El Hassan Essaky
Faculté Polydisciplinaire de Safi
Département Maths-Info

Introduction au calcul des probabilités

SMA Semestre 6
First draft

Contents

1 Introduction **5**
1.1 Espace probabilisable . . . 5
1.1.1 Tribus . . . 5
1.1.2 Classe monotone . . . 8
1.1.3 Variables aléatoires . . . 9
1.1.4 Loi de probabilité d'une variable aléatoire . . . 9
1.2 Indépendance en probabilité . . . 9
1.2.1 Indépendance d'événements . . . 10
1.2.2 Indépendance de tribus . . . 11
1.2.3 Indépendance de variables aléatoires . . . 11
1.2.4 Lien entre les différents types d'indépendance . . . 12
1.2.5 Loi conjointe d'un n-uplet de variables aléatoires indépendantes . . . 12

2 Lois des v.a **15**
2.1 Fonction de répartition . . . 15
2.1.1 Exemples des lois des variables aléatoires . . . 18
2.1.2 Changement de variables . . . 19
2.2 Vecteurs aléatoires . . . 20
2.2.1 Fonction de répartition . . . 20
2.2.2 Changement de variables . . . 21

3 Moments des v.a **23**
3.1 Variables aléatoires intégrables . . . 23
3.2 Espérance d'une v.a. . . . 23
3.3 Calcul du loi d'une variable aléatoire et loi de la somme de deux variables aléatoires . 27
3.4 Inégalités . . . 29

4 Convergence des v.a **31**
4.1 Fonctions caractéristiques . . . 31
4.2 Lemme de Borel-Cantelli . . . 34
4.3 Types de convegence . . . 36
4.3.1 Convergence presque sûre . . . 36
4.3.2 Convergence L^p . . . 38
4.3.3 Convergence en probabilité . . . 39
4.3.4 Lien entre différentes convergences . . . 41
4.3.5 Convergence en loi . . . 42
4.4 Loi des grands nombres . . . 44
4.5 Théorème central-limite . . . 47

5 Exercices **49**

Chapitre 1

Introduction au calcul de probabilités

Ce chapitre d'introduction a pour but d'introduire les objets mathématiques de base utilisés pour le calcul des probabilités.

1.1 Espace probabilisable et loi de variable aléatoire

Tout phénomène aléatoire fait appel à deux ensembles de type différent.
• Un ensemble Ω, appelé espace fondamental ou univers, qui contient l'ensemble de tous les résultats possibles. Ces derniers sont également appelés épreuves.
• Une famille $\mathcal{B}$ de parties (i.e. de sous ensembles) de Ω. Ces parties sont appelées des événements. On dit que l'événement A s'est réalisé si et seulement si le résultat ω de Ω qui s'est produit appartient à A.

Rappelons tout d'abord quelques résultats d'intégration. Pour un ensemble Ω, on note $\mathcal{P}(\Omega) = \{A : A \subset \Omega\}$ l'ensemble de ses parties.

1.1.1 Tribus

Définition 1.

On dit qu'une famille $\mathcal{B} \subset \mathcal{P}(\Omega)$ est une tribu si

1. $\Omega \in \mathcal{B}$,
2. $\mathcal{B}$ est stable par passage au complémentaire, i.e. $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$,
3. $\mathcal{B}$ est stable par réunion dénombrable, i.e. $(\forall n \in \mathbb{N} : A_n \in \mathcal{B}) \Rightarrow \bigcup_{n=0}^{+\infty} A_n \in \mathcal{B}$.

Si la condition sur l'union est remplacée par la stabilité par union finie, on définit alors une algèbre (ou algèbre de Boole).

Définition 2.

Lorsque $\mathcal{B}$ est une tribu sur Ω, le couple $(\Omega, \mathcal{B})$ est appelé espace probabilisable (ou mesurable).

Proposition 3.

1. L'image réciproque d'une tribu par une application f est une tribu.
2. Soit I une partie de $\mathbb{N}$ et $(\mathcal{B}_i)_{i\in I}$ une famille de tribus sur le même espace fondamental Ω. La famille de parties $\mathcal{B} = \cap_{i\in I}\mathcal{B}_i$ est une tribu.

On a le théorème suivant.

Théorème 4.

Soit $\mathcal{F}$ une famille de parties de Ω. Il existe une plus petite tribu sur Ω qui contient $\mathcal{F}$. On l'appelle tribu engendrée par $\mathcal{F}$ et on la note $\sigma(\mathcal{F})$.

Preuve. Comme $\mathcal{P}(\Omega)$ est une tribu contenant $\mathcal{F}$, l'ensemble des tribus contenant $\mathcal{F}$ n'est pas vide. L'intersection de ces tribus est encore une tribu. Elle contient $\mathcal{F}$ et c'est forcément la plus petite tribu contenant $\mathcal{F}$. ∎

Définition 5.

Soient $(\Omega, \mathcal{A})$ et $(E, \mathcal{B})$ deux espaces probabilisables. Une application f de Ω vers E est dite mesurable (ou $\mathcal{A}$–mesurable) si

$$\forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{B}.$$

On sait que $f^{-1}(\mathcal{B})$ est une tribu. Dire que la fonction f est mesurable revient donc à dire que $f^{-1}(\mathcal{B}) \subset \mathcal{A}$. Ainsi, pour tout événement B, l'ensemble : $f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\}$, est un événement de la tribu initiale.

Notons en particulier que toute fonction continue est mesurable. De même, pour tout événement A de la tribu $\mathcal{A}$, la fonction $\mathbf{1}_A$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Proposition 6.

1. Si f et g sont deux fonctions mesurables de $(\Omega; \mathcal{A})$ vers $(\mathbb{R}; \mathcal{B}_{\mathbb{R}})$, alors les fonctions $f + g$ et fg sont encore mesurables.
2. Si $(f_n)_n$ est une suite de fonctions mesurables de $(\Omega; \mathcal{A})$ vers $(\mathbb{R}; \mathcal{B}_{\mathbb{R}})$, alors les fonctions

$$\sup_n f_n, \inf_n f_n; \lim\sup_n f_n \text{ et } \lim\inf_n f_n$$

sont mesurables, à condition qu'elles ne prennent pas de valeurs infinies.

Définition 7.

Soit f une application de Ω dans $(E, \mathcal{B})$, espace mesurable. La tribu engendrée par f sur Ω, notée $\sigma(f)$, est la plus petite tribu rendant $f : (\Omega, \sigma(f)) \to (E, \mathcal{B})$ mesurable. C'est la tribu formée des $\{f^{-1}(B) : B \in \mathcal{B}\}$. i.e $\sigma(f) = \{f^{-1}(B) : B \in \mathcal{B}\}$.

Définition 8.

On appelle tribu borélienne sur $\mathbb{R}$, la tribu engendrée par les intervalles ouverts de la forme $]-\infty, x[$, pour tout x dans $\mathbb{R}$. On la note $\mathcal{B}_{\mathbb{R}}$. .

Remarque 1. *On peut montrer que l'on a le résultat suivant. La tribu borélienne est également engendrée par les intervalles de la forme* $]-\infty; x]; [x; +\infty[;]x; +\infty[; [x; y];]x; y[$ *etc...*

Définition 9.

On appelle mesure positive sur l'espace probabilisable $(\Omega, \mathcal{B})$ toute application μ de $\mathcal{B}$ dans $\mathbb{R}_+$ telle que d'une part l'on ait $\mu(\varnothing) = 0$ et que d'autre part pour toute suite $(A_n)_n$ d'éléments de $\mathcal{B}$, deux à deux disjoints, on ait :

$$\mu\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \sum_{n=0}^{+\infty} \mu(A_n), \quad \sigma - \text{additivité}.$$

Le triplet $(\Omega, \mathcal{B}, \mu)$ est appelé espace mesuré.

Définition 10.

Une mesure $\mathbb{P}$ sur $(\Omega, \mathcal{B})$ telle que $\mathbb{P}(\Omega) = 1$ est dite une probabilité. Le triplet $(\Omega, \mathcal{B}, \mathbb{P})$ est appelé espace probabilisé.

Proposition 11.

Une probabilité vérifie les assertions suivantes :

1. $\forall A \in \mathcal{B}, \mathbb{P}(A) = 1 - \mathbb{P}(A^c)$,
2. $\mathbb{P}(\varnothing) = 0$,
3. $\forall (A, B) \in \mathcal{B}^2, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$,
4. Si A et B, deux éléments de A, sont tels que $A \subset B$, on a alors
$$\mathbb{P}(A) \leq \mathbb{P}(B),$$
5. Si $A_1, ..., A_n$ sont des événements de $\mathcal{B}$, on a :
$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Proposition 12.

Si $\mathbb{P}$ est une probabilité sur $(\Omega, \mathcal{B})$ alors

1. Pour toute suite $(A_n)_n$ d'événements croissante, on a :
$$\mathbb{P}\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n) = \sup_n \mathbb{P}(A_n).$$
2. Pour toute suite $(A_n)_n$ d'événements décroissante, on a
$$\mathbb{P}\Big(\bigcap_{n=0}^{\infty} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n) = \inf_n \mathbb{P}(A_n).$$

.

En effet, dans le cas croissant, on pose $C_n = A_n \backslash A_{n-1}$, $n \geq 2$ et $C_1 = A_1$, donc $\cup_{k=1}^n C_k = A_n$. Comme les $(C_n)_n$ sont 2 à 2 disjoints alors

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = \lim_{n\to+\infty} \mathbb{P}(\cup_{k=1}^n C_k) = \lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{P}(C_k) = \sum_{n=1}^{\infty} \mathbb{P}(C_n) = \mathbb{P}(\cup_{n=1}^{\infty} C_n) = \mathbb{P}(\cup_{n=1}^{\infty} A_n).$$

1.1.2 Classe monotone

Dans cette section on rappelle un procédé d'extension des définitions de certains objets sur les tribus après les avoir définis sur des classes restreintes d'ensemble.

Définition 13.

(Classe monotone ou λ-système) Une famille $\mathcal{M}$ de parties de Ω est appelée classe monotone si

1. $\Omega \in \mathcal{M}$;
2. lorsque $A, B \in \mathcal{M}$ et $B \subset A$, alors $A \setminus B \in \mathcal{M}$;
3. $\mathcal{M}$ est stable par réunion croissante ($A_j \in \mathcal{M}, j \in \mathbb{N}, A_j \subset A_{j+1} \Rightarrow \bigcup_{j\in\mathbb{N}} A_j \in \mathcal{M}$).

Pour $\mathcal{E} \subset \mathcal{P}(\Omega)$, on appelle classe monotone engendrée par $\mathcal{E}$, la plus petite classe monotone contenant $\mathcal{E}$, c'est à dire l'intersection de toutes les classes monotones contenant $\mathcal{E}$. On la note $\mathcal{M}(E)$.

Remarque 2. *- Une classe monotone est stable par complémentaire : il suffit d'écrire $A^c = \Omega \setminus A$ pour $\Omega, A \in \mathcal{M}$.*
- Une classe monotone est stable par intersection décroissante.
- Une tribu est une classe monotone. Il suffit pour cela de voir que $A \setminus B = A \cap B^c$.
- Une classe monotone stable par intersection finie est une tribu. En effet cette classe sera aussi stable par réunion finie. On utilise alors la reécriture d'une réunion dénombrable comme une réunion croissante $\left(\bigcup_{j\in\mathbb{N}} A_j = \bigcup_{j\in\mathbb{N}} \bigcup_{k\leq j} A_j \text{ pour toute famille } A_j\ ,\ j \in \mathbb{N}\right)$.

Théorème 14.

(des classes monotones) Soit $\mathcal{E}$ une famille de parties de Ω stable par intersection finie. Alors $\mathcal{M}(\mathcal{E}) = \sigma(\mathcal{E})$.

Théorème 15.

Si deux probabilités P et Q coincident sur $\mathcal{A}$ stable par intersections finies alors P et Q coincident sur $\sigma(\mathcal{A})$.

1.1.3 Variables aléatoires

Définition 16.

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(E, \mathcal{B})$ un espace probabilisable. Une application mesurable X de $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(E, \mathcal{B})$ est appelée variable aléatoire.
Une variable aléatoire (v.a) est donc tout simplement une application mesurable sur un espace de probabilité.
• Lorsque $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, on parle d'une variable aléatoire réel (v.a.r).
• Lorsque $(E, \mathcal{B}) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, on parle d'un vecteur aléatoire.

Tous les résultats sur les fonctions mesurables restent donc vrais pour les variables aléatoires. Ainsi, on pourra parler du supremum sur une famille infinie de variables aléatoires et de limite de variables aléatoires. On sera assuré qu'il s'agit encore de variables aléatoires.

En pratique, lorsque la tribu de l'espace d'arrivée $\mathcal{B}$ est engendrée par un système de générateurs, pour vérifier qu'une fonction est mesurable (en particulier, variable aléatoire), il suffit de vérifer la propriété caractéristique sur les générateurs.

Exercice 1. *Soient $(\Omega, \mathcal{F})$ et $(E, \mathcal{E})$ deux espaces probabilisables et $\mathcal{C}$ une classe de parties de E.*

1. *Montrer que : $\sigma(X^{-1}(C)) = X^{-1}(\sigma(C))$.*

2. *On suppose que : $\sigma(\mathcal{C}) = \mathcal{E}$. Montrer que X est une varibales aléatoire de $(\Omega, \mathcal{F})$ à valeurs dans $(E, \mathcal{E})$ si et seulement si $X^{-1}(\mathcal{C}) \subset \mathcal{F}$.*

3. *En déduire qu'une fonction $X : \Omega \to \mathbb{R}$ est une variable aléatoire réelle ssi pour tout $t \in \mathbb{R}$, $\{X \geq t\} \in \mathcal{F}$.*

4. *On suppose que E un espace topologique et $\mathcal{E}$ la tribu engendrée par la classe des ouverts de E. Montrer que toute fonction continue $X : (E, \mathcal{E}) \longrightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est mesurable.*

1.1.4 Loi de probabilité d'une variable aléatoire

Soit X une variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(E, \mathcal{B})$. Définissons une application $\mathbb{P}_X$ de $\mathcal{B}$ vers $[0,1]$ par :

$$\forall B \in \mathcal{B}, \mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega : X(\omega) \in B\}).$$

La définition précédente a bien un sens puisque l'on a $X^{-1}(B) \in \mathcal{A}$, par mesurabilité de X. On peut donc prendre la probabilité de cet événement.

Définition 17.

$\mathbb{P}_X$ est appelée probabilité image de $\mathbb{P}$ par X ou encore loi de probabilité de la variable aléatoire X. On note $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$.

1.2 Indépendance en probabilité

On introduit dans cette section une notions fondamentale du calcul des probabilités : l'indépendance. L'objective est de caractériser l'indépendance des v.a et les tribus et étudier quelques propriétés liées à cette notion.

1.2.1 Indépendance d'événements

Définition 18.

Deux événements A et B sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sont dits indépendants en probabilité (noté $A \perp B$ ou $A \perp\!\!\!\perp B$) si

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

Rappelons que les événements A et B sont dits disjoints si $A \cap B = \emptyset$ et donc $\mathbb{P}(A \cap B) = 0$. Remarquons enfin que deux événements disjoints ne peuvent en fait être indépendants que si l'un ou l'autre des deux événements est de probabilité nulle, i.e.

$$\mathbb{P}(A \cap B) = 0 = \mathbb{P}(A)\mathbb{P}(B) \Leftrightarrow \mathbb{P}(A) = 0 \text{ ou } \mathbb{P}(B) = 0.$$

En fait, deux événements A et B de probabilités non nulles ne peuvent être à la fois incompatibles et indépendants.

Proposition 19.

1. Si A et B sont indépendants, alors $A \perp B^c$, $A^c \perp B$ et $A^c \perp B^c$
2. Si l'événement A est tel que sa probabilité est soit nulle soit égale à 1, alors

$$\forall B \in \mathcal{A}, A \perp B.$$

Preuve. 1. On a

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap B^c)$$

d'où

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c),$$

ce qui prouve l'indépendance entre A et B^c.
2. Si l'événement A est tel que $\mathbb{P}(A) = 0$, il vient

$$\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0 = \mathbb{P}(A)\mathbb{P}(B).$$

donc $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Si $\mathbb{P}(A) = 1$, on travaille avec $\mathbb{P}(A^c)$.

Définition 20.

Soit $(A_i)_{i=1,\ldots,n}$ une famille d'événements de $\mathcal{A}$. Ces événements sont dits (mutuellement) indépendants en probabilité si :

$$\forall I \subset \{1, \ldots, n\}, \quad \mathbb{P}(\bigcap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i).$$

L'indépendance mutuelle entraîne clairement l'indépendance deux à deux ($A_i \cap A_j$ si $i \neq j$) mais la réciproque est fausse. en effet,

Exemple 1. *On lance deux dés équilibrés et de manière indépendante. Soit A l'événement "le premier dé amène un nombre pair", B l'événement" le "second dé amène un nombre impair" et C l'événement "les deux dés amènent des nombres de même parité". On calcule facilement les probabilités suivantes :*

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2},$$

et

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \cap C) = \mathbb{P}(A \cap C) = \frac{1}{4}.$$

Ainsi, les événement A, B et C sont indépendants deux à deux. En revanche ils ne sont pas mutuellement indépendants puisque l'on a :

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0 \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

Remarquons enfin que l'on peut généraliser cette notion d'indépendance mutuelle pour une famille non nécessairement finie d'éléments.

Définition 21.

Soit I un ensemble d'indices quelconque. Une famille $(A_i)_{i\in I}$ d'événements est une famille d'événements mutuellement indépendants si, pour tout ensemble d'indices K fini et dans I, la famille $(A_i)_{i\in K}$ forme une famille d'événements mutuellement indépendants. Autrement, si pour toute sous-famille $A_{i_1}, \cdots, A_{i_p}$ avec $i_k \in I$, on a

$$\mathbb{P}(\bigcap_{k=1}^{p} A_{i_k}) = \prod_{k=1}^{p} \mathbb{P}(A_{i_k}).$$

1.2.2 Indépendance de tribus

Soit $(\mathcal{A}_i)_{i=1,\ldots,n}$ une famille de sous tribus de $\mathcal{A}$,

Définition 22.

