Probabilités

LMA3

HDHIRI Ibtissem

كلية العلوم بقابس
Faculté des Sciences de Gabès

Chapter 1

Introduction aux probabilités

1.1 Introduction

La théorie des probabilités est la science qui vise à modéliser les phénoménes aléatoires qui sont des phénoménes dont on ne peut prévoir le résultat à l'avance.
La modélisation d'une experience aléatoire à travers le calcul des probabilités a été introduite par Kolmogorov en 1933. Elle est essentiellement basée sur quatres objets fondamentaux:

a- L'espace d'états ou univers: C'est l'ensemble de tous les résultats possibles de l'experience aléatoire étudiée. Il est habituellement noté Ω et ses éléments, dits événements élémentaires sont notés par ω.

Exemples 1.1.1
- *Pour un tirage de pile ou face, l'univers est donné par:* $\Omega = \{P, F\}$
- *Pour le lancer de deux dés,* $\Omega = \{(i, j);\ 1 \leq i \leq 6;\ 1 \leq j \leq 6\}$

b- Les événements: Ce sont des parties de l'univers Ω. On notera $\mathcal{A}$ l'ensemble des tous les evenements. Cet ensemble coincide souvent mais pas toujours avec l'ensemble des parties de Ω. D'une maniére générale, $\mathcal{A}$ doit avoir la structure d'une tribu.

Il est utile de savoir que les termes utilisés pour décrire les événement admettent des traductions en language ensembliste:

- Ω : événement certain
- $\emptyset$: événement impossible
- $A \cup B$ Au moins A ou B est réalisé
- $A \cap B$: A et B sont réalisés
- $A^c = \Omega \backslash A$: événement contraire de A.

c- La probabilité : A tout événement A de $\mathcal{A}$; on associe un nombre $\mathcal{P}(A)$ compris entre 0 et 1. Moralement, ce nombre mésure la vraissemblance de la réalisation de l'événement A.

Définition 1.1.1 *Étant donnés un espace d'états Ω et une tribu d'événements $\mathcal{A}$; une probabilité est une mesure $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ vérifiant $\mathbb{P}(\Omega) = 1$.*

Le triplet $(\Omega; \mathcal{A}, \mathbb{P})$ est dit espace probabilisé ou espace de probabilité.

Exemple:

(i) La mesure de Lebesgue sur $[0, 1]$

(ii) Probabilité de Bernouilli $\mathbb{P} = p\delta_1 + (1-p)\delta_0$.

Voici quelques propriétés utiles dont certaines ne sont que des reformulations de la définition. Soient $A, B \in \mathcal{A}$

- Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}(A \cap B) = min(\mathbb{P}(A), \mathbb{P}(B))$
- $\mathbb{P}(A \cup B) = max(\mathbb{P}(A), \mathbb{P}(B))$
- Si $(An)_{n\in\mathbb{N}}$ est une suite croissante au sens de l'inclusion ($i.e. A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$) alors la suite $(\mathbb{P}(An))_{n\in\mathbb{N}}$ est croissante et $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n)$
- Si $(An)_{n\in\mathbb{N}}$ est une suite décroissante au sens de l'inclusion alors la suite $(\mathbb{P}(An))_{n\in\mathbb{N}}$ est décroissante et $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\bigcap_{n\in\mathbb{N}} A_n)$

d- Variable aléatoire : (v.a. en abrégé) C'est une grandeur qui dépend du résultat de l'experience aléatoire en question. En termes mathématiques, une v.a. X est une application mesurable de Ω dans E ($E = \mathbb{R}^d,' d \geq 1$).

Maintenant, on peut définir une nouvelle probabilité $\mathbb{P}_X$ en posant:

$$\mathbb{P}_X(B) = \mathbb{P}[X^{-1}(B)],\ B \in \mathcal{B}(E).$$

Cette probabilité est dite loi de la v.a. X.

1.2 Indépendance d'événements

Intuitivement, dire que deux événements sont indépendants revient à dire que la réalisation de l'un ne donne aucune information sur celle de l'autre.
Exemple : On lance un dé deux fois successives; on note A: 'Avoir un nombre pair au premier lancer' et B:'Avoir 2 au deuxiéme lancer'.

Définition 1.2.1 *1. Deux événements A et B sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$*

2. Soit $(A_i)_{i\in I}$ une famille d'événements. Ces événements sont dits indépendants ou mutuellement indépendants si pour toute famille finie J de I, on a :

$$\mathbb{P}(\bigcap_{j\in J} A_j) = \prod_{j\in J} \mathbb{P}(A_j).$$

Remarque 1.2.1 *Des événements indépendants sont deux à deux indépendants mais la réciproque n'est pas nécesserement vraie.*

Proposition 1.2.1 *Si A et B sont deux événements indépendants alors A^c et B sont également indépendants. En effet; $B = (A \cap B) \cup (A^c \cap B)$; donc $\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A^c)$.*

1.3 conditionnement

Définition 1.3.1 *Soit* $(\Omega;\ \mathcal{A}, \mathbb{P})$ *un espace de probabilité et* B *un événement vérifiant* $\mathbb{P}(B) > 0$. *On définit une nouvelle probabilité* $\mathbb{P}_B$ *pour tout événement* A *par:*

$$\mathbb{P}_B(A) = \mathbb{P}[A|B] = \frac{\mathbb{P}(B)}{\mathbb{P}(A \cap B)};\ .$$

Cette quantité est dite probabilité conditionnelle de A *sachant* B.

Définition 1.3.2 *Une famille* $(B_i)_{i\in\mathbb{N}}$ *est dite partition de* Ω *si* $\bigcup_{i\in\mathbb{N}} B_i = \Omega$ *et* $B_i \cap B_j = \emptyset,\ \forall i \neq j$.

Proposition 1.3.1 *Soient* I *un ensemble au plus dénombrable de* $\mathbb{N}$ *et* $(B_i)_{i\in I}$ *une partition de* Ω *telle que* $\mathbb{P}(B_i) > 0;\ \forall i I$

(i) *Pour tout* $A \in \mathcal{A}$; *on a* $\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)$

(ii) *Pour tout* $A \in \mathcal{A}$ *tel que* $\mathbb{P}(A) > 0$, *on a*

$$\mathbb{P}_A(B_j) = \frac{\mathbb{P}_{B_j}(A)\mathbb{P}(B_j)}{\sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)},\ j \in I.$$

La derniére formule est dite formule de Bayes.

Preuve :

(i) On a $A = \dot{\bigcup}_{i\in I} A \cap B_i$ donc $\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}_{A\cap B_i} = \sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)$

(ii) On a pour tout $j \in I$;

$$\mathbb{P}_A(B_j) = \frac{\mathbb{P}(A \cap B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}_{B_j}(A)\mathbb{P}(B_j)}{\sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}.$$

Chapter 2

Probabilités sur les ensembles au plus dénombrables

Dans tout ce chapitre, Ω désignera un espace de probabilité au plus dénombrable et on utilisera la convention que si on somme sur un ensemble vide on obtient zéro.

