$. on a $Z(\Omega) = \{1, 2, 3, 4, 6, 8, 9, 12\}$

$$\mathbb{P}(Z = 1) = \mathbb{P}(X = 1, Y = 1) = 0.08, \mathbb{P}(Z = 2) = \mathbb{P}(X = 1, Y = 2) + \mathbb{P}(X = 2, Y = 1) = 0.08$$

$$\mathbb{P}(Z = 3) = \mathbb{P}(X = 1, Y = 3) + \mathbb{P}(X = 3, Y = 1) = 0.24, \mathbb{P}(Z = 4) = \mathbb{P}(X = 1, Y = 4) + \mathbb{P}(X = 2, Y = 2) = 0.14,$$

$$\mathbb{P}(Z = 6) = \mathbb{P}(X = 2, Y = 3) + \mathbb{P}(X = 3, Y = 2) = 0.12, \mathbb{P}(Z = 8) = \mathbb{P}(X = 2, Y = 4) = 0.06,$$

$$\mathbb{P}(Z = 9) = \mathbb{P}(X = 3, Y = 3) = 0.16, \mathbb{P}(Z = 12) = \mathbb{P}(X = 3, Y = 4) = 0.12,$$

on a $\sum_{z \in Z(\Omega)} \mathbb{P}_Z(z) = 1$, donc on a défini la loi de probabilité de $Z = XY$.

Cas particuliers des transformations des variables aléatoires

On s'intéresse à la loi de la somme et du produit de deux variables aléatoires continues X et Y .

1- Somme de variables aléatoires

Théorème 2.4.4 Soit (X, Y) un couple de variables aléatoires discrètes, admettant une loi de probabilité conjointe $\mathbb{P}_{(X,Y)}$. On définit la somme de ces deux variables aléatoire comme suit

$$Z = X + Y.$$

La loi de probabilité de la variable aléatoire discrète Z est donnée pour tout $z \in Z(\Omega) = (X + Y)(\Omega)$ par

$$\mathbb{P}(Z = z) = \sum_{x \in X(\Omega), z-x \in Y(\Omega)} \mathbb{P}(X = x, Y = z - x).$$

Exemple 2.4.5 Soient X et Y deux variables aléatoires dont on suppose que la loi du couple est donnée

X/Y	1	2	3
1	2α	3α	3α
2	3α	2α	3α
3	3α	3α	2α

1- Déterminer la valeur de α .

2- Calculer la loi de probabilité de $Z = X + Y$.

Exemple 2.4.6 1- On détermine la valeur de α : on a par supposition $\mathbb{P}_{(X,Y)}$ est une loi de probabilité conjointe, alors $\mathbb{P}_{(X,Y)}(x, y) \geq 0$ et $\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} \mathbb{P}_{(X,Y)}(x, y) = 1$. Par conséquent $\alpha \geq 0$ et $\alpha = \frac{1}{24}$.

2- La loi de $Z = X + Y$. On détermine d'abord l'ensemble des valeurs possibles de la variable aléatoire Z , $Z(\Omega) = \{2, 3, 4, 5, 6\}$, puis on applique le théorème (), pour calculer les probabilités $\mathbb{P}_Z(z)$:

$$\mathbb{P}(Z = 2) = \mathbb{P}(X = 1, Y = 1) = \frac{1}{12}, \mathbb{P}(Z = 3) = \mathbb{P}(X = 1, Y = 2) + \mathbb{P}(X = 2, Y = 1) = \frac{1}{4},$$

$$\mathbb{P}(Z = 4) = \mathbb{P}(X = 1, Y = 3) + \mathbb{P}(X = 3, Y = 1) + \mathbb{P}(X = 2, Y = 2) = \frac{1}{3},$$

$$\mathbb{P}(Z = 5) = \mathbb{P}(X = 2, Y = 3) + \mathbb{P}(X = 3, Y = 2) = \frac{1}{4}, \mathbb{P}(Z = 6) = \mathbb{P}(X = 3, Y = 3) = \frac{1}{12}.$$

on voit que

$$\sum_{z \in Z(\Omega)} \mathbb{P}_Z(z) = 1.$$

donc on a défini la loi de probabilité de Z .

Proposition 2.4.1 *Si X et Y sont deux variables aléatoires discrètes et indépendantes, la loi de la somme $Z = X + Y$ est donné par*

$$\mathbb{P}(Z = z) = \sum_{x \in X(\Omega), z-x \in Y(\Omega)} \mathbb{P}(X = x) \mathbb{P}(Y = z - x).$$

Exemple 2.4.7 *Soient X et Y deux variables aléatoires discrètes indépendantes telles que $X \sim B(n, p)$ et $Y \sim B(m, p)$ c'est à dire $X(\Omega) = \{0, 1, 2, \dots, n\}$ et $Y(\Omega) = \{0, 1, 2, \dots, m\}$ avec*

$$\begin{aligned} \forall x \in X(\Omega), \mathbb{P}(X = x) &= C_n^x \left(\frac{1}{4}\right)^x \left(1 - \frac{1}{4}\right)^{n-x} \\ \forall y \in Y(\Omega), \mathbb{P}(Y = y) &= C_m^y \left(\frac{1}{4}\right)^y \left(1 - \frac{1}{4}\right)^{m-y} \end{aligned}$$

et soit $Z = X + Y$. La loi de la variable aléatoire discrète Z est définie pour tout $z \in Z(\Omega) = \{0, 1, 2, 3, 4, 5, \dots, n + m\}$ par

$$\begin{aligned} \mathbb{P}(Z = z) &= \sum_{x=0}^z \mathbb{P}(X = x) \mathbb{P}(Y = z - x) \\ &= \sum_{x=0}^z C_n^x \left(\frac{1}{4}\right)^x \left(1 - \frac{1}{4}\right)^{n-x} C_m^{z-x} \left(\frac{1}{4}\right)^{z-x} \left(1 - \frac{1}{4}\right)^{m-(z-x)} \\ &= \sum_{x=0}^z C_n^x C_m^{z-x} \left(\frac{1}{4}\right)^{x+z-x} \left(1 - \frac{1}{4}\right)^{(n+m)-(x+z-x)} \\ &= \sum_{x=0}^z C_n^x C_m^{z-x} \left(\frac{1}{4}\right)^z \left(1 - \frac{1}{4}\right)^{(n+m)-z} = \left(\frac{1}{4}\right)^z \left(1 - \frac{1}{4}\right)^{(n+m)-z} \sum_{x=0}^z C_n^x C_m^{z-x} \\ &= C_{n+m}^z \left(\frac{1}{4}\right)^z \left(1 - \frac{1}{4}\right)^{(n+m)-z}, \end{aligned}$$

d'où

$$Z \sim B(n + m, p).$$

2.4.3 Covariance et Corrélation linéaire de deux variables aléatoires

Dans ce paragraphe on s'intéresse à la relation qui peut exister entre les deux variables aléatoires X et Y . En particulier, la relation linéaire entre X et Y .

Définition 2.4.1 (*Covariance*)

Soient X et Y deux variables aléatoires de carré intégrable. on définit la covariance de X et Y par

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

La covariance $\text{Cov}(X, Y)$ est une quantité réelle, lorsque elle est positive ($\text{Cov}(X, Y) > 0$) on dira que X et Y varient dans le même sens et quand elle est négative ($\text{Cov}(X, Y) < 0$) on dira que X et Y varient dans deux sens différents. Quand elle est nulle ($\text{Cov}(X, Y) = 0$) on dira que X et Y ne sont pas corrélés.

Propriétés de covariance :

Soient X, Y et Z trois variables aléatoires réelles admettant un moment d'ordre 2, alors

1- **La positivité :** $Cov(X, X) = Var(X) \geq 0$.

2- **La symétrie :** $Cov(X, Y) = Cov(Y, X)$.

3- **La bilinéarité :** pour tous réels a et b , on a

la linéarité à gauche : $Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$.

la linéarité à droite : $Cov(X, aY + bZ) = aCov(X, Y) + bCov(X, Z)$.

4- Si X ou Y est centrée, alors $Cov(X, Y) = \mathbb{E}(X, Y)$.

5- $Cov(X, a) = 0$;

6- Formule de **Huygens** : $Cov(X, Y) = \mathbb{E}(X.Y) - \mathbb{E}(X)\mathbb{E}(Y)$.

7- Si X et Y sont indépendants, alors $Cov(X, Y) = 0$.

6) $|Cov(X, Y)| \leq \sqrt{Var(X)Var(Y)}$.