On dit que la famille $(\mathcal{A}_i)_{i=1,\ldots,n}$ est une famille indépendante de sous tribus si pour toute famille d'événements $(A_i)_{i=1,\ldots,n}$ où $A_i \in \mathcal{A}_i$, pour tout i, on a :

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

1.2.3 Indépendance de variables aléatoires

Nous avons, jusqu'à présent, abordé la notion d'indépendance pour des tribus, nous allons transposer cette notion pour les variables aléatoires.
Soit une famille $(X_i)_{i=1,\ldots,n}$ de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs respectivement dans l'espace probabilisable $(E_i, \mathcal{B}_i)$.

Définition 23.

Une famille $(X_i)_{i=1,\ldots,n}$ de variables aléatoires est dite indépendante en probabilité si :
$\forall (B_i)_{i=1,\ldots,n}$ où $B_i \in \mathcal{B}_i$, pour tout i, on a :

$$\mathbb{P}(\bigcap_{i=1}^{n} \{X_i \in B_i\} = \prod_{i=1}^{n} \mathbb{P}(\{X_i \in B_i\}).$$

Théorème 24.

Si, pour tout i, les fonctions φ_i sont des fonctions mesurables de $(E_i, \mathcal{B}_i)$ vers $(E'_i, \mathcal{B}'_i)$, alors l'indépendance des variables aléatoires $(X_i)_{i=1,\ldots,n}$ entraîne celle des $(\varphi_i(Xi))_{i=1,\ldots,n}$.

Preuve. Pour toute famille $(B'_i)_{i=1,\ldots,n}$ où $B'_i \in \mathcal{B}_i$, pour tout i, on a $\varphi_i^{-1}(B'_i) = B_i \in \mathcal{B}_i$ par mesurabilité des φ_i. Il vient alors :

$$[\varphi_i(X_i)]^{-1}(B'_i) = X_i^{-1}(\varphi_i^{-1})(B'_i) = X_i^{-1}(B_i).$$

Maintenant appliquer la défintion d'indépendance et l'indépendance de X_i. ■

1.2.4 Lien entre les différents types d'indépendance

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(\mathcal{A}_i)_{i=1,\ldots,n}$ une famille de sous tribus de $\mathcal{A}$. On a le théorème suivant :

Théorème 25.

On a l'équivalence entre les assertions suivantes :

1. $\perp\!\!\!\perp_{i=1}^n \mathcal{A}_i$,
2. $\perp\!\!\!\perp_{i=1}^n X_i$ pour toute famille $(X_i)_{i=1,\ldots,n}$ de v.a. où X_i est $\mathcal{A}_i$-mesurable,
3. $\perp\!\!\!\perp_{i=1}^n \mathbf{1}_{A_i}$ pour toute famille $(A_i)_{i=1,\ldots,n}$ de v.a. où $A_i \in \mathcal{A}_i$.
4. $\perp\!\!\!\perp_{i=1}^n A_i$ pour toute famille $(A_i)_{i=1,\ldots,n}$ de v.a. où $A_i \in \mathcal{A}_i$.
5. $\perp\!\!\!\perp_{i=1}^n \sigma(A_i)$ pour toute famille $(A_i)_{i=1,\ldots,n}$ de v.a. où $A_i \in \mathcal{A}_i$.

1.2.5 Loi conjointe d'un n-uplet de variables aléatoires indépendantes

Soit $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$, $i = 1, \ldots, n$ des espaces probabilisés. On a le théorème suivant.

Théorème 26.

Il existe une probabilité unique $\mathbb{P}$ sur $(\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i)$, où $\otimes_{i=1}^n \mathcal{A}_i$ est la tribu engendrée par les pavés mesurables $A_1 \times A_2 \times \ldots \times A_n$, $A_i \in \mathcal{A}_i$, telle que pour tout A_i dans $\mathcal{A}_i$, pour $i = 1, \ldots, n$, on ait :

$$\mathbb{P}(A_1 \times A_2 \times \ldots \times A_n) = \prod_{i=1}^n \mathbb{P}_i(A_i).$$

Cette probabilité $\mathbb{P}$ est appelée probabilité produit des $\mathbb{P}_i$ et est notée $\mathbb{P} = \otimes_{i=1}^n \mathbb{P}_i$.

Soit $X_1, \ldots, X_n$ des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs vers respectivement $(\Omega_i, \mathcal{A}_i)_{i=1,\ldots,n}$. On admet que le vecteur $(X_1, \ldots, X_n)$ est encore une variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ vers

$$(\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i).$$

On peut en effet montrer qu'une fonction h à valeurs dans l'espace mesurable $(\prod_{i=1}^{n}\Omega_i, \otimes_{i=1}^{n}\mathcal{A}_i)$ est mesurable si, et seulement si, $\pi_i \circ h$ est $(\Omega_i, \mathcal{A}_i)$-mesurable, où π_i est la projection sur la i-ième coordonnée.

Définition 27.

On appelle loi conjointe du vecteur $X = (X_1, ..., X_n)$ la loi P_X de X sur

$$(\prod_{i=1}^{n}\Omega_i, \otimes_{i=1}^{n}\mathcal{A}_i).$$

La loi P_{X_i} de chacune des variables aléatoires X_i est alors appelée loi marginale.

Théorème 28.

Les variables aléatoires $(X_i)_{i=1,...,n}$ sont indépendantes si et seulement si on a :

$$\mathbb{P}_X = \otimes_{i=1}^{n}\mathbb{P}_{X_i}.$$

Preuve. D'après la définition de variables aléatoires indépendantes on a les équivalences suivantes :

$$\begin{aligned}
&\perp\!\!\!\perp_{i=1}^{n} X_i \Leftrightarrow \forall A_i \in \mathcal{A}_i, i = 1, ..., n : \mathbb{P}(\cap_{i=1}^{n} X_i^{-1}(A_i)) = \prod_{i=1}^{n}\mathbb{P}(X_i^{-1}(A_i))\\
&\Leftrightarrow \forall A_i \in \mathcal{A}_i, i = 1, ..., n : \mathbb{P}(X^{-1}(A_1 \times ... \times A_n)) = \prod_{i=1}^{n}\mathbb{P}(X_i^{-1}(A_i))\\
&\Leftrightarrow \mathbb{P}_X = \otimes_{i=1}^{n}\mathbb{P}_{X_i}.
\end{aligned}$$

Chapitre 2

Lois des variables aléatoires

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une v.a. de $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(E, \mathcal{B})$. On sait déjà que la loi de X est l'application $\mathbb{P}_X$ de $\mathcal{B}$ vers $[0,1]$ définie par :

$$\forall B \in \mathcal{B}, \mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega : X(\omega) \in B\}).$$

2.1 Fonction de répartition

Définition 29.

On appelle fonction de répartition de la v.a.r. X, la fonction F_X définie sur $\mathbb{R}$ par

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

Théorème 30.

La fonction de répartition F_X d'une v.a.r. X satisfait les propriétés suivantes :

1. $0 \leq F_X(x) \leq 1$ pour tout x dans $\mathbb{R}$,
2. $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.
3. La fonction F_X est croissante,
4. la fonction F_X est continue à droite,
5. On a $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.

F_X vérifiant les propriétée 1-4, est appelée la fonction de répartition de la v.a.r. X.

Preuve. La propriété 1 est évidente puisque la probabilité de n'importe quel événement est toujours positive et inférieure à 1.
Pour établir 2, on a

$$\mathbb{P}(a < X \leq b) = 1 - \mathbb{P}(\{X \leq a\} \cup \{X > b\}) = 1 - (\mathbb{P}(\{X \leq a\}) + \mathbb{P}(\{X > b\})) = \mathbb{P}(\{X \leq b\} - \mathbb{P}(\{X \leq a\} = F_X(b) - F_X(a).$$

Pour la propriété 3, considérons x et x' deux réels tels que $x \leq x'$. On a bien sûr l'inclusion

$$]-\infty, x] \subset]-\infty, x'],$$

et alors

$$\mathbb{P}_X(]-\infty, x]) \subset \mathbb{P}_X(]-\infty, x']).$$

Pour prouver 4, considérons une suite $(h_n)_n$ de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a :

$$\mathbb{P}_X(]x, x+h_n]) = F_X(x+h_n) - F_X(x).$$

Or la suite d'intervalles $(]x, x+h_n])_n$ est décroissante avec n. Ainsi il vient :

$$\lim_{n \to +\infty} \mathbb{P}_X(]x, x+h_n]) = \mathbb{P}_X(\bigcap_{n=1}^{+\infty}]x, x+h_n]) = \mathbb{P}_X(\emptyset) = 0.$$

F est donc continue à droite.
Pour 5, considérons la suite d'intervalles $(]-\infty,-n])_n$ décroissante vers $\emptyset$ quand n tend vers $+\infty$. On a :

$$\lim_{x\to-\infty} F_X(x) = \lim_{n\to+\infty} F_X(-n) = \mathbb{P}_X(\bigcap_{n=1}^{+\infty}]-\infty,-n]) = \mathbb{P}_X(\emptyset) = 0.$$

On admettra le théorème suivant :

Théorème 31.

Toute fonction F définie sur $\mathbb{R}$ et vérifiant les propriétés 1-4 est une fonction de répartition d'une v.a.r.

Lois continues

Définition 32.

On dit qu'une v.a.r. $X : \Omega \to E$ est de loi continue si sa loi P_X est une mesure de probabilité continue i.e. pour tout $x \in E$, on a $\mathbb{P}_X(\{x\}) = 0$.

Une v.a.r. continue est donc telle que : $\forall x \in \mathbb{R}$, $\mathbb{P}_X(\{x\}) = 0$.

Proposition 33.

Une v.a.r. est continue si et seulement si sa fonction de répartition est continue.

Preuve. Soit X une v.a.r. continue. On a F_X est continue à droite, il suffit de montrer qu'elle est continue à gauche. En effet, considérons une suite $(h_n)_n$ de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a la suite d'intervalles $(]x-h_n,x])_n$ est croissante avec n vers $\{x\}$. Ainsi il vient :

$$\lim_{n\to+\infty} (F_X(x) - F_X(x-h_n)) = \lim_{n\to+\infty} \mathbb{P}_X(]x-h_n,x]) = \mathbb{P}_X(\bigcap_{n=1}^{+\infty}]x-h_n,x]) = \mathbb{P}_X(\{x\}) = 0.$$

F est donc continue à gauche. Et on a

$$\mathbb{P}_X(\{x\}) = F_X(x) - \lim_{y\to x^-} F_X(y).$$

Définition 34.

Une application borélienne $f : \mathbb{R} \to \mathbb{R}$ est appelée une fonction densité si

1. f est positive,
2. $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1.$

On dit alors qu'une v.a.r. X, a pour densité f si la loi de X, P_X, est de densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}$, i.e. $dP_X = f(x)dx$, où dx est la mesure de Lebesgue sur $\mathbb{R}$. Dans ce cas, on dit ue X (ou sa loi P_X) est une v.a.r. absolument continue.

Lorsqu'elle existe la densité f est naturellement reliée à la fonction de répartition F_X.

Proposition 35.

Si X est une v.a.r. de densité f, sa fonction de répartition F_X vérifie

1. $\forall x \in \mathbb{R}$, $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)dt$.
2. F_X est continue sur $\mathbb{R}$.
3. Si f est continue au point x_0, alors F_X est dérivable en x_0 de dérivée $F_X'(x_0) = f(x_0)$.

Proof. 1. Soit $x \in \mathbb{R}$, et soit $A = \{\omega \in \Omega : X(\omega) \in]-\infty, x]\}$ et $A_n = \{\omega \in \Omega : X(\omega) \in]-n, x]\}$

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \int_{-n}^{x} f(t)dt = \int_{-\infty}^{x} f(t)dt$$

2. F_X est continue à droite, il suffit de montrer qu'elle est continue à gauche. En effet, considérons une suite $(h_n)_n$ de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a la suite d'intervalles $(]-\infty, x-h_n])_n$ est croissante avec n vers $]-\infty, x[$. Ainsi il vient :

$$\lim_{n\to+\infty} F_X(x-h_n) = \lim_{n\to+\infty} \mathbb{P}_X(]-\infty, x-h_n]) = \mathbb{P}_X(\bigcup_{n=1}^{+\infty}]-\infty, x-h_n]) = \mathbb{P}_X(X < x) = F_X(x).$$

F est donc continue à gauche.

3. Comme par hypothèse f est continue en x_0, elle est définie sur tout un voisinage de x_0. Par continuité de f en x_0, on a:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in]x_0 - \delta, x_0 + \delta[: |f(x) - f(x_0)| < \varepsilon.$$

Pour tout h tel que $0 < |h| < \delta$, on a alors

$$F_X(x_0 + h) - F_X(x_0) = \int_{x_0}^{x_0+h} f(t)dt.$$

D'où

$$|F_X(x_0 + h) - F_X(x_0) - hf(x_0)| = |\int_{x_0}^{x_0+h} (f(t) - f(x_0))dt| \leq |\int_{x_0}^{x_0+h} |f(t) - f(x_0)|dt| \leq |h|\varepsilon.$$

En divisant par h puis en faisant h vers 0, on constate que F_X est dérivable en x_0, de dérivée $F_X'(x_0) = f(x_0)$.

□

Remarque 3. *La fonction de répartition d'une variable aléatoire X est donc continue si et seulement si X est continue, i.e. si et seulement si*

$$\forall x \in \mathbb{R},\ \mathbb{P}[X = x] = 0.$$

Proposition 36.

Si X est une v.a.r. de fonction de répartition F telle que F est dérivable, alors X a une densité f qui est égale à $f(x) = F'(x)$. Si F est dérivable partout sauf en un nombre fini de point, X est encore continue et elle a pour densité $f = F'$ (que l'on peut calculer partout sauf en un nombre fini de points, on met n'importe quelle valeur pour f aux points où F n'est pas dérivable).

2.1.1 Exemples des lois des variables aléatoires

Lois discrètes

Définition 37.

La mesure de Dirac δ_a en a est définie par

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

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ une variable aléatoire discrète avec $X(\Omega) = \{x_i : i \in I\}$ et I une partie dénombrable, alors sa loi est une somme de mesures de Dirac en ses atomes:

$$\mathbb{P}_X = \sum_{i \in I} \mathbb{P}(X = x_i)\delta_{x_i}$$

1. Loi de Dirac : Soit $x_0 \in \mathbb{R}$. Une v.a.r. X est dite de loi de Dirac δ_{x_0} si elle est à valeurs dans $\mathbb{R}$ et telle que $P_X = \delta_{x_0}$.
2. Loi de Bernoulli : Une v.a.r. X est dite variable aléatoire de Bernoulli de paramètre p, (pour $p \in [0,1]$) si elle est à valeurs dans $\{0,1\}$ et si
$$\mathbb{P}_X(\{1\}) = \mathbb{P}(X = 1) = p, \mathbb{P}_X(\{0\}) = \mathbb{P}(X = 0) = 1 - p.$$
3. Loi binomiale Une v.a.r. X est dite de loi Binomiale de paramètres n et p (pour $n \in \mathbb{N}^*$ et $p \in [0,1]$) si elle est à valeurs dans $\{0, 1, ..., n\}$ et si
$$\mathbb{P}(X = k) = \mathbb{P}_X(\{k\}) = C_n^k p^k (1-p)^{n-k},$$
pour $k = 0, 1, ..., n$. On écrit $X \sim B(n, p)$.
4. Loi géométrique : Une v.a.r. X est dite de loi géométrique de paramètre p, pour p compris entre 0 et 1, si elle est à valeurs dans N^* et si
$$\mathbb{P}(X = k) = (1-p)^{k-1} p.$$
On note $X \sim G(p)$.
5. Loi de Poisson : Une v.a.r. X est dite de loi de Poisson de paramètre λ, si elle est à valeurs dans $\mathbb{N}$ et si
$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$
On note $X \sim \mathcal{P}(\lambda)$.

Lois continues

1. Loi uniforme sur $[a,b]$: Une v.a.r. X à valeurs dans $[a,b]$ est dite de loi uniforme sur cet intervalle si elle est absolument continue et admet pour densité
$$f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x).$$
On note $X \sim \mathcal{U}([a,b])$.
2. Loi normale $N(\mu, \sigma^2)$ Une v.a.r. X à valeurs dans $\mathbb{R}$ est dite de loi normale de moyenne μ et de variance σ^2 si elle est absolument continue et admet pour densité
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

3. Loi exponentielle : Soit λ un réel strictement positif. Une v.a.r. X à valeurs dans R_+^* est dite de loi exponentielle de paramètre λ si elle est absolument continue et admet pour densité

$$f(x) = \lambda e^{-\lambda x}\mathbf{1}_{]0,+\infty[}(x).$$

   On note $X \sim \mathcal{E}(\lambda)$.

4. Loi Gamma : Une v.a.r. X à valeurs dansR_+^* est dite de loi Gamma $\gamma(\alpha, \beta)$, où α et β sont des réels strictement positifs, si elle est absolument continue et admet pour densité

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha-1}\mathbf{1}_{]0,+\infty[}(x).$$

   On note $X \sim \gamma(\alpha, \beta)$.

5. Loi log-normale : Une v.a.r. X à valeurs dans $]0, +\infty[$ est dite de loi log-normale de paramètre μ et σ^2 si la v.a.r. $Y = \ln(X)$ est de loi normale $N(\mu, \sigma^2)$. On note $X \sim LN(\mu, \sigma^2)$.

