

Filière Informatique,
ENSEIRB-MATMECA

Probabilités et statistiques

Semestre 5

2018

Questions de cours

- Définition d'une fonction génératrice (définition 10, page 20).
- La fonction génératrice d'une somme est le produit des fonctions génératrices lorsque les variables sont indépendantes (proposition 8, page 21) (+ démo?).
- Savoir calculer les fonction génératrices d'une loi binomiale, Bernoulli et de Poisson (exemple page 22).
- La définition des différents modes de convergence (définition 14, page 27).
- L'énoncé et la démonstration de la Loi faible des grands nombres (proposition 20, page 29). Penser à insister sur le fait que la suite de variable est *IID* et de *carré intégrable*.

Table des matières

Cours	7
1 Probabilités sur un espace fini	9
1.1 Probabilités sur un espace fini, évènement	9
1.1.1 Définitions	9
1.1.2 Probabilité uniforme	10
1.2 Probabilité conditionnelle et indépendance	10
1.2.1 Loi de Bayes	10
1.2.2 Indépendance	10
2 Variables aléatoires discrètes	13
2.1 Espace de probabilité	13
2.2 Variable aléatoires discrètes	13
2.2.1 Définitions	13
2.2.2 Lois discrètes usuelles	14
2.2.3 Loi marginale	16
2.3 Espérance et variance	17
2.3.1 Espérance	17
2.3.2 Variance	19
2.4 Fonctions génératrices des variables aléatoires entières	20
3 Variables aléatoires à densité	23
3.1 Variable aléatoire réelles	23
3.1.1 Espérance et variance	24
3.2 Vecteurs aléatoires à densité	24
3.2.1 Définition	24
3.2.2 Densité marginale	25
3.3 Changement de variable	25
3.4 Indépendance	25
3.5 Covariance	26
4 Convergence et théorèmes limites	27
4.1 Convergence	27

4.2	Loi des grands nombres	29
4.2.1	Loi faible des grands nombres	29
4.2.2	Loi forte des grands nombres	30
4.3	Fonction caractéristique et convergence en loi	31
4.3.1	Fonction caractéristique	31
4.3.2	Convergence en loi (Théorème limite centrale)	32
4.4	Le théorème de la limite centrale	32
4.4.1	Théorèmes	32
4.4.2	Intervalle de confiance et méthode de Monte-Carlo	33

Annexes **36**

Questions de cours **39**

Fonction génératrice	39
Convergence	39
Loi faible des grands nombres	40

Cours

Chapitre 1

Probabilités sur un espace fini

1.1 Probabilités sur un espace fini, évènement

1.1.1 Définitions

On s'intéresse à une expérience aléatoire qui conduit à la validation d'un *seul résultat* parmi un nombre *fini* de résultats possibles notés $\omega_1, \dots, \omega_n$. On note $\Omega = \{\omega_1, \dots, \omega_n\}$ l'ensemble des résultats possibles. Par exemple, pour le jet d'une pièce à deux faces on a $\Omega = \{P, F\}$, pour celui d'un dé à six faces $\Omega = \{1, 2, 3, 4, 5, 6\}$.

On appelle *évènement* tout sous-ensemble de Ω .

Définition 1. Une probabilité \mathbb{P} sur un ensemble fini $\Omega = \{\omega_1, \dots, \omega_n\}$ est une famille de nombre p_1, \dots, p_n tels que $\forall i \in \{1, \dots, n\} p_i \geq 0$ et $\sum_{i=1}^n p_i = 1$. On attribue à tout évènement $A \subset \Omega$ le nombre

$$\mathbb{P}(A) = \sum_{i: \omega_i \in A} p_i .$$

Terminologie

- Si $\mathbb{P}(A) = 0$, on dit que l'évènement A est *négligeable*.
- Si $\mathbb{P}(A) = 1$, l'évènement A est dit *presque sûr*.
- L'évènement contraire de A est noté A^C , c'est le *complémentaire* de A dans Ω .
- Si A et B sont deux évènements, l'évènement *A et B* vaut $A \cap B$.
- *A ou B* vaut $A \cup B$. On a que :

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

La fonction indicatrice d'un évènement est une fonction définie sur Ω qui vaut 0 ou 1. Soit $A \subset \Omega$, I_A est la fonction indicatrice de A .

$$I_A : \Omega \longrightarrow \{0, 1\}$$

$$\omega \longmapsto I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

De plus on a $I_{A \cap B} = I_A \times I_B$ et $I_{A \cup B} = I_A + I_B - I_{A \cap B}$.

1.1.2 Probabilité uniforme

Dans le cas où les résultats possibles de Ω jouent le même rôle, ces résultats doivent avoir la même pondération : $\mathbb{P}(\omega_1) = \dots = \mathbb{P}(\omega_n)$ (soit $p_1 = \dots = p_n$). $\forall i \in \{1, \dots, n\} p_i = p_1$.

$$\sum_{i=1}^n p_i = 1 \quad \Rightarrow \quad p_i = \frac{1}{n} = \frac{1}{\text{card}(\Omega)}$$

$$\mathbb{P}(\{\omega_i\}) = p_i = \frac{1}{\text{card}(\Omega)}$$

Soit $A \in \Omega$.

$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i = \sum_{i:\omega_i \in A} p_1 = p_1 \cdot \text{card}(A) \quad \Rightarrow \quad \mathbb{P}(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

1.2 Probabilité conditionnelle et indépendance

La notion de probabilité conditionnelle permet de prendre en compte l'information dont on dispose (à savoir qu'un évènement B s'est réalisé) pour actualiser la probabilité d'un évènement A .