Proposition 2.4.2 Soient X et Y deux variables aléatoires de carré intégrable, on a

$$var(X + Y) = var(X) + var(Y) + 2Cov(X, Y)$$

$$var(X - Y) = var(X) + var(Y) - 2Cov(X, Y)$$

1- Si X et Y sont indépendants alors

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

par conséquent

$$1 - Cov(X, Y) = 0.$$

$$2 - var(X + Y) = var(X - Y) = var(X) + var(Y)$$

Définition 2.4.2 (Espérance et matrice de covariance)

- Soient X et Y deux variables aléatoires intégrables, alors l'espérance du couple (X, Y) , notée $\mathbb{E}(X, Y)$ est le vecteur donné par

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

- Soient X et Y deux variables aléatoires de carré intégrables. La matrice de covariance du couple (X, Y) est la matrice définie par

$$K_{(X,Y)} = \begin{pmatrix} cov(X, X) & cov(X, Y) \\ cov(Y, X) & cov(Y, Y) \end{pmatrix} = \begin{pmatrix} var(X) & cov(X, Y) \\ cov(X, Y) & var(Y) \end{pmatrix}.$$

Définition 2.4.3 (coefficient de corrélation linéaire)

On appelle coefficient de corrélation linéaire de X et Y , le réel, noté $\rho(X, Y)$, défini par

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

Le coefficient de corrélation linéaire mesure la dépendance linéaire entre deux variables aléatoires.

Remarque 2.4.2 1- $|\rho(X, Y)| \leq 1$.

2- Si le coefficient de corrélation linéaire $|\rho(X, Y)| = 1$, on dit que X et Y sont corrélées linéairement. C'est à dire qu'il existe $(a, b) \in \mathbb{R}_*^2$ pour lequel $X = aY + b$.

3- Si $\rho(X, Y)$ est proche de 1 ou de -1 , la dépendance linéaire entre X et Y est importante. Par exemple X = nombre de remises promotionnelles, Y = volume des ventes.

4- Le signe du coefficient de corrélation linéaire $\rho(X, Y)$ indique si les variations vont dans le même sens ou dans le sens opposé. Exemple pour $\rho(X, Y) > 0$: X = température extérieure, Y = consommation de crèmes glacées. Exemple pour $\rho(X, Y) < 0$: X = température extérieure, Y = facture de chauffage.

4- Si le coefficient de corrélation linéaire $\rho(X, Y) = 0$, on dit que X et Y sont dites "**non corrélées linéairement**".

Exemple 2.4.8 Soient X et Y deux variables aléatoires définies dans l'exemple et on calcule $\text{cov}(X, Y)$.

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}_X(x) = \sum_{x=1}^3 x \mathbb{P}(X = x) = 2.$$

$$\mathbb{E}(Y) = \sum_{y \in Y(\Omega)} y \mathbb{P}_Y(y) = \sum_{y=1}^4 y \mathbb{P}(Y = y) = 2.8.$$

$$\mathbb{E}(X.Y) = \mathbb{E}(Z) = \sum_{z \in Z(\Omega)} z \mathbb{P}_Z(z) = \sum_{z \in Z(\Omega)} z \mathbb{P}(Z = z) = 5.4$$

$$\text{Cov}(X, Y) = \mathbb{E}(X.Y) - \mathbb{E}(X)\mathbb{E}(Y) = -0.2$$

On remarque que $\text{Cov}(X, Y) < 0$, donc la relation entre X et Y est négative. on a le vecteur des moyennes est

$$\mathbb{E}(X, Y) = (\mathbb{E}(X), \mathbb{E}(Y)) = (2, 2.8)$$

et la matrice de covariance est

$$K_{(X,Y)} = \begin{pmatrix} \text{var}(X) & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \text{var}(Y) \end{pmatrix} = \begin{pmatrix} 0.8 & -0.2 \\ -0.2 & 1.16 \end{pmatrix}.$$

Maintenant, on calcule le coefficient de corrélation linéaire

$$\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2 \mathbb{P}_X(x) = \sum_{x=1}^3 x^2 \mathbb{P}(X = x) = 4.8 \times$$

$$\mathbb{E}(Y^2) = \sum_{y \in Y(\Omega)} y^2 \mathbb{P}_Y(y) = \sum_{y=1}^4 y^2 \mathbb{P}(Y = y) = 9$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = 0.8$$

$$\text{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) = 1.16$$

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-0.2}{\sqrt{0.8 \times 1.16}} =$$

on remarque que

2.5 Principales lois de probabilité

Chapitre 3

Fonctions caractéristiques et génératrices

Ce chapitre est consacré à l'étude des deux fonctions importantes dans la théorie des probabilités, caractérisant la loi d'une variable aléatoire. On parle de la fonction caractéristique et la fonction génératrice des moments.

3.1 Fonctions caractéristiques

On va définir et donner les propriétés d'une nouvelle fonction qui permettra de caractériser la loi d'une variable aléatoire, mais de façon plus intéressante que la fonction de répartition. En théorie des probabilités, la fonction caractéristique d'une variable aléatoire permet de calculer plus simplement les moments de la variable ; en utilisant simplement les dérivées successives de cette fonction au point 0. La force des fonctions caractéristiques est que deux variables aléatoires ont la même loi si et seulement si leurs fonctions caractéristiques coïncident.

Définition 3.1.1 (fonction caractéristique)

Soit X une variable aléatoire. On appelle fonction caractéristique de X , et on la note φ_X , la fonction définie

sur \mathbb{R} , à valeurs complexes, par :

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}(\cos(tX)) + i\mathbb{E}(\sin(tX)), \forall t \in \mathbb{R}.$$

- Pour X une v.a. discrète la fonction caractéristique est définie comme suit :

$$\varphi_X(t) = \sum_{x_j \in X(\Omega)} e^{itx_j} \mathbb{P}(X = x_j).$$

- Pour X une v.a. absolument continue de densité de probabilité f_X , la fonction caractéristique est définie comme suit :

$$\varphi_X(t) = \int_{\Omega} e^{itX(\omega)} d\mathbb{P}(\omega) = \int_{\mathbb{R}} e^{itx} f_X(x) dx.$$

Exemple 3.1.1 Soit X une v.a. de loi de Bernoulli de paramètre p , c'est à dire

$$\forall k \in \{0, 1\}, \mathbb{P}(X = k) = p^k (1 - p)^{1-k}.$$

La fonction caractéristique de la v.a. discrète X est

$$\begin{aligned}\varphi_X(t) &= \mathbb{E}(e^{itX}) = \sum_{x_j \in X(\Omega)} e^{itx_j} \mathbb{P}(X = x_j) \\ &= e^{it(0)} \mathbb{P}(X = 0) + e^{it(1)} \mathbb{P}(X = 1) \\ &= (1 - p) + pe^{it}.\end{aligned}$$

Exemple 3.1.2 Soit X une v.a. de loi de Uniforme sur l'intervalle $[a, b]$, de densité

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b] \\ 0, & \text{sinon} \end{cases}.$$

La fonction caractéristique de X est

$$\begin{aligned}\varphi_X(t) &= \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx} f_X(x) dx \\ &= \int_a^b \frac{1}{b-a} e^{itx} dx \\ &= \frac{e^{itb} - e^{ita}}{it(b-a)}.\end{aligned}$$

Remarque 3.1.1 Dans le chapitre précédent on a vu que l'espérance mathématique n'existe pas toujours. Or, dans le présent on peut montrer que la fonction caractéristique est toujours définie, car

- Si X est une v.a. discrète avec $X(\Omega)$ est un ensemble dénombrable, alors

$$\begin{aligned}|\varphi_X(t)| &= \left| \sum_{x_j \in X(\Omega)} e^{itx_j} \mathbb{P}(X = x_j) \right| \\ &\leq \sum_{j=1}^{\infty} |e^{itx_j}| \mathbb{P}(X = x_j) \\ &\leq \sum_{j=1}^{\infty} \mathbb{P}(X = x_j) = 1,\end{aligned}$$

c'est à dire que la série est uniformément convergente, par conséquent elle est convergente.

- Si X est une v.a. continue, alors

$$|\varphi_X(t)| = \left| \int_{\mathbb{R}} e^{itx} f_X(x) dx \right| \leq \int_{\mathbb{R}} |e^{itx}| f_X(x) dx \leq \int_{\mathbb{R}} f_X(x) dx = 1,$$

donc cette intégrale est uniformément convergente.

Proposition 3.1.1 Soient X une variable aléatoire, a et b deux réels. Alors les propriétés suivantes sont toujours vraies :

- 1- La fonction φ_X existe pour toute distribution de la v.a. X .
- 2- $\varphi_X(0) = 1$.
- 3- $|\varphi_X(t)| \leq 1$ pour tout $t \in \mathbb{R}$.
- 4- φ_X est uniformément continue, c'est-à-dire $\forall \varepsilon > 0, \exists \delta > 0$ tel que $|\varphi_X(t) - \varphi_X(s)| < \varepsilon, |t - s| \leq \delta$.
- 5- Pour tout $t \in \mathbb{R}, \varphi_X(-t) = \overline{\varphi_X(t)}$.
- 6- Pour tout $t \in \mathbb{R}, \varphi_{aX+b}(t) = e^{itb} \varphi_X(at)$.
- 7- Pour tout $t \in \mathbb{R}, \varphi_{-X}(t) = \overline{\varphi_X(t)}$, c'est-à-dire φ_{-X} est le conjugué de φ_X .
- 8- La fonction caractéristique φ_X est à valeurs réelles si et seulement si la distribution de X est symétrique autour de 0, c'est-à-dire ssi $\mathbb{P}(X > x) = \mathbb{P}(X < -x), \forall x \geq 0$.