2.1 Construction de probabilité

Une mesure de probabilité $\mathbb{P}$ définie sur Ω est entiérement caractérisée par ses valeurs sur les singletons. En effet, on : $\forall A \in \mathcal{A}$;

$$\mathbb{P}(A) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}).$$

Ceci découle du fait que tout événement A est une réunion disjointe de ses singletons, *i.e.*; $A = \dot{\bigcup}_{\omega \in A} \{\omega\}$.
Inversement, si $\omega \mapsto p(\omega)$ est une application de Ω dans $\mathbb{R}_+$ vérifiant $\sum_{\omega \in \Omega} p(\omega) = 1$, alors on peut facilement montrer que l'application définie sur $P : \mathcal{P}(\Omega)$ par

$$P(A) = \sum_{\omega \in A} p(\omega);\ A \in \mathcal{P}(\Omega)$$

est une mesure de probabilité.

2.2 Variables aléatoires discrètes

Les v.a. seront representées par des lettres majiscules telles que X; Y; Z...

2.2.1 Loi d'une variable aléatoire discrète

Une v.a. X est une application mesurable de Ω dans un espace mesurable (E, τ). Étant donné que l'univers Ω est fini ou dénombrable; il en est de même pour $X(\Omega)$. Dans ce cas, on dit que X est une v.a.discrète (v.a.d. en abrégé).
La loi de X est la probabilité définie sur $X(\Omega)$ dont la valeur sur les singletons est donnée par: $\forall j \in X(\Omega)$,

$$\mathbb{P}_X(\{j\}) = \mathbb{P}[X = j] = \sum_{\omega,\, X(\omega) = j} \mathbb{P}(\omega).$$

On supposera dans la suite que X est une v.a. réelle, *i.e.*, $E = \mathbb{R}$.

Définition 2.2.1 *On appelle fonction de répartition (fdr en abrégé) de la v.a. X et on note F_X la fonction allant de $\mathbb{R}$ dans $[0,1]$ définie par:*

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

Cette fonction vérifie les propriétés suivantes:

Proposition 2.2.1 **(i)** *La fonction F_X est croissante et vérifie:* $\lim_{n\to-\infty} F_X(x) = 0$ *et* $\lim_{n\to+\infty} F_X(x) = 1$.

(ii) *C'est une fonction continue à droite limitée à gauche (càdlàg).*

Preuve :

(i) On a pour $x \leq x'$; *ona* $]-\infty, x] \subset]-\infty, x']$. La monotonie est donc obtenue en appliquant la probabilité $\mathbb{P}_X$.

Maintenant, on a

$$1 = \mathbb{P}_X(\mathbb{R}) = \mathbb{P}_X(\bigcup_{n\in\mathbb{N}}]-\infty, n]) = \lim_{n\to\infty} \mathbb{P}_X(]-\infty, n]) = \lim_{n\to+\infty} F_X(n).$$

De même; on a :

$$1 = \mathbb{P}_X(\mathbb{R}) = \mathbb{P}_X(\bigcup_{n\in\mathbb{N}} [-n, +\infty[) = \lim_{n\to\infty} \mathbb{P}_X([-n, +\infty[) = \lim_{n\to+\infty} (1 - F_X(-n)).$$

D'où le résultat.

(ii) Soit $x \in \mathbb{R}$ et $(u_n)_{n\geq 0}$ (resp. $(v_n)_{n\geq 0}$) une suite réele qui converge en décroissant (resp. en croissant) vers x. On a pour tout n, $]-\infty, u_{n+1}] \subset]-\infty, u_n]$ et $\cap_{n\geq 0}]-\infty, u_n] =]-\infty, x]$ donc, par la limite décroissante de la mesure $\lim_{n\to+\infty} \mathbb{P}(]-\infty, u_n]) = \mathbb{P}(\cap_{n\geq 0}]-\infty, u_n]) = \mathbb{P}(]-\infty, x])$. Ainsi, $\lim_{n\to+\infty} F(u_n) = F(x)$ et F est donc continue à droite.

De même; ona pour tout n, $]-\infty, v_n] \subset]-\infty, v_{n+1}]$, $\cup_{n\geq 0}]-\infty, v_n] =]-\infty, x[$, donc par la propriété de la limite croissante, $\lim_{n\to+\infty} F(v_n) = \mathbb{P}(]-\infty, x[)$. Par conséquent, F admet une limite à gauche.

2.3 Moments d'une v.a.d

Soient $\mathbb{P}$ une probabilité caractérisée par $p_\omega = \mathbb{P}(\{\omega\})$ et X une v.a.d. définie sur $(\Omega, \mathcal{A})$.

Définition 2.3.1 *Si $\sum_{\omega\in\Omega} |X(\omega)| p_\omega < \infty$, on dit que X admet un moment d'ordre 1 et on pose*

$$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega) p_\omega.$$

Cette quantité est dite esperance mathématique ou moyenne de X.
D'une maniére générale; pour $k > 0$; on dit que X admet un moment d'ordre k si $|X|^k$ admet un moment d'ordre 1. On note $\mathcal{L}^k$ l'ensembles des v.a. admettant des moments d'ordre k.

Les propriétés suivantes sont immédiates:

($\mathcal{P}_1$) $\mathcal{L}^1$ est un espace vectoriel sur lequel l'esperance est linéaire

($\mathcal{P}_2$) Si X est bornée alors $X \in \mathcal{L}^k$; $\forall k > 0$

($\mathcal{P}_3$) Si $X(\omega) = \alpha$; $\forall \omega \in \Omega$; alors $\mathbb{E}[X] = \alpha$.

($\mathcal{P}_4$) Si X est une v.a. positive de $\mathcal{L}^1$; alors $\mathbb{E}[X] \geq 0$.

($\mathcal{P}_5$) Si l'univers Ω est fini alors toutes les v.a. sont dans $\mathcal{L}^1$.

($\mathcal{P}_6$) Pour tout $A \in \mathcal{A}$; $\mathbb{E}[1_A] = \mathbb{P}(A)$.

Proposition 2.3.1 *$\mathcal{L}^2$ est un sous-espace vectoriel de $\mathcal{L}^1$. De plus; pour toute $X \in \mathcal{L}^2$; on a*

$$\mathbb{E}[|X|] \leq \mathbb{E}[X^2]^{\frac{1}{2}}.$$

Preuve : Soient X, Y deux v.a. de $\mathcal{L}^2$ et $\alpha \in \mathbb{R}$. On a $(\alpha X + Y)^2 \leq 2(\alpha^2 X^2 + Y^2)$, ce qui ipmlique que $\alpha X + Y \in \mathcal{L}^2$.
Par ailleurs, le fait que $|X| \leq 1 + |X|^2$ nous mène à dire que $\mathcal{L}^2 \subset \mathcal{L}^1$.
Maintenant, $\forall X \in \mathcal{L}^2$, on a :

$$0 \leq \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - 2\mathbb{E}[X\mathbb{E}[X]] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

D'où le résultat.

Définition 2.3.2 *Soit $X \in \mathcal{L}^2$, La variance de X; notée $Var(X)$ est définie par :*

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

D'après Proposition 2.3.1, cette quantité est positive. On note σ_X sa racine carrée dite écart type de la v.a. X.