Définition 2. Soit Ω muni d'une proba \mathbb{P} et $A \subset \Omega$, $B \subset \Omega$. La probabilité conditionnelle de A sachant B , notée $\mathbb{P}(A|B)$ est définie par :

$$\mathbb{P}(A|B) = \begin{cases} \mathbb{P}(A) & \text{si } \mathbb{P}(B) = 0 \\ \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} & \text{si } \mathbb{P}(B) > 0 \end{cases} .$$

1.2.1 Loi de Bayes

Proposition 1. Soit B_1, \dots, B_n une partition de Ω ($B_i \cap B_j = \emptyset$ pour $j \neq i$ et $\Omega = \bigcup_{i=1}^n B_i$) et $A \subset \Omega$ tel que $\mathbb{P}(A) > 0$. Alors $\forall 1 \leq i \leq n$, on a que :

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}{\sum_{j=1}^n \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)} .$$

On a au numérateur : $\mathbb{P}(A|B_i)\mathbb{P}(B_i) = \mathbb{P}(A \cap B_i)$ et au dénominateur :

$$\sum_{j=1}^n \mathbb{P}(A \cap B_j) = \mathbb{P}\left(\bigcap_{j=1}^n (A \cap B_j)\right) = \mathbb{P}(A \cap \Omega) = \mathbb{P}(A)$$

1.2.2 Indépendance

Définition 3. Soit Ω un espace muni d'une probabilité \mathbb{P} . deux évènements sont indépendants si :

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

Remarque : l'indépendance est aussi définie par les relations

$$\mathbb{P}(A|B) = \mathbb{P}(A) \quad \text{ou} \quad \mathbb{P}(B|A) = \mathbb{P}(B).$$

Définition 4. Soient m évènements A_1, \dots, A_m . Ils sont indépendants si :

$$\forall I \subset \{1, \dots, m\} \quad \mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

Remarques :

- Il ne suffit pas que $\mathbb{P}(\bigcap_{i=1}^m A_i) = \prod_{i=1}^m \mathbb{P}(A_i)$ pour que A_1, \dots, A_m soient indépendants.
- Pour que trois évènements soient indépendants, il ne suffit pas qu'ils soient indépendants deux à deux.

Exemple Jet de deux pièces à pile ou face : $\Omega = \{(P, P), (F, F), (P, F), (F, P)\}$. \mathbb{P} est la probabilité uniforme sur Ω .

- $A = \{\text{La première pièce donne Pile.}\} : A = \{(P, P), (P, F)\}$.
- $B = \{\text{La deuxième pièce donne Face.}\} : A = \{(P, F), (F, F)\}$.
- $C = \{\text{Les deux pièces donnent le même résultat.}\} : C = \{(P, P), (F, F)\}$.

On a donc : $A \cap B = \{(P, F)\}$, $A \cap C = \{(P, P)\}$ et $C \cap B = \{(F, F)\}$.

- $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$
- $\mathbb{P}(A \cap B) = \frac{1}{4} = \mathbb{P}(A) \cdot \mathbb{P}(B)$
- $\mathbb{P}(A \cap C) = \frac{1}{4} = \mathbb{P}(A) \cdot \mathbb{P}(C)$
- $\mathbb{P}(B \cap C) = \frac{1}{4} = \mathbb{P}(B) \cdot \mathbb{P}(C)$

Or $\mathbb{P}(A \cap B \cap C) = 0 \neq \frac{1}{8} = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$.

Chapitre 2

Variables aléatoires discrètes

2.1 Espace de probabilité

Définition 5. Une tribu \mathcal{A} sur Ω est un ensemble de parties (sous-ensemble) de Ω qui vérifie les trois propriétés suivantes :

1. $\emptyset \in \mathcal{A}$ et $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$.
3. Si $(A_i)_{i \in \mathbb{N}}$ est une famille d'éléments de \mathcal{A} : $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$ et $\bigcap_{i \in \mathbb{N}} A_i \in \mathcal{A}$

Tout élément de \mathcal{A} sera appelé un évènement.

Exemples Ω est un ensemble.

- $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω est une tribu.
- $\{\emptyset, \Omega\}$ est aussi une tribu.
- $\{A, A^C, \emptyset, \Omega\}$ (avec $A \subset \Omega$) la plus petite tribu qui contient A .

Définition 6. Soit Ω muni d'une tribu \mathcal{A} . On appelle probabilité sur (Ω, \mathcal{A}) une application $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ qui vérifie :

- $\mathbb{P}(\Omega) = 1$.
- Si $(A_i)_{i \in \mathbb{N}}$ est une famille d'évènements de \mathcal{A} deux à deux disjoints : $A_i \cap A_j = \emptyset$ pour $i \neq j$, alors

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i).$$

Tout ce qui a été vu au chapitre précédent reste valable comme la notion de probabilité conditionnelle et d'indépendance que l'on complète de la façon suivante.

Une famille quelconque d'évènements est dite indépendante si toute sous-famille finie est indépendante.

2.2 Variable aléatoires discrètes

2.2.1 Définitions

Définition 7. On appelle variable aléatoire discrète toute application $X : \Omega \rightarrow F$ où F est dénombrable. Pour $x \in F$, on note $\{X = x\}$ l'évènement défini par $\{\omega \in \Omega : X(\omega) = x\}$.

La famille des nombres $(\mathbb{P}(\{X = x\}))_{x \in F}$ s'appelle la loi de X .

Cette définition n'a de sens que si on suppose :

$$\forall x \in F \quad \{\omega \in \Omega : X(\omega) = x\} = \{X = x\} \in \mathcal{A} \quad (\text{mesurabilité de } X).$$

Soit $A \in \mathcal{A}$. $\mathbb{P}(\{I_A = 1\}) = \mathbb{P}(A)$. On rappelle que

$$I_A : \Omega \longrightarrow \{0, 1\}$$

$$\omega \longmapsto I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

d'où, $\{I_A = 1\} = \{\omega \in \Omega : I_A(\omega) = 1\} = A$.

Indépendance de variables aléatoires discrètes

Soient n variables aléatoires notées $X_i : \Omega \longrightarrow F_i$ pour $\{1, \dots, n\}$ où F_i est dénombrable. $(X_i)_{i \in \{1, \dots, n\}}$ est une famille aléatoire discrète *indépendante* si $\forall (x_1, \dots, x_n) \in F_1 \times \dots \times F_n$

$$\mathbb{P}(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}) = \prod_{i=1}^n \mathbb{P}(\{X_i = x_i\}).$$

Une famille quelconque de variables aléatoires discrètes est dite indépendante *ssi* toute sous-famille est indépendante.

2.2.2 Lois discrètes usuelles

Loi de Bernoulli

On dit que X suit la loi de Bernoulli de paramètre $p \in [0, 1]$ et on note $X \sim \mathcal{B}(p)$ si $\mathbb{P}(\{X = 1\}) = p$ et $\mathbb{P}(\{X = 0\}) = 1 - p$.

$X : \Omega \longrightarrow \{0, 1\}$. Jeu de Pile ou Face.

Loi de Poisson de paramètre $\lambda > 0$

X suit une loi de Poisson ($\lambda > 0$).

$X : \Omega \longrightarrow \mathbb{N}$

$$\boxed{\mathbb{P}(\{X = k\}) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}} \quad \forall k \in \mathbb{N}$$

Loi géométrique de paramètre $p \in]0, 1]$

C'est la loi du temps du premier succès dans une suite d'expériences aléatoires indépendantes dont le succès est de probabilité p .