2.1.2 Changement de variables

Le problème que l'on se propose d'étudier dans cette partie est la détermination de la loi de fonctions d'une v.a.r. dont on connaît la loi. Soit donc X une v.a.r. de loi P X et de fonction de répartition F_X . Soit φ une application mesurable de $\mathbb{R}$ vers $\mathbb{R}$. La v.a.r. $Y = \varphi \circ X$ est donc encore une v.a.r. et on cherche à déterminer sa loi. Une première méthode, convenant autant aux variables discrètes que continues, consiste à déterminer la fonction de répartition F_Y de Y . On a, pour tout y dans $\mathbb{R}$,

$$F_Y(y) = \mathbb{P}(\varphi \circ X \in]-\infty, y]) = \mathbb{P}_X(\varphi^{-1}(]-\infty, y])).$$

Voyons un exemple d'application de cette méthode.

Exemple 2. *Supposons que la v.a.r. X suive une loi $N(0,1)$ et posons $Y = X^2$. On a :*

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y).$$

On constate déjà que l'on a : $F_Y(y) = 0$ si $y \leq 0$. Par ailleurs,

$$F_Y(y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

De plus , comme cette dernière est absolument continue sur $\mathbb{R}$ et de densité f_X continue sur $\mathbb{R}$, sa f.d.r. F_X est dérivable (de dérivée f_X). Par composition de fonctions dérivables, la f.d.r. F_Y est dérivable sur $\mathbb{R}^+$ (la v.a.r. Y est donc absolument continue) et la densité de Y est donc, pour $y > 0$:

$$f_Y(y) = \frac{1}{2\sqrt{y}} F_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} F_X(-\sqrt{y}) = \frac{1}{\sqrt{y}} \frac{1}{2\pi} e^{-\frac{y}{2}}.$$

Ainsi la loi de Y est $\gamma(\frac{1}{2}, \frac{1}{2})$.
Une deuxième méthode pour calculer la loi de $\varphi(X) = Y$ est donnée par le théorème suivant et ne convient que pour des variables aléatoires absolument continues.

Théorème 38.

Soient S et T deux ouverts de $\mathbb{R}$ et X une v.a.r. absolument continue à valeurs dans S et de densité f_X. Soit φ une bijection de S vers $T = Im(\varphi)$, continûment différentiable ainsi que son inverse (φ^{-1} est C^1-difféomorphisme). Alors, la v.a.r. $Y = \varphi(X)$ est absolument continue, à valeurs dans T et de densité :

$$f_Y(y) = f_X(\varphi^{-1}(y))|(\varphi^{-1})'(y)|\mathbf{1}_{Im(\varphi)}(y).$$

Preuve. On a

$$F_Y(y) = \mathbb{P}_X(\varphi^{-1}(]-\infty, y])) = \int_{\{x:\varphi(x)\le y\}} f_X(x)dx$$

Puisque φ est inversible et φ^{-1} continûment différentiable, le changement de variable $x = \varphi^{-1}(u)$ dans l'intégrale donne

$$F_Y(y) = \int_{-\infty}^{y} f_X(\varphi^{-1}(u))|(\varphi^{-1})'(u)|du.$$

Exemple 3. *Appliquons cette formule pour le calcul de la densité de la loi log-normale. Une v.a.r.* X *est de loi* $LN(\mu, \sigma^2)$ *si la v.a.r.* $Y = \ln X$ *est de loi* $N(\mu, \sigma^2)$*. La fonction* $\varphi = exp$ *est clairement un* C^1*-difféomorphisme de* $\mathbb{R}$ *dans* $\mathbb{R}^+$ *et on a*

$$f_Y(x) = f_X(\varphi^{-1}(x))|(\varphi^{-1})'(x)|\mathbf{1}_{Im(\varphi)}(x).$$

2.2 Vecteurs aléatoires

On a déjà vu que l'on appelle vecteur aléatoire toute variable aléatoire à valeurs dans $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})^{\otimes n}$. On notera X_i , la i-ième coordonnée du vecteur X et rappelons que celle-ci est encore une variable aléatoire.

2.2.1 Fonction de répartition

Définition 39.

On appelle fonction de répartition (conjointe) du vecteur aléatoire $X = (X_1, ..., X_n)$ l'application F_X définie sur $\mathbb{R}^n$ et à valeurs dans $[0,1]$ par :

$$F_X(x) = \mathbb{P}_X(]-\infty, x_1] \times ... \times]-\infty, x_n]),$$

où $x = (x_1, ..., x_n)$ est un vecteur de $\mathbb{R}^n$.

Définition 40.

On dit que la loi $\mathbb{P}_X$ d'une v.a.r. X (ou le vecteur $X = (X_1, ..., X_n)$) absolument continue densité s'il existe une fonction $f : (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ positive et telle que pour tout $x = (x_1, ..., x_n) \in \mathbb{R}^n$ on ait

$$F_X(x) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_X(x_1, ..., x_n)dx_1....dx_n.$$

Théorème 41.

Une fonction f sur $\mathbb{R}^n$ est une densité de probabilité si et seulement si elle vérifie les trois assertions suivantes :

1. f est positive
2. f est mesurable.
3. f est intégrable et $\int_{-\infty}^{+\infty} f(x_1, ..., x_n)dx_1...dx_n = 1.$

2.2.2 Changement de variables

La question que l'on se pose dans ce paragraphe est la même que celle vue précédemment en unidimensionnel. Soit X un vecteur aléatoire dans $\mathbb{R}^n$ et φ une application mesurable de $\mathbb{R}^n$ vers $\mathbb{R}^n$. On veut déterminer la loi du vecteur aléatoire $\varphi(X)$.

Théorème 42.

Soient S et T deux ouverts de $\mathbb{R}^n$ et X une v.a.r. absolument continue à valeurs dans S et de densité f_X. Soit φ une bijection de S vers $T = Im(\varphi)$, continûment différentiable ainsi que son inverse (φ est C^1-difféomorphisme). Alors, la vecteur aléatoire $Y = \varphi(X)$ est absolument continue, à valeurs dans T et de densité :

$$f_Y(y) = f_X(\varphi^{-1}(y))|J_{\varphi^{-1}}(y)|\mathbf{1}_{Im(\varphi)}(y), y \in \mathbb{R}^n.$$

où $J_{\varphi^{-1}}$ est le jacobien de la fonction φ^{-1}.

Théorème 43.

Il y a équivalence entre les assertions suivantes :

1. La famille $(X_i)_{i=1,\dots,n}$ est une famille de v.a.r. indépendantes.
2. La fonction de répartition conjointe F_X est le produit des fonctions de répartitions marginales, i.e. :

$$F_X = \prod_{i=1}^{n} F_{X_i}.$$

   Si de plus la v.a.r. X est absolument continue sur $\mathbb{R}^n$ de densité f_X (continue), les assertions précédentes sont encore équivalentes à

3. La densité conjointe est le produit des densités marginales, i.e. :

$$f_X(x_1,\dots,x_n) = \prod_{i=1}^{n} f_{X_i}(x_i).$$

Preuve. $i) \Rightarrow ii)$: On a

$$F_X(x) = \prod_{i=1}^{n} \mathbb{P}(\{X_i \le x_i\}) = \prod_{i=1}^{n} F_{X_i}(x_i).$$

$ii) \Rightarrow iii)$: Pour tout $x = (x_1,\dots,x_n)$ de $\mathbb{R}^n$ on a :

$$\mathbb{P}_X(]-\infty,x_1]\times\dots\times]-\infty,x_n]) = \prod_{i=1}^{n} \mathbb{P}_{X_i}(]-\infty,x_i])i = 1$$

et on peut montrer que cela est suffisant pour avoir $\mathbb{P}_X = \mathbb{P}_{X_i}$.
$ii) \Rightarrow iii)$ Par hypothèse, on a

$$F_X(x) = \prod_{i=1}^{n} F_{X_i}.$$

Si les densités $(f_{X_i})_{i=1,\dots,n}$ sont continues, on a :

$$\frac{\partial^n F_X}{\partial x_1 \dots \partial x_n} = f_{X_1} \dots f_{X_n}.$$

$iii) \Rightarrow ii)$: Soit $x = (x_1, ..., x_n)$ un réel de $\mathbb{R}^n$, on a :

$$\mathbb{P}(\cap_{i=1}^n \{X_i \in]-\infty, x_i]) = \mathbb{P}_X(]-\infty, x_1] \times ... \times]-\infty, x_n])$$
$$= \int_{]-\infty,x_1]} f_{X_1}(u_1) du_1 \times ... \times \int_{]-\infty,x_n]} f_{X_n}(u_n) du_n$$
$$\mathbb{P}(X_1 \in]-\infty, x_1]) \times ... \times \mathbb{P}(X_n \in]-\infty, x_n]),$$

et, par définition, la famille $(X_i)_{i=1,...,n}$ est donc bien indépendante. Enfin pour $n = 2$, on a l'équivalence suivante :

$$X_1 \perp\!\!\!\perp X_2 \Leftrightarrow f_{(X_1,X_2)}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2).$$

Exemple 4. *Soit (X, Y) un couple de variables aléatoires indépendantes dont chacune suit une loi $N(0,1)$. Posons*

$$U = \frac{X+Y}{\sqrt{2}}, \qquad V = \frac{X-Y}{\sqrt{2}}.$$

Le couple (U, V) est alors formé de variables aléatoires réelles indépendantes et dont chacune suit une loi $N(0,1)$. En effet, puisque les v.a.r. X et Y sont indépendantes et de même loi $N(0,1)$, la densité du couple (X, Y) est

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

La fonction φ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par

$$\varphi(x, y) = (u = \frac{x+y}{\sqrt{2}}, v = \frac{x-y}{\sqrt{2}}),$$

est C^1-difféomrphisme. Le jacobien $J_{\varphi^{-1}} = -1$. On applique alors le théorème du changement de variable et la densité du couple (U, V) est :

$$f_{(U,V)}(u, v) = f_{(X,Y)}(\varphi^{-1}(u, v)) |J_{\varphi^{-1}}(u, v)| \mathbf{1}_{Im(\varphi)}(u, v)$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \mathbf{1}_{\mathbb{R}}(u) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2} \mathbf{1}_{\mathbb{R}}(v).$$

Les variables aléatoires U et V sont donc indépendantes puisque la densité conjointe se factorise et elles sont toutes les deux de loi $N(0,1)$.

Chapitre 3

Moments des variables aléatoires

3.1 Variables aléatoires réelles intégrables et espérance mathématique

On se donne un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

3.2 Espérance d'une v.a.

Comme nous venons de le voir au chapitre précédent, une v.a. dans $(\Omega, \mathcal{A}, \mathbb{P})$ (disons positive pour l'instant) n'est rien d'autre qu'une application mesurable à valeurs dans $\mathbb{R}^+$, on peut alors définir son intégrale.

Définition 44.

L'intégrale de X par rapport à la mesure $\mathbb{P}$ est appelée son espérance. Elle se définit comme suit :

1. Indicatrice : si $X = 1_A$, $\mathbb{E}[X] = \mathbb{P}(A)$.
2. Variable étagée positive : si $X = \sum_{i=1}^{n} a_i 1_{A_i}$ alors $\mathbb{E}[X] = \sum_{i=1}^{n} a_i P(A_i)$. On vérifie que $\mathbb{E}[X]$ ne dépend pas de l'expression de X.
3. Variable positive : si X est une variable aléatoire positive,
$$\mathbb{E}[X] = sup\{\mathbb{E}[Y] : Y \text{ variable aléatoire étagée } \leq X\}.$$
4. Variable quelconque : Si X est une variable aléatoire quelconque telle que $\mathbb{E}[|X|] < +\infty$ (au sens précédent) alors on définit l'espérance de X par
$$\mathbb{E}[X] = \mathbb{E}[X^+] - E[X^-].$$
En effet comme $0 \leq X^+ \leq |X|$ et $0 \leq X^- \leq |X|$, les espérances de variable aléatoires positives $\mathbb{E}[X^+]$ et $\mathbb{E}[X^-]$ sont bien définies. La quantité $\mathbb{E}[|X|]$ s'appelle le moment d'ordre 1 de X. L'espérance d'une variable aléatoire X est définie si son moment d'ordre 1 est fini : $\mathbb{E}[|X|] < \infty$. L'espérence de X est notée :
$$\mathbb{E}(X) = \int_{\Omega} X(\omega)\mathbb{P}(d\omega).$$

Une variable aléatoire X intégrable est dite centrée si $\mathbb{E}[X] = 0$.

De façon générale, si $X = (X_1, ..., X_d)$ et un vecteur aléatoire, on définit son espérance comme étant le vecteur des espérances de ses marginales X_i, $1 \leq i \leq d$:

$$\mathbb{E}[X] = (\mathbb{E}[X1], ..., \mathbb{E}[X_d]).$$

Toutes les propriétés de linéarité, de monotonie, de convergence pour les intégrales restent donc vraies pour l'espérance mathématique.

Proposition 45.

1. Croissance : si X et Y sont deux v.a.r. telles que $X \leq Y$ (c'est à dire $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
2. Si $A \subseteq B$ et $X \geq 0$ alors $\mathbb{E}[X 1_A] \leq \mathbb{E}[X 1_B]$.
3. $\mathbb{E}[X] \leq \mathbb{E}[|X|]$.
4. Linéarité : si X et Y sont deux v.a.r. et $a, b \in \mathbb{R}$, $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
5. $\mathbb{E}[X] = 0$ ssi $X = 0$ presque sûrement.

Preuve. 1. Il suffit de remarquer que :

$$\left\{\mathbb{E}[Z] : Z \text{ variable aléatoire étagée } \leq X\right\} \subset \left\{\mathbb{E}[Z] : Z \text{ variable aléatoire étagée } \leq Y\right\}$$

2. On a $A \subseteq B$ et $X \geq 0$, donc $X 1_A \leq X 1_B$. Le résultat découle par 1.
3. Appliquer 1 car $X \leq |X|$.
4. La preuve se fait tout d'abord pour les fonctions étagées puis l'utilisation le théorème de convergence monotone pour généraliser aux fonctions positives et de signe quelconque.

5. Si $X = 0$ p.s. Si on note $A = \{X \neq 0\}$, alors $\mathbb{P}(A) = 0$. Si $\omega \notin A$, alors $X(\omega) = 0$, et donc

$$\mathbb{E}(X) = \mathbb{E}(X 1_A) + \mathbb{E}(X 1_{A^c}) = 0.$$

Il faut remarquer que, en utilisant la définition de l'espérance et la forme de la fonction en étagée, on a si $\mathbb{P}(N) = 0$ alors $\mathbb{E}(X 1_N) = 0$. ■

Remarque 4. *Pour utiliser correctement les lois des variables aléatoires discrètes, il est essentiel de noter que, pour une fonction f mesurable $\int f d\delta_c = f(c)$: cela est clair pour $f = 1_A$ puis par linéarité pour f étagée et par convergence monotone pour f positive et enfin par linéarité pour f quelconque.*
Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ avec $X(\Omega) = \{x_i : i \in I\}$ dénombrable. La loi de X est donnée par la mesure de probabilité discrète

$$\mathbb{P}_X = \sum_{i \in I} p_i \delta_{x_i}$$

où $p_i = \mathbb{P}(X = x_i), i \in I$. La loi est une somme (dénombrable) de mesures de Dirac : en chaque atome $x \in X(\Omega)$, il y a la masse $\mathbb{P}(X = x)$. Alors X est intégrable ssi

$$E[|X|] = \sum_{i \in I} |x_i| p_i < +\infty,$$

et dans ce cas $E[X] = \sum_{i \in I} x_i p_i$.

De même, la notion de négligeabilité (Un ensemble $A \in \mathcal{A}$ est dit $\mathbb{P}$−négligeable si $\mathbb{P}(A) = 0$.) est conservée mais, dans le langage des probabilités, on ne dit plus "μ-p.p." mais "$\mathbb{P}$-presque sûrement" ou simplement "presque sûrement", noté p.s., quand il n'y a pas de risque de confusion.

On a vu précédemment que bien souvent on ignore ce qu'est l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ et on ne connaît bien que l'espace probabilisé $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathbb{P}_X)$. C'est pourquoi, en particulier dans le calcul de l'espérance mathématique une v.a.r., on utilise le théorème suivant qui permet de transformer le calcul de l'intégrale sur $(\Omega, \mathcal{A}, \mathbb{P})$ en un calcul d'une intégrale sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathbb{P}_X)$.

Théorème 46.

(Théorème de transfert)
Soit X une v.a.r. de $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ de loi $\mathbb{P}_X$ et h une fonction mesurable de $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ vers lui même, positive ou $\mathbb{P}_X$-intégrable. On a alors :

$$\mathbb{E}(h(X)) = \int_\Omega h(X) d\mathbb{P} = \int_\mathbb{R} h(x) d\mathbb{P}_X(x).$$

En particulier : X est $\mathbb{P}$-intégrable ssi

$$\int_\mathbb{R} |x| dP_X(x) < +\infty$$

et son espérance est alors

$$E[X] = \int_\mathbb{R} x dP_X(x)$$

Si X est absolument continue de densité f alors

$$E[X] = \int_\mathbb{R} x f(x) dx \text{ et } E[h(X)] = \int_\mathbb{R} h(x) f(x) dx$$

Preuve. La stratégie est de nouveau de commencer par $h = 1_A$ une indicatrice, puis par linéarité de généraliser aux fonctions étagées et par convergence monotone aux fonctions mesurables positives; enfin par linéarité aux fonctions mesurables quelconques (méthode standard : voir cours d'intégration).