Proposition 2.3.2 *Soient X une v.a.d. définie sur Ω et f une fonction allant de $X(\Omega)$ dans $\mathbb{R}$. La v.a. $Y = f(X)$ admet un moment d'ordre 1 sur $(\Omega; \mathbb{P})$ si et seulement si f admet un moment d'ordre 1 sur $(X(\Omega); \mathbb{P}_X)$ et on a :*

$$\mathbb{E}[Y] = \sum_{\omega \in \Omega} f[X(\omega)]\mathbb{P}(\{\omega\}) = \sum_{k \in X(\Omega)} f(k)\mathbb{P}_X(\{k\}) \tag{2.1}$$

Preuve :On a

$$\begin{aligned} \sum_{\omega \in \Omega} |f[X(\omega)]| p_\omega &= \sum_{k \in X(\Omega)} |f(k)| \sum_{\omega \in X^{(-1)}(\{k\})} \mathbb{P}(\{\omega\}) \\ &= \sum_{k \in X(\Omega)} |f(k)| \mathbb{P}_X(\{k\}). \end{aligned}$$

Ainsi, Y est intégrable sur $(\Omega; \mathbb{P})$ si et seulement si f est integrable sur $(X(\Omega); \mathbb{P}_X)$ et on a:

$$\mathbb{E}[Y] = \sum_{\omega \in \Omega} f[X(\omega)]\mathbb{P}(\{\omega\}) = \sum_{k \in X(\Omega)} f(k)\mathbb{P}_X(\{k\})$$

2.4 Fonction génératrice des v.a. entières

Soit X une v.a. é valeurs dans $\mathbb{N}$ (v.a. entiére). On pose

$$p_n = \mathbb{P}[X = n];\ n \in \mathbb{N}.$$

Définition 2.4.1 *La fonction génératrice de X est la fonction définie sur $[0,1]$ par:*

$$g_X(x) := \mathbb{E}[x^X] = \sum_{n\geq 0} x^n p_n.$$

La fonction g est continue et indéfiniment dérivable. Elle sera notée g s'il ya pas risque d'ambiguité.
D'autre part, on a

$$g^{(n)}(0) = n!\, p_n;\ n \geq 1.$$

Ainsi la fonction génératrice caractérise la loi d'une v.a. entiére X.

Proposition 2.4.1 *Soit X une v.a. entière et g sa fonction génératrice alors on a équivalence entre:*

(i) $X \in L^1$

(ii) *La v.a. X est dérivable à gauche en 1.*

De plus on a $\mathbb{E}[X] = g'(1)$.

Preuve : on a

$$\begin{aligned}
\lim_{s\uparrow 1} \frac{g(s)-g(1)}{s-1} &= \lim_{s\uparrow 1} \sum_n \frac{s^n-1}{s-1} p_n \\
&= \lim_{s\uparrow 1} \sum_n (1+s+...s^{n-1}) p_n \\
&= \sum_n n p_n = \mathbb{E}[X].
\end{aligned}$$

Les termes considérés étant positifs, on peut permuter somme et limite grace au théoréme de la convergence monotonne.
On peut également montrer que si la v.a. X admet un moment d'ordre r, alors on a :

$$g^{(r)}(1) = \mathbb{E}[X(X-1)...(X-r+1)].$$

2.5 Quelques lois usuelles

2.5.1 Loi de Bernouilli $\mathcal{B}(p);\ 0 < p < 1$

Elle est généralement associée aux expériences aléatoires ayant seulement deux issues (succès-echec). On dit que X suit la loi de Bernouilli de paramétre p et on note $X \rightsquigarrow \mathcal{B}(p)$ si l'on a

$$\mathbb{P}[X = 1] = p = 1 - \mathbb{P}[X = 0].$$

On a

$$\mathbb{E}[X] = \sum_{k\in\{0,1\}} k\mathbb{P}[X=k] = 0 + p = p,$$

$$\mathbb{E}[X^2] = \sum_{k\in\{0,1\}} k^2\mathbb{P}[X=k] = p$$

Ainsi

$$Var(X) = p - p^2 = p(1-p).$$

On peut retrouver le même résultat en utilisant la fonction génératrice des moments; on a $\forall x \in [0;1]$

$$g(x) := \mathbb{E}[x^X] = \sum_{n\in\{0;\,1\}} x^n p_n = p_0 + xp_1 = px + 1 - p$$

D'où $\mathbb{E}[X] = g'(1) = p$ et

$$Var(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = g''(1) + p - p^2 = p - p^2.$$

2.5.2 Loi Binomiale $\mathcal{B}(n,p)$; $n \geq 1$, $0 < p < 1$

On dit que X suit la loi Binomiale de paramétres n et p et on note $X \rightsquigarrow \mathcal{B}(n,p)$ si l'on a

$$\mathbb{P}[X=k] = C_n^k p^k (1-p)^{n-k}.;\ k \in \{0,1,...;n\}.$$

C'est la loi de la somme de n v.a de Bernoulli indépendantes.
On a

$$g(X) = \sum_{k\in\{0;\,1;...;n\}} x^k p_k = C_n^k (px)^k (1-p)^{n-k} = (px+1-p)^n.$$

Ainsi,

$$\mathbb{E}[X] = g'(1) = np$$

et

$$Var(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = g''(1) + np - (np)^2 = n(n-1)p + np - n^2\,p^2 = np(1-p)$$

2.5.3 Loi de Poisson $\mathcal{P}(\lambda)$; $(\lambda > 0)$

On dit que X suit la loi de Poisson de paramètre λ et on note $X \rightsquigarrow \mathcal{P}(\lambda)$ si l'on a

$$\mathbb{P}[X=k] = \frac{\lambda^k}{k!} e^{-\lambda k};\ k = 0;1;2;....$$

On a pour $x \in [0;1]$

$$g(x) = \sum_{k\geq 0} x^k p_k = \sum_{k\geq 0} \frac{(\lambda x)^k}{k!} e^{-\lambda k} = e^{\lambda(x-1)}.$$

D'où

$$\mathbb{E}[X] = g'(1) = \lambda$$

et

$$Var(X) = g''(1) + g'(1) - (g'(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

2.5.4 Loi géométrique $\mathcal{G}(p)$; $(0 < p < 1)$

On dit que X suit la loi géométrique de paramètre p et on note $X \rightsquigarrow \mathcal{G}(p)$ si l'on a

$$\mathbb{P}[X = k] = p(1-p)^{k-1};\ k = 1; 2; ....$$

On a

$$g(x) = \sum_{k\geq 1} x^k p_k = p \sum_{k\geq 1} x^k (1-p)^{k-1} = px \sum_{k\geq 0} x^k (1-p)^k = \frac{px}{1-x(1-p)}.$$

Ce qui implique que

$$\mathbb{E}[X] = g'(1) = \frac{1}{p}$$

et

$$Var(X) = g''(1) + g'(1) - (g'(1))^2 = \frac{1-p}{p^2}.$$

Chapter 3

Espaces de probabilité : cas général

Definition 3.0.1 *Soit X une variable aléatoire. On appelle loi de X la mesure $\mathbb{P}_X$ sur $(\Omega, \mathcal{A})$ définie par*

$$\mathbb{P}_X(A) \quad = \quad \mathbb{P}(X^{-1}(A)).$$

Remarque 3.0.1 *La mesure $\mathbb{P}_X$ est la mesure image de $\mathbb{P}$ par X.*
La mesure $\mathbb{P}_X$ est une mesure de probabilité.

Definition 3.0.2 *Soit X une v.a. à valeurs dans $\mathbb{R}$. On appelle fonction de répartition de X la fonction définie par :*

$$\begin{array}{rcl} F & : & \mathbb{R} \quad \to \quad \mathbb{R}_+ \\ & & t \quad \mapsto \quad F_X(t) = \mathbb{P}_X(]-\infty, t]) = \mathbb{P}(X \leq t) \end{array}$$

Cette fonction sera notée F_X ou F s'il n'y a pas de risque d'ambiguité.