X suit une loi géométrique de paramètre $p \in]0, 1]$.

$$\boxed{\mathbb{P}(X = k) = (1 - p)^{k-1} \cdot p} \quad \forall k \in \mathbb{N}$$

On peut modéliser ce jeu par une suite de variables aléatoires $(X_j)_{j \in \mathbb{N}^*}$ où X_j suit une loi de Bernoulli de paramètre $p \in]0, 1]$. La suite $(X_j)_{j \in \mathbb{N}^*}$ est indépendante. Si on note T la variable aléatoire qui modélise le temps du premier succès :

$$T = \inf\{j \in \mathbb{N}^* : X_j = 1\}.$$

Pour $k \in \mathbb{N}^*$, on a

$$\mathbb{P}(\{T = k\}) = \mathbb{P}(\{X_1 = 0\} \cap \{X_2 = 0\} \cap \dots \cap \{X_{k-1} = 0\} \cap \{X_k = 1\})$$

d'où par indépendance des $(X_j)_{j \in \mathbb{N}^*}$:

$$\mathbb{P}(\{T = k\}) = \mathbb{P}(\{X_1 = 0\}) \times \mathbb{P}(\{X_2 = 0\}) \times \dots \times \mathbb{P}(\{X_{k-1} = 0\}) \times \mathbb{P}(\{X_k = 1\}).$$

On obtient donc :

$$\boxed{\mathbb{P}(\{T = k\}) = (1 - p)^{k-1} \cdot p} \quad \forall k \in \mathbb{N}^*.$$

Loi binomiale de paramètre $n \in \mathbb{N}^*$ et $p \in [0, 1]$

C'est le nombre de succès dans une suite de n expériences aléatoires indépendantes dont le succès à pour probabilité $p \in [0, 1]$.

Soit $(X_j)_{j=1, \dots, n}$ n variables aléatoires indépendantes qui suivent la loi de Bernoulli de paramètre p . Si on note B la variable aléatoire représentant le nombre de succès. On a

$$B = \sum_{j=1}^n X_j$$

donc $B : \Omega \longrightarrow \{0, 1, \dots, n\}$. On cherche la probabilité que $B = k$ pour $k \in \{0, \dots, n\}$.

$$\{B = k\} = \left\{ \sum_{j=1}^n X_j = k \right\} = \bigcup_{\substack{(x_1, \dots, x_n) \in \{0, 1\}^n \\ \sum_{i=1}^n x_i = k}} [\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}]$$

$$\mathbb{P}(\{B = k\}) = \mathbb{P} \left(\bigcup_{\substack{(x_1, \dots, x_n) \in \{0, 1\}^n \\ \sum_{i=1}^n x_i = k}} [\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}] \right)$$

La famille d'évènements $[\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}]$, pour (x_1, \dots, x_n) variant dans $\{0, 1\}^n$, est disjointe deux à deux. D'où

$$\mathbb{P}(\{B = k\}) = \sum_{\substack{(x_1, \dots, x_n) \in \{0, 1\}^n \\ \sum_{i=1}^n x_i = k}} \mathbb{P}(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}),$$

mais

$$\begin{aligned} \mathbb{P}(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}) &= \mathbb{P}(\{X_1 = x_1\}) \times \dots \times \mathbb{P}(\{X_n = x_n\}) \\ &= p^{x_1} (1 - p)^{1-x_1} \dots p^{x_n} (1 - p)^{1-x_n} \quad \text{par indépendance} \\ &= p^{\sum_{j=1}^n x_j} (1 - p)^{n - \sum_{j=1}^n x_j} \end{aligned}$$

enfin

$$\mathbb{P}(\{B = k\}) = \sum_{\substack{(x_1, \dots, x_n) \in \{0,1\}^n \\ \sum_{i=1}^n x_i = k}} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Au final, on obtient

$$\mathbb{P}(\{B = k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

2.2.3 Loi marginale

Soient X et Y deux variables aléatoires discrètes à valeurs respectivement dans F et G (ensembles dénombrables). Le couple $Z = (X, Y)$ est aussi une variable aléatoire discrète à valeurs dans $F \times G$.

La loi de X est $(\mathbb{P}(\{X = x\}))_{x \in F}$ et celle de Y : $(\mathbb{P}(\{Y = y\}))_{y \in G}$. La loi du couple $Z = (X, Y)$ est $(\mathbb{P}(\{(X, Y) = (x, y)\}))_{(x, y) \in F \times G}$.

À partir de la loi du couple $Z = (X, Y) : (\mathbb{P}(\{(X, Y) = (x, y)\}))_{(x, y) \in F \times G}$, on peut calculer la loi de X :

$$\{X = x\} = \bigcup_{y \in G} \{(X, Y) = (x, y)\} = \bigcup_{y \in G} [\{X = x\} \cap \{Y = y\}] = \{X = x\} \cap \underbrace{\left(\bigcup_{y \in G} \{Y = y\} \right)}_{\Omega}.$$

$$\mathbb{P}(\{X = x\}) = \mathbb{P} \left(\bigcup_{y \in G} \{(X, Y) = (x, y)\} \right)$$

Or $\emptyset = \{(X, Y) = (x, \hat{y})\} \cap \{(X, Y) = (x, \tilde{y})\}$ avec $\hat{y} \neq \tilde{y}$.

On obtient donc la *formule de la loi marginale* :

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

La réciproque est valable dans des cas particuliers comme par exemple si X et Y sont *indépendantes* alors $\mathbb{P}(\{(X, Y) = (x, y)\}) = \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(\{X = x\}) \cdot \mathbb{P}(\{Y = y\})$.

Supposons que X suit une loi de Bernoulli de paramètre 0.5 et on pose $Y = 1 - X$. Donc Y suit une loi de Bernoulli de paramètre 0.5.

$$\mathbb{P}(\{Y = 1\}) = \mathbb{P}(\{Y = 0\}) = 0.5$$

$$\mathbb{P}(\{(X, X) = (0, 0)\}) = \mathbb{P}(\{X = 0\}) = 0.5$$

$$\mathbb{P}(\{(X, Y) = (0, 0)\}) = \mathbb{P}(\{X = 0\} \cap \{Y = 0\}) = \mathbb{P}(\{X = 0\} \cap \{X = 1\}) = \mathbb{P}(\emptyset)$$

Les couples (X, X) et (X, Y) n'ont pas la même loi.