1. Si $h = \mathbf{1}_A$, avec $A \in \mathcal{B}_\mathbb{R}$. On a

$$\mathbb{E}(h(X)) = \mathbb{E}(\mathbf{1}_A(X)) = \mathbb{E}(\mathbf{1}_{\{X \in A\}}) = \mathbb{P}(X \in A) = \mathbb{P}_X(A) = \int_\mathbb{R} h(x) d\mathbb{P}_X(x).$$

2. Le résultat passe sans difficulté au cas des fonctions positive en étagé.

3. 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. Si maintenant h est mesurable positive, on sait que d'après le théorème d'approximation qu'elle existe une suite de fonctions $(h_n)n$ croissnate positive en étagé qui converge vers h. Par application du théorème de convergence monotone et l'étape précédente, on a

$$\mathbb{E}(\lim_{n\to+\infty} h_n(x)) = \lim_{n\to+\infty} \mathbb{E}(h_n(x)) = \lim_{n\to+\infty} \int_\mathbb{R} h_n(x) d\mathbb{P}_X(x) = \int_\mathbb{R} h(x) d\mathbb{P}_X(x).$$

4. Si h une fonction mesurable, on pose $f = f^+ - f^-$ et on applique l'étape 3 pour les fonctions f^+ et f^-. ■

Pour l'espérance, on applique le thèorème de transfert avec $h(x) = x$.
La variance d'une v. a. peut se définir comme suit.

Définition 47.

Si X est une v.a.r. telle que X^2 est intégrable alors la variance de X est la quantité

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

Lemme 48.

La variance vérifie la propriété suivante:

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

Preuve.On a

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

■

Exemple 5. *Soit X une variable aléatoire réelle de densité g et $B \in \mathbb{B}(\mathbb{R})$.*

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

Exemple 6. *Soit X v.a.r. de densité $x \mapsto e^{-x}\mathbf{1}_{x\geq 0}$, par une intégration par parties on obtient*

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

Exemple 7. *Soit X v.a. à valeurs dans $\{0,\dots,n\}$ (n un entier fixé) avec $\forall 0 \leq k \leq n$,*

$$\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k},$$

($p \in [0,1]$ fixé). Alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k\mathbb{P}(X=k) \\
&= \sum_{k=0}^{n} kC_n^k p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-1-(k-1))!} p^k (1-p)^{n-k} \\
&= n\sum_{q=0}^{n-1} C_{n-1}^q p^{q+1} (1-p)^{n-1-q} \\
&= np(p+1-p)^{n-1-q} = np .
\end{aligned}$$

Rappel sur le binôme de Newton : $(a+b)^n = \sum_{i=0}^{n} C_n^i a^i b^{n-i}$.

Exemple 8. *Soit X une v.a. à valeurs dans $\mathbb{N}$ avec $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ ($\lambda > 0$ fixé) (X suit une loi exponentielle de paramètre λ). Alors*

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{+\infty} \frac{k\lambda^k e^{-\lambda}}{k!} \\ &= \sum_{k=1}^{+\infty} \frac{1}{(k-1)!}\lambda^k e^{-\lambda} \\ &= \sum_{q=0}^{+\infty} \frac{\lambda^{q+1} e^{-\lambda}}{q!} = \lambda. \end{aligned}$$

3.3 Calcul du loi d'une variable aléatoire et loi de la somme de deux variables aléatoires

Une autre méthode pour déterminer la loi d'un vecteur aléatoire, est basée sur la caractérisation suivante.

Proposition 49.

La loi d'une variable aléatoire X à valeurs dans $\mathbb{R}^d$ est uniquement déterminée par le calcul de $\mathbb{E}(f(X))$ pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ borélienne bornée telle que $f(X)$ soit intégrable ou positive. Autrement dit :
Soit X variable aléatoire à valeurs dans $\mathbb{R}^d$. S'il existe $g : \mathbb{R}^d \to \mathbb{R}$ telle que $\forall f : \mathbb{R}^d \to \mathbb{R}$ borélienne bornée telle que $f(X)$ soit intégrable ou positive,

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

alors g est la densité de X.

Preuve. La preuve est une conséquence de ce qui précède. ■

Notation 50.

On note $\mathcal{B}_b^+(\mathbb{R}^d)$ (resp. $\mathcal{B}_b^+(\mathbb{R})$) l'ensemble des fonctions boréliennes bornées de $\mathbb{R}^d$ (resp. $\mathbb{R}$) à valeurs dans $\mathbb{R}^+$.

Exemple 9. *Soit X v.a.r. de densité $x \mapsto \frac{e^{-x^2/2}}{\sqrt{2\pi}}$. Soient $(a,b) \in \mathbb{R}^* \times \mathbb{R}$ (X est de loi normale). Calculons la loi de $aX+b$. Soit $f \in \mathcal{B}_b^+(\mathbb{R})$. Nous avons*

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

Donc, par la proposition 49, la variable $aX+b$ a une loi de densité $y \mapsto \frac{\exp\left(-\frac{1}{2}\left(\frac{y-b}{a}\right)^2\right)}{\sqrt{2\pi}\times a}$.

Exemple 10. *Soit (X,Y) v.a. à valeurs dans $\mathbb{R}^2$ de densité $(x,y)\mapsto \frac{1}{\pi}\mathbf{1}_{x^2+y^2\leq 1}$. Calculons la loi de $X+Y$. Soit Soit $f\in\mathcal{B}_b^+(\mathbb{R})$. Soit $F:(x,y)\in\mathbb{R}^2\mapsto f(x+y)\in\mathbb{R}^+$. Alors,*

$$\begin{aligned}\mathbb{E}(f(X+Y)) &= \mathbb{E}(F(X,Y))\\ &= \int_{\mathbb{R}^2} F(x,y)\frac{1}{\pi}\mathbf{1}_{x^2+y^2\leq 1}dxdy\,.\end{aligned}$$

Opérons un changement de variable

$$\begin{cases} u &= x+y\\ v &= x-y\end{cases},\ \begin{cases} x &= \dfrac{u+v}{2}\\ y &= \dfrac{u-v}{2}\end{cases}$$

Difféomorphisme $\phi:(x,y)\in\mathbb{R}^2\mapsto(x+y,x-y)\in\mathbb{R}^2$. Matrice jacobienne :

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

de déterminant $-1/2$. Donc par application du théorème de Fubini-Tonelli,

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

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

Théorème 51.

La loi de la somme de deux variables aléatoires X et Y indépendantes, absolument continues et de densité f_X et f_Y respectivement est de densité

$$f_{X+Y}(u)=\int_{\mathbb{R}} f_X(u-v)f_Y(v)dv.$$

pour tout u dans $\mathbb{R}$. Cette densité est appelée la convoluée de f_X et f_Y. On note $f_X * f_Y$ le produit de convolution.

Preuve. Soit (X,Y) un couple aléatoire sur $\mathbb{R}^2$, absolument continu et de densité $f_{(X,Y)}$. La fonction φ définie sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}^2$ par

$$\varphi(x,y)=(u=x+y, v=y).$$

Ainsi, d'après le théorème du changement de variable

$$f_{(U,V)}(u,v)=f_{(X,Y)}(\varphi^{-1}(u,v))|J_{\varphi^{-1}}(u,v)|\mathbf{1}_{Im(\varphi)}(u,v)=f_{(X,Y)}(u-v,v)\mathbf{1}_{\mathbb{R}^2}(u,v)$$

On peut alors déterminer la densité de la loi marginale de U

$$f_U(u)=\int_{\mathbb{R}} f_{(U,V)}(u,v)dv=\int_{\mathbb{R}} f_{(X,Y)}(u-v,v)dv$$

pour tout u dans $\mathbb{R}$. De plus, si les variables aléatoires X et Y sont indépendantes de densité respective f_X et f_Y , on a :

$$f_{X+Y}(u)=\int_{\mathbb{R}} f_X(u-v)f_Y(v)dv.$$

3.4 Inégalités

Théorème 52.

Inégalité de Jensen
Soit $f : \mathbb{R} \to \mathbb{R}$ mesurable convexe. Soit X v.a.r. intégrable telle que $f(X)$ est intégrable. Alors

$$f(\mathbb{E}(X)) \le \mathbb{E}(f(X)) .$$

Lemme 53.

Soit f une application d'un intervalle I de $\mathbb{R}$ dans $\mathbb{R}$. Les f convexe (resp. strictement convexe) sur I si et seulement si Pour tout $a \in I$, l?application Δ_a de $I \setminus \{a\}$ dans $\mathbb{R}$ définie par

$$\Delta_a(x) = \frac{f(x) - f(a)}{x - a},$$

est croissante (resp. strictement croissante). En pariculier on a
si $x < u < y$ dans I alors

$$\frac{f(u) - f(x)}{u - x} \le \frac{f(y) - f(x)}{y - x} \le \frac{f(y) - f(u)}{y - u}.$$

Comme conséquence on a la proposition suivante.

Corollaire 1. *Soit f une fonction convexe sur un intervalle ouvert I. La fonction f est continue sur I et possède en chaque point a de I une dérivée à droite et une dérivée à gauche telles que*

$$f'_g(a) \le f'_d(a).$$

Preuve du théorème 52. Par application du lemme 53 et son corollaire pour $I = \mathbb{R}$ on a $\forall x, y \in \mathbb{R}$

$$f(x) \ge f(y) + f'_g(y)(x - y)$$

On pose $x = X(\omega)$ et $y = \mathbb{E}(X(\omega))$ on alors

$$f(X(\omega)) \ge f(\mathbb{E}(X(\omega))) + f'_g(\mathbb{E}(X(\omega)))(X(\omega) - \mathbb{E}(X(\omega)))$$

En prenant l'espérance, on trouve le résultat. ■

Théorème 54.

Inégalité de Bienaymé-Tchebychev
Soit X v.a.r. positive, intégrable. Soit $\lambda > 0$. Alors

$$\mathbb{P}(X \ge \lambda) \le \frac{1}{\lambda}\mathbb{E}(X) .$$

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

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

Preuve.[Démonstration du théorème 54] Pour tout ω, $X(\omega) \ge \lambda \mathbf{1}_{X(\omega) \ge \lambda}$ donc, par la propriété de croissance de l'espérance mathématique

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

■

Preuve.[Démonstration du corollaire 2]

$$\begin{aligned} \mathbb{P}(|X-\mathbb{E}(X)| \geq \lambda) &= \mathbb{P}((X-\mathbb{E}(X))^2 \geq \lambda^2) \\ \text{(par inégalité de Bienaymé-Tchebichev)} &\leq \frac{1}{\lambda^2}\mathbb{E}((X-\mathbb{E}(X))^2)\,. \end{aligned}$$

∎

Théorème 55.

Inégalité de Markov
Si X v.a.r. avec X^2 intégrable et si $\lambda > 0$ alors

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

Preuve.

$$\begin{aligned} \mathbb{P}(|X| \geq \lambda) &= \mathbb{P}(X^2 \geq \lambda^2) \\ \text{(par inégalité de Bienaymé-Tchebichev)} &\leq \frac{\mathbb{E}(X^2)}{\lambda^2}\,. \end{aligned}$$

∎

Chapitre 4

Convergence de variables aléatoires

4.1 Fonctions caractéristiques

On a déjà vu que la fonction de répartition d'une variable aléatoire X caractérise sa loi. Elle est aussi uniquement déterminée par le calcul de $\mathbb{E}(f(X))$ pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ borélienne bornée telle que $f(X)$ soit intégrable ou positive. Dans ce chapitre, on va plus loin encore et on montre qu'il suffit de ne considérer que deux types de fonction : les sinus et les cosinus. Bien sûr, tout cela est lié à l'analyse de Fourier.

Définition 56.

Soit μ une probabilité (ou une mesure finie) sur $\mathbb{R}^d$. On appelle transformée de Fourier de μ la fonction complexe définie par

$$\forall t \in \mathbb{R}^d,\ \hat{\mu}(t) = \int_{\mathbb{R}^d} e^{it.x} d\mu(x),$$

où $t.x =< t, x >$ désigne le produit scalaire de t et x.

Proposition 57.

Si deux mesures finies μ et ν sur $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$ ont même transformée de Fourier ($\hat{\mu} = \hat{\nu}$), elles sont égales ($\mu = \nu$).

Définition 58.

Soit X est une variable aléatoire à valeurs dans $\mathbb{R}^d$, on appelle fonction caractéristique de X, notée φ_X, la transformée de Fourier de la loi $\mathbb{P}_X$, *i.e.*

$$\forall t \in \mathbb{R}^d,\ \varphi_X(t) = \hat{\mathbb{P}}_X(t) = \int_{\mathbb{R}^d} e^{it.x} d\mathbb{P}_X(x) = \mathbb{E}(e^{it.X}).$$

Exemple 11. *1. Soit X une v.a.r. qui suit une loi de Poisson $P(\lambda)$. Comme $\mathbb{P}_X = \sum_{k\geq 0} \mathbb{P}(X = k)\delta_k$ On a*

$$\begin{aligned}
\forall t \in \mathbb{R},\ \varphi_X(t) &= \mathbb{E}(e^{it.X}) = \int_{\mathbb{R}} e^{it.x} d\mathbb{P}_X(x) \\
&= \sum_{k=0}^{+\infty} e^{itk}\mathbb{P}(X = k) = e^{-\lambda} \sum_{k=0}^{+\infty} e^{itk} \frac{\lambda^k}{k!} \\
&= e^{\lambda(e^{it}-1)}
\end{aligned}$$

2. Si X une v.a.r. qui suit une loi normale $\mathcal{N}(0,1)$, alors sa fonction caractéristique est définie pour tout t dans $\mathbb{R}$ par :

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

On a

$$\begin{aligned}\forall t \in \mathbb{R},\ \varphi_X(t) &= \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{it.x} d\mathbb{P}_X(x) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{it.x} e^{-\frac{x^2}{2}} dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} \cos(tx) e^{-\frac{x^2}{2}} dx,\end{aligned}$$

On remarque que le théorème de dérivabilité sous le signe s'applique et une intégration par parties donnent que

$$\begin{aligned}\forall t \in \mathbb{R},\ \varphi_X'(t) &= -\frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x \sin(tx) e^{-\frac{x^2}{2}} dx \\ &= -t\varphi_X(t).\end{aligned}$$

Ce qui implique que $\varphi_X(t) = Ce^{-\frac{t^2}{2}}$. Comme $\varphi_X(0) = 1$, on en déduit que $\varphi_X(t) = e^{-\frac{t^2}{2}}$.

Soit maintenant une variable aléatoire Y de loi $\mathcal{N}(\mu, \sigma^2)$. En notant $X = \dfrac{Y-\mu}{\sigma}$, on a:

$$\varphi_Y(t) = \mathbb{E}(e^{itY}) = \mathbb{E}(e^{it(\sigma X+\mu)}) = e^{it\mu}\mathbb{E}(e^{it\sigma X}) = e^{it\mu}\varphi_X(\sigma t) = e^{it\mu - \frac{t^2\sigma^2}{2}}.$$

Théorème 59.

Pour $X = (X_1, \ldots, X_d)$ un vecteur de $\mathbb{R}^d$, notons φ_X et φ_{X_i} , respectivement les fonctions caractéristiques de X et de X_i . On a alors les résultats suivants :

1. $\forall t \in \mathbb{R}^d$, $|\varphi_X(t)| \leq 1 = \varphi_X(0)$.
2. $\forall j \in \{1, \ldots, n\}$, et $\forall t_j \in \mathbb{R}$: $\varphi_{X_j}(t_j) = \varphi_X(0, \ldots, 0, t_j, 0, \ldots, 0)$.
3. $\forall t \in \mathbb{R}^d$, $\overline{\varphi_X(t)} = \varphi_X(-t)$.
4. La fonction φ_X est continue sur $\mathbb{R}^d$.
5. Si X est une v.a.r. dans L^p , où p est un entier, alors φ_X est dérivable p-fois et on a

$$\forall t \in \mathbb{R}^d,\ \forall k \leq p : \varphi_X^{(k)}(t) = i^k \mathbb{E}(X^k e^{itX}).$$

En particulier

$$\forall k \leq p : \varphi_X^{(k)}(0) = i^k \mathbb{E}(X^k).$$

Preuve. Les propriétés 1, 2 et 3 sont évidentes. La propriété 4 de la continuité est une application immédiate du théorème de continuité sous le signe somme avec pour fonction dominante la constante 1. La propriété 5 est une conséquence du théorème de dérivation sous le signe somme. ∎

Proposition 60.

Si la mesure μ est une mesure produit $\mu = \mu_1 \otimes \mu_2 \cdots \otimes \mu_d$ de mesures finies sur $\mathbb{R}$, alors pour tout $t = (t_1, \ldots, t_d) \in \mathbb{R}^d$

$$\hat{\mu}(t) = \prod_{i=1}^{d} \hat{\mu}_i(t_i)$$

Comme son nom l'indique, la fonction caractéristique caractérise la loi d'un vecteur aléatoire.

Théorème 61.

La fonction caractéristique caractérise la loi d'une variable aléatoire. Autrement dit, si deux variables aléatoires X et Y ont même fonction caractéristique, elles ont même loi.

Preuve. Il suffit d'appliquer la Proposition 57. ∎

Théorème 62.

Les variables aléatoires $X_1, \dots, X_d$ sont indépendantes si et seulement si pour tout $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, on a :

$$\varphi_{(X_1,\dots,X_d)}(t) = \prod_{i=1}^{d} \varphi_{X_i}(t_i).$$

Preuve. $\Rightarrow$) On a par indépendance des variables aléatoires $X_1, \dots, X_d$,

$$\varphi_{(X_1,\dots,X_d)}(t) = \mathbb{E}(e^{i\sum_{i=1}^{d} t_i X_i}) = \mathbb{E}(e^{itX_1} \dots e^{itX_d}) = \prod_{i=1}^{d} \mathbb{E}(e^{itX_i}) = \prod_{i=1}^{d} \varphi_{X_i}(t_i).$$

$\Leftarrow$) Notons ψ la fonction caractéristique de la loi $\nu = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \cdots \otimes \mathbb{P}_{X_d}$. Par la proposition 60, on a

$$\psi(t) = \prod_{i=1}^{d} \varphi_{X_i}(t_i),$$

donc, par hypothèse $\psi = \varphi_X$. Ce qui donne que $\hat{\nu} = \hat{\mathbb{P}}_X$ et donc $\nu = \mathbb{P}_X$ par la Proposition 57. ∎

Proposition 63.