Preuve. 1- Voir la remarque

- 2- On a $\varphi_X(0) = \mathbb{E}(e^{i(0)X}) = \mathbb{E}(1) = 1$.
- 3- Soit $t \in \mathbb{R}, |\varphi_X(t)| = |\mathbb{E}(e^{itX})| \leq \mathbb{E}|e^{itX}| = 1$, car $|e^{itx}| = \sqrt{\cos^2(tx) + \sin^2(tx)} = 1$.
- 4- TD.
- 5- Soit $t \in \mathbb{R}, \varphi_X(-t) = \mathbb{E}(e^{i(-t)X}) = \mathbb{E}(\cos(-tX)) + i\mathbb{E}(\sin(-tX)) = \mathbb{E}(\cos(tX)) - i\mathbb{E}(\sin(tX)) = \overline{\varphi_X(t)}$.
- 6- Soit $t \in \mathbb{R}, \varphi_{aX+b}(t) = \mathbb{E}(e^{it(aX+b)}) = \mathbb{E}(e^{itaX} \cdot e^{itb}) = e^{itb} \varphi_X(at)$.
- 7- Soit $t \in \mathbb{R}, \varphi_{-X}(t) = \mathbb{E}(e^{it(-X)}) = \mathbb{E}(\cos(-tX)) + i\mathbb{E}(\sin(-tX)) = \mathbb{E}(\cos(tX)) - i\mathbb{E}(\sin(tX)) = \overline{\varphi_X(t)}$.
- 8-TD. ■

3.1.1 Fonctions caractéristiques des lois de probabilité usuelles

Maintenant on va donner les expressions des fonctions caractéristiques de quelques lois de probabilité déjà vues dans le chapitre précédent.

lois de probabilité de X	fonctions caractéristiques φ_X
loi uniforme $\mathcal{U}(n)$	$\varphi_X(t) = \begin{cases} \frac{1-e^{itn}}{n(e^{-it}-1)}, & \text{si } t \neq 2\pi k \text{ pour tout } k \in \mathbb{Z} \\ 1, & \text{si } t = 2\pi k, k \in \mathbb{Z} \end{cases}$
loi de Bernoulli $\mathcal{B}(p)$	$\varphi_X(t) = pe^{it} + 1 - p.$
loi binomiale $\mathcal{B}(n, p)$	$\varphi_X(t) = (pe^{it} + 1 - p)^n.$
loi de Poisson $\mathcal{P}(\lambda)$	$\varphi_X(t) = e^{\lambda(e^{it}-1)}.$
loi géométrique $\mathcal{G}(p)$	$\varphi_X(t) = \frac{p}{1-(1-p)e^{it}}.$
loi uniforme $\mathcal{U}([a, b])$	$\varphi_X(t) = \frac{e^{ibt}-e^{iat}}{it(b-a)} \text{ si } t \neq 0, \varphi_X(0) = 1.$
loi exponentielle $\mathcal{E}(\lambda)$	$\varphi_X(t) = \frac{\lambda}{\lambda-it}.$
loi de Cauchy	$\varphi_X(t) = e^{- t }.$
loi normale $\mathcal{N}(\mu, \sigma)$	$\varphi_X(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}}.$

Remarque 3.1.2 On a les cas particuliers suivants, pour tout $t \in \mathbb{R}$

- 1- $X \sim \mathcal{U}([-a, a])$, alors $\varphi_X(t) = \frac{e^{iat}-e^{-iat}}{-2it} = \frac{\sin(at)}{at}$ si $t \neq 0, \varphi_X(0) = 1$.
- 2- $X \sim \mathcal{N}(0, 1)$, alors $\varphi_X(t) = e^{-\frac{t^2}{2}}$.

3.1.2 Propriétés des fonctions caractéristiques

Théorème 3.1.1 (théorème d'unicité)

Soient X et Y deux variables aléatoires. Si $\varphi_X = \varphi_Y$ alors $X \stackrel{L}{=} Y$.

Le théorème suivant donne la formule d'inversion qui permet d'obtenir la fonction de répartition F_X d'une variable aléatoire X à partir de sa fonction caractéristique φ_X .

Théorème 3.1.2 (Formule inverse)

Soit X une variable aléatoire de fonction de répartition F_X et de fonction caractéristique φ_X . Pour $a, b \in \mathbb{R}$, tels que $a < b$ on a

$$F_X(b) - F_X(a) + \frac{\mathbb{P}(X=a) - \mathbb{P}(X=b)}{2} = \lim_{T \rightarrow \infty} \int_{-T}^{+T} \frac{e^{-ibt} - e^{-iat}}{-it} \varphi_X(t) dt.$$

En particulier, pour $\mathbb{P}(X=a) = \mathbb{P}(X=b) = 0$ (c'est à dire si a et b sont deux points de continuité de F_X), on a

$$F_X(b) - F_X(a) = \lim_{T \rightarrow \infty} \int_{-T}^{+T} \frac{e^{-ibt} - e^{-iat}}{-it} \varphi_X(t) dt.$$

Pour les deux types de v.a. discret ou continu, le théorème qui suit donne les formules inverses pour obtenir la loi de probabilité de la v.a. X .

Théorème 3.1.3 Soit φ_X la fonction caractéristique de la variable aléatoire X . Si φ_X est intégrable, c'est à dire

$$\int_{\mathbb{R}} |\varphi_X(t)| dt < +\infty,$$

alors

Si X a une distribution absolument continue de densité de probabilité bornée $f = F'$, cette densité est définie sur \mathbb{R} par

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t) dt, \text{ pour tout } x \in \mathbb{R}.$$

Exemple 3.1.3 Soit φ_X la fonction caractéristique de la loi normale centrée réduite $\mathcal{N}(0,1)$, donnée par

$$\varphi_X(t) = e^{-\frac{t^2}{2}}.$$

Cette fonction est intégrable car

$$\int_{\mathbb{R}} |\varphi_X(t)| dt = \int_{\mathbb{R}} \left| e^{-\frac{t^2}{2}} \right| dt = \int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} < \infty.$$

Donc la densité de probabilité de la v.a. X est

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t) dt = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} e^{-\frac{t^2}{2}} dt = \frac{1}{2\pi} e^{-\frac{x^2}{2}} \int_{\mathbb{R}} e^{-\frac{(t-ix)^2}{2}} dt = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Remarque 3.1.3 La fonction caractéristique X peut ne pas être intégrable bien que X est a densité. C'est le cas de la loi exponentielle de paramètre $\lambda > 0$.

Théorème 3.1.4 Soit φ_X la fonction caractéristique de la variable aléatoire discrète X d'ensemble de valeurs possible $X(\Omega)$, sa loi de probabilité est définie pour tout $x \in X(\Omega)$ par

$$\mathbb{P}(X = x) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} e^{-itx} \varphi_X(t) dt.$$

Proposition 3.1.2 Si le moment d'ordre k d'une variable aléatoire X existe (c'est à dire si $\mathbb{E}|X|^k < \infty$), alors la fonction caractéristique de X est k fois dérivable et :

$$\mathbb{E}(X^k) = \frac{1}{i^k} \varphi_X^{(k)}(0).$$

Remarque 3.1.4 Dans certains cas où les calculs directs sont complexes, cette proposition permet d'obtenir rapidement $\mathbb{E}(X^k)$. En particulier $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$

$$\mathbb{E}(X) = -i\varphi_X^{(1)}(0) \text{ et } \mathbb{E}(X^2) = -\varphi_X^{(2)}(0).$$

Exemple 3.1.4 On considère X une v.a. qui suit une loi binomiale $B(n, p)$. À partir de sa fonction caractéristique on calculera son espérance mathématique et sa variance.

$$\begin{aligned} \varphi_X(t) &= \mathbb{E}(e^{itX}) = \sum_{k=0}^n e^{itk} \mathbb{P}(X = k) = \sum_{k=0}^n e^{itk} C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n C_n^k (e^{it}p)^k (1-p)^{n-k} = (e^{it}p + (1-p))^n \end{aligned}$$

La première dérivée $\varphi_X^{(1)}(t)$ s'écrit

$$\varphi_X^{(1)}(t) = n(e^{it}p + (1-p))^{n-1} e^{it}pi$$

d'où

$$\mathbb{E}(X) = \frac{1}{i} \varphi_X^{(1)}(0) = np$$

La dérivée seconde $\varphi_X^{(2)}(t)$ est

$$\varphi_X^{(2)}(t) = n(n-1)(e^{it}p + (1-p))^{n-2} p^2 i^2 + npi^2 \cdot (e^{it}p + (1-p))^{n-1}.$$

donc

$$\varphi_X^{(2)}(0) = n(n-1)p^2 i^2 + npi^2 = -n(n-1)p^2 - np,$$

le moment d'ordre 2 est

$$\mathbb{E}(X^2) = \frac{1}{i^2} \varphi_X^{(2)}(0) = n(n-1)p^2 + np,$$

par conséquent la variance de X est

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

Exemple 3.1.5 On considère Z une v.a. qui suit une loi normale $\mathcal{N}(0, 1)$ et X une v.a. qui suit une loi normale $\mathcal{N}(\mu, \sigma^2)$. À partir de la fonction caractéristique de Z , on calculera la fonction caractéristique de X et on en retrouvera l'espérance mathématique et la variance de X .