2.3 Espérance et variance

2.3.1 Espérance

Définition 8. Soit $X : \Omega \rightarrow F$ une variable discrète à valeurs dans F (dénombrable) et on suppose $F \subset \mathbb{R}$. Elle est dite intégrable si $\sum_{x \in F} |x| \cdot \mathbb{P}(\{X = x\}) < \infty$ (si la série est convergente). dans ce cas on note l'espérance de X par $\mathbb{E}[X]$ et

$$\mathbb{E}[X] = \sum_{x \in F} x \cdot \mathbb{P}(\{X = x\}).$$

Remarques :

- Le caractère intégrable de l'espérance d'une variable aléatoire ne dépendent que de la loi de cette variable aléatoire.
- Soit A un évènement, $\mathbb{E}[I_A] = 0 \times \mathbb{P}(\{I_A = 0\}) + 1 \times \mathbb{P}(\{I_A = 1\}) = \mathbb{P}(A)$.

Exemple L'espérance d'une variable aléatoire X qui suit¹ :

1. une loi de Bernoulli de paramètre p vaut $\mathbb{E}[X] = p$;
2. une loi de Poisson de paramètre $\lambda > 0$ vaut $\mathbb{E}[X] = \lambda$;
3. une loi de géométrique de paramètre $p \in]0, 1]$ vaut $\mathbb{E}[X] = \frac{1}{p}$.

Proposition 2 (Propriétés de l'espérance).

1. Linéarité : Si X et Y sont deux variables aléatoires discrètes à valeurs discrètes intégrables

$$\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda \cdot \mathbb{E}[Y] \quad \text{pour } \lambda \in \mathbb{R}.$$

2. Positivité : Si X est une variable aléatoire discrète intégrable et presque sûrement positive (c'est-à-dire que $\mathbb{P}(\{X \geq 0\}) = 1$) alors

$$\mathbb{E}[X] \geq 0 \quad \text{et} \quad [\mathbb{E}[X] = 0 \Rightarrow \mathbb{P}(\{X = 0\}) = 1].$$

3. Croissance : Si X et Y sont deux variables aléatoires discrètes intégrables et $\mathbb{P}(\{X \leq Y\}) = 1$ alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

4. Condition suffisante d'intégrabilité : Si X et Y sont deux variables aléatoires discrètes telles que $\mathbb{P}(\{|X| \leq Y\}) = 1$ et Y intégrable alors X est intégrable.

Démonstration.

1. cf TD pour les démos

1. On note $Z = X + \lambda Y$ et $H = \{z : z = x + \lambda y, x \in F \text{ et } y \in G\}$.

$$\begin{aligned}
\sum_{z \in H} |z| \cdot \mathbb{P}(\{Z = z\}) &= \sum_{z \in H} |z| \sum_{(x,y) \in F \times G} I_{z=x+\lambda y} \cdot \mathbb{P}(\{(X, Y) = (x, y)\}) \\
&= \sum_{z \in H} \sum_{(x,y) \in F \times G} |z| \cdot I_{z=x+\lambda y} \cdot \mathbb{P}(\{(X, Y) = (x, y)\}) \\
&= \sum_{z \in H} \sum_{(x,y) \in F \times G} I_{z=x+\lambda y} \cdot |x + \lambda y| \cdot \mathbb{P}(\{(X, Y) = (x, y)\}) \\
&= \sum_{(x,y) \in F \times G} \sum_{z \in H} I_{z=x+\lambda y} \cdot |x + \lambda y| \cdot \mathbb{P}(\{(X, Y) = (x, y)\}) \\
&= \sum_{(x,y) \in F \times G} |x + \lambda y| \cdot \mathbb{P}(\{X = x, Y = y\})
\end{aligned}$$

d'où

$$\begin{aligned}
\sum_{z \in H} |z| \cdot \mathbb{P}(\{Z = z\}) &\leq \sum_{(x,y) \in F \times G} [|x| + \lambda |y|] \cdot \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\
&= \underbrace{\sum_{x \in F} |x| \cdot \mathbb{P}(\{X = x\})}_{< \infty} + \lambda \underbrace{\sum_{y \in G} |y| \cdot \mathbb{P}(\{Y = y\})}_{< \infty} \\
&\Rightarrow \mathbb{E}[(X + \lambda Y)] = \mathbb{E}[X] + \lambda \mathbb{E}[Y]
\end{aligned}$$

□

Le résultat suivant exprime que l'espérance de $f(x)$ se calcule en fonction de la loi de X .

Théorème 3. Soit $X : \Omega \longrightarrow F$ une variable aléatoire discrète à valeurs dans F et $f : F \longrightarrow \mathbb{R}$ une application. Alors la variable discrète $f(x)$ est intégrable ssi

$$\sum_{x \in F} |f(x)| \mathbb{P}(\{X = x\}) < \infty$$

et dans ce cas

$$\mathbb{E}[f(x)] = \sum_{x \in F} f(x) \mathbb{P}(\{X = x\}) .$$

Remarque : si f est bornée alors $f(x)$ est intégrable si f est bornée alors

$$\sup_{x \in F} |f(x)| < \infty$$

d'où

$$\sum_{x \in F} |f(x)| \cdot \mathbb{P}(\{X = x\}) \leq \sum_{x \in F} \sup_{x \in F} |f(x)| \cdot \mathbb{P}(\{X = x\}) \leq \sup_{x \in F} |f(x)| \cdot \underbrace{\sum_{x \in F} \mathbb{P}(\{X = x\})}_1 .$$

La proposition va caractériser l'indépendance de variables aléatoires en fonction de leurs espérances.

Proposition 4. Soient X et Y deux variables aléatoires à valeurs dans F et G et deux fonctions $f : F \rightarrow \mathbb{R}$ et $g : G \rightarrow \mathbb{R}$. Si X et Y sont indépendantes et si $f(X)$ et $g(Y)$ sont intégrables alors $f(X) \cdot g(Y)$ est intégrable et

$$\mathbb{E}[f(X) \cdot g(Y)] = \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)].$$

Démonstration. On va vérifier que cette somme est finie.

$$\begin{aligned} & \sum_{(x,y) \in F \times G} |f(x) \cdot g(y)| \cdot \mathbb{P}(\{X = x, Y = y\}) \\ &= \sum_{x \in F} \sum_{y \in G} |f(x)| \cdot |g(y)| \cdot \mathbb{P}(\{X = x\}) \cdot \mathbb{P}(\{Y = y\}) \\ &= \underbrace{\left(\sum_{x \in F} |f(x)| \cdot \mathbb{P}(\{X = x\}) \right)}_{\mathbb{E}[|f(X)|]} \cdot \underbrace{\left(\sum_{y \in G} |g(y)| \cdot \mathbb{P}(\{Y = y\}) \right)}_{\mathbb{E}[|g(Y)|]} \end{aligned}$$

□

Proposition 5. Soient X et Y deux variables aléatoires à valeurs discrètes dans F et G . Si pour toutes fonctions f et g bornées $f : F \rightarrow \mathbb{R}$ et $g : G \rightarrow \mathbb{R}$, on a que $\mathbb{E}[f(X) \cdot g(Y)] = \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)]$ alors X et Y sont indépendantes.