Theorem 3.0.1 *Soit X une v.a.r.. Alors*

1. *F_X est croissante*
2. $\lim_{t\to-\infty} F_X(t) = 0$, $\lim_{t\to+\infty} F_X(t) = 1$
3. *F_X est càdlàg et $\lim_{t\to t_0^-} F_X(t) = \mathbb{P}(X < t_0)$*
   F_X est continue en t_0 si et seulement si $\mathbb{P}(X = t_0) = 0$.

Preuve 1

- *Soient $x \leq y$. Comme $]-\infty, x] \subset]-\infty, y]$, alors $F(x) = \mathbb{P}(]-\infty, x]) \leq \mathbb{P}(]-\infty, y]) = F(y)$.*
- *Soit $x \in \mathbb{R}$ et $(u_n)_{n\geq 0}$ suite réele qui converge en décroissante vers x. On a pour tout n, $]-\infty, u_{n+1}] \subset]-\infty, u_n]$ et $\bigcap_{n\geq 0}]-\infty, u_n] =]-\infty, x]$ donc, par la limite décroissante de la mesure $\lim_{n\to+\infty} \mathbb{P}(]-\infty, u_n]) = \mathbb{P}(\bigcap_{n\geq 0}]-\infty, u_n]) = \mathbb{P}(]-\infty, x])$. Ainsi, $\lim_{n\to+\infty} F(u_n) = F(x)$ et F est donc continue à droite.*
- *soit $x \in \mathbb{R}$ et $(v_n)_{n\geq 0}$ une suite réele qui converge en croissante vers x. Pour tout n, $]-\infty, v_{n+1}] \supset]-\infty, v_n]$, $\bigcup_{n\geq 0}]-\infty, v_n] =]-\infty, x[$, donc par la propriété de la limite croissante, $\lim_{n\to+\infty} F(v_n) = \mathbb{P}(]-\infty, x[)$. Ceci prouve que F a une limite à gauche.*
- *On trouve également la limite de F en $\pm\infty$ en utilisant la propriété de la limite croissante et décroissante de la mesure.*

3.1 V.a. à densité

Definition 3.1.1 *Soit X una v.a. à valeurs dans $\mathbb{R}^d$. On dit que X admet une densité $f_X : \mathbb{R}^d \to \mathbb{R}^+$ si $\forall \phi$ mesurable $\mathbb{R}^d \to \mathbb{R}$,*

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

Remarque 3.1.1 *Pour $\phi = 1$, on a*

$$\int_{\mathbb{R}^d} f_X(x) dx \ = 1.$$

Proposition 3.1.1 *Soit X una v.a. à valeurs dans $\mathbb{R}^d$ de densité f_X. On a*

- *La loi de X vérifie :*
$$\mathbb{P}(X \in B) = \int_{\mathbb{R}^d} f_X(x) \mathbf{1}_B(x) dx \ .$$
- *La f.d.r. de X est donnée par :*
$$F_X(t) = \int_{-\infty}^{t} f_X(u) du \ .$$

Remarque 3.1.2 *La loi d'une variable aléatoire à densité est entièrement déterminée par cette dernière.*

Example 3.1.1 *Soit X une v.a.r. de densité $f_X : x \in \mathbb{R} \mapsto e^{-x} \mathbf{1}_{x>0}$.*
Calculons la fonction de répartition de X. Si $t < 0$,

$$\mathbb{P}(X \leq t) \ = \ \int_{\mathbb{R}} e^{-x} \mathbf{1}_{\{x \geq 0\}} \mathbf{1}_{\{x \leq t\}} dx = 0 \ .$$

Si $t \geq 0$,

$$\begin{aligned} \mathbb{P}(X \leq t) &= \int_{\mathbb{R}} e^{-x} \mathbf{1}_{\{x \geq 0\}} \mathbf{1}_{\{x \leq t\}} dx \\ &= \int_0^t e^{-x} dx \\ &= 1 - e^{-t} \ . \end{aligned}$$

3.2 Espérance d'une v.a.

Definition 3.2.1 *Soit X une v.a.r. on dit que*

- *X est integrable si $\int_\Omega |X| d\mathbb{P} < \infty$*
- *X admet une espérance ou moyenne si $\int_\Omega X d\mathbb{P}$ existe.*

Cette espérance est donnée par : On note

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

- On note pour $k \geq 1$; $\mathcal{L}^k$ l'ensemble des v.a. telles que X^k est integrable, i.e., qui vérifient $\mathbb{E}(|X|^k) < \infty$.
- Si $X \in \mathcal{L}^2$, on dit que X est de carré integrable.

Remarque 3.2.1 *L'espérance est une intégrale. Toutes les propriétés de l'intégrale restent valables pour l'ésperance* $\mathbb{E}$.

Remarque 3.2.2 *Soient X et Y sont deux v.a.r. et $a, b \in \mathbb{R}$*

i- *Si $X = a$, alors $\mathbb{E}(X) = a$.*

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

3i- *Si $X \leq Y$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.*

4i- *Si $X = Y$ p.p. alors $\mathbb{E}(X) = \mathbb{E}(Y)$.*

5i- *Si $X \geq 0$ et $\mathbb{E}(X) < \infty$ alors X est finie $\mathbb{P}-$p.s.*

6i-] Si $X \geq 0$ et $\mathbb{E}(X) = 0$ alors $X = 0$ $\mathbb{P}-$p.s.

Proposition 3.2.1 *Soit X une variable aléatoire à valeurs dans $(F, \mathcal{F})$ et $f : F \to [0, +\infty]$ mesurable. La v.a. $f(X)$ vérifie :*

$$\mathbb{E}(f(X)) = \int_F f d\mathbb{P}_X \ .$$

En particulier, si $E = \mathbb{R}^d$ et X a une densité g alors

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

Example 3.2.1 *Soit X une variable aléatoire réelle de densité g*

- *Soit $B \in \mathcal{B}(\mathbb{R})$, on a :*

$$\begin{aligned} \mathbb{P}(X \in B) &= \mathbb{E}(\mathbf{1}_B(X)) \\ &= \int_{\mathbb{R}} \mathbf{1}_B(x)g(x)dx \\ &= \int_B g(x)dx \ . \end{aligned}$$

- $\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 \mathbb{P}_X(dx) = \int_{\mathbb{R}} x^2 g(x)dx$.