Proposition 3.1.3 Soient X_1, \dots, X_n des variables aléatoires indépendantes, alors

$$\varphi_{X_1 + \dots + X_n}(t) = \prod_{j=1}^n \varphi_{X_j}(t).$$

Si de plus les X_1, \dots, X_n sont identiquement distribuées, alors

$$\varphi_{X_1 + \dots + X_n}(t) = [\varphi_{X_1}(t)]^n.$$

Preuve. 1- Soient X_1, \dots, X_n des v.a. indépendantes et $Y = X_1 + \dots + X_n$. la fonction caractéristique φ_Y est définie pour tout $t \in \mathbb{R}$, par

$$\varphi_Y(t) = \mathbb{E}(e^{itY}) = \mathbb{E}(e^{it(X_1 + \dots + X_n)}) = \mathbb{E}\left(\prod_{j=1}^n e^{itX_j}\right),$$

comme les v.a. X_1, \dots, X_n sont indépendantes alors les v.a. $e^{itX_1}, \dots, e^{itX_n}$ sont indépendantes aussi, alors

$$\varphi_Y(t) = \prod_{j=1}^n \mathbb{E}(e^{itX_j}) = \prod_{j=1}^n \varphi_{X_j}(t).$$

2- Si X_1, \dots, X_n des v.a. indépendantes et identiquement distribuées, alors

$$\varphi_Y(t) = [\varphi_{X_1}(t)]^n.$$

■

Remarque 3.1.5 La proposition précédente permet de démontrer les résultats suivants :

- 1- Soient $X \sim \mathcal{B}(n_1, p)$ et $Y \sim \mathcal{B}(n_2, p)$. Si $X \perp Y$, alors $X + Y \sim \mathcal{B}(n_1 + n_2, p)$.
- 2- Soient $X \sim \mathcal{P}(\lambda_1)$ et $Y \sim \mathcal{P}(\lambda_2)$. Si $X \perp Y$, alors $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
- 3- Soient $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Si $X \perp Y$, alors $X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Maintenant on peut obtenir la fonction caractéristique à partir des moments d'ordre k .

Proposition 3.1.4 Soit X une variable aléatoire. Si $\mathbb{E}(X^k) < \infty$ pour tout $k \in \mathbb{N}$, alors

$$\varphi_X(t) = \sum_{k=0}^{\infty} \frac{(ik)^k}{k!} \mathbb{E}(X^k).$$

c'est à dire que la fonction caractéristique est entièrement déterminée par la suite $(\mathbb{E}(X^k), k \in \mathbb{N})$.

3.1.3 Fonction caractéristique d'un couple de variables aléatoires

Définition 3.1.2 La fonction caractéristique d'un couple de variables aléatoires (X_1, X_2) est la fonction notée $\varphi_{(X_1, X_2)}$ définie de \mathbb{R}^2 dans \mathbb{C} par

$$\varphi_{(X_1, X_2)}(t_1, t_2) = \mathbb{E}(e^{i(t_1 X_1 + t_2 X_2)}) = \mathbb{E}(\cos(t_1 X_1 + t_2 X_2)) + i \mathbb{E}(\sin(t_1 X_1 + t_2 X_2)), \forall (t_1, t_2) \in \mathbb{R}^2.$$

- Pour (X_1, X_2) un couple de v.a. discrètes, la fonction caractéristique est définie comme suit

$$\varphi_{(X_1, X_2)}(t_1, t_2) = \sum_{x_i \in X_1(\Omega)} \sum_{x_j \in X_2(\Omega)} e^{i(t_1 x_i + t_2 x_j)} \mathbb{P}(X_1 = x_i, X_2 = x_j).$$

- Pour (X_1, X_2) un couple de v.a. absolument continu de densité de probabilité conjointe $f_{(X_1, X_2)}$, la fonction caractéristique est définie comme suit

$$\varphi_{(X_1, X_2)}(t_1, t_2) = \int_{\mathbb{R}} e^{i(t_1 x_1 + t_2 x_2)} f_{(X_1, X_2)}(x_1, x_2) dx_1 dx_2.$$

Proposition 3.1.5 Soit (X_1, X_2) un couple de variables aléatoires. Les marginales X_1 et X_2 sont indépendantes si et seulement si

$$\varphi_{(X_1, X_2)}(t_1, t_2) = \varphi_{X_1}(t_1) \cdot \varphi_{X_2}(t_2).$$

3.2 Fonctions génératrices des moments

Le fonction génératrice des moments est un outil puissant en théorie des probabilités et en statistique qui nous permet d'étudier les propriétés de variables aléatoires. Il fournit un moyen de générer des moments d'une variable aléatoire en prenant la dérivées de la fonction.

Définition 3.2.1 (fonction génératrice des moments)

La fonction génératrice des moments d'une variable aléatoire discrète, à valeurs possibles entières non négative, X est la fonction notée G_X , définie de $[-1, 1]$ dans $[0, 1]$ par

$$G_X(s) = \mathbb{E}(s^X),$$

avec la convention $0^0 = 1$.

Exemple 3.2.1 Soit X la v.a. de loi de poisson de paramètre λ , alors $X(\Omega) = \{0, 1, 2, \dots\}$ et

$$\forall k \in X(\Omega), \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

La fonction génératrice des moments de X est

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) = \sum_{k=0}^{\infty} s^k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{(s\lambda)^k}{k!} = e^{\lambda(s-1)}, |s| < 1.$$

Exemple 3.2.2 Soit X la v.a. de loi de Bernoulli de paramètre p , alors $X(\Omega) = \{0, 1\}$ et

$$\forall k \in X(\Omega), \mathbb{P}(X = k) = p^k (1 - p)^{1-k}.$$

La fonction génératrice des moments de X est

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^1 s^k \mathbb{P}(X = k) = s^0 \mathbb{P}(X = 0) + s^1 \mathbb{P}(X = 1) = 1 + p(s - 1).$$

Proposition 3.2.1 Soit X une variable aléatoire discrète à valeurs dans \mathbb{N} , sa fonction génératrice des moments est G_X définie par

$$G_X(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k).$$

On a

1- Cette série est absolument convergente pour tout $s \in [-1, 1]$.

2- $G_X(1) = 1$.

3- Deux variables aléatoires $X, Y : \Omega \rightarrow \mathbb{N}$ ont même loi si et seulement si $G_X(s) = G_Y(s)$ pour tout $s \in [0, \delta]$ pour un certain $\delta > 0$.

Preuve. 1- La série $\sum_{k=0}^{\infty} s^k \mathbb{P}(X = k)$ est absolument convergente si

$$\left| \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) \right| < \infty,$$

en effet,

$$\begin{aligned} \left| \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) \right| &\leq \sum_{k=0}^{\infty} |s^k \mathbb{P}(X = k)| \leq \sum_{k=0}^{\infty} |s^k| \mathbb{P}(X = k) \\ &\leq \sum_{k=0}^{\infty} \mathbb{P}(X = k) = 1. \end{aligned}$$

car $|s| \leq 1$.

2- $G_X(1) = \mathbb{E}(1^X) = 1$.

3-TD ■

La fonction G_X est appelée fonction génératrice des moments du fait que tous les moments d'ordre k de X peuvent être calculés en dérivant k fois G_X puis en évaluant la dérivée au point $s = 1$.

Théorème 3.2.1 Soit X une variable aléatoire discrète de Ω dans \mathbb{N} , sa fonction génératrice est G_X .

1- Si X admet une espérance, alors $G_X(s)$ admet une dérivée à gauche $G'_X(1)$ en $s = 1$, et l'on a

$$\mathbb{E}(X) = G'_X(1).$$

2- Si X admet un moment d'ordre 2, alors $G_X(s)$ admet une dérivée seconde à gauche $G_X^{(2)}(1)$ en $s = 1$, et l'on a

$$\text{Var}(X) = G_X^{(2)}(1) + G'_X(1) - \left[G'_X(1)\right]^2.$$

3- Inversement si $G_X(s)$ est une fois (resp. 2 fois) dérivable en $s = 1$ alors X est intégrable et $\mathbb{E}(X) = G'_X(1)$ (resp. 2 fois dérivable et $\text{Var}(X) = G_X^{(2)}(1) + G'_X(1) - \left[G'_X(1)\right]^2$).