Démonstration. X et Y sont indépendantes ssi $\forall (\bar{x}, \bar{y}) \in F \times G, \mathbb{P}(\{X = \bar{x}\})\mathbb{P}(\{Y = \bar{y}\}) = \mathbb{P}(\{(X, Y) = (\bar{x}, \bar{y})\})$.

$$\text{Appliquons l'hypothèse pour } \begin{array}{l} x : F \rightarrow \{0, 1\} \text{ et } g : G \rightarrow \{0, 1\} \\ x \mapsto I_{\{\bar{x}\}}(x) \qquad y \mapsto I_{\{\bar{y}\}}(y) \end{array} .$$

$$\begin{aligned} \mathbb{E}[f(X) \cdot g(Y)] &= E[I_{\{\bar{x}\}}(X) \cdot I_{\{\bar{y}\}}(Y)] = E[I_{\{(\bar{x}, \bar{y})\}}(X, Y)] = E[I_{\{(X, Y) = (\bar{x}, \bar{y})\}}] \\ &= \mathbb{P}(\{(X, Y) = (\bar{x}, \bar{y})\}) \end{aligned}$$

$$\begin{aligned} \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)] &= E[I_{\{\bar{x}\}}(X)] \cdot E[I_{\{\bar{y}\}}(Y)] = E[I_{\{X = \bar{x}\}}] \cdot E[I_{\{Y = \bar{y}\}}] \\ &= \mathbb{P}(\{X = \bar{x}\}) \cdot \mathbb{P}(\{Y = \bar{y}\}) \end{aligned}$$

$\Rightarrow X$ et Y sont indépendants.

□

2.3.2 Variance

Définition 9. Soit $X : \Omega \rightarrow F$ une variable discrète à valeurs dans $F \subset \mathbb{R}$. X est dite de carré intégrable si

$$\sum_{x \in F} x^2 \cdot \mathbb{P}(\{X = x\}) < \infty.$$

Dans ce cas, la variance vaut $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$.

Enfin, l'écart-type est la racine carrée de la variance.

Il existe une autre formule pour la variance :

$$\text{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\right] = \mathbb{E}[X^2] - \underbrace{\mathbb{E}[2X\mathbb{E}[X]]}_{2(\mathbb{E}[X])^2} + \mathbb{E}[\mathbb{E}[X]^2]$$

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

Par conséquent, on obtient l'inégalité suivante : $\boxed{\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2}$.

Proposition 6. Soient X_1, \dots, X_n n variables aléatoires discrètes de carré intégrable à valeurs réelles et indépendantes. Alors $\sum_{i=1}^n X_i$ est de carré intégrable et

$$\boxed{\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)}.$$

Démonstration.

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= \mathbb{E}\left[\left(\sum_{i=1}^n X_i - \mathbb{E}\left[\sum_{i=1}^n X_i\right]\right)^2\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right)^2\right] \\ &= \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n (X_i - \mathbb{E}[X_i]) \cdot (X_j - \mathbb{E}[X_j])\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] \end{aligned}$$

Si $i \neq j$, $\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \mathbb{E}[X_i - \mathbb{E}[X_i]] \cdot \mathbb{E}[X_j - \mathbb{E}[X_j]] = 0$.

$$\Rightarrow \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}[(X_i - \mathbb{E}[X_i])^2] = \sum_i \text{Var}(X_i)$$

□

2.4 Fonctions génératrices des variables aléatoires entières

Définition 10.² Soit X une variable aléatoire $X: \Omega \rightarrow \mathbb{N}$. On appelle fonction génératrice de X la fonction $G_X: [-1] \rightarrow \mathbb{R}$ et elle est définie par

$$\boxed{G_x(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbb{P}(\{X = k\})}.$$

2. Question de cours.

Proposition 7. La fonction génératrice caractérise la loi d'une variable aléatoire. X et Y ont même loi, c'est-à-dire $\mathbb{P}(\{X = k\}) = \mathbb{P}(\{Y = k\})$, $\forall k \in \mathbb{N}$ (on note cela $X \stackrel{\mathcal{L}}{=} Y$) ssi $G_X(s) = G_Y(s) \forall s \in [-1, 1]$.

Démonstration. « avec les mains » □

Proposition 8.³ Si $X: \Omega \rightarrow \mathbb{N}$ et $Y: \Omega \rightarrow \mathbb{N}$, X et Y sont indépendantes alors

$$\boxed{G_{X+Y}(s) = G_X(s)G_Y(s)}$$

Démonstration. $G_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y]$ par indépendance on obtient $G_{X+Y}(s) = \mathbb{E}[f(x)f(y)] = \mathbb{E}[f(x)]\mathbb{E}[f(y)] = G_X(s)G_Y(s)$ où $f(x) = s^X$ □

Exercice

Soit $(X_i)_{i \geq 1}$ des variables aléatoires entières IID (Indépendantes et Identiquement Distribuées) et N une variable aléatoire à valeurs dans \mathbb{N} . N et $(X_i)_{i \geq 1}$ sont indépendantes. On définit

$$S = \begin{cases} X_1 + X_2 + \dots + X_N & \text{si } N > 1 \\ 0 & \text{si } N = 0 \end{cases} .$$

1. Calculer la fonction génératrice de S .
2. En déduire la loi de S si X_1 suit une loi géométrique de paramètre q et N suit une loi géométrique de paramètre p .