Soit (X, Y) un couple de variables aléatoires indépendantes. Alors, pour tout réel t, on a :

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t).$$

Preuve. On a par indépendance des variables aléatoires X et Y,

$$\varphi_{X+Y}(t) = \mathbb{E}(e^{i(X+Y)t}) = \mathbb{E}(e^{itX} e^{itY}) = \varphi_X(t)\varphi_Y(t).$$

∎

4.2 Lemme de Borel-Cantelli

Définition 64.

Rappelons qu'un ensemble $A \in \mathcal{A}$ est dit $\mathbb{P}$−négligeable si $\mathbb{P}(A) = 0$.

Une propriété est vraie $\mathbb{P}$-p.s., ou simplement p.s. si l'ensemble des points où elle n'est pas vraie est $\mathbb{P}$-négligeable.

Par exemple, $X = Y$ p.s. si l'ensemble $\Big\{X \neq Y\Big\}$ est inclu dans une partie $\mathbb{P}$-négligeable, i.e. $\Big\{X \neq Y\Big\} \subset A$ avec $A \in \mathcal{A}$ et $\mathbb{P}(A) = 0$. i.e. Il existe un ensemble N $\mathbb{P}$-négligeable tel que $\forall \omega \in N^c$, on a $X(\omega) = Y(\omega)$.

Si $\Big\{X \neq Y\Big\} \in \mathcal{A}$ alors on peut écrire $\mathbb{P}\Big(\Big\{X \neq Y\Big\}\Big) = 0$ ou $\mathbb{P}\Big(\Big\{X = Y\Big\}\Big) = 1$.

Notation: l'ensemble des parties $\mathbb{P}$−négligeable est noté $\mathcal{N}$.

Rappelons aussi que la notion des limites supérieures et inférieures pour les ensembles. Si $(A_n)_{n\in\mathbb{N}}$ est une suite de parties de $\mathcal{A}$, ses limites supérieures et inférieures sont les événements définis par

$$\limsup A_n = \overline{\lim} A_n = \bigcap_{n\geq 0}\bigcup_{k\geq n} A_k \quad \text{et} \quad \liminf A_n = \underline{\lim} A_n = \bigcup_{n\geq 0}\bigcap_{k\geq n} A_k.$$

$\limsup A_n$signifie que ω appartient à une infinité de A_n et $\liminf A_n$ signifie que ω appartient à tous les A_n sauf un nombre fini. On a aussi

$$\Big(\limsup A_n\Big)^c = \liminf A_n^c \quad \text{et} \quad \Big(\liminf A_n\Big)^c = \limsup A_n^c.$$

Voici un résultat très utile en pratique connu sous le nom de lemme de Borel-Cantelli.

Théorème 65.

(Lemme de Borel-Cantelli)

1. Soient $A_1, A_2, \ldots$ une famille dénombrable d'événements telle que $\sum_{n\geq 1} \mathbb{P}(A_n) < \infty$. Alors
$$\mathbb{P}(\limsup A_n) = 0.$$
Ce qui s'énonce : p.s., seul un nombre fini d'événements A_n est réalisé. Ce qui signifie aussi, $\mathbb{P}(\{\omega : \omega \in \text{ une infinité de } A_n\}) = 0$.

2. Si on a $A_1, A_2, \ldots$ une famille dénombrable d'événements <u>indépendants</u> tels que
$$\sum_{n\geq 1} \mathbb{P}(A_n) = \infty,$$
alors
$$\mathbb{P}(\limsup A_n) = 1.$$
Ce qui s'énonce aussi : p.s., une infinité d'événements A_n est réalisée
$$\mathbb{P}(\{\omega : \omega \in \text{ une infinité de } A_n\}) = 1 .$$

Preuve

1. Nous avons
$$\mathbb{P}(\limsup A_n) = \mathbb{P}\Big(\bigcap_{n\geq 1}\bigcup_{k\geq n} A_k\Big) \leq \mathbb{P}\Big(\bigcup_{k\geq n} A_k\Big) \leq \sum_{k\geq n} \mathbb{P}(A_k).$$
Comme la série $\sum_{n\geq 1} \mathbb{P}(A_n)$ converge, le reste $\sum_{k\geq n} \mathbb{P}(A_k)$ converge vers 0, et donc $\mathbb{P}(\limsup A_n) = 0$. D'où le résultat.

2. Soit n_0 fixé, par intersection décroissante , on a
$$\mathbb{P}\left(\bigcap_{k\geq n_0} A_k^c\right) = \lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right).$$
Le but est de démontrer que $\mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right) = 0$. Car, par réunion,
$$\mathbb{P}\left[\left(\limsup A_n\right)^c\right] = \mathbb{P}\left(\bigcup_{n_0\geq 1}\bigcap_{n_0\leq k} A_k^c\right) \leq \sum_{n_0\geq 1} \mathbb{P}(\bigcap_{n_0\leq k} A_k^c) = 0\ .$$
Nous avons donc par indépendance $\forall n \geq n_0$
$$\begin{aligned}\mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right) &= \prod_{n_0\leq k\leq n} \mathbb{P}\left(A_k^c\right)\\ &= \prod_{n_0\leq k\leq n} (1-\mathbb{P}(A_k)),\end{aligned}$$
donc $\ln\left(\mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right)\right) = \sum_{n_0\leq k\leq n} \ln(1-\mathbb{P}(A_k))$. Comme $\forall x \in [0,1]$
$$\ln(1-x) \leq -x,$$
alors nous avons
$$\ln(1-\mathbb{P}(A_k)) \leq -\mathbb{P}(A_k).$$
Par conséquent, la série précédente diverge. Donc
$$\sum_{k=n_0}^{+\infty} \ln(1-\mathbb{P}(A_k)) = \lim_{n\to+\infty} \sum_{n_0\leq k\leq n} \ln(1-\mathbb{P}(A_k)) = -\infty.$$
Ce qui donne
$$\lim_{n\to+\infty} \ln\left(\mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right)\right) = -\infty.$$
Donc
$$\lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right) = 0.$$
Pour tout $n \geq n_0$, $\bigcap_{n_0\leq k\leq n+1} A_k^c \subset \bigcap_{n_0\leq k\leq n} A_k^c$. Donc par intersection décroissante
$$\mathbb{P}\left(\bigcap_{n_0\leq k} A_k^c\right) = \lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{n_0\leq k\leq n} A_k^c\right) = 0\ .$$

Et donc par réunion,

$$\mathbb{P}\left(\bigcup_{n_0\geq 1}\bigcap_{n_0\leq k} A_k^c\right) \leq \sum_{n_0\geq 1}\mathbb{P}\left(\bigcap_{n_0\leq k} A_k^c\right) = 0\ .$$

Ce qui implique que

$$\mathbb{P}\left(\bigcup_{n_0\geq 1}\bigcap_{n_0\leq k} A_k^c\right) = 0\ .$$

Par passage au complémentaire, nous obtenons

$$\mathbb{P}\left(\bigcap_{n_0\geq 1}\bigcup_{n_0\leq k} A_k\right) = 1\ .$$

Ce qui est équivalent à

$$\mathbb{P}(\limsup A_n) = 1.$$

$\blacksquare$

Exemple 12. *Considérons une suite $(X_n)_{n\geq 1}$ de variables aléatoires réelles, telle que, pour tout $n \geq 1$*

$$\mathbb{P}(X_n = 0) = \frac{1}{n^2}.$$

On pose $A_n := \{X_n = 0\}$. On a $\sum_{n\geq 1}\mathbb{P}(A_n) = \sum_{n\geq 1}\frac{1}{n^2} < \infty$, donc d'après le lemme de Borel-Cantelli la probabilité que $A_n = \{X_n = 0\}$ se produise pour une infinité d'indices n est 0. En d'autres termes, avec une probabilité de 1, X_n est non nul à partir d'un certain rang (aléatoire) n_0.

4.3 Les différentes notions de convergence

On se donne un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Le but de ce paragraphe est d'étudier les différents types de convergence pour une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ disons pour simplifier à valeurs dans $\mathbb{R}^d$. Nous distinguerons quatre type de convergence qui se divisent en deux catégories. La première concerne les modes de convergence "trajectorielle" pour lesquels on peut déduire une information plus ou moins précise sur la convergence de la suite $(X_n(\omega))_{n\in\mathbb{N}}$ à ω fixé. Dans ce cas de figure, les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. La convergence en loi, quant à elle, ne concerne pas directement la convergence de la suite d'applications mesurables $(X_n(\omega))_{n\in\mathbb{N}}$ mais traite de la convergence de la suite de mesures de probabilité $\mathbb{P}_{X_n}$. On se donne $(X_n)_{n\geq 0}$ une suite des v.a. à valeurs dans $\mathbb{R}^d$.

4.3.1 Convergence presque sûre

On commence par une des notions que vous connaissez déjà.

Définition 66.

(Convergence presque sûre)
On dit que La suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X et on note $X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$ ou $X_n \underset{n\to+\infty}{\longrightarrow} X$ $\mathbb{P}$-p.s.si

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

ce qui est équivalent à

$$\mathbb{P}\Big(\Big\{\omega\in\Omega : X(\omega) \neq \lim_{n\to+\infty} X_n(\omega)\Big\}\Big) = 0\ ,$$

ou encore il existe un N $\mathbb{P}$-négligeable tel que tel que $\forall\omega\in N^c$, $\lim\limits_{n\to+\infty} X_n(\omega) = X(\omega)$.

Bien évidemment $X_n = (X_n^1,\cdots,X_n^d)$ converge $\mathbb{P}$-p.s. vers $X = (X^1,\cdots,X^d)$ si et seulement si, pour tout $i = 1,\cdots,d$, X_n^i converge vers X^i presque sûrement.
La convergence presque sûre est facile à caractériser.

Proposition 67.

La suite $(X_n)_{n\in\mathbb{N}}$ converge vers X presque sûrement si et seulement si

$$\forall\varepsilon>0,\quad \mathbb{P}\Big(\limsup\Big\{|X_n - X| > \varepsilon\Big\}\Big) = 0.$$

Preuve. $\Rightarrow$) supposons que $X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$, alors il existe un N $\mathbb{P}$-négligeable tel que tel que $\forall\omega\in N^c$, $\lim\limits_{n\to+\infty} X_n(\omega) = X(\omega)$. Soit alors $\omega\in N^c$, alors, $\forall\varepsilon>0$ $\exists k\in\mathbb{N}$ tel que $\forall n\geq k$, $|X_n(\omega) - X(\omega)|\leq\varepsilon$. Donc $\forall\varepsilon>0, \omega\in\liminf\{|X_n(\omega)-X(\omega)|\leq\varepsilon\}$. En d'autres termes $\forall\varepsilon>0$

$$\limsup\Big\{|X_n(\omega)-X(\omega)|>\varepsilon\Big\}\subseteq\Big\{X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X\Big\}^c.$$

Ce qui donne le résultat.
$\Leftarrow$) Suposons que

$$\forall\varepsilon>0, \mathbb{P}\Big(\limsup\{|X_n - X|>\varepsilon\}\Big) = 0,$$

et montrons que $X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$. Posons

$$N = \bigcup_{r\in\mathbb{N}}\limsup\Big\{|X_n(\omega)-X(\omega)|>\frac{1}{2^r}\Big\} = \bigcup_{r\in\mathbb{N}}\bigcap_{k\geq 0}\bigcup_{n\geq k}\Big\{|X_n(\omega)-X(\omega)|>\frac{1}{2^r}\Big\}.$$

On a $N\in\mathcal{F}$ (comme réunion dénombrable des ensembles dans $\mathcal{F}$) et $\mathbb{P}(N=0)$. Soit $\omega\in N^c$. Donc $\forall r\in\mathbb{N}\ \exists k\in\mathbb{N}$ tel que $\forall n\geq k$, $|X_n(\omega)-X(\omega)|\leq\frac{1}{2^r}$. En passant à la limite premiérement quand n tend vers $+\infty$ et après on fait tendre r vers $+\infty$, on trouve le résultat. ∎

Corollaire 3. *Si, pour tout $\varepsilon>0$,*

$$\sum_{n=0}^{+\infty}\mathbb{P}\Big(|X_n - X|>\varepsilon\Big) < +\infty,$$

alors $(X_n)_{n\in\mathbb{N}}$ converge vers X presque sûrement.

Preuve. Il suffit d'appliquer le lemme de Borel-Cantelli. ∎

Le critère de Cauchy de convergence presque sûre a l'avantage de s'exprimer sans la limite X. On peut donc déterminer la convergence presque sûre sans connaître la limite.

Proposition 68.

(Critère de Cauchy) $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement si et seulement si

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\bigcup_{m\geq 0}\bigcap_{n\geq m}\left\{|X_n - X_m| \leq \varepsilon\right\}\right) = 1.$$

Exercice 2. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes.*
Montrer que la suite $(X_n)_{n\in\mathbb{N}}$ converge vers 0 presque sûrement si et seulement si

$$\forall \varepsilon > 0, \quad \sum_{n=0}^{+\infty}\mathbb{P}\Big(|X_n| > \varepsilon\Big) < +\infty.$$

Voici quelques propriétés immédiates de la convergence presque sûre.

Proposition 69.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires.

1. Si $(X_n)_{n\in\mathbb{N}}$ converge vers X presque sûrement et si $f : \mathbb{R}^d \to \mathbb{R}^q$ est une fonction continue alors $(f(X_n))_{n\in\mathbb{N}}$ converge vers $f(X)$ presque sûrement (la convergence p.s. est conservée par continuité).
2. Si $(X_n)_{n\in\mathbb{N}}$ converge vers X et Y presque sûrement alors X et Y sont égales presque sûrement (unicité de la limite p.s.).

Preuve. Il suffit d'appliquer la définition de la convergence presque sûre. ∎

4.3.2 Convergence L^p

Définition 70.

Soit $p \geq 1$ un réel, on dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge dans L^p vers X et on note $X_n \underset{n\to+\infty}{\overset{L^p}{\longrightarrow}} X$ si $\mathbb{E}(|X - X_n|^p) \underset{n\to+\infty}{\longrightarrow} 0$.
Lorsque $p = 1$ on parle de convergence en moyenne, et pour $p = 2$ de convergence en moyenne quadratique.

Remarque 5. *Rappelons le résultat central que vous avez vu en intégration :*
Pour $X \in L^p$, on note $\|X\|_p = \mathbb{E}[|X|^p]^{\frac{1}{p}}$ la norme dans L^p. Pour tout $p \geq 1$, $(L^p, \|.\|_p)$ est un espace de Banach.

Proposition 71.

Soit $(X_n)_{n\in\mathbb{N}} \subset L^q$. Si $(X_n)_{n\in\mathbb{N}}$ converge vers X dans L^q alors$(X_n)_{n\in\mathbb{N}}$ converge vers X dans L^p pour tout $p \leq q$.

Preuve. D'après l'inégalité de Hölder, on a

$$\mathbb{E}\Big(|X - X_n|^p\Big) \leq \Big(\mathbb{E}\Big(|X - X_n|^q\Big)\Big)^{\frac{q}{p}}.$$

ce qui donne le résultat. ∎

Signalons également que le théorème de convergence dominée permet de passer de la convergence presque sûre à la convergence en moyenne.

Théorème 72.

(Convergence dominée). Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires. Supposons que $(X_n)_{n\in\mathbb{N}}$ converge vers X presque sûrement. S'il existe une variable aléatoire Y telle que,

$$\forall n \in \mathbb{N}, |X_n| \leq Y \ p.s., \text{ et } Y \in L^p,$$

alors $(X_n)_{n\in\mathbb{N}} \subset L^p$ et converge vers X dans L^p.

Exemple 13. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables réelles indépendantes. La loi de X_n est donnée par*

$$\mathbb{P}(X_n = 0) = 1 - p_n, \mathbb{P}(X_n = x_n) = p_n,$$

où $0 < p_n < 1$ et $x_n \geq 1$.

1. *Les variables étant indépendantes, X_n converge presque sûrement vers $X = 0$ si et seulement si, $\forall \varepsilon > 0, \sum_{n=0}^{+\infty} \mathbb{P}(|X_n| > \varepsilon) < +\infty$, c'est à dire si et seulement si $\sum_{n=0}^{+\infty} p_n < +\infty$ ($x_n \geq 1$).*
2. *$(X_n)_{n\in\mathbb{N}}$ converge dans L^p vers 0 si et seulement si $\mathbb{E}[|X_n|^p] = x_n^p p_n \to 0$.*
3. *Si $p_n = 2^{-n}, x_n = 2^n$: X_n converge vers 0 presque sûrement mais pas dans L^1;*
4. *Si $p_n = n^{-1}, x_n = 1$: X_n converge vers 0 dans L^p pour tout $p \geq 1$ mais ne converge pas vers 0 presque sûrement;*
5. *$p_n = n^{-2}, x_n = n$: X_n converge vers 0 dans L^1 mais pas dans L^2.*

4.3.3 Convergence en probabilité

Nous en venons à présent à une notion nouvelle de convergence : la convergence en probabilité.

Définition 73.

On dit que X_n converge en probabilité vers X et on note $X_n \underset{n\to+\infty}{\xrightarrow{\mathbb{P}}} X$ si

$$\forall \varepsilon > 0, \quad \mathbb{P}\Big(|X - X_n| > \varepsilon\Big) \underset{n\to+\infty}{\longrightarrow} 0.$$

Proposition 74.