Definition 3.2.2 *Si X est une v.a.r. de carré intégrable alors la variance de X est la quantité*

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

Lemma 3.2.1 $Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$

Preuve 2 *On a*

$$\begin{aligned}
\mathbb{E}((X-\mathbb{E}(X))^2) &= \mathbb{E}(X^2+\mathbb{E}(X)^2-2X\mathbb{E}(X)) \\
&= \mathbb{E}(X^2)+\mathbb{E}(\mathbb{E}(X^2))-2\mathbb{E}(X\mathbb{E}(X)) \\
&= \mathbb{E}(X^2)+\mathbb{E}(X)^2-2\mathbb{E}(X)^2 \\
&= \mathbb{E}(X^2)-\mathbb{E}(X)^2
\end{aligned}$$

Example 3.2.2 *Soit X v.a.r. de densité $x \mapsto e^{-x}\mathbf{1}_{x\geq 0}$, on a :*

$$\begin{aligned}
\mathbb{E}(X) &= \int_{\mathbb{R}} xe^{-x}\mathbf{1}_{x\geq 0}dx \\
&= \int_0^{+\infty} xe^{-x}dx \\
\textit{(intégration par parties)} &= [-xe^{-x}]_0^{+\infty}+\int_0^{+\infty} e^{-x}dx \\
&= 0+[-e^{-x}]_0^{+\infty}=1\ .
\end{aligned}$$

3.3 Calcul de loi

Proposition 3.3.1 *Soit X variable aléatoire à valeurs dans $\mathbb{R}^d$. S'il existe une fonction positive g définie sur $\mathbb{R}^d$ vérifiant :*
($\mathcal{P}_l$): Pour toute fonction continue bornée $f:\mathbb{R}^d\to\mathbb{R}_+$, on a :

$$\mathbb{E}(f(X))=\int_{\mathbb{R}^d} f(x)g(x)dx\ ,$$

alors g est la densité de X.

Example 3.3.1 *Soit X v.a.r. de densité $x\mapsto \frac{e^{-x^2/2}}{\sqrt{2\pi}}$. Soient $(a,b)\in\mathbb{R}^*\times\mathbb{R}$. Calculons la loi de $aX+b$. Soit $f:\mathbb{R}\to\mathbb{R}_+$ continue et bornée. Par la proposition précédante, on a :*

$$\begin{aligned}
\mathbb{E}(f(aX+b)) &= \int_{-\infty}^{+\infty} f(ax+b)\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx \\
\textit{(changement de variable } y=ax+b) &= \int_{-\infty}^{+\infty} f(y)\frac{e^{-\left(\frac{y-b}{a}\right)^2\frac{1}{2}}}{\sqrt{2\pi}\times a}dy
\end{aligned}$$

Donc, par la proposition 3.3.1, la variable $aX+b$ a une loi de densité $y\mapsto \frac{\exp\left(-\frac{1}{2}\left(\frac{y-b}{a}\right)^2\right)}{\sqrt{2\pi}\times a}$.

Exercice : Soit (X,Y) v.a. à valeurs dans $\mathbb{R}^2$ de densité $(x,y)\mapsto f(x,y)=\frac{1}{\pi}\mathbf{1}_D$, avec $D=\{(x,y);\ x^2+y^2\leq 1\}$. Calculer la loi de $X+Y$.
Indication : On pose $U=X+Y$ et $V=X-Y$ et on utilise le changement de variable : changement de variable

$$\left\{\begin{array}{lcl} u &=& x+y \\ v &=& x-y \end{array}\right.,\ \left\{\begin{array}{lcl} x &=& \frac{u+v}{2} \\ y &=& \frac{u-v}{2} \end{array}\right.$$

Solution : Soit $g : \mathbb{R}^2 \to \mathbb{R}^+$ continue. On a :

$$\mathbb{E}(g(X+Y)) = \int_{\mathbb{R}^2} g(x+y)f(x,y)dxdy = \int_{\mathbb{R}^2} g(x+y)\frac{1}{\pi}\mathbf{1}_D dxdy \ .$$

On considère la fonction $\phi : (x,y) \mapsto (x+y, x-y)$ qui est un difféomorphisme entre D et $D' = \{(u,v) \in \mathbb{R}^2;\ u^2+v^2 \leq 2\}$. la matrice jacobienne de ϕ^{-1} est:

$$\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} ,$$

son déterminant est $-1/2$. Donc

$$\begin{aligned} \mathbb{E}(g(X+Y)) &= \int_{\mathbb{R}^2} g(u)\frac{1}{\pi}\mathbf{1}_{D'}\left|\frac{-1}{2}\right| dudv \\ &= \int_{\mathbb{R}} \frac{g(u)}{2\pi}\left(\int_{\mathbb{R}} \mathbf{1}_{u^2+v^2\leq 2}dv\right) du \text{ (Fubini-Tonelli)} \\ &= \int_{\mathbb{R}} \frac{g(u)}{\pi}\sqrt{2-u^2}\mathbf{1}_{|u|\leq 2}du \end{aligned}$$

Donc $X+Y$ a pour densité $u \mapsto \frac{1}{\pi}\sqrt{2-u^2}\mathbf{1}_{|u|\leq 2}$.

Definition 3.3.1 *Pour une fonction $f : E \to \mathbb{C}$ avec $(E, \mathcal{A}, \mu)$ un espace mesuré quelconque, on note*

$$\int_E f(x)\mu(dx) = \int_E Re(f)(x)\mu(dx) + i\int_E Im(f)(x)\mu(dx) \ .$$

3.4 Quelques Inégalités

3.4.1 Inégalité de Jensen

Theorem 3.4.1 *Soit X v.a.r. intégrable et $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable convexe telle que $g(X)$ est intégrable. Alors*

$$g(\mathbb{E}(X)) \leq \mathbb{E}(g(X)) \ .$$

Pour la démonstration de ce théorème, voir cours de calcul intégral.

3.4.2 Inégalité de Markov

Theorem 3.4.2 *Soit X v.a.r. positive, intégrable. Soit $\lambda > 0$. Alors*

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

Preuve 3 *On a $X = X\mathbf{1}_{X\geq\lambda} + X\mathbf{1}_{X<\lambda} \geq \lambda\mathbf{1}_{X\geq\lambda}$ donc, par la propriété de croissance*

$$\begin{aligned} \mathbb{E}(X) &\geq \mathbb{E}(\lambda\mathbf{1}_{X\geq\lambda}) \\ &= \lambda\mathbb{P}(X \geq \lambda) \ . \end{aligned}$$

3.4.3 Inégalité de Bienaymé-Tchebichev

Corollary 3.4.1 *Soit X v.a.r. telle que X^2 est intégrable. Alors*

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

Il suffit d'appliquer l'inégalité de Markov à la v.a. $|X - \mathbb{E}(X)|$ pour avoir

Preuve 4

$$\begin{aligned} \mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) &= \mathbb{P}((X - \mathbb{E}(X))^2 \geq \lambda^2) \\ &\leq \frac{1}{\lambda^2}\mathbb{E}((X - \mathbb{E}(X))^2) \ . \end{aligned}$$

3.5 Fonctions caractéristiques

Definition 3.5.1 *(Fonction caractéristique) La fonction caractéristique d'une v.a. X à valeurs dans $\mathbb{R}^d$est donnée par : $\forall u \in \mathbb{R}^d$,*

$$\Phi_X(u) = \mathbb{E}(e^{i<u,X>}) \ .$$

Theorem 3.5.1 *La fonction caractéristique d'une v.a.r. caractérise entièrement la loi de cette variable. C'est à dire que si X et Y des v.a.r. de même fonction caractéristique alors X et Y ont même loi.*