1. $G_S(s) = \mathbb{E}[s^S]$ or $1 = \sum_{k=1}^{\infty} I_{\{N=k\}}(\omega)$ d'où

$$G_S(s) = \mathbb{E} \left[s^S \sum_{k=0}^n I_{\{N=k\}} \right] = \mathbb{E} \left[\sum_{k=0}^n s^S I_{\{N=k\}} \right] = \sum_{k=0}^n \mathbb{E} [s^S I_{\{N=k\}}].$$

On a $\mathbb{E}[s^S I_{\{N=0\}}] = \mathbb{E}[s^0 I_{\{N=0\}}] = \mathbb{P}(\{N = 0\})$ et pour $k > 0$

$$\mathbb{E}[s^S I_{\{N=k\}}] = \mathbb{E} \left[s^{\sum_{i=1}^k X_i} I_{\{N=k\}}(\omega) \right] = \mathbb{E} \left[s^{\sum_{i=1}^k X_i} \right] \mathbb{E} [I_{\{N=k\}}(\omega)]$$

car N et $(X_i)_{i \geq 1}$ sont indépendantes, d'où

$$\mathbb{E}[s^S I_{\{N=k\}}] = G_{X_1}(s) \dots G_{X_k}(s) \mathbb{P}(\{N = k\}) = [G_{X_1}(s)]^k \cdot \mathbb{P}(\{N = k\}).$$

Mais

$$\begin{aligned} G_S(s) &= \sum_{k=0}^{\infty} \mathbb{E} [s^S I_{\{N=k\}}] \\ \Rightarrow G_S(s) &= \mathbb{P}(\{N = 0\}) + \sum_{k=1}^{\infty} (G_{X_1}(s))^k \mathbb{P}(\{N = k\}) \\ &= G_N(G_{X_1}(s)) \end{aligned}$$

Si Y suit une loi géométrique de paramètre r : $g_Y(s) = \frac{rs}{1-(1-r)s}$. En effet,

$$g_Y(s) = \sum_{k=1}^{\infty} s^k \mathbb{P}(\{Y = k\}) = \sum_{k=1}^{\infty} s^k r(1-r)^{k-1} = sr = \sum_{k=1}^{\infty} [s(1-r)]^{k-1} = \frac{rs}{1-(1-r)s}$$

3. Question de cours.

$$G_S(s) = \frac{pG_{X_1}(s)}{1 - (1-p)G_{X_1}(s)} \quad \text{et} \quad G_{X_1}(s) = \frac{1-qs}{1-(1-q)s}$$

$$G_S(s) = \frac{pqs}{1-(1-pq)s}$$

Exemple de fonctions génératrices ⁴

Si X suit une loi binomiale de paramètre $n \in \mathbb{N}^*$ $p \in [0, 1]$: $G_X(s) = (sp + (1-p))^n$.
 $X = \sum_{i=1}^n X_i$ où X_i suit une loi de Bernoulli de paramètre p pour $i = 1, \dots, n$, $(X_i)_{i=1, \dots, n}$ IID. On a alors $G_X(s) = \prod_{i=1}^n G_{X_i}(s)$ où

$$G_{X_i}(s) = \mathbb{E}[s^{X_i}] = s^0\mathbb{P}(X_i = 0) + s\mathbb{P}(X_i = 1) = (1-p) + ps$$

donc on a $G_X(s) = \prod_{i=1}^n [(1-p) + ps]$.

Si N suit une loi de Poisson de paramètre $\lambda > 0$.

$$G_N(s) = \sum_{k=0}^{\infty} \mathbb{P}(N = k) \cdot s^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda + \lambda s} = e^{\lambda(s-1)}.$$

4. Question de cours.

Chapitre 3

Variables aléatoires à densité

3.1 Variable aléatoire réelles

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. On dit qu'une variable aléatoire $X: \Omega \rightarrow \mathbb{R}$ possède la densité $p: \mathbb{R} \rightarrow \mathbb{R}_+$ si $\forall a < b, \forall (a, b) \in \mathbb{R} \cup \{-\infty, +\infty\}$

$$\mathbb{P}(\{X \in]a, b]\}) = \int_{]a, b]} p(x) dx .$$

Remarque : en toute rigueur pour considérer la probabilité de l'évènement $\{\omega \in \Omega: X(\omega) \in]a, b]\}$ il faut supposer que $\{\omega \in \Omega: X(\omega) \in]a, b]\} \in \mathcal{A}$.

- $\int_{\mathbb{R}} p(x) dx = 1 = \mathbb{P}(\{X \in \mathbb{R}\})$ car $\{\omega \in \Omega: X(\omega) \in \mathbb{R}\} = \Omega$.
- Soit $x \in \mathbb{R}$ qui vaut

$$\mathbb{P}(\{X = x\}) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\left\{X \in \left]x - \frac{1}{n}, x\right]\right\}\right) = \int_{]x - \frac{1}{n}, x]} p(x) dx .$$

Exemple de densité

On dit que X suit

- la loi uniforme sur $[a, b]$ pour $(a, b)^2 \in \mathbb{R}^2$ ssi pour $a \neq b$

$$p(x) = \frac{1}{b-a} I_{[a, b]}(x)$$

- la loi exponentielle de paramètre $\lambda > 0$ ssi

$$p(x) = \lambda \cdot e^{-\lambda x} I_{\mathbb{R}_+}(x)$$

- la loi gaussienne (ou normale) de paramètre $\mu \in \mathbb{R}, \sigma^2 > 0$ ssi

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

3.1.1 Espérance et variance

Définition 11. La variable aléatoire $X: \Omega \rightarrow \mathbb{R}$ qui possède la densité $p: \mathbb{R} \rightarrow \mathbb{R}_+$ est dite

- *intégrable* si $\int_{\mathbb{R}} |x| \cdot p(x) dx < \infty$ et on définit dans ce cas $\mathbb{E}[X] = \int_{\mathbb{R}} x \cdot p(x) dx$
- *de carré intégrable* si $\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 \cdot p(x) dx < \infty$ et dans ce cas on définit la variance par $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$.w

Proposition 9 (Propriétés de l'espérance).

1. **Linéarité :** $\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda \cdot \mathbb{E}[Y]$ pour $\lambda \in \mathbb{R}$. X et Y variables aléatoires à densité intégrable.
2. **Condition suffisante d'intégrabilité** Si $\mathbb{P}(\{|X| \leq Y\}) = 1$ et Y intégrable alors X est intégrable.
3. **Croissance :** Si X et Y sont intégrables et $\mathbb{P}(\{X \geq Y\}) = 1$ alors $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

Fonction de répartition La fonction de répartition d'une variable aléatoire X à densité p est définie par

$$F_x(t) = \mathbb{P}(\{X \leq t\}) = \int_{]-\infty, t]} p(x) dx.$$

3.2 Vecteurs aléatoires à densité

3.2.1 Définition

On dit que le vecteur aléatoire $X = (X_1, X_2, \dots, X_d): \Omega \rightarrow \mathbb{R}^d$ possède la densité $p: \mathbb{R}^d \rightarrow \mathbb{R}_+$ si $\forall O$ ouvert de \mathbb{R}^d

$$\mathbb{P}(\{X \in O\}) = \int_O p(x_1, \dots, x_d) dx_1 dx_2 \dots dx_d.$$

Théorème 10. Le vecteur aléatoire $X: \Omega \rightarrow \mathbb{R}^d$ possède la densité $p: \mathbb{R}^d \rightarrow \mathbb{R}_+$ ssi

$$\mathbb{E}[f(x)] = \int_{\mathbb{R}^d} f(x_1, \dots, x_d) \cdot p(x_1, \dots, x_d) dx_1 dx_2 \dots dx_d.$$

Pour toute fonction $f: \mathbb{R}^d \rightarrow \mathbb{R}$ bornée.