On a l'équivalence suivante:

$$X_n \underset{n\to+\infty}{\xrightarrow{\mathbb{P}}} X \quad \text{si et seulement si} \quad \lim_{n\to+\infty} \mathbb{E}\Big(\frac{|X_n - X|}{1 + |X_n - X|}\Big) = 0.$$

Preuve. $\Rightarrow$) On suppose sans perdre de généralité que $X = 0$. On a

$$\frac{|X_n|}{1+|X_n|} \leq \frac{|X_n|}{1+|X_n|} 1_{\{|X_n|>\varepsilon\}} + \varepsilon 1_{\{|X_n|\leq\varepsilon\}} \leq 1_{\{|X_n|>\varepsilon\}} + \varepsilon.$$

Donc

$$\mathbb{E}\Big(\frac{|X_n|}{1+|X_n|}\Big) \leq \mathbb{P}(|X_n| > \varepsilon) + \varepsilon.$$

Comme ε est arbitraire, en passant à la limite quand n tend vers $+\infty$, on trouve le résultat.
$\Leftarrow$) Puisque la fonction $x \longrightarrow \dfrac{x}{1+x}$ est strictement croissante, alors

$$\frac{\varepsilon}{1+\varepsilon} 1_{\{|X_n|>\varepsilon\}} \leq \frac{|X_n|}{1+|X_n|} 1_{\{|X_n|>\varepsilon\}} \leq \frac{|X_n|}{1+|X_n|},$$

alors

$$\frac{\varepsilon}{1+\varepsilon} \mathbb{P}(\{|X_n| > \varepsilon) \leq \mathbb{E}(\frac{|X_n|}{1+|X_n|}),$$

Puisque $\varepsilon > 0$ est fixé, on fait tendre n tend vers $+\infty$, on a le résultat.

Remarque 6. *On a montré que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ si et seulement si $\lim_{n\to+\infty} \mathbb{E}(f(|X_n - X|)) = 0$, avec $f(x) = \dfrac{x}{1+x}$. En fait, on peut remplacer f par n'importe quelle fontion g bornée, strictement croissante sur $[0,+\infty[$, continue et $f(0) = 0$, par exmple $g(x) = \arctan(x)$ ou $g(x) = x \wedge 1$. En considérant la distance $d(X,Y) = \mathbb{E}(|X - Y| \wedge 1)$, alors on a*

$$X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X \Leftrightarrow \lim_{n\to+\infty} d(X_n, X) = 0,$$

la convergence en probabilité est alors métrisable.

Proposition 75.

Soit $f : \mathbb{R}^d \longrightarrow \mathbb{R}^q$ une application continue. Si $(X_n)_{n\in\mathbb{N}}$ converge vers X en probabilité alors $(f(X_n))_{n\in\mathbb{N}}$ converge vers $f(X)$ en probabilité.

Preuve. La preuve utilise l'uniforme continuité de f sur les compacts. En effet, soit $\varepsilon > 0$ et $a > 0$; il existe $\eta_{\varepsilon,a} > 0$ tel que

$$|x| \leq a \ \text{ et } |x - y| \leq \eta_{\varepsilon,a} \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

En particulier,

$$\Big\{|X| \leq a\Big\} \bigcap \{|X_n - X| \leq \eta_{\varepsilon,a}\} \subset \Big\{|f(X_n) - f(X)| \leq \varepsilon\Big\},$$

et par passage au complémentaire

$$\Big\{|f(X_n) - f(X)| > \varepsilon\Big\} \subset \Big\{|X| > a\Big\} \bigcup \Big\{|X_n - X| > \eta_{\varepsilon,a}\Big\}.$$

Par conséquent, on a l'inégalité

$$\mathbb{P}\Big(|f(X_n) - f(X)| > \varepsilon\Big) \leq \mathbb{P}\Big(|X| > a\Big) + \mathbb{P}\Big(|X_n - X| > \eta_{\varepsilon,a}\Big).$$

Ce qui donne

$$\lim_{n\to+\infty} \mathbb{P}\Big(|f(X_n) - f(X)| > \varepsilon\Big) \leq \mathbb{P}\Big(|X| > a\Big) + \lim_{n\to+\infty} \mathbb{P}\Big(|X_n - X| > \eta_{\varepsilon,a}\Big).$$

Comme X_n converge vers X en probabilité alors $\lim_{n\to+\infty} \mathbb{P}\{|X_n - X| > \eta_{\varepsilon,a}\} = 0$ mais aussi puisque X est à valeurs dans $\mathbb{R}^d$, $\lim_{a\to+\infty} \mathbb{P}\{|X| > a\} = 0$. Donc

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

D'où, $f(X_n)$ converge vers $f(X)$ en probabilité. ∎

Remarque 7. *Si $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X et vers Y , alors $X = Y$ presque sûrement. En effet, si $\varepsilon > 0$, on a, pour tout $n \in \mathbb{N}$,*

$$\mathbb{P}\Big(|X-Y| > \varepsilon\Big) \le \mathbb{P}\Big(|X_n - X| > \frac{\varepsilon}{2}\Big) + P\Big(|X_n - Y| > \frac{\varepsilon}{2}\Big).$$

Comme le majorant de cette inégalité tend vers 0 si n tend vers $+\infty$, on en déduit que, pour tout

$$\forall \varepsilon > 0, \mathbb{P}\Big(|X-Y| > \varepsilon\Big) = 0$$

ce qui donne l'égalité presque sûre de X et Y. ∎

Donnons des exemples d'application de ce résultat. On considère deux suites X_n et Y_n convergeant en probabilité vers X et Y et α_n une suite de réels convergeant vers α. Alors $X_n Y_n$ converge vers XY en probabilité et $X_n + \alpha_n Y_n$ converge en probabilité vers $X + \alpha Y$.

4.3.4 Lien entre différentes convergences

Proposition 76.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. Alors

1. $X_n \underset{n\to+\infty}{\overset{\text{p.s.}}{\longrightarrow}} X \Longrightarrow X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X.$
2. $X_n \underset{n\to+\infty}{\overset{L^p}{\longrightarrow}} X \Longrightarrow X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X.$

Preuve. 1. Supposons que X_n converge vers X presque sûrement, alors par the théorème de convergence dominée

$$\lim_{n\to+\infty} \mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) = \mathbb{E}\Big(\lim_{n\to+\infty} \frac{|X_n - X|}{1+|X_n - X|}\Big) = 0.$$

Ce qui donne le résultat.

2. Soit $\varepsilon > 0$. On a par l'inégalité de Markov

$$\mathbb{P}(|X_n - X| > \varepsilon) \le \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p},$$

et le résultat en découle. ∎

La réciproque de la proposition précédente est fausse c.à.d. la convergence en probabilité n'implique pas la convergence presque sûre. Toutefois, on peut montrer le résultat suivant.

Proposition 77.

Si $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X, alors il existe une sous-suite $(X_{n_k})_{k\in\mathbb{N}}$ qui converge vers X presque sûrement.

Preuve. On a $X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X$, donc $\lim\limits_{n\to+\infty} \mathbb{E}\Big(\dfrac{|X_n - X|}{1+|X_n - X|}\Big) = 0.$, on peut alors choisir une sous-suite n_k telle que

$$\mathbb{E}\Big(\frac{|X_{n_k} - X|}{1+|X_{n_k} - X|}\Big) \le \frac{1}{2^k},$$

et donc

$$\sum_{k=0}^{+\infty} \mathbb{E}\Big(\frac{|X_{n_k} - X|}{1+|X_{n_k} - X|}\Big) < +\infty.$$

Par le théorème de Fubini, on a

$$\mathbb{E}\Big(\sum_{k=0}^{+\infty}\frac{|X_{n_k}-X|}{1+|X_{n_k}-X|}\Big)<+\infty.$$

Donc

$$\sum_{k=0}^{+\infty}\frac{|X_{n_k}-X|}{1+|X_{n_k}-X|}<\infty \quad p.s.$$

Puisque le terme général d'une série convergente doit tendre vers 0, on a donc

$$\lim_{n\to+\infty}\Big(\frac{|X_{n_k}-X|}{1+|X_{n_k}-X|}\Big)=0.$$

Par conséquent, on a forcément $X_{n_k}\xrightarrow[k\to+\infty]{\text{p.s.}}X$. ∎

4.3.5 Convergence en loi

Définition 78.

On dit que X_n converge en loi vers X et on note $X_n\xrightarrow[n\to+\infty]{\text{loi}}X$ si et seulement si pour toute fonction $\phi\in C_b(\mathbb{R}^d)$,

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

Proposition 79.

Soit $f:\mathbb{R}^d\to\mathbb{R}^q$ une application continue. Si $(X_n)_{n\in\mathbb{N}}$ converge vers X en loi alors $(f(X_n))_{n\in\mathbb{N}}$ converge vers $f(X)$ en loi.

Preuve. Il suffit d'utiliser la définition. ∎

Remarque 8. *On peut noter que, pour cette convergence, les variables aléatoires ne sont pas nécessairement définies sur le même espace de probabilite $(\Omega,\mathcal{A},\mathbb{P})$. Cela est par contre nécessaire pour les convergences presque sûre, en probabilité et dans L^p.*

Définition 80.

Soit (μ_n) une suite de mesures de probabilité sur $\mathbb{R}^d$. On dit que (μ_n) converge étroitement vers μ et on note $\mu_n\xrightarrow[n\to+\infty]{\text{étr.}}\mu$ si $\forall\phi\in\mathcal{C}_b(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d}\phi(x)\mu_n(dx)\underset{n\to+\infty}{\longrightarrow}\int_{\mathbb{R}^d}\phi(x)\mu(dx)\,.$$

Remarque 9. *Pour une suite de v.a. à valeurs dans $\mathbb{R}^d$,*

$$\Big[X_n\xrightarrow[n\to+\infty]{loi}X\Big]\Leftrightarrow\Big[\mathbb{P}_{X_n}\xrightarrow[n\to+\infty]{étr.}\mathbb{P}_X\Big]$$

Nous passons à résultat reliant la convergence étroite à celle des transformées de Fourier des probabilités. Nous admettrons le théorème suivant dû à Paul Lévy.

Théorème 81.

(Théorème de Paul Lévy)
Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de probabilités sur $\mathbb{R}^d$. Si la suite de fonctions $(\hat{\mu}_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction φ continue au point 0, il existe une probabilité μ sur $\mathbb{R}^d$ telle que $\varphi = \hat{\mu}$ et $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement vers μ. De plus, la convergence de $\hat{\mu}_n$ vers $\hat{\mu}$ est uniforme sur tout compact de R^d.

En terme de loi : Soit $(\varphi_{X_n})_{n\in\mathbb{N}}$ la fontion caractéristique d'une une suite de v.a. $(X_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^d$. Si φ_{X_n} converge vers une une fonction φ continue au point 0, alors il existe une v.a X sur $\mathbb{R}^d$ telle que φ est la fontion caractéristique de X. De plus $X_n \underset{n\to+\infty}{\overset{\text{loi}}{\longrightarrow}} X$.

Remarque 10. *Le théorème de Paul Lévy nous dit que la convergence en loi est équivalente à la convergence simple des fonctions caractéristiques soit :*

$$X_n \underset{n\to+\infty}{\overset{loi}{\longrightarrow}} X \Leftrightarrow \forall t \in \mathbb{R}^d, \varphi_{X_n}(t) \longrightarrow \varphi_X(t).$$

Théorème 82.

Pour des variables dans $\mathbb{R}$, nous avons l'équivalence

$$\left[X_n \underset{n\to+\infty}{\overset{\text{loi}}{\longrightarrow}} X\right] \quad \Leftrightarrow \quad \begin{array}{l}[F_{X_n}(t) \underset{n\to+\infty}{\overset{\text{loi}}{\longrightarrow}} F_X(t) \text{ en tout point } t \text{ où } F_X \text{ est continue} \\ (\text{c'est à dire en tout point } t \text{ tel que } \mathbb{P}(X=t)=0).] \ .\end{array}$$

Théorème 83.

On a l'implications suivante: $\left[X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X\right] \Rightarrow \left[X_n \underset{n\to+\infty}{\overset{\text{loi}}{\longrightarrow}} X\right]$

Preuve. On se contente de faire la démonstration pour des variables à valeurs réelles. Soit t un point où F_X est continue. Soit $\varepsilon > 0$ quelconque. Par la propriété d'additivité et la propriété de croissance :

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

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

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

Donc $\liminf_{n\to+\infty} \mathbb{P}(X_n \le t) \ge \mathbb{P}(X \le t-\varepsilon) = F_X(t-\varepsilon)$. Tous ces calculs sont vrais $\forall\varepsilon$ et F_X est continue en t donc $\lim_{n\to+\infty} F_{X_n}(t) = F_X(t)$. Par conséquent $X_n \underset{n\to+\infty}{\overset{\text{loi}}{\longrightarrow}} X$. ∎

Autre Preuve du théorème 83 : Signalons que toutes les variables sont définies sur le même espace et montrons que

$$\forall t \in \mathbb{R}^d, \varphi_{X_n}(t) \longrightarrow \varphi_X(t).$$

Soit $t \in \mathbb{R}^d$,

$$|\varphi_{X_n}(t) - \varphi_X(t)| \le \mathbb{E}|e^{it.X_n} - e^{it.X}| \le \mathbb{E}[min(2, |t||X_n - X|)],$$

de sorte que, pour tout $\varepsilon > 0$, en écrivant

$$\mathbf{1} = \mathbf{1}_{\{[0,\varepsilon]\}}(|X_n - X|) + \mathbf{1}_{\{\varepsilon,+\infty[\}}(|X_n - X|),$$

on a

$$|\varphi_{X_n}(t) - \varphi_X(t)| \leq \varepsilon|t|\mathbb{P}(|X_n - X| \leq \varepsilon) + 2\mathbb{P}(|X_n - X| > \varepsilon) \leq \varepsilon|t| + 2\mathbb{P}(|X_n - X| > \varepsilon),$$

par suite, pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} |\varphi_{X_n}(t) - \varphi_X(t)| \leq \varepsilon$$

ce qui donne le résultat. ■

Proposition 84.

On suppose les variables $(X_n)_n$ sont définies sur le même espace probabilisé. Soit $c \in \mathbb{R}^d$, si $X_n \xrightarrow[n\to+\infty]{\text{loi}} c$, alors $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} c$.

Preuve. On se ramène au cas réel en considérant les composantes de X_n . Soit $\varepsilon > 0$, on a

$$\mathbb{P}(|X_n - c| > \varepsilon) = \mathbb{P}(X_n < c - \varepsilon) + \mathbb{P}(X_n > c + \varepsilon) \leq F_{X_n}(c - \varepsilon) + 1 - F_{X_n}(c + \varepsilon).$$

Puisque X_n converge vers c en loi, d'après la Proposition 82, pour tout $t \neq c$, $F_{X_n}(t)$ converge vers $\mathbf{1}_{\mathbb{R}_+}(t - c)$. Ceci montre que $\mathbb{P}(|X_n - c| > \varepsilon)$ tend vers 0 pour tout $\varepsilon > 0$. ■

Finalement, on a le diagramme suivant :

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

Toutes les autres implications sont fausses.

4.4 Loi des grands nombres

L'objectif de cette section est d'établir "la loi forte des grands nombres" de Kolmogorov qui est l'un des résultats fondamentaux de la théorie des probabilités.

Notation 85.

Soient $X_1, X_2, \ldots$ des variables indépendantes et de même loi. On dira que ces variables sont indépendantes et identiquement distribuées et on utilisera la notation i.i.d..

Théorème 86.

(Loi faible des grands nombres L^2)
Soient $X_1, X_2, \dots$ des v.a.r. i.i.d. Si $\mathbb{E}(X_1^2) < \infty$, alors

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{L^2} \mathbb{E}(X_1)\,.$$

La convergence aura lieu aussi en probabilité

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}(X_1)\,.$$

On a donc la convergence de la moyenne arithmétique (dite aussi moyenne empirique) vers la moyenne probabiliste (l'espérance probabiliste). C'est grâce à ce résultat qu'on peut estimer une proportion dans une population par une proportion dans un ?échantillon (représentatif).

Preuve. On a

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{X_1 + \cdots + X_n}{n} - \mathbb{E}(X_1)\right)^2\right) &= \mathbb{E}\left(\left(\frac{(X_1 - \mathbb{E}(X_1)) + \cdots + (X_n - E(X_n))}{n}\right)^2\right) \\
&= \frac{1}{n^2}\mathrm{Var}(X_1 + \cdots + X_n) \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(X_k) \qquad \text{(par indépendance)} \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(X_1) \qquad \text{(les v.a. ont même loi)} \\
&= \frac{\mathrm{Var}(X_1)}{n} \xrightarrow[n\to+\infty]{} 0.
\end{aligned}$$

En ce qui concerne la convergence en probabilité : soit $\varepsilon > 0$. Par l'inégalité de Markov

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mathbb{E}(X_1)\right| > \varepsilon\right) \leq \frac{\mathbb{E}\left(\left(\frac{X_1+\cdots+X_n}{n} - \mathbb{E}(X_1)\right)^2\right)}{\varepsilon^2} = \frac{\mathrm{Var}(X_1)}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

■

Remarque 11. *Plus que la convergence, nous avons obtenu la vitesse de convergence, elle est en* $O(\frac{1}{n})$.

Théorème 87.