Exemple 3.2.3 Soit X la v.a. de loi de poisson de paramètre λ . Calculer l'espérance et la variance.

Proposition 3.2.2 Soient $X : \Omega \rightarrow \mathbb{N}$ et $Y : \Omega \rightarrow \mathbb{N}$ deux variables aléatoires discrètes de fonctions génératrices G_X et G_Y . Si X et Y sont indépendantes alors

$$\forall s \in]-1, 1[, G_{X+Y}(s) = G_X(s) G_Y(s).$$

Preuve. On a $G_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X s^Y) = \mathbb{E}(s^X) \mathbb{E}(s^Y) = G_X(s) G_Y(s)$ par indépendance des deux v.a. ■

Proposition 3.2.3 Soit X une variable aléatoire discrète d'ensemble de valeurs possibles $X(\Omega) \subset \mathbb{N}$. La fonction génératrice des moments de X caractérise la loi de X . Plus précisément :

$$\forall k \in X(\Omega), \mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}.$$

Exemple 3.2.4 Soit X la v.a. à valeurs dans \mathbb{N} et de fonction génératrice G_X de la forme

$$G_X(s) = e^{\lambda(s-1)}, \lambda > 0, s \in [-1, 1].$$

- Déterminer la loi de probabilité de X .

En effet, on a d'après la proposition précédente

$$\forall k \in X(\Omega), \mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!},$$

donc, pour $k = 0$ on obtient

$$\mathbb{P}(X = 0) = \frac{G_X^{(0)}(0)}{0!} = e^{\lambda(0-1)} = e^{-\lambda},$$

pour $k = 1$ on obtient

$$\mathbb{P}(X = 1) = \frac{G_X^{(1)}(0)}{1!} = \lambda e^{\lambda(0-1)} = \lambda e^{-\lambda},$$

pour $k = 2$ on obtient

$$\mathbb{P}(X = 2) = \frac{G_X^{(2)}(0)}{2!} = \frac{\lambda^2}{2} e^{\lambda(0-1)} = \frac{\lambda^2}{2} e^{-\lambda},$$

pour $k = 3$ on obtient

$$\mathbb{P}(X = 3) = \frac{G_X^{(3)}(0)}{3!} = \frac{\lambda^3}{3!} e^{\lambda(0-1)} = \frac{\lambda^3}{3!} e^{-\lambda},$$

donc

$$\mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!} = \frac{\lambda^k}{k!} e^{\lambda(0-1)} = \frac{\lambda^k}{k!} e^{-\lambda}$$

par conséquent $X \sim P(\lambda)$.

3.2.1 Fonction génératrice d'un couple de variables aléatoires.

Définition 3.2.2 La fonction génératrice des moments d'un couple de variables aléatoires (X_1, X_2) est la fonction notée $G_{(X_1, X_2)}$, définie par

$$G_{(X_1, X_2)}(s_1, s_2) = \mathbb{E}(s_1^{X_1} s_2^{X_2}), \forall (s_1, s_2) \in [-1, 1]^2.$$

Exemple 3.2.5 Soit (X_1, X_2) un couple de variables aléatoires dont on suppose que la loi conjointe est

X_1/X_2	1	2	3
1	$\frac{2}{24}$	$\frac{3}{24}$	$\frac{3}{24}$
2	$\frac{3}{24}$	$\frac{2}{24}$	$\frac{3}{24}$
3	$\frac{3}{24}$	$\frac{3}{24}$	$\frac{2}{24}$

- Calculer $G_{(X_1, X_2)}$ et $\varphi_{(X_1, X_2)}$
 On a par définition

$$\begin{aligned} G_{(X_1, X_2)}(s_1, s_2) &= \mathbb{E}(s_1^{X_1} s_2^{X_2}) = \sum_{i \in X_1(\Omega)} \sum_{j \in X_2(\Omega)} s_1^i s_2^j \mathbb{P}(X = i, Y = j) \\ &= \frac{2}{24} (s_1 s_2 + s_1^2 s_2^2 + s_1^3 s_2^3) + \frac{3}{24} (s_1^1 s_2^2 + s_1^1 s_2^3 + s_1^2 s_2^1 + s_1^2 s_2^3 + s_1^3 s_2^1 + s_1^3 s_2^2). \end{aligned}$$

Pour la fonction caractéristique $\varphi_{(X_1, X_2)}$

$$\begin{aligned} \varphi_{(X_1, X_2)}(t_1, t_2) &= \sum_{i \in X_1(\Omega)} \sum_{j \in X_2(\Omega)} e^{i(t_1 i + t_2 j)} \mathbb{P}(X_1 = i, X_2 = j) \\ &= \frac{2}{24} (e^{i(t_1 + t_2)} + e^{i(2t_1 + 2t_2)} + e^{i(3t_1 + 3t_2)}) \\ &\quad + \frac{3}{24} (e^{i(t_1 + 2t_2)} + e^{i(t_1 + 3t_2)} + e^{i(2t_1 + t_2)} + e^{i(2t_1 + 3t_2)} + e^{i(3t_1 + t_2)} + e^{i(3t_1 + 2t_2)}). \end{aligned}$$

Chapitre 4

Modes de convergence des suites de variables aléatoires

4.1 Modes de convergence usuels

4.1.1 Convergence en loi

Soient $(X_n)_{n \geq 0}$ une suite de variables aléatoires et X une autre variable aléatoire
On note $(F_{X_n})_{n \in \mathbb{N}}$ la suite des fonctions de répartition de $(X_n)_{n \geq 0}$ et F_X celle de X .
On note $(\varphi_{X_n})_{n \geq 0}$ la suite des fonctions caractéristiques de $(X_n)_{n \geq 0}$ et φ_X celle de X .
On note $(G_{X_n})_{n \geq 0}$ la suite des fonctions caractéristiques de $(X_n)_{n \geq 0}$ et G_X celle de X .

Définition 4.1.1 On dit que la suite $(X_n)_{n \geq 0}$ converge en loi vers X si et seulement si pour toute fonction réelle bornée et continue Ψ de \mathbb{R}^n dans \mathbb{R} , on a

$$\lim_{n \rightarrow \infty} \mathbb{E}[\Psi(X_n)] = \mathbb{E}[\Psi(X)].$$

Remarque 4.1.1

Les v.a. $(X_n)_{n \geq 0}$ et X ne sont pas nécessairement définies sur les mêmes espaces de probabilités, mais peuvent être définies sur des espaces probabilisés tous différents, disons, par exemples $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ et $(\Omega, \mathcal{A}, \mathbb{P})$, car **la convergence en loi est une convergence d'une suite de lois de probabilités** des v.a. $(X_n)_{n \geq 0}$ vers la loi de probabilité de la v.a. X .

Exemple 4.1.1 Soit $(X_n)_{n \geq 0}$ une suite de v.a. de loi $\mathcal{B}(n, \frac{\lambda}{n})$. Alors, lorsque n tend vers l'infini,

$$X_n \rightarrow^{\mathcal{L}} X, \text{ où } X \text{ suit une loi poisson } P(\lambda).$$

En effet, soit $k \in \{0, 1, \dots, n\}$

$$\begin{aligned} \mathbb{P}(X_n = k) &= C_n^k \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1) \dots (n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} \\ &= \frac{\lambda^k}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} \end{aligned}$$

lorsque n tends vers l'infini on obtient

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) &= \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda} \frac{\lambda^k}{k!}. \text{Maintenant}\end{aligned}$$

Le théorème suivant donne l'équivalence entre la définition de la convergence en loi et la convergence des fonctions caractérisant la loi de probabilité des v.a.

Théorème 4.1.1 Soient $(X_n)_{n \geq 0}$ une suite de v.a. et X une v.a, alors on a

1- $X_n \xrightarrow{\mathcal{L}} X \iff F_{X_n}(x) \rightarrow_{n \rightarrow \infty} F_X(x)$ pour chaque x point de continuité de F_X .

2- $X_n \xrightarrow{\mathcal{L}} X \iff \forall x, \varphi_{X_n}(x) \rightarrow_{n \rightarrow \infty} \varphi_X(x)$ et Φ_X est continue en 0.