3.6 Quelques Lois usuelles continues

3.6.1 Loi uniforme $\mathcal{U}(.)$

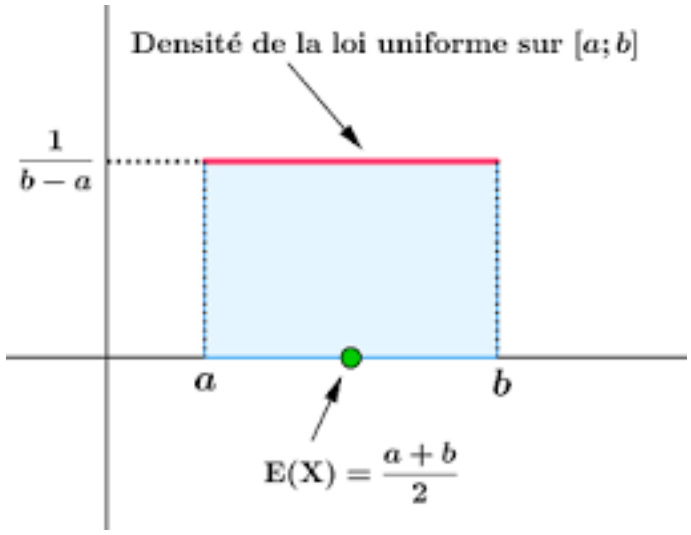
On dit qu'une v.a. X suit la loi uniforme sur $[a, b]$ $(a < b)$, notée $\mathcal{U}([a, b])$: si sa densité est donnée par : $x \mapsto \frac{1}{b-a}\mathbf{1}_{[a,b]}(x)$.

a- Sa f.d.r. est donnée par :

$$F_X(t) = \begin{cases} 0 & \text{si } t < a \\ \frac{t-a}{b-a} & \text{si } a \leq t \leq b \\ 1 & \text{si } t > b \end{cases}$$

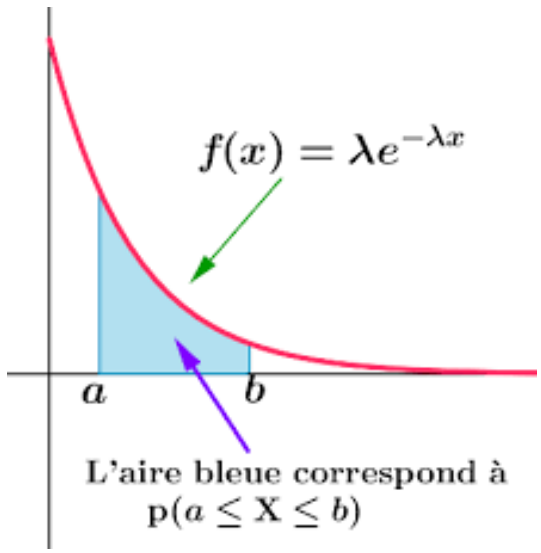
b- $\mathbb{E}[X] = \frac{a+b}{2}$, $Var[X] = \frac{(a-b)^2}{12}$

c- Sa fonction caractéristique est donnée par : $\Phi_X(u) = \frac{e^{iub} - e^{iua}}{iu(b-a)}$



3.6.2 Loi exponentielle

On dit qu'une v.a. X suit la loi exponentielle de paramètre λ ($\lambda > 0$), notée $\mathcal{E}(\lambda)$ si sa densité est donnée par : $x \mapsto \lambda e^{-\lambda x}\mathbf{1}_{\mathbb{R}^+}(x)$.



a- Sa f.d.r. est donnée par :

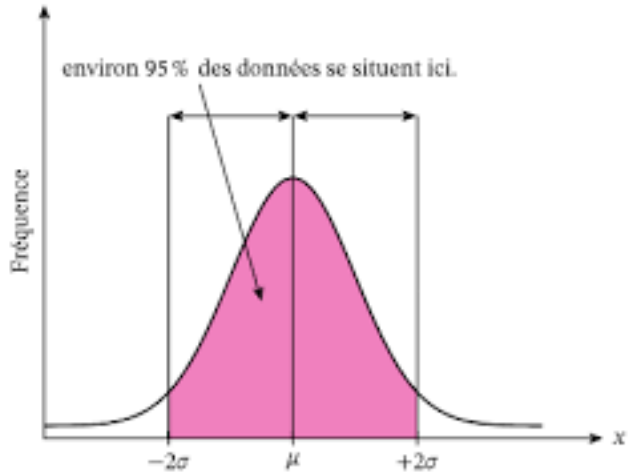
$$F_X(t) = \begin{cases} 0, & \text{si } t < 0 \\ 1 - e^{-\lambda t} & \text{si } t \geq 0. \end{cases}$$

b- $\mathbb{E}[X] = \frac{1}{\lambda}$, $Var[X] = \frac{1}{\lambda^2}$

c- Sa fonction caractéristique est donnée par : $\Phi_X(u) = (1 - \frac{iu}{\lambda})^{-1}$

3.6.3 la loi gaussienne (ou normale)

On dit que X suit la loi gaussienne de moyenne $m \in \mathbb{R}$ et de variance $\sigma^2 \in \mathbb{R}^+$, notée $\mathcal{N}(m, \sigma^2)$ si sa densité est donnée par : $x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$. Si $m = 0$ et $\sigma^2 = 1$, on parle de loi normale centrée réduite.



On a :

a- $\mathbb{E}[X] = m$, $Var[X] = \sigma^2$

b- Sa fonction caractéristique est donnée par : $\Phi_X(u) = \exp\left(ium - \frac{\sigma^2 u^2}{2}\right)$

Preuve : Nous ferons la démonstration dans le cas $m = 0, \sigma = 1, u \in \mathbb{R}$. Nous avons

$$\begin{aligned}
\Phi_X(u) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{iux} dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \mathrm{Re}(e^{-x^2/2} e^{iux}) dx + i \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \mathrm{Im}(e^{-x^2/2} e^{iux}) dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(xu) dx + i \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sin(xu) dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(xu) dx + 0
\end{aligned}$$

car l'intégrale d'une fonction impaire sur $\mathbb{R}$ est nulle.
Pour tout u, $\left|\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(xu)\right| \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ qui est intégrable sur $\mathbb{R}$. Pour tout $x \in \mathbb{R}$, $u \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(xu)$ est dérivable, de dérivée $u \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}(-x \sin(xu))$. Pour tous u, x, $\left|\frac{1}{\sqrt{2\pi}} e^{-x^2/2}(-x \sin(xu))\right| \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2}|x|$ qui est intégrable sur $\mathbb{R}$. En effet, par symétrie,

$$\begin{aligned}
\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} |x| dx &= 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} x dx \\
&= \left[\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right]_0^{+\infty} \\
&= \frac{1}{\sqrt{2\pi}} .
\end{aligned}$$

Donc par théorème de dérivation globale

$$\begin{aligned}
\Phi'_X(u) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}(-x\sin(xu))dx \\
&= \left[e^{-x^2/2}\sqrt{2\pi}\sin(xu)\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} e^{-x^2/2}\sqrt{2\pi}u\cos(xu)dx \\
&= 0 - u\Phi_X(u) \ .
\end{aligned}$$

Nous avons donc l'équation

D'où

$$\begin{aligned}
\log(\Phi_X(u)) - \log(\Phi_X(0)) &= -\frac{u^2}{2} \\
\Phi_X(u) &= \Phi_X(0)e^{-u^2/2} \ .
\end{aligned}$$

Remarquons que $\Phi_X(0) = \mathbb{E}(1) = 1$. Nous avons donc $\Phi_X(u) = e^{-u^2/2}$.

Chapter 4

Convergence de variables aléatoires

Dans toute la suite, $(\Omega, \mathcal{A}, \mathbb{P})$ désignera un espace de probabilité et X, $(X_n)_{n\geq 0}$ des v.a. à valeurs dans $\mathbb{R}$.