(Loi forte des grands nombres p.s.)
Soient $X_1, X_2, \dots$ des v.a.r. i.i.d. Si $\mathbb{E}(|X_1|) < \infty$ (en d'autres termes, X_1 est intégrable) alors

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

Preuve. Nous ne ferons la démonstration que dans le cas $\mathbb{E}(X_1^4) < \infty$. Nous voulons montrer que

$$\frac{(X_1 - \mathbb{E}(X_1)) + \cdots + (X_n - \mathbb{E}(X_n))}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} 0\,.$$

Posons pour tout i, $X_i' = X_i - \mathbb{E}(X_i)$. Calculons

$$\mathbb{E}\left(\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4\right) = \frac{1}{n^4} \sum_{i_1,i_2,i_3,i_4 \in \{1,\ldots,n\}} \mathbb{E}(X_{i_1}' X_{i_2}' X_{i_3}' X_{i_4}') .$$

Remarquons que dans cette dernière somme, certains termes sont nuls. Par exemple, en utilisant les propriétés des variables indépendantes

$$\begin{aligned} \mathbb{E}(X_1' X_2' X_2' X_2') &= \mathbb{E}(X_1')\mathbb{E}((X_2')^3) = 0 \\ \mathbb{E}(X_1' X_2' X_3' X_3') &= \mathbb{E}(X_1')\mathbb{E}(X_2')\mathbb{E}((X_3')^2) = 0 . \end{aligned}$$

Après regroupement des termes identiques, nous obtenons

$$\begin{aligned} \mathbb{E}\left(\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4\right) &= \frac{1}{n^4}(n\mathbb{E}((X_1')^4 + 6n(n-1)\mathbb{E}((X_1)^2(X_2)^2) \\ &\leq \frac{7}{n^2} . \end{aligned}$$

Et donc $\sum_{n\geq 1} \mathbb{E}\left(\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4\right) < \infty$. Par Fubini-Tonelli, on obtient

$$\mathbb{E}\left(\sum_{n\geq 1}\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4\right) = \sum_{n\geq 1} \mathbb{E}\left(\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4\right) < \infty .$$

Donc la variable $\sum_{n\geq 1}\left(\frac{X_1' + \cdots + X_n'}{n}\right)^4$ est finie p.s. Donc le terme général de la série converge vers 0, p.s. ■

Exemple 14. *Soient $U_1, U_2, \ldots$ i.i.d. de loi $\mathcal{U}([0,1])$. Soient $0 \leq a < b \leq 1$. Soit pour tout i, $X_i = \mathbf{1}_{[a,b]}(U_i)$. Les variables $X_1, X_2, \ldots$ sont i.i.d. de loi de Bernoulli de paramètre $b-a$, $\mathcal{B}(b-a)$. En effet, les (X_i) se sont des fonctions mesurables de U_i qui sont indépendantes. Donc les (X_i) sont indépendantes. De plus*

$$\mathbb{P}\Big(X_i = 1\Big) = \mathbb{P}\Big(U_i \in [a,b]\Big) = F_{U_i}(b) - F_{U_i}(a) = b - a.$$

et

$$\mathbb{P}\Big(X_i = 0\Big) = \mathbb{P}\Big(U_i \notin [a,b]\Big) = 1 - (F_{U_i}(b) - F_{U_i}(a)) = 1 - (b-a).$$

Les (X_i) vérifient aussi $\mathbb{E}(|X_i|) < \infty$ puiqu'elles sont bornées. Par la loi des grands nombres

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

Ce qui veut dire que la proportion de points tombant dans $[a,b]$ converge vers $b-a$.

Application : Méthode de Monte Carlo.

Soient B un domaine mesurable bornée, $f : B \to \mathbb{R}$ une application mesurable et $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur B (de densité $g(x) = \frac{1}{\lambda(B)} 1_B(x)$, λ est la mesure de Lebesgue). On suppose que $f(X_1) \in L^1$ (par exemple, il suffit d'avoir f bornée sur B). On a alors

$$\frac{f(X_1) + \cdots + f(X_n)}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(f(X_1)) = \frac{1}{\lambda(B)} \int_B f(x) d\lambda(x).$$

La méthode de Monte-Carlo permet donc de faire un calcul approché d'intégrale en utilisant la loi forte des grands nombres. L'avantage par rapport aux méthodes classiques d'analyse numérique (trapèze, Simpson, interpolation) est qu'aucune régularité n'est à supposer pour f. L'inconvénient est que la convergence n'est que presque sûre (pas vraiment gênant en pratique) mais surtout qu'il n'y a pas facilement de contrôle de l'erreur commise dans l'approximation.
Si $B = [0,1]$, alors

$$\frac{f(X_1) + \cdots + f(X_n)}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(f(X_1)) = \int_0^1 f(x) dx.$$

Exercice 3. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. indépendantes de même loi telle que $\mathbb{E}(X_1)^2 < +\infty$. On pose $S_n = \sum_{k=1}^{n} X_k$, $\quad M_n = \frac{S_n}{n}$ et $q_n = [\sqrt{n}]$, où $[x]$ désigne la partie entière de x.*

1. *On suppose que pour tout $n \in \mathbb{N}^*$, $X_n \geq 0$.*

   a. *Montrer que $(Z_n)_{n\in\mathbb{N}^*} = (M_{n^2})_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.*

   b. *Prouver que pour tout $n \in \mathbb{N}^*$, on a*

   $$\frac{q_n^2}{n} Z_{q_n} \leq M_n \leq \frac{(q_n+1)^2}{n} Z_{q_n+1}.$$

   c. *En déduire que $(M_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.*

2. *Montrer que $(M_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.*

4.5 Théorème central-limite

Nous démontrons dans cette section un résultat fondamental du calcul des probabilités : le "Théorème Limite Central". Signalons que l'adjectif "central" se reporte au "théorème" et non à la "limite", il s'agit donc d'un théorème limite qui joue un rôle central en théorie des probabilités.

Il explique les fluctuations autour de leur moyenne des effets cumulés de phénomènes répétés indépendamment, après normalisation, se comporte comme la loi $\mathbb{N}(0;1)$. Autrement dit, la somme d'un grand nombre de variables aléatoires indépendantes (et de variance finie) suit, à peu près, une loi normale.

Théorème 88.

Théorème central-limite (aussi noté TCL)
Soit (X_n) une suite de v.a.r. i.i.d. avec $\mathbb{E}(X_1) = m$ et $\mathrm{Var}(X_1) = \sigma^2$ $(m, \sigma^2 < \infty)$. Alors

$$\frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{\text{loi}} Z \text{ de loi } \mathcal{N}(0,1)\,,$$

(où $\sigma > 0$ est la racine carrée de la variance).

Il existe des résultats raffinés sur la vitesse de cette convergence en loi. Voir, par exemple, le théorème de Berry-Esseen.

Remarque 12. *Sous les hypothèses du théorème précédent, prenons $a < b$, $f(x) = \mathbf{1}_{[a,b]}(x)$. Par la remarque 9,*

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

c'est à dire

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

C'est cette propriété qui sera le plus souvent utilisée dans les exercices.

Preuve.[Démonstration du théorème 88] Posons $\forall n$, $Y_n = X_n - m$. Soient

$$S'_n = Y_1 + \cdots + Y_n,\ Z_n = \frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} = \frac{S'_n}{\sigma\sqrt{n}}\,.$$

Nous avons

$$\begin{aligned}
\Phi_{Z_n}(t) &= \mathbb{E}\left(\exp\left(\frac{itS'_n}{\sigma\sqrt{n}}\right)\right) \\
&= \mathbb{E}\left(\exp\left(\frac{it}{\sigma\sqrt{n}}(Y_1+\dots+Y_n)\right)\right) \\
\text{(par indépendance des } Y_j) &= \prod_{1\le j\le n}\mathbb{E}\left(\exp\left(\frac{it}{\sigma\sqrt{n}}Y_j\right)\right) \\
\text{(car les } Y_j \text{ sont identiquement distribués)} &= \Phi_{Y_1}\left(\frac{t}{\sigma\sqrt{n}}\right)^n .
\end{aligned}$$

Regardons 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, $|Y_1e^{iuY_1(\omega)}|\le|Y_1(\omega)|$ qui est intégrable (et qui ne dépend pas de u). 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 $\Phi'_{Y_1}(0)=\mathbb{E}(iY_1)=i\mathbb{E}(Y_1)=0$, $\Phi''_{Y_1}(0)=-\mathbb{E}(Y_1^2)=-\sigma^2$. Supposons que Φ_{Y_1} admette un développement limité en 0 (ce n'est pas toujours le cas). Ce développement est alors :

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

Donc

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

par continuité de l'exponentielle. ∎

Chapitre 5

Exercices

Exercice 4. Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle bornée. Pour tout $n \in \mathbb{N}$, on pose

$$A_n = \{u_m,\ m \geq n\}.$$

1. Pour tout $n \in \mathbb{N}$, on pose $V_n = \inf(A_n)$ et $W_n = \sup(A_n)$. Montrer que la suite $(V_n)_{n\in\mathbb{N}}$ est croissante et la suite $(W_n)_{n\in\mathbb{N}}$ est décroissante. En déduire qu'elles sont convergentes.

2. On note : $\lim\limits_{n\to+\infty} V_n = \liminf\limits_{n\to+\infty} u_n = \sup\limits_{n\geq 0}\inf\limits_{m\geq n} u_m$ et $\lim\limits_{n\to+\infty} W_n = \limsup\limits_{n\to+\infty} u_n = \inf\limits_{n\geq 0}\sup\limits_{m\geq n} u_m$.

   Calculer $\liminf\limits_{n\to+\infty} u_n$ et $\limsup\limits_{n\to+\infty} u_n$ pour les suites $(u_n)_n = ((-1)^n)_{n\in\mathbb{N}}$ et $(u_n)_n = (\frac{1}{n})_{n\in\mathbb{N}^*}$.

Exercice 5. Soient x un nombre réel et $(u_n)_n$ la suite définie par : $\forall n \in \mathbb{N}$, $u_n = \dfrac{\left[10^n x\right]}{10^n}$, où $[x]$ représente la partie entière d'un réel x.

1. Montrer que la suite $(u_n)_n$ converge vers x.
   On pose $a_0 = u_0 = [x]$ et $\forall n \in \mathbb{N}$, $a_{n+1} = 10^{n+1}(u_{n+1} - u_n)$.

2. Montrer que pour tout entier $n \in \mathbb{N}$ on a : $0 \leq a_{n+1} \leq 9$ et $u_n = \displaystyle\sum_{k=0}^{n} \frac{a_k}{10^k}$.

3. On pose pour tout entier $n \in \mathbb{N}$, $v_n = u_n + \dfrac{1}{10^n}$. Montrer que les deux suites $(u_n)_n$ et $(v_n)_n$ sont adjacentes qui convergent vers x avec $u_n \leq x \leq v_n$ pour tout $n \in \mathbb{N}$.
   u_n est appelée l'approximation décimale par défaut à 10^n près de x et v_n l'approximation décimale par excès à 10^n près.

Exercice 6. Soient $(\Omega, \mathcal{F})$ et $(E, \mathcal{E})$ deux espaces probabilisables, $\mathcal{C}$ une classe de parties de E et X une application de $(\Omega, \mathcal{F})$ à valeurs dans $(E, \mathcal{E})$.

1. Montrer que : $\sigma(X^{-1}(C)) = X^{-1}(\sigma(C))$.

2. On suppose que : $\sigma(\mathcal{C}) = \mathcal{E}$. Montrer que X est une variables aléatoire si et seulement si $X^{-1}(\mathcal{C}) \subset \mathcal{F}$.

3. On suppose que : $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Déduire que X est une variable aléatoire réelle ssi pour tout $t \in \mathbb{R}$, $\{X \geq t\} \in \mathcal{F}$.

4. On suppose que E un espace topologique et $\mathcal{E}$ la tribu engendrée par la classe des ouverts de E. Montrer que toute fonction continue $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est mesurable.

Exercice 7. Soient $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées et F la fonction de répartition de X_1. Pour tout $k \in \mathbb{N}^*$, on note Z_k la variable aléatoire définie par

$$Z_k = \max(X_i : 1 \leq i \leq k) = \max(X_1, X_2, ..., X_k).$$

1. Pour $k \in \mathbb{N}$, exprimer la fonction de répartition H_k de Z_k en fonction de F et k.

2. Soit N une variable aléatoire à valeurs dans $\mathbb{N}^*$ indépendante des $(X_n)_{n\geq 1}$. On considère la variable aléatoire Z définie par $Z = \max(X_i : 1 \leq i \leq N)$

   (a) Justifier l'égalité, pour $x \in \mathbb{R}$,

   $$\{Z \leq x\} = \bigcup_{k\geq 1}\Big(\{Z_k \leq x\} \bigcap \{N = k\}\Big).$$

   (b) En déduire que la fonction de répartition H de Z est donnée par

   $$\forall x \in \mathbb{R}, H(x) = \sum_{k\geq 1} (F(x))^k\, \mathbb{P}(N = k).$$

(c) Déterminer H dans le cas où X_1 suit la loi uniforme sur $[0,1]$ et N suit la loi géométrique de paramètre $p \in]0,1[$, *i.e.* $\mathbb{P}(N=k) = p(1-p)^{k-1}$.

Exercice 8. 1. Soient X une variable aléatoire positive et $a > 0$. Démontrer l'inégalité de Markov

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

2. Soient $(X_n)_{n\geq 1}$, une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi exponentielle $\mathcal{E}(1)$ de paramètre 1, et $r > 0$. On pose $S_n = \sum_{k=1}^{n} X_k$

(a) Montrer, en utilisant l'inégalité de Markov, que pour tout $0 < \lambda < 1$,

$$\mathbb{P}\left[\frac{S_n}{n} \geq 1+r\right] \leq \mathbb{P}\left[e^{\lambda S_n} \geq e^{n\lambda(1+r)}\right] \leq \left(\frac{\mathbb{E}e^{\lambda X_1}}{e^{\lambda(1+r)}}\right)^n.$$

(b) En déduire que $\mathbb{P}\left[\frac{S_n}{n} \geq 1+r\right] \leq (1+r)^n e^{-nr}$.

Exercice 9. 1. Soit X une variables aléatoire de loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$. Déterminer la loi $\mathbb{P}_Z$ de la variable aléatoire $Z = [X]+1$ où $[x]$ désigne la partie entière de x. Identifier la loi trouvée avec une loi usuelle.

2. Soit U une variable aléatoire de loi uniforme $\mathcal{U}([-1,1])$. Déterminer la loi $\mathbb{P}_V$ de la variable aléatoire $V = \arcsin(U)$.

3. Soit X une variables aléatoire de loi normale $\mathcal{N}(0,1)$. Déterminer la loi $\mathbb{P}_R$ de la variable aléatoire $R = |X|$.

4. Soit X, Y deux variables aléatoires indépendantes de loi normale $\mathcal{N}(0,1)$.

(a) Déterminer la loi $\mathbb{P}_Z$ de la variable aléatoitre $Z = \frac{X}{Y}$. Identifier la loi trouvée avec une loi usuelle.

(b) Déterminer la loi $\mathbb{P}_H$ de la variable aléatoitre $H = \frac{1}{Z}$.

Exercice 10. Soient U une variable aléatoire de loi uniforme sur $[-1,1]$ et X la variable aléatoire définie par $X = \frac{1+U}{1-U}$.

1. Déterminer la densité de la variable aléatoire X.

2. Calculer la fonction de répartition de X.

3. Déterminer la loi de la variable aléatoire $Z = [X]$ où $[x]$ désigne la partie entière de x.

Exercice 11. Soit X une variable aléatoire à valeurs dans $\mathbb{R}$ de loi de densité

$$x \mapsto f(x) = \frac{1}{\pi(1+x^2)}.$$

Montrer $\frac{1}{X}$ est de même loi que X.

Exercice 12. Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes de loi exponentielle $]\S_{\surd}(\lambda)$ de paramètre $\lambda > 0$.

1. Calculer la loi de la variable aléatoire $Y = \max_{1\leq i\leq n} X_i$.

2. Calculer la loi de la variable aléatoire $Z = \min_{1\leq i\leq n} X_i$.

Exercice 13. 1. Soit X une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n,p)$ de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$. Calculer la fonction caractéristique φ_X de X.

2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et de même loi de Bernoulli de paramètre $p \in]0,1[$. On pose $S_0 = T_0 = 0$ et, pour tout $n \geq 1$,

$$S_n = \sum_{k=1}^{n} X_k, \quad T_n = n - S_n.$$

   (a) Pour $n \geq 1$, préciser la loi des variables aléatoires S_n et T_n.

   (b) Pour $n \geq 1$, calculer $\mathbb{P}(S_n = 0, T_n = 0)$. S_n et T_n sont-elles indépendantes?

3. Soit N une variable aléatoire indépendante des variables aléatoires $(X_n)_{n\geq 1}$ suivant la loi de Poisson de paramètre $\lambda > 0$. On pose

$$\forall \omega \in \Omega, \quad U(\omega) = S_{N(\omega)}(\omega), \quad V(\omega) = T_{N(\omega)}(\omega) = N(\omega) - U(\omega).$$

   (a) Justifier l'égalité, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(U = k) = \sum_{n\geq k} \mathbb{P}(S_n = k, N = n).$$

   (b) En déduire que U suit la loi de Poisson de paramètre λp.

   (c) Déterminer la loi de $1 - X_1$ puis, préciser la loi de V .

   (d) Montrer que, pour $(k,l) \in \mathbb{N}^2$, $\mathbb{P}(U = k, V = l) = \mathbb{P}(N = k + l)\mathbb{P}(S_{k+l} = k)$.

   (e) En déduire que les variables aléatoires U et V sont indépendantes.

Exercice 14. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi de Poisson $\mathbb{P}(\lambda)$, $\lambda > 0$. On rappelle que, $\mathbb{P}[X = k] = e^{-\lambda}\dfrac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$.

1. Calculer la fonction caractéristique φ_{X_1} de X_1.

2. En déduire la loi de la somme $S = X_1 + X_2 + ... + X_n$.

Exercice 15. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi normale $\mathcal{N}(0,1)$ et N une variable aléatoire à valeurs dans $\mathbb{N}^*$ indépendante des variables aléatoires $(X_n)_{n\geq 1}$. On note, pour $n \geq 1$,

$$Y_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k, \quad Y = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} X_k.$$

1. Calculer, pour tout $n \geq 1$, la fonction caractéristique de Y_n et préciser sa loi.

2. Déterminer la fonction caractéristique de Y ainsi que sa loi.

Exercice 16. Soient X et Y deux variables aléatoires réelles définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

1. On suppose que $X = Y$ *p.s.* Montrer que X et Y ont la même loi. Montrer que la réciproque est fausse.

2. On suppose que X et Y ont la même loi. Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction borélienne. Montrer que les variables aléatoires $f(X)$ et $f(Y)$ ont la même loi.