3- Soient $(X_n)_{n \geq 0}$ à valeurs entières positives et $(G_{X_n})_{n \geq 0}$ et G_X les fonctions génératrices de $(X_n)_{n \geq 0}$ et de X respectivement. Alors

$$X_n \xrightarrow{\mathcal{L}} X \iff \forall s, |s| \leq 1, G_{X_n}(s) \rightarrow_{n \rightarrow \infty} G_X(s).$$

Exemple 4.1.2 Soit $(X_n)_{n \geq 1}$ une suite de v.a. de densité de probabilité donnée par

$$f_{X_n}(x) = \begin{cases} \alpha x_n^{-\alpha-1}, & \text{si } x_n > 1, \alpha > 0 \\ 0, & \text{sinon} \end{cases}$$

et soit $(Y_n)_{n \geq 1}$ une suite de v.a définie pour tout $n \geq 1$ par

$$Y_n = n^{-\frac{1}{\alpha}} \max_{1 \leq k \leq n} X_k.$$

-Montrer que $(Y_n)_{n \geq 1}$ converge en loi vers une limite Y .

Maintenant on donne quelques propriétés de la convergence en loi

Théorème 4.1.2 (Théorème de Slutsky)

Soient $(X_n)_{n \geq 0}$ et $(Y_n)_{n \geq 0}$ deux suites de variables aléatoires. Alors pour tout $c \in \mathbb{R}$, tel que $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} c$, on a

- a) $(X_n, Y_n) \xrightarrow{\mathcal{L}} (X, c)$,
- b) $X_n + Y_n \xrightarrow{\mathcal{L}} X + c$,
- c) $X_n \cdot Y_n \xrightarrow{\mathcal{L}} X \cdot c$,
- d) $X_n / Y_n \xrightarrow{\mathcal{L}} X / c, c \neq 0$.

Remarque 4.1.2 $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} Y \not\Rightarrow X_n + Y_n \xrightarrow{\mathcal{L}} X + Y$.

4.1.2 Convergence en probabilité

Considérons une suite de v.a.r. $(X_n)_{n \geq 0}$ et une autre v.a. X toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 4.1.2 La suite $(X_n)_{n \geq 0}$ converge en probabilité vers la v.a. X (notation $X_n \xrightarrow{\mathbb{P}} X$) si et seulement si

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0.$$

où bien si et seulement si

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}) = 1.$$

Exemple 4.1.3 Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires telles que pour tout $n \geq 1$, $X_n \sim \Gamma(n, \frac{1}{n})$.

- Montrer que $X_n \xrightarrow{\mathbb{P}} 1$.

En effet, on a $\forall n \geq 1$, $X_n \sim \Gamma(n, \frac{1}{n})$, donc $\mathbb{E}(X_n) = 1$ et $\text{Var}(X_n) = \frac{1}{n}$ pour tout $n \geq 1$. En appliquant l'inégalité de Chibyshev on obtient pour tout $\varepsilon >$

$$\mathbb{P}(|X_n(\omega) - 1| > \varepsilon) \leq \frac{1}{n\varepsilon^2} \Rightarrow \lim_{n \rightarrow \infty} \mathbb{P}(|X_n(\omega) - 1| > \varepsilon) \leq \lim_{n \rightarrow \infty} \frac{1}{n\varepsilon^2} = 0.$$

D'où $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n(\omega) - 1| > \varepsilon) = 0 \Leftrightarrow X_n \xrightarrow{\mathbb{P}} 1$.

Exemple 4.1.4 Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires dont la loi est définie par

$$\mathbb{P}(X_n = \sqrt{n}) = \frac{1}{n}, \mathbb{P}(X_n = 0) = 1 - \frac{1}{n}.$$

- Montrer que $X_n \xrightarrow{\mathbb{P}} 0$.

En effet, on a pour $\varepsilon > 0$ fixé

$$\begin{aligned} \mathbb{P}(|X_n - 0| > \varepsilon) &= \mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(\{|X_n| > \varepsilon \cap X_n = \sqrt{n}\} \cup \{|X_n| > \varepsilon \cap X_n = 0\}) \\ &= \mathbb{P}(\{|X_n| > \varepsilon \cap X_n = \sqrt{n}\}) + \mathbb{P}(\{|X_n| > \varepsilon \cap X_n = 0\}) \\ &= \mathbb{P}\{|X_n| > \varepsilon \cap X_n = \sqrt{n}\} \end{aligned}$$

Or, pour n assez grand, $\{|X_n| > \varepsilon\} \subset \{X_n = \sqrt{n}\}$, donc

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n| > \varepsilon) = \lim_{n \rightarrow \infty} \mathbb{P}\{|X_n| > \varepsilon \cap X_n = \sqrt{n}\} \leq \lim_{n \rightarrow \infty} \mathbb{P}(X_n = \sqrt{n}) = 0$$

d'où $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n| > \varepsilon) = 0 \Leftrightarrow X_n \xrightarrow{\mathbb{P}} 0$.

Théorème 4.1.3 (Condition suffisante de convergence en probabilité)

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoire telle que

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = c \text{ et } \lim_{n \rightarrow \infty} \text{Var}(X_n) = 0, c \in \mathbb{R}$$

Alors $X_n \xrightarrow{\mathbb{P}} c$.

4.1.3 La convergence presque sûre

Définition 4.1.3 La suite de v.a. $(X_n)_{n \geq 0}$ converge presque sûrement (converge avec une probabilité égale à 1) vers la v.a. X (notation $X_n \xrightarrow{p.s.} X$) si et seulement si

$$\mathbb{P}\left(\left\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

c'est à dire l'ensemble de tous les $\omega \in \Omega$ pour lesquels $X_n(\omega)$ converge vers $X(\omega)$ est de probabilité 1.

Propriétés (même propriétés pour la convergence en probabilité)

4.1.4 Convergence dans L^p

Considérons l'espace $L^p = \{X, X \text{ est une v.a.r., telle que } \int_{\Omega} |X|^p d\mathbb{P} < \infty, 0 < p \leq \infty\}$ muni de la norme $\|X\|_p = (\mathbb{E} |X|^p)^{\frac{1}{p}}$. Pour $p \geq 1$, L^p est un espace normé complet (espace de Banach). De plus L^2 est un espace de Hilbert.

Définition 4.1.4 Pour $1 \leq p < \infty$, la suite de v.a. $(X_n)_{n \geq 0}$ converge dans L^p (ou en moment d'ordre p) vers la v.a. X (notation $X_n \rightarrow^{L^p} X$) si et seulement si

$$(X_n)_{n \geq 0} \text{ et } X \in L^p, \text{ et } \lim_{n \rightarrow \infty} \|X_n - X\|_p = 0$$

Remarque 4.1.3 1- Pour $p = 2$, la convergence dans L^2 est dite convergence quadratique ou en moment d'ordre 2.

$$2- X_n \rightarrow^{L^2} X \Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E} |X_n - X|^2 = 0.$$

Exemple 4.1.5 Soit $(X_n)_{n \geq 1}$ une suite de v.a. indépendantes et toutes de carré intégrables.

1- Montrer que pour tout $n \geq 1$ et $a \in \mathbb{R}$, on a

$$\mathbb{E} [(X_n - a)^2] = (\mathbb{E} (X_n) - a)^2 - \text{var} (X_n)$$

2- En déduire que la suite $(X_n)_{n \geq 1}$ converge en moyenne quadratique (dans L^2) vers une constante a si et seulement si on a les convergences

$$\lim_{n \rightarrow \infty} \mathbb{E} (X_n) = a \text{ et } \lim_{n \rightarrow \infty} \text{var} (X_n) = 0.$$

Solution 4.1.1 1- On montre que pour tout $n \geq 1$ et $a \in \mathbb{R}$, $\mathbb{E} [(X_n - a)^2] = (\mathbb{E} (X_n) - a)^2 - \text{var} (X_n)$.

On a

$$\begin{aligned} (\mathbb{E} (X_n) - a)^2 - \text{var} (X_n) &= \mathbb{E}^2 (X_n) + a^2 - 2a\mathbb{E} (X_n) - \mathbb{E}^2 (X_n) + \mathbb{E} (X_n^2) \\ &= \mathbb{E} (X_n^2) + a^2 - 2a\mathbb{E} (X_n) \\ &= \mathbb{E} [(X_n - a)^2]. \end{aligned}$$

2- On déduit que $X_n \rightarrow^{L^2} a \Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E} (X_n) = a$ et $\lim_{n \rightarrow \infty} \text{var} (X_n) = 0$.

On a :

$$\mathbb{E} [(X_n - a)^2] = (\mathbb{E} (X_n) - a)^2 - \text{var} (X_n) \Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E} [(X_n - a)^2] = \lim_{n \rightarrow \infty} ((\mathbb{E} (X_n) - a)^2 - \text{var} (X_n)).$$

Donc $\lim_{n \rightarrow \infty} \mathbb{E} [(X_n - a)^2] = 0 \Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E} (X_n) - a = 0$ et $\lim_{n \rightarrow \infty} \text{var} (X_n) = 0$.

4.1.5 La convergence presque complète

Définition 4.1.5 La suite de variables aléatoire $(X_n)_{n \geq 0}$ converge presque complètement vers une variable aléatoire X , si et seulement si pour tout $\varepsilon > 0$, on a

$$\sum_{n \geq 0} P(|X_n - X| > \varepsilon) < \infty.$$

On note ce mode de convergence comme suit $X_n \rightarrow^c X$.