4.1 Les différents types de convergence

Definition 4.1.1 *(Convergence presque sûre)*
On dit que X_n converge presque sûrement vers X et on note $X_n \xrightarrow[n\to+\infty]{p.s.} X$ si

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = \lim_{n\to+\infty} X_n(\omega)\}) = 1 \ .$$

Definition 4.1.2 *(Convergence en probabilité) On dit que X_n converge en probabilité vers X et on note $X_n \xrightarrow[n\to+\infty]{proba.} X$ ou $\xrightarrow[n\to+\infty]{\mathbb{P}}$ si $\forall \varepsilon > 0$,*

$$\mathbb{P}(|X - X_n| \geq \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0.$$

Definition 4.1.3 *(Convergence dans L^p) Soit $p > 0$, on dit que X_n converge dans L^p vers X et on note $X_n \xrightarrow[n\to+\infty]{L^p} X$ si les v.a. $X, (X_n)_{n\geq 0}$ sont dans L^p et*

$$\mathbb{E}(|X - X_n|^p) \underset{n\to+\infty}{\longrightarrow} 0.$$

En particulier, on dit que X_n converge dans L^1 (resp. L^2) vers X si les v.a. $X, (X_n)_{n\geq 0}$ sont dans L^1 (resp. L^2) et $\mathbb{E}(|X - X_n|) \underset{n\to+\infty}{\longrightarrow} 0$. (resp. $\mathbb{E}(|X - X_n|^2) \underset{n\to+\infty}{\longrightarrow} 0$.)

Definition 4.1.4 *(Convergence en loi) On dit que X_n converge en loi vers X et on note $X_n \xrightarrow[n\to+\infty]{loi} X$ si $\forall h$ continue bornée,*

$$\mathbb{E}(h(X_n)) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(h(X)).$$

Theorem 4.1.1 *Nous avons équivalence:*

(i) *La suite $[X_n]$ converge en loi vers X*

(ii) *La fonction $F_{X_n}(t)$ converge vers $F_X(t)$ sur son domaine de continuité càd en tout point t vérifiant $\mathbb{P}(X = t) = 0$*

(iii) *La fonction caractéristique* Φ_{X_n} *converge vers* Φ_X

Theorem 4.1.2 i) *La convergence presque sûre entraîne la convergence en probabilité*

ii) *La convergence* L^1 *entraîne la convergence en probabilité*

iii) *La convergence en probabilité entraîne la convergence en loi*

iv) *La convergence* L^2 *entraîne la convergence* L^1.

Preuve 5 *On se contentera de donner la preuve pour des variables aléartoires réelles.*

i) *Montrons que si* $\left[X_n \xrightarrow[n\to+\infty]{p.s.} X\right]$ *alors* $\left[X_n \xrightarrow[n\to+\infty]{proba.} X\right]$. *Soit* $\varepsilon > 0$, *on a :*

$$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{E}(\mathbf{1}_{\{|X_n - X| > \varepsilon\}}) = \mathbb{E}(\mathbf{1}_{]\varepsilon;+\infty[}(|X_n - X|)) \ .$$

Or on a Pour presque tout ω, $|X_n(\omega) - X(\omega)| \underset{n\to+\infty}{\longrightarrow} 0$ *alors on a :*

$$\mathbf{1}_{]\varepsilon;+\infty[}(|X_n(\omega) - X(\omega)|) \underset{n\to+\infty}{\longrightarrow} 0.$$

Comme pour tout n, *la v.a.* $\mathbf{1}_{]\varepsilon;+\infty[}(|X_n - X|)$ *est majoréé par* $1 \in L^1$; *donc par le théorème de convergence dominée, on obtient :* $\mathbb{E}(\mathbf{1}_{]\varepsilon;+\infty[}(X_n - X)) \underset{n\to+\infty}{\longrightarrow} 0$.

ii) *Par l'inégalité de Markov, on a :* $\forall \varepsilon > 0$,

$$\begin{aligned} \mathbb{P}[|X_n - X| > \varepsilon] \quad &\leq \quad \frac{\mathbb{E}(|X_n - X|)}{\varepsilon} \\ &\to_{n\to+\infty} 0 \end{aligned}$$

iii) *Soit* t *un point où* F_X *est continue. Soit* $\varepsilon > 0$ *quelconque. On a :*

$$\begin{aligned} \mathbb{P}(X_n \leq t) \quad &= \quad \mathbb{P}(X_n \leq t, |X - X_n| \leq \varepsilon) + \mathbb{P}(X_n \leq t, |X - X_n| > \varepsilon) \\ &\leq \quad \mathbb{P}(X \leq t + \varepsilon) + \mathbb{P}(|X - X_n| > \varepsilon) \ . \end{aligned}$$

Comme $\mathbb{P}(|X - X_n| > \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0$ *alors* $\limsup_{n\to+\infty} \mathbb{P}(X_n \leq t) \leq \mathbb{P}(X \leq t + \varepsilon) = F_X(t + \varepsilon)$. *De même :*

$$\begin{aligned} \mathbb{P}(X \leq t - \varepsilon) \quad &= \quad \mathbb{P}(X \leq t - \varepsilon, |X - X_n| \leq \varepsilon) + \mathbb{P}(X \leq t - \varepsilon, |X - X_n| > \varepsilon) \\ &\leq \quad \mathbb{P}(X_n \leq t) + \mathbb{P}(|X - X_n| > \varepsilon) \ . \end{aligned}$$

Donc $\liminf_{n\to+\infty} \mathbb{P}(X_n \leq t) \geq \mathbb{P}(X \leq t - \varepsilon) = F_X(t - \varepsilon)$. *Tous ces calculs sont vrais* $\forall \varepsilon$ *et* F_X *est continue en* t *donc* $\lim_{n\to+\infty} F_{X_n}(t) = F_X(t)$.

iv) *C'est une conséquence de l'inégalité de Jensen.*

Proposition 4.1.1 i) *Si* $X_n \xrightarrow[n\to+\infty]{proba.} X$ *alors il existe une sous-suite* $(X_{u(n)})$ *vérifiant* $X_{u(n)} \xrightarrow[n\to+\infty]{p.s.} X$

ii) *Si* $X_n \xrightarrow[n\to+\infty]{loi} X$ *avec* $X = a \in \mathbb{R}$ *alors* $X_n \xrightarrow[n\to+\infty]{proba.} X$

4.1.1 Diagramme de convergence

$$\begin{array}{ccccc} \text{convergence } L^p & \Rightarrow & \text{convergence en probabilité} & \Leftarrow & \text{convergence p.s.} \\ & & \Downarrow & & \\ & & \text{convergence en loi.} & & \end{array}$$

4.2 Loi des grands nombres

Notation 4.2.1 *(i.i.d.) Soient $X_1, X_2, \ldots$ des variables indépendantes de même loi. On dira que ces variables sont indépendantes et identiquement distribuées (i.i.d. en abrégé). On désignera pas S_n la somme des n premiers termes*

$$S_n = X_1 + ... + X_n.$$

Il est facile à voir que si $X_1 \in L^2$, alors $\mathbb{E}(\frac{S_n}{n}) = \mathbb{E}(X_1)$ et $Var(\frac{S_n}{n}) = \frac{1}{n}Var(X_1)$.