Exercice 17. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On considère la variable aléatoire réelle positive X de fonction de répartition F_X.

1. Montrer que la fonction $\begin{array}{llll} f : & \Omega \times \mathbb{R}_+ \longrightarrow \mathbb{R} \\ & (\omega\ ,\ t) \mapsto f(\omega,t) = 1_{\{X(\omega)>t\}}(\omega,t) \end{array}$ est mésurable.

2. Prouver que pour tout $n \in \mathbb{N}^*$, on a
$$\mathbb{E}(X^n) = \int_0^{+\infty} nt^{n-1}\mathbb{P}[X > t]dt = \int_0^{+\infty} nt^{n-1}(1 - F_X(t))dt.$$

Exercice 18. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi de Poisson $\mathbb{P}(\lambda)$, $\lambda > 0$. On rappelle que, $\mathbb{P}[X = k] = e^{-\lambda}\dfrac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$.

1. Calculer la fonction caractéristique φ_{X_1} de X_1.
2. En déduire la loi de la somme $S = X_1 + X_2 + ... + X_n$.

Exercice 19. On considère la fonction Gamma définie sur $\mathbb{R}_+^*$ par : $\Gamma(x) = \displaystyle\int_0^{+\infty} e^{-t}t^{x-1}dt$.
On appelle loi $\gamma(a,\beta)$ de paramètres a et β ($a > 0$ et $\beta > 0$) la loi sur $\mathbb{R}$ de densité
$$f_{a,\beta}(x) = \frac{\beta^a}{\Gamma(a)}e^{-\beta x}x^{a-1}1_{\mathbb{R}_+}(x).$$

1. Soit X une variable aléatoire de loi $\gamma(a,\beta)$. Calculer $\mathbb{E}(X)$ et $Var(X)$.
2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et de même loi exponentielle $\mathcal{E}(\beta)$, $\beta > 0$. Montrer par récurrence que la loi de la somme $X_1 + X_2 + ... + X_n$ est la loi $\gamma(n,\beta)$.
3. Soit X et Y deux v.a. réelles indépendantes de loi $\gamma(a,\beta)$ et $\gamma(b,\beta)$ respectivement.
   a. On pose $B(a,b) = \displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}dx$. Montrer que $\Gamma(a)\Gamma(b) = \Gamma(a+b)B(a,b)$.
   b. En déduire que : $\forall u > 0$, $\displaystyle\int_0^u x^{a-1}(u-x)^{b-1}dx = u^{a+b-1}\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.
   c. Déterminer la loi de $X + Y$.
4. Soit X et Y deux v.a. réelles indépendantes de loi $\gamma(a,\beta)$.
   a. Déterminer la loi de βX et vérifier que la v.a. $\dfrac{X}{X+Y}$ est bien définie.
   b. Montrer que $X + Y$ et $\dfrac{X}{X+Y}$ sont des v.a. indépendantes.
5. Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de v.a. réelles indépendantes et de même loi normale $\mathbb{N}(0,1)$.
   a. Montrer que Y_1^2 suit la loi gamma $\gamma(\frac{1}{2},\frac{1}{2})$.
   b. Montrer que $Y_1^2 + Y_2^2 + ... + Y_n^2$ suit une loi $\gamma(\frac{n}{2},\frac{1}{2})$.
   La loi $\gamma(\frac{n}{2},\frac{1}{2})$ est également appelée loi Khi-deux à n degrés de liberté et notée $\chi^2(n)$.

Exercice 20. X une variable aléatoire dans $\mathbb{R}$ est dite symétrique si $-X$ a même loi que X.

1. Si X a une densité f, montrer que : X est symétrique si et seulement si $f(x) = f(-x)$ pour presque tout $x \in \mathbb{R}$.
2. Donner un exemple de loi symétrique.
3. Montrer que X est symétrique si et seulement si le nombre $E(e^{iuX})$ est réel pour tout $u \in \mathbb{R}$.
4. Si Y et Z sont deux variables aléatoires réelles de même loi et indépendantes, montrer que $Y - Z$ est symétrique.

Exercice 21. Soient X et Y deux variables aléatoires indépendantes de loi $\mathbb{N}(0,1)$. Montrer que $X + Y$ et $X - Y$ sont indépendantes.

Exercice 22. Soient U une variable aléatoire de loi uniforme sur $[-1,1]$ et X la variable aléatoire définie par $X = \dfrac{1+U}{1-U}$.

1. Déterminer la densité de la variable aléatoire X.
2. Calculer la fonction de répartition de X.
3. Déterminer la loi de la variable aléatoire $Z = [X]$ où $[x]$ désigne la partie entière de x.

Exercice 23. Soit (X,Y) une variable aléatoire à valeurs dans $\mathbb{R}^2$ de loi de densité

$$(x,y) \mapsto f(x,y) = \frac{3}{4}\exp(-|x+2y|-|x-y|).$$

Calculer la densité de la loi de $(X+2Y, X-Y)$ puis les densités des lois de X et Y.

Exercice 24. Soit X une variable aléatoire à valeurs dans $\mathbb{R}$ de loi de densité

$$x \mapsto f(x) = \frac{1}{\pi(1+x^2)}.$$

Montrer $\dfrac{1}{X}$ est de même loi que X.

Exercice 25. Soit $X_1, ..., X_n$ des variables aléatoires indépendantes de loi exponentielle]§$\surd(\lambda)$ de paramètre $\lambda > 0$.

1. Calculer la loi de la variable aléatoire $Y = \max\limits_{1\le i\le n} X_i$.
2. Calculer la loi de la variable aléatoire $Z = \min\limits_{1\le i\le n} X_i$.

Exercice 26. Soient X et Y deux variables aléatoires réelles définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

1. On suppose que $X = Y$ *p.s.* Montrer que X et Y ont la même loi. Montrer que la réciproque est fausse.
2. On suppose que X et Y ont la même loi. Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction borélienne. Montrer que les variables aléatoires $f(X)$ et $f(Y)$ ont la même loi.

Exercice 27. 1. Soit X une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n,p)$ de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$. Calculer la fonction caractéristique φ_X de X.

2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et de même loi de Bernoulli de paramètre $p \in]0,1[$. On pose $S_0 = T_0 = 0$ et, pour tout $n \ge 1$,

$$S_n = \sum_{k=1}^{n} X_k, \quad T_n = n - S_n.$$

   (a) Pour $n \ge 1$, préciser la loi des variables aléatoires S_n et T_n.
   (b) Pour $n \ge 1$, calculer $\mathbb{P}(S_n = 0, T_n = 0)$. S_n et T_n sont-elles indépendantes?

3. Soit N une variable aléatoire indépendante des variables aléatoires $(X_n)_{n\ge1}$ suivant la loi de Poisson de paramètre $\lambda > 0$. On pose

$$\forall \omega \in \Omega, \quad U(\omega) = S_{N(\omega)}(\omega), \quad V(\omega) = T_{N(\omega)}(\omega) = N(\omega) - U(\omega).$$

   (a) Justifier l'égalité, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(U = k) = \sum_{n\ge k} \mathbb{P}(S_n = k, N = n).$$

(b) En déduire que U suit la loi de Poisson de paramètre λp.

(c) Déterminer la loi de $1 - X_1$ puis, préciser la loi de V .

(d) Montrer que, pour $(k,l) \in \mathbb{N}^2$, $\mathbb{P}(U = k, V = l) = \mathbb{P}(N = k+l)\mathbb{P}(S_{k+l} = k)$.

(e) En déduire que les variables aléatoires U et V sont indépendantes.

Exercice 28. $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi de Poisson $\mathbb{P}(\lambda)$, $\lambda > 0$. On rappelle que, $\mathbb{P}[X = k] = e^{-\lambda}\dfrac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$.

1. Calculer la fonction caractéristique φ_{X_1} de X_1.
2. En déduire la loi de la somme $S = X_1 + X_2 + ... + X_n$.

Exercice 29. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi normale $\mathcal{N}(0,1)$ et N une variable aléatoire à valeurs dans $\mathbb{N}^*$ indépendante des variables aléatoires $(X_n)_{n\geq 1}$. On note, pour $n \geq 1$,

$$Y_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k, \quad Y = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} X_k.$$

1. Calculer, pour tout $n \geq 1$, la fonction caractéristique de Y_n et préciser sa loi.
2. Déterminer la fonction caractéristique de Y ainsi que sa loi.

Exercice 30. On considère la fonction Gamma définie sur $\mathbb{R}_+^*$ par : $\Gamma(x) = \displaystyle\int_0^{+\infty} e^{-t}t^{x-1}dt$.
On appelle loi $\gamma(a,\beta)$ de paramètres a et β ($a > 0$ et $\beta > 0$) la loi sur $\mathbb{R}$ de densité

$$f_{a,\beta}(x) = \frac{\beta^a}{\Gamma(a)}e^{-\beta x}x^{a-1}1_{\mathbb{R}_+}(x).$$

1. Soit X une variable aléatoire de loi $\gamma(a,\beta)$. Calculer $\mathbb{E}(X)$ et $Var(X)$.
2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et de même loi exponentielle $]\S_{\surd}(\beta)$, $\beta > 0$. Montrer par récurrence que la loi de la somme $X_1 + X_2 + ... + X_n$ est la loi $\gamma(n,\beta)$.
3. Soit X et Y deux v.a. réelles indépendantes de loi $\gamma(a,\beta)$ et $\gamma(b,\beta)$ respectivement.

   a. On pose $B(a,b) = \displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}dx$. Montrer que $\Gamma(a)\Gamma(b) = \Gamma(a+b)B(a,b)$.

   b. En déduire que : $\forall u > 0$, $\displaystyle\int_0^u x^{a-1}(u-x)^{b-1}dx = u^{a+b-1}\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

   c. Déterminer la loi de $X + Y$.

4. Soit X et Y deux v.a. réelles indépendantes de loi $\gamma(a,\beta)$.

   a. Déterminer la loi de βX et vérifier que la v.a. $\dfrac{X}{X+Y}$ est bien définie.

   b. Montrer que $X + Y$ et $\dfrac{X}{X+Y}$ sont des v.a. indépendantes.

5. Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de v.a. réelles indépendantes et de même loi normale $\mathbb{N}(0,1)$.

   a. Montrer que Y_1^2 suit la loi gamma $\gamma(\dfrac{1}{2},\dfrac{1}{2})$.

   b. Montrer que $Y_1^2 + Y_2^2 + ... + Y_n^2$ suit une loi $\gamma(\dfrac{n}{2},\dfrac{1}{2})$.

   La loi $\gamma(\dfrac{n}{2},\dfrac{1}{2})$ est également appelée loi Khi-deux à n degrés de liberté et notée $\chi^2(n)$.

Exercice 31. Soient X une variable aléatoire réelle définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, F sa fonction de répartition et U est une variable aléatoire de loi uniforme sur $[0,1]$.

1. Si F est continue et strictement croissante, quelle est la loi de la variable aléatoire $F^{-1}(U)$?

2. Dans le cas général on définit F^{-1}, l'inverse généralisé de F par :

$$F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}, \quad 0 < u < 1.$$

   Quelle est la loi de la variable aléatoire $F^{-1}(U)$?

Exercice 32. Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes de loi exponentielle de paramètre $\lambda > 0$. On pose

$$Y = \limsup_n \frac{X_n}{\ln n}.$$

Le but est de démontrer que $Y = \dfrac{1}{\lambda}$ p.s.

1. Montrer que

$$\mathbb{P}\left(\limsup_n \left\{\frac{X_n}{\ln n} \geq \frac{1}{\lambda}\right\}\right) \leq \mathbb{P}\left(\limsup_n \frac{X_n}{\ln n} \geq \frac{1}{\lambda}\right).$$

2. Montrer que $\mathbb{P}\left(\limsup_n \left\{\dfrac{X_n}{\ln n} \geq \dfrac{1}{\lambda}\right\}\right) = 1$. En déduire que $\mathbb{P}\left(Y \geq \dfrac{1}{\lambda}\right) = 1$.

3. Montrer que pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\limsup_n \frac{X_n}{\ln n} > \frac{1+\varepsilon}{\lambda}\right) \leq \mathbb{P}\left(\limsup_n \left\{\frac{X_n}{\ln n} > \frac{1+\varepsilon}{\lambda}\right\}\right).$$

4. Soit $\varepsilon > 0$, prouver que

   (a) $\mathbb{P}\left(\limsup_n \left\{\dfrac{X_n}{\ln n} > \dfrac{1+\varepsilon}{\lambda}\right\}\right) = 0$.

   (b) $\mathbb{P}\left(Y > \dfrac{1+\varepsilon}{\lambda}\right) = 0$.

5. En déduire que $\mathbb{P}\left(Y = \dfrac{1}{\lambda}\right) = 1$.

6. Montrer que la suite $\left(\dfrac{X_n}{\ln n}\right)_{n\geq 2}$ converge vers 0 en probabilité.

Exercice 33. Soit N une variable aléatoire à valeurs dans $\mathbb{N}$ et $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes intégrables de même loi. On suppose que la suite $(X_n)_{n\geq 1}$ est indépendante de la variable N. On pose

$$Z = \begin{cases} 0 & \text{si} \quad N = 0, \\ \displaystyle\sum_{i=1}^{N} X_i & \text{si} \quad N \geq 1. \end{cases}$$

1. Montrer que Z est une variable aléatoire.

2. Calculer $\mathbb{E}(Z)$ en fonction de $\mathbb{E}(N)$ et $\mathbb{E}(X_1)$.

3. On suppose que les variables X_i sont de même loi de Bernoulli de paramètre $p \in]0,1[$ ($P\{X_i = 1\} = p$) et N suit la loi de Poisson de paramètre $\lambda > 0$.

   Déterminer la loi de Z.

Exercice 34. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On suppose que la loi de X_n est donnée par :

$$\mathbb{P}(X_n = 0) = 1 - p_n, \ \mathbb{P}(X_n = x_n) = p_n \quad \text{où} \quad 0 < p_n < 1 \quad \text{et} \quad x_n \geq 1.$$

1. Donner une condition nécessaire et suffisante pour que la suite $(X_n)_{n\in\mathbb{N}^*}$:
   a. converge vers 0 presque sûrement.
   b. converge vers 0 dans L^p, $p \geq 1$.
   c. converge vers 0 en probabilité.
   d. converge vers 0 en loi.
2. Donner un exemple de suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires tels que :
   a. $(X_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers 0 mais pas dans L^1.
   b. $(X_n)_{n\in\mathbb{N}^*}$ converge vers 0 dans L^p pour tout $p \geq 1$ mais ne converge pas vers 0 presque sûrement.
   c. $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers 0 sans converger presque sûrement.

Exercice 35. 1. Soit $(p_n)_n$ une suite de réels dans $]0,1[$ telle que $\lim\limits_{n\to+\infty} np_n = \lambda > 0$. Soit $(X_n)_n$ une suite de v.a. telles que pour tout n, X_n est une v.a. de loi binomiale $\mathcal{B}(n,p_n)$ et X une v.a. de loi de poisson de paramètre λ. Montrer que $(X_n)_n$ converge en loi vers X.

2. Soit (X_n) une suite de v.a. réelles indépendantes de même loi $\mathcal{N}(0,1)$. Déterminer la limite en loi de la suite $Y_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}\sqrt{k}X_k$.

Exercice 36. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. indépendantes de même loi telle que $\mathbb{E}(X_1)^2 < +\infty$. On pose $S_n = \displaystyle\sum_{k=1}^{n} X_k$, $\quad M_n = \dfrac{S_n}{n}$ et $q_n = [\sqrt{n}]$, où $[x]$ désigne la partie entière de x.

1. On suppose que pour tout $n \in \mathbb{N}^*$, $X_n \geq 0$.
   a. Montrer que $(Z_n)_{n\in\mathbb{N}^*} = (M_{n^2})_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.
   b. Prouver que pour tout $n \in \mathbb{N}^*$, on a
   $$\frac{q_n^2}{n} Z_{q_n} \leq M_n \leq \frac{(q_n+1)^2}{n} Z_{q_n+1}.$$
   c. En déduire que $(M_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.
2. Montrer que $(M_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.

Exercice 37. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes définie sur un espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ de loi exponentielle de paramètre $\lambda > 0$. On pose $S_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ et $Z_n = \dfrac{1}{S_n}$.

1. Montrer que Z_n converge presque sûrement lorsque n tend vers l'infini vers une limite que l'on précisera.
2. Montrer que la suite $(\dfrac{nS_n - n\mathbb{E}X_1}{\sqrt{n}})_{n\geq 1}$ converge en loi vers une variable aléatoire Y dont on donnera la loi.
3. Soit N une variable aléatoire de loi $\mathbb{N}(0,1)$, montrer qu'il existe un unique $\phi \in \mathbb{R}_+$ tel que $\mathbb{P}(|N| \leq \phi) = 0.95$.
4. Déterminer β tel que $\mathbb{P}(S_n \in I = [\dfrac{1}{\lambda} - \beta, \dfrac{1}{\lambda} + \beta]) = 0.95$.
5. Déterminer α_1 et α_2 tels que $\mathbb{P}(Z_n \in J = [\lambda\alpha_1, \lambda\alpha_2]) = 0.95$
6. Donner J, en fonction de λ inconnu, pour $n = 10000$ et $\phi = 1.96$.

Exercice 38. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. indépendantes de loi de Poisson $\mathbb{P}(1)$.

1. On pose $S_n = \sum_{k=1}^{n} X_k$. Calculer la loi de S_n.

2. Montrer que la suite $(\frac{1}{\sqrt{n}}(S_n - n))_{n\geq 1}$ converge en loi vers Y dont on donnera la loi.

3. Calculer $\mathbb{P}(\frac{1}{\sqrt{n}}(S_n - n)) \leq 0)$.

4. En déduire $\lim_{n\to+\infty} \sum_{k=0}^{n} e^{-n} \frac{n^k}{k!}$.