4.2 Liens entre les modes de convergence usuels

Dans ce paragraphe, on s'intéresse aux différentes relations d'implication qui peuvent exister entre les modes de convergence de suites des variables aléatoires réelles, définies précédemment. Ainsi que les conditions suffisantes pour obtenir les implications inverses.

Théorème 4.2.1 *(la convergence presque complète implique la convergence presque sûre)*

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. qui converge presque complètement vers la v.a.r. X . Alors $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X .

Pour démontrer cette implication on utilise l'équivalence suivante :

Théorème 4.2.2 *La suite de v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers la v.a.r. X si et seulement si*

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\}) = 1$$

Preuve. On suppose que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers X et on montre sa convergence presque sûre vers la même limite. On note que, pour $\omega \in \Omega$,

$$\{|X_m(\omega) - X| \leq \varepsilon, \text{ pour tout } m \geq n\}^C = \{|X_m(\omega) - X| > \varepsilon, \text{ pour quelques } m \geq n\},$$

donc pour $\varepsilon > 0$

$$\mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\}) = \mathbb{P}(\cup_{m=n}^{\infty} \{|X_m(\omega) - X(\omega)| > \varepsilon\}).$$

Ce qui implique

$$\mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\}) \leq \sum_{m=n}^{\infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon\}).$$

On a par supposition $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers X , c'est à dire pour tout $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) < \infty.$$

Pour que cette somme converge, il faut que pour tout $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \sum_{m=n}^{\infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon\}) = 0,$$

c'est-à-dire

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\}) = 0.$$

Comme $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$ pour tout événement A de \mathcal{A} , alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\}) = 1,$$

ce qui est équivalent à la convergence presque sûre de $(X_n)_{n \in \mathbb{N}^*}$ vers X . ■

Proposition 4.2.1 (la convergence presque sûre implique la convergence en probabilité)

Si la suite $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X alors elle converge en probabilité, c-à-d

$$X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X.$$

Preuve. Supposons que $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X et montrons la convergence en probabilité. On a d'après de théorème () pour tout $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\}) = 1.$$

Notons que

$$\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\} \subset \{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\},$$

donc

$$\mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\}) \leq \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}) \text{ pour chaque } n \in \mathbb{N}.$$

Par conséquent

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\}) \leq \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}),$$

ce qui implique que

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}) = 1.$$

■

Proposition 4.2.2 (la convergence en probabilité implique la convergence en loi)

Si la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers X alors elle converge en loi. C'est-à-dire

$$X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X \implies X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X.$$

Preuve. On utilise la définition () pour démontrer la convergence en loi de la suite $(X_n)_{n \in \mathbb{N}^*}$ vers la v.a. X . Soit Ψ une fonction continue et bornée et rappelons que ça implique que Ψ est uniformément continue. Soient \mathbb{P}_{X_n} et \mathbb{P}_X représentent les lois images de la probabilité \mathbb{P} par les v.a.r. X_n et X respectivement. On a par définition

$$\mathbb{E}[\Psi(X_n)] = \int_{\Omega} \Psi d\mathbb{P}_{X_n} \text{ et } \mathbb{E}[\Psi(X)] = \int_{\Omega} \Psi d\mathbb{P}_X,$$

et on montre que

$$\lim_{n \rightarrow \infty} \mathbb{E}[\Psi(X_n)] = \mathbb{E}[\Psi(X)], \text{ c'est à dire } \lim_{n \rightarrow \infty} |\mathbb{E}[\Psi(X_n)] - \mathbb{E}[\Psi(X)]| = 0.$$

$$\begin{aligned} |\mathbb{E}[\Psi(X_n)] - \mathbb{E}[\Psi(X)]| &= \left| \int_{\Omega} \Psi d\mathbb{P}_{X_n} - \int_{\Omega} \Psi d\mathbb{P}_X \right| = \left| \int_{\Omega} \Psi(X_n) d\mathbb{P} - \int_{\Omega} \Psi(X) d\mathbb{P} \right| \\ &= \left| \int_{\Omega} (\Psi(X_n) - \Psi(X)) d\mathbb{P} \right| \leq \int_{\Omega} |\Psi(X_n) - \Psi(X)| d\mathbb{P}, \end{aligned}$$

et

$$\begin{aligned} \int_{\Omega} |\Psi(X_n) - \Psi(X)| d\mathbb{P} &= \int_{\{\omega: |X_n(\omega) - X(\omega)| < \delta\}} |\Psi(X_n) - \Psi(X)| d\mathbb{P} \\ &+ \int_{\{\omega: |X_n(\omega) - X(\omega)| \geq \delta\}} |\Psi(X_n) - \Psi(X)| d\mathbb{P} \end{aligned}$$

Ψ est uniformément continue, c'est à dire $\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in \mathbb{R} / |x - y| < \delta \implies |\Psi(x) - \Psi(y)| < \varepsilon$. Donc

$$\int_{\Omega} |\Psi(X_n) - \Psi(X)| d\mathbb{P} \leq 2 \sup_{x \in \mathbb{R}} |\Psi(x)| \mathbb{P}(|X_n - X| \geq \delta) + \varepsilon \mathbb{P}(|X_n - X| < \delta).$$

Par hypothèse de la convergence en probabilité on a $\forall \delta > 0, \mathbb{P}(|X_n - X| \geq \delta) \xrightarrow{n \rightarrow \infty} 0$. Donc

$$\int_{\Omega} \Psi d\mathbb{P}_{X_n} \xrightarrow{n \rightarrow \infty} \int_{\Omega} \Psi d\mathbb{P}_X. \text{ Alors } X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X.$$

■

Proposition 4.2.3 (la convergence dans L^p implique la convergence dans L^q , pour $0 < q < p$)

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. qui converge vers X dans L^p . Alors $(X_n)_{n \in \mathbb{N}^*}$ converge vers X dans L^q , pour $0 < q < p$.

Preuve. On suppose que $X_n \xrightarrow[n \rightarrow \infty]{L^p} X$, c'est à dire $\lim_{n \rightarrow \infty} \|X_n - X\|_p = 0$ et on montre que $X_n \xrightarrow[n \rightarrow \infty]{L^q} X$.

On a $\lim_{n \rightarrow \infty} \|X_n - X\|_p = 0 \Leftrightarrow \lim_{n \rightarrow \infty} (\mathbb{E} |X_n - X|^p)^{\frac{1}{p}} = 0 \Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E} |X_n - X|^p = 0$. On sait que la fonction $f(x) = x^{\frac{p}{q}}$ est une fonction convexe, alors en appliquant l'inégalité de Jensen, on obtient

$$\lim_{n \rightarrow \infty} (\mathbb{E} |X_n - X|^q)^{\frac{p}{q}} \leq \lim_{n \rightarrow \infty} \mathbb{E} (|X_n - X|^q)^{\frac{p}{q}} = \lim_{n \rightarrow \infty} \mathbb{E} |X_n - X|^p = 0,$$

ce qui implique

$$\lim_{n \rightarrow \infty} \mathbb{E} |X_n - X|^q = 0.$$

D'où $X_n \xrightarrow[n \rightarrow \infty]{L^q} X$. ■

Proposition 4.2.4 (la convergence dans L^p implique la convergence en probabilité)

Si les v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ et X sont dans L^p , alors la convergence en moment d'ordre p de la suite $(X_n)_{n \in \mathbb{N}^*}$ vers X implique sa convergence en probabilité vers la même limite.

$$X_n \xrightarrow[n \rightarrow \infty]{L^p} X \implies X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X.$$

Preuve. Il suffit d'appliquer l'inégalité de Markov. Soit $\varepsilon > 0$,

$$0 \leq \lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = \lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X|^p > \varepsilon^p) \leq \lim_{n \rightarrow \infty} \frac{1}{\varepsilon^p} \mathbb{E}(|X_n - X|^p) = 0,$$

donc $X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$. ■

On a démontré que la convergence presque complète entraîne la convergence presque sûre. En suite, on a prouvé que la convergence en probabilité est un résultat de la convergence presque sûre et entraîne la convergence en loi. De plus la convergence dans L^p est plus forte que celle en probabilité. Maintenant on cherche les conditions sous lesquelles on obtient les implications inverses.

Proposition 4.2.5 *Si la suite $(X_n)_{n \in \mathbb{N}^*}$ est convergente en loi vers une constante a dans \mathbb{R} , alors elle converge également en probabilité vers a , c'est-à-dire*

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} a \Rightarrow X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} a.$$

Proposition 4.2.6 *(la convergence en probabilité et celle presque sûre)*

La suite $(X_n)_{n \in \mathbb{N}^}$ des v.a.r. converge en probabilité vers X si et seulement si pour toute suite croissante d'entiers $(n_k)_{k \in \mathbb{N}^*}$ on peut extraire une sous suite $(n_k)_{k \in \mathbb{N}^*}$ telle que, $(X_{n_k})_{k \in \mathbb{N}^*}$ converge presque sûrement vers X*

L'intégrabilité uniforme est un concept important qui joue un rôle essentiel dans la théorie des probabilité, dans la suite nous avons montré une autre relation entre deux modes de convergence par application du théorème associé.