Remarque : soit X une variable aléatoire à valeurs dans \mathbb{R}^d $X: \Omega \rightarrow \mathbb{R}^d$ et de densité $p: \mathbb{R}^d \rightarrow \mathbb{R}_+$ et $f: \mathbb{R}^d \rightarrow \mathbb{R}$. La variable aléatoire $f(x)$ est intégrable ssi $\int_{\mathbb{R}^d} |f(x_1, \dots, x_d)| \cdot p(x_1, \dots, x_d) dx_1 dx_2 \dots dx_d < \infty$ et dans ce cas

$$\mathbb{E}[f(x)] = \int_{\mathbb{R}^d} f(x_1, \dots, x_d) \cdot p(x_1, \dots, x_d) dx_1 dx_2 \dots dx_d.$$

3.2.2 Densité marginale

Soit $X = (X_1, X_2)$ un vecteur aléatoire qui admet la densité $p: \mathbb{R}^2 \rightarrow \mathbb{R}_+$, X_1 possède la densité $p_1: \mathbb{R} \rightarrow \mathbb{R}_+$.

$$\begin{aligned} \mathbb{P}(\{X_1 \in]a, b]\}) &= \int_{]a, b]} p_1(x_1) dx_1 = \mathbb{P}(\{(X_1, X_2) \in]a, b] \times \mathbb{R}\}) \\ &= \int_{]a, b] \times \mathbb{R}} p_1(x_1, x_2) dx_1 dx_2 = \int_{]a, b]} \int_{\mathbb{R}} p_1(x_1, x_2) dx_1 dx_2 \end{aligned}$$

On en conclut que $p_1(x_1) = \int_{\mathbb{R}} p(x_2, x_1) dx_2$, c'est la formule de la *densité marginale*.

3.3 Changement de variable

Proposition 11. Soit $X: \Omega \rightarrow \mathbb{R}^d$ est un vecteur aléatoire qui possède la densité $p(x)$ portée par un ouvert O :

$$\int_O p(x_1, \dots, x_d) dx_1 \dots dx_d = 1$$

et φ est une bijection de O sur O' de classe C^1 ainsi que son inverse φ^{-1} . Alors le vecteur $Y = \varphi(X)$ possède la densité :

$$q(y) = I_{O'}(y) \cdot (\varphi^{-1}(y)) \cdot |\text{Jac } \varphi^{-1}(y)| \quad \text{où} \quad |\text{Jac } \varphi^{-1}(y)| = \det \left(\frac{\partial \varphi_i^{-1}}{\partial y_j}(y) \right)_{1 \leq i, j \leq d}$$

3.4 Indépendance

Définition 12. Les vecteurs aléatoires $X_1: \Omega \rightarrow \mathbb{R}^{d_1} \dots X_n: \Omega \rightarrow \mathbb{R}^{d_n}$ qui possèdent les densités p_1, \dots, p_n sont dits *indépendants* si $X = (X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^{d_1 + \dots + d_n}$ possède la densité produit $p_1(x_1) \dots p_n(x_n)$ avec $x_i \in \mathbb{R}^{d_i}$

La proposition suivante va caractériser l'indépendance de vecteurs aléatoires en utilisant l'espérance.

Proposition 12. Soient $X_1: \Omega \rightarrow \mathbb{R}^{d_1} \dots X_n: \Omega \rightarrow \mathbb{R}^{d_n}$.

1. Si ces vecteurs aléatoires sont indépendants alors pour toutes fonctions $f_1: \mathbb{R}^{d_1} \rightarrow \mathbb{R} \dots f_n: \mathbb{R}^{d_n} \rightarrow \mathbb{R}$ telles que $f_1(X_1); \dots; f_n(X_n)$ sont intégrables alors $\prod_{i=1}^n f_i(X_i)$ est intégrable et

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

2. Inversement, si pour toutes fonctions $f_1: \mathbb{R}^{d_1} \rightarrow \mathbb{R} \dots f_n: \mathbb{R}^{d_n} \rightarrow \mathbb{R}$ bornées

$$\mathbb{E} \left[\prod_{i=1}^n f_i(X_i) \right] = \prod_{i=1}^n \mathbb{E} [f_i(X_i)]$$

alors (X_1, \dots, X_n) sont indépendantes.

3.5 Covariance

Définition 13. Soient Y et Z deux variables aléatoires à valeurs réelles et de carré intégrable. On appelle covariance de Y et Z , le réel noté $\text{Cov}(Y, Z)$ défini par

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

Soit $X = (X_1, \dots, X_d)$ un vecteur aléatoire à valeurs dans \mathbb{R}^d dont les composantes sont de carré intégrable. On appelle matrice de covariance du vecteur X la matrice K^X de dimension $d \times d$ et d'élément $K_{i,j}^X = \text{Cov}(X_i, X_j)$.

- Propriétés 13.**
1. $\text{Cov}(Y, Y) = \text{Var}(Y)$ où $Y: \Omega \rightarrow \mathbb{R}$.
 2. $\text{Cov}(Y, Z) = \mathbb{E}[YZ] = \mathbb{E}[Y]\mathbb{E}[Z]$ où $Y: \Omega \rightarrow \mathbb{R}$ et $Z: \Omega \rightarrow \mathbb{R}$.
 3. Si Y et Z sont indépendantes, $\text{Cov}(Y, Z) = 0$.
 4. K^X est symétrique et positive.

Chapitre 4

Convergence et théorèmes limites

Dans ce chapitre, on va introduire différents modes de convergence pour une suite de variables aléatoires. On va alors établir deux résultats.

1. *La loi forte des grands nombres* : la moyenne empirique d'une suite IID de variables aléatoires intégrables $\frac{1}{n} \sum_{i=1}^n X_i$ « converge » vers $\mathbb{E}[X_i]$ lorsque $n \rightarrow \infty$.
2. *Le théorème limite centrale* indique à quelle vitesse cette convergence a lieu sous l'hypothèse supplémentaire que les X_j sont de carré intégrable.