Theorem 4.2.1 *(Loi faible des grands nombres)*
Soient $X_1, X_2, \ldots$ des v.a.r. i.i.d. de L^2, on a

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{L^2} \mathbb{E}(X_1) \ .$$

Preuve 6

$$\begin{aligned} \mathbb{E}\left(\left(\frac{X_1 + \cdots + X_n}{n} - \mathbb{E}(X_1)\right)^2\right) &= Var\left(\frac{S_n}{n}\right) \\ &= \frac{Var(X_1)}{n} \xrightarrow[n\to+\infty]{} 0 \end{aligned}$$

Theorem 4.2.2 *(Loi forte des grands nombres)*
Soient $X_1, X_2, \ldots$ des v.a.r. i.i.d. de L^1. On a:

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{p.s.} \mathbb{E}(X_1) \ .$$

4.3 Théorème central-limite

Theorem 4.3.1 *(Théorème central-limite(TCL en abrégé))*
Soit (X_n) une suite de v.a.r. i.i.d. avec $\mathbb{E}(X_1) = m$ et $Var(X_1) = \sigma^2$ ($m \in \mathbb{R}$; $\sigma^2 \in \mathbb{R}_+^*$). Alors

$$\frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{loi} Z \ ,$$

ou encore

$$\frac{\frac{S_n}{n} - m}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n\to+\infty]{loi} Z,$$

avec Z une v.a. de loi $\mathcal{N}(0,1)$ et σ l'écart type de X_1.

Preuve 7 *On pose pour tout $n \in \mathbb{N}^*$, $Y_n = X_n - m$, $S'_n = Y_1 + \cdots + Y_n$, et $Z_n = \frac{X_1+\cdots+X_n-nm}{\sigma\sqrt{n}}$. On a*

$$Z_n = \frac{S'_n}{\sigma\sqrt{n}}.$$

Par conséquent, par indépendance des v.a. Y_n on a :

$$\begin{aligned}
\Phi_{Z_n}(t) &= \mathbb{E}\left(\exp\left(\frac{itS'_n}{\sigma\sqrt{n}}\right)\right) \\
&= \mathbb{E}\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(Y_1+\cdots+Y_n)\right)\right) \\
&= \prod_{1\leq k\leq n} \mathbb{E}\left(\exp\left(\frac{it}{\sigma\sqrt{n}}Y_k\right)\right) \\
&= \Phi_{Y_1}\left(\frac{t}{\sigma\sqrt{n}}\right)^n .
\end{aligned}$$

la dernière égalité provient du fait que les v.a. Y_n; $n \in \mathbb{N}^*$ sont de même loi.
Regardons maintenant la fonction $\Phi_{Y_1}(u) = \mathbb{E}(e^{iuY_1})$ pour $u \in \mathbb{R}$. Pour tout u, $\mathbb{E}(|e^{iuY_1}|) = 1 < \infty$. Pour tout ω, $u \mapsto e^{iuY_1(\omega)}$ est dérivable et de dérivée $u \mapsto iY_1e^{iuY_1(\omega)}$. Pour tous u, ω, $|Y_1e^{iuY_1(\omega)}| \leq |Y_1(\omega)|$ avec $Y_1 \in L^1$. Donc, par théorème de dérivation

$$\Phi'_{Y_1}(u) = \mathbb{E}(iY_1e^{iY_1u}) .$$

De même, $\Phi''_{Y_1}(u) = \mathbb{E}(-Y_1^2e^{iY_1u})$. Donc $\Phi'_{Y_1}(0) = \mathbb{E}(iY_1) = i\mathbb{E}(Y_1) = 0$, $\Phi''_{Y_1}(0) = -\mathbb{E}(Y_1^2) = -\sigma^2$. Le développement limité en 0 de Φ_{Y_1} est donc donné par :

$$\begin{aligned}
\Phi_{Y_1}(u) &= \Phi_{Y_1}(0) + u\Phi'_{Y_1}(0) + \frac{u^2}{2}\Phi''_{Y_1}(0) + o(u^2) \\
&= 1 - \frac{u^2\sigma^2}{2} + o(u^2) .
\end{aligned}$$

Donc

$$\begin{aligned}
\Phi_{Z_n}(t) &= \left(1 - \frac{t^2}{\sigma^2 n} + o\left(\frac{1}{n}\right)\right)^n \\
&= \exp\left(n\log\left(1 - \frac{t^2}{\sigma^2 n} + o\left(\frac{1}{n}\right)\right)\right) \\
&= \exp\left(-\frac{t^2}{\sigma^2} + o(1)\right) \\
&\underset{n\to+\infty}{\longrightarrow} e^{-t^2/\sigma^2}
\end{aligned}$$

par continuité de l'exponentielle.

La grande force de ce théorème réside dans le fait qu'en manipulant n'importe quelle loi de probabilité, on peut toujours se ramener à une loi normale d'où l'importance de cette dernière.

Remarque 4.3.1 *Sous les hypothèses du théorème précédent, prenons $a < b$ et $h = \mathbf{1}_{[a,b]}$. Comme $Z_n convloi Z$ alors*

$$\mathbb{E}\left(h\left(\frac{X_1+\cdots+X_n-nm}{\sigma\sqrt{n}}\right)\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(h(Z)).$$

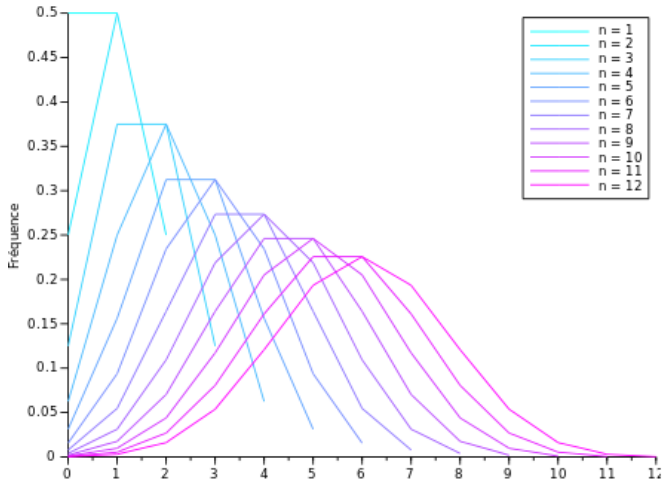
Autrement dit,

$$\mathbb{P}\left(a \leq \frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \leq b\right) \underset{n\to+\infty}{\longrightarrow} \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = F_Z(b) - F_Z(a) \ .$$

Etant donné qu'on peut lire directement sur la table de loi normale centrée réduite, cette dernière propriété sera souvent utilisée dans les exercices.

4.3.1 Illustration

Supposons qu'on est entrain d'effectuer des lancers indépendants d'une pièce de monnaie équilibrée et notons X_k; $k \in \mathbb{N}^*$ les v.a. qui prennent la valeur 1 si on obtient pile au lancer numéro k et 0 sinon. Les v.a. $X_1, \ldots, X_n$ sont i.i.d. et suivent la loi de Bernouilli de paramètre $\frac{1}{2}$.
Ci-dessous une réprésentation graphique de la loi de S_n pour $n \in \{1, 2 \ldots 12\}$



On constate que plus le nombre de tirages augmente, plus la courbe de fréquence se rapproche d'une densité d'une loi normale.