Définition 4.2.1 *Une famille $(X_n)_{n \in \mathbb{N}^*}$ de v.a.r. sur $(\Omega, \mathcal{A}, \mathbb{P})$ est uniformément intégrable (U.I) ou équi-intégrable, si pour tout c un réel positif.*

$$\lim_{c \rightarrow +\infty} \sup_{n \in \mathbb{N}^*} \mathbb{E}(|X_n| 1_{\{|X_n| > c\}}) = 0.$$

Proposition 4.2.7 *On suppose que $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers X . Si $(X_n)_{n \in \mathbb{N}^*}$ est uniformément intégrable, alors $X \in L^1$ et $(X_n)_{n \in \mathbb{N}^*}$ converge dans L^1 vers X .*

La convergence presque complète entraîne la convergence presque sûre, et la réciproque reste vraie avec une condition, comme le montre le théorème suivant.

Théorème 4.2.3 *Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. indépendantes et a une constante réelle. Si $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} a$, alors*

$$X_n \xrightarrow[n \rightarrow +\infty]{c} a.$$

Théorèmes importants de la convergence

La question qui se pose maintenant est sous quelles conditions on peut écrire

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X).$$

La réponse est donnée dans les théorèmes suivants.

Théorème 4.2.4 *(Théorème de la convergence dominée)*

Pour $(X_n)_{n \geq 0}$ une suite de variable aléatoire telle que $X_n \xrightarrow{p.s.} X$, s'il existe une v.a. Y intégrable ($\mathbb{E}|Y| < \infty$) telle que pour tout $n \geq 0$, $|X_n| \leq Y$, alors

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X)$$

Théorème 4.2.5 (*Théorème de la convergence monotone*)

Pour $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires positive et $X_n \xrightarrow{p.s} X$. Alors

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X).$$

Théorème 4.2.6 (*Lemme de Fatou*)

Pour $(X_n)_{n \in \mathbb{N}}$ une suite de variable aléatoire positives, on a

$$\mathbb{E}(\liminf_{n \rightarrow \infty} X_n) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(X_n).$$

4.2.1 Propriétés des convergence en probabilité, presque sûre, dans L^p et presque complète

Soient $(X_n)_{n \geq 0}$ et $(Y_n)_{n \geq 0}$ deux suites de variables aléatoires qui convergent respectivement vers les variables aléatoires X et Y . Soit $(a, b) \in \mathbb{R}^2$, les résultats suivant sont

Proposition 4.2.8 1- La limite en probabilité, presque sûre, dans L^p et presque complètement, si elle existe, elle est unique presque surement.

2- La limite en loi, si elle existe, elle est unique en loi.

3- Si $\forall n, X_n = a$, alors $X_n \rightarrow a$.

4- Si $X_n \rightarrow X$ et $Y_n \rightarrow Y$ et si $(a, b) \in \mathbb{R}^2$, alors

$$\begin{aligned} aX_n + bY_n &\rightarrow aX + bY, \\ X_n \cdot Y_n &\rightarrow X \cdot Y, \\ \text{si } \mathbb{P}(X_n \neq 0) &= 1, \frac{1}{X_n} \rightarrow \frac{1}{X} \end{aligned}$$

5- $(X_n)_{n \geq 0}$ est convergente ssi elle est une suite de Cauchy en probabilité, c'est à dire

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} \mathbb{P}(|X_n - X_m| > \varepsilon) = 0.$$

6- Si $X_n \rightarrow X$, alors $g(X_n) \rightarrow g(X)$ pour toute fonction continue g .

Chapitre 5

Théorèmes limites

Chapitre 6

Vecteurs aléatoires réels de dimension n

On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et n variables aléatoires X_1, X_2, \dots, X_n définies sur Ω à valeurs dans \mathbb{R} .

Définition 6.0.1 On appelle vecteur aléatoire réel, de dimension n , toute application **mesurable** $X = (X_1, X_2, \dots, X_n)$ de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Proposition 6.0.1 $X = (X_1, X_2, \dots, X_n)$ est un vecteur aléatoire si et seulement si

$$\{\omega \in \Omega, X_1(\omega) \leq t_1, X_2(\omega) \leq t_2, \dots, X_n(\omega) \leq t_n\} \in \mathcal{A}, \forall t_1, t_2, \dots, t_n \in \mathbb{R}.$$

Remarque 6.0.1 Si X est un vecteur aléatoire (vec.aléa), alors chaque X_i est une v.a. Pour tout i , X_i s'appelle la i -ème marginale du vecteur aléatoire X .

Définition 6.0.2 (Loi de probabilité d'un vecteur aléatoire)

La loi de probabilité du vec.aléa. X , notée \mathbb{P}_X , est la mesure image, sur \mathbb{R}^n , de la probabilité \mathbb{P} par X , c'est à dire, pour tout $(A_1, A_2, \dots, A_n) \in (\mathcal{B}(\mathbb{R}^n))^n$

$$\begin{aligned} \mathbb{P}_X(A_1 \times A_2 \times \dots \times A_n) &= \mathbb{P}(X \in A_1 \times A_2 \times \dots \times A_n) \\ &= \mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n). \end{aligned}$$

Définition 6.0.3 La fonction de répartition F_X du vec.aléa. X est définie pour tout $x \in \mathbb{R}^n$ comme suit

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) \\ &= \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n), x = (x_1, x_2, \dots, x_n) \\ &= \mathbb{P}(\cap_{i=1}^n \{X_i \leq x_i\}). \end{aligned}$$

Définition 6.0.4 (vecteur aléatoire discret)

X est un vec.aléa. discret, s'il prend ses valeurs dans un ensemble dénombrable. Dans ce cas, la distribution de X est définie pour tout $x \in X(\Omega)$

$$\mathbb{P}_X(x) = \mathbb{P}(X = x) = \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n), x = (x_1, x_2, \dots, x_n)$$

et la loi marginale \mathbb{P}_{X_i} de X_i est définie, pour tout $x_i \in X_i(\Omega)$ par

$$\mathbb{P}_{X_i}(x_i) = \mathbb{P}(X_i = x_i) = \sum_{x_1 \in X_1(\Omega)} \dots \sum_{x_i \in X_{i-1}(\Omega)} \sum_{x_{i+1} \in X_{i+1}(\Omega)} \dots \sum_{x_n \in X_n(\Omega)} \mathbb{P}_X(x_1, x_2, \dots, x_n).$$

Définition 6.0.5 Soit X un vecteur aléatoire à valeurs discrètes. Toute fonction discrète est une loi de probabilité conjointe à condition que les deux propriétés suivantes soient vérifiées :

$$1 - \forall x \in \mathbb{R}^n, \mathbb{P}_X(x) \geq 0,$$

$$2 - \sum_{x \in X(\Omega)} \mathbb{P}_X(x) = \sum_{x_1} \sum_{x_2} \dots \sum_{x_n} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = 1.$$

Définition 6.0.6 2- X est un vec.aléa. continu, s'il existe une fonction borélienne $f_X : \mathbb{R}^n \rightarrow \mathbb{R}^+$, appelée densité jointe ; telle que

$$\begin{cases} \int_{\mathbb{R}^n} f_X(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1. \\ \text{et } \forall B \in \mathcal{B}(\mathbb{R}^n), \mathbb{P}(\{X \in B\}) = \int_B f_X(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n. \end{cases}$$

définie la loi de probabilité de X . Et dans ce cas, les densités marginales f_{X_i} de X_i sont définies par

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f_X(x_1, x_2, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n.$$

Remarque 6.0.2 1- X est un vec.aléa discret ssi chaque X_i est une v.a.d.

2- Si X est un vec. aléa. continu, alors chaque X_i est une v.a.continue, mais la réciproque n'est pas vraie.

Définition 6.0.7 (espérance et matrice de covariance)

Soit $X = (X_1, X_2, \dots, X_n)$ un vec.aléa. tel que $\forall i = 1, 2, \dots, n, X_i \in L^2(\Omega)$.

1- l'espérance du vec. X , notée $\mathbb{E}(X)$, est le vecteur des espérances de ces marginales, c'est à dire

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

2- La matrice de covariance du vec. X est la matrice carrée symétrique définie positive, notée $K(\text{cov}(X_i, X_j))_{i,j=1,\dots,n}$, avec $\forall i, j = 1, \dots, n$,

$$\text{cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \cdot \mathbb{E}(X_j),$$

ou simplement $\text{cov}(X)$, donnée par :

$$\text{cov}(X) = K = \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \dots & \dots & \text{var}(X_n) \end{pmatrix}.$$