4.1 Convergence

Définition 14.¹ Pour $n \rightarrow \infty$, on dit qu'une suite $(X_n)_{n \geq 1}$ de variables à valeurs dans \mathbb{R}^d converge vers la variable aléatoire X à valeurs dans \mathbb{R}^d .

1. *Presque sûrement* si

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

pour presque tous les points ω (excepté pour un ensemble de probabilité égale à 0) la suite de réels $X_n(\omega)$ converge vers $X(\omega)$.

2. *En probabilité* si

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

3. Dans L^1 si les variables $(X_n)_{n \geq 1}$ et X sont intégrables et

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

4. Dans L^2 si les variables $(X_n)_{n \geq 1}$ et X sont de carré intégrable et

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

Remarque : Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles qui convergent vers X alors $\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \mathbb{E}[X]$. On peut montrer facilement que

$$|\mathbb{E}[X_n] - \mathbb{E}[X]| \leq \mathbb{E}[|X_n - X|] \Rightarrow \lim_{n \rightarrow \infty} |\mathbb{E}[X_n] - \mathbb{E}[X]| = 0$$

1. Question de cours.

Théorème 14 (Convergence dominée). Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles qui convergent presque sûrement vers X . On suppose de plus que la suite est dominée au sens où il existe une variable aléatoire Y intégrable telle que $\forall n \geq 1, \mathbb{P}(|X_n| \leq Y) = 1$. Alors X est intégrable et $(X_n)_{n \geq 1}$ converge dans L^1 vers X ce qui entraîne que

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

Pour comparer les différents modes de convergence on va établir des *inégalités*.

Proposition 15 (Inégalité de Markov.). Soit X une variable aléatoire réelle. Si $\mathbb{E}[|X|] < \infty$ alors

$$\forall a > 0 \quad \mathbb{P}(\{|X| \geq a\}) \geq \frac{\mathbb{E}[|X|]}{a}.$$

Démonstration.

$$\begin{cases} \text{Si } |X(\omega)| < a \Rightarrow I_{\{|X| \geq a\}}(\omega) = 0 \leq \frac{|X(\omega)|}{a}. \\ \text{Si } |X(\omega)| \geq a \Rightarrow I_{\{|X| \geq a\}}(\omega) = 1 \leq \frac{|X(\omega)|}{a}. \end{cases}$$

$$\Rightarrow I_{\{|X| \geq a\}} \leq \frac{|X(\omega)|}{a} \Rightarrow \mathbb{E}[I_{\{|X| \geq a\}}] \leq \frac{\mathbb{E}[|X(\omega)|]}{a}$$

□

Proposition 16 (Inégalité de Bienaymé-Tchebychev.). Soit X une variable aléatoire réelle. Si $\mathbb{E}[|X|^2] < \infty$ alors

$$\forall a > 0 \quad \mathbb{P}(\{|X| \geq a\}) \geq \frac{\mathbb{E}[X^2]}{a^2}.$$

Démonstration.

$$\begin{cases} \text{Si } |X(\omega)| < a \Rightarrow I_{\{|X| \geq a\}}(\omega) = 0 \leq \frac{|X(\omega)|^2}{a^2}. \\ \text{Si } |X(\omega)| \geq a \Rightarrow I_{\{|X| \geq a\}}(\omega) = 1 \leq \frac{|X(\omega)|^2}{a^2}. \end{cases}$$

$$\Rightarrow I_{\{|X| \geq a\}} \leq \frac{|X(\omega)|^2}{a^2} \Rightarrow \mathbb{E}[I_{\{|X| \geq a\}}] \leq \frac{\mathbb{E}[|X(\omega)|^2]}{a^2}$$

□

Proposition 17 (Inégalité de Cauchy Schwartz.). Soient X et Y deux variables aléatoires réelles. Si $\mathbb{E}[|X|^2] < \infty$ et $\mathbb{E}[|Y|^2] < \infty$ alors

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]}.$$

Proposition 18. La convergence dans L^2 implique la convergence dans L^1 qui implique la convergence en probabilité.

Démonstration.

— « La convergence dans L^2 implique la convergence dans L^1 ». $(X_n)_{n \geq 1}$ converge vers Y dans L^2 $\lim_{n \rightarrow \infty} \mathbb{E}[|X - X_n|^2] = 0$.

— « La convergence dans L^1 implique la convergence en probabilité ». On cherche à démontrer l'implication suivante :

$$\lim_{n \rightarrow \infty} \mathbb{E}[|X - X_n|] = 0 \quad \Rightarrow \quad \forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(\{|X - X_n| \geq \epsilon\}) = 0.$$

X et X_n sont intégrables, $X - X_n$ est intégrable et on a Markov :

$$\epsilon > \mathbf{a} \quad \mathbb{P}(\{|X - X_n| \geq \epsilon\}) \leq \frac{\mathbb{E}[|X| - |X_n|]}{\epsilon}$$

$$\Rightarrow \quad 0 \leq \lim_{n \rightarrow +\infty} \mathbb{P}(\{|X - X_n| \geq \epsilon\}) \leq \frac{1}{2} \lim_{n \rightarrow +\infty} \mathbb{E}[|X - X_n|] = 0$$

D'où le résultat. □

Proposition 19. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles qui convergent vers X dans L^2 . Alors $\mathbb{E}[X_n]$, $\mathbb{E}[X_n^2]$ et $\text{Var}(X_n)$ converge respectivement vers $\mathbb{E}[X]$, $\mathbb{E}[X^2]$ et $\text{Var}(X)$.

4.2 Loi des grands nombres

Cette loi porte sur le comportement de la moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ pour une suite $(X_k)_{k \geq 1}$ IID de variables aléatoires, sous certaines hypothèses.

4.2.1 Loi faible des grands nombres

Proposition 20.² La moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ d'une suite $(X_k)_{k \geq 1}$ de variables aléatoires réelles IID et de carré intégrable converge dans L^2 (et donc dans L^1 et en probabilité) vers $\mathbb{E}[X_1]$.

Démonstration. On suppose que $(X_j)_{j \geq 1}$ vérifie $\mathbb{E}[X^2] < \infty$ pour tout $j \geq 1$. On veut montrer que

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[(\bar{X}_n - \mathbb{E}[X_1])^2\right] = 0.$$

$$\mathbb{E}\left[(\bar{X}_n - \mathbb{E}[X_1])^2\right] = \mathbb{E}\left[\left(\frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}[X_1]\right)^2\right]$$

Mais $\mathbb{E}[X_1] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}[X_1] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}[X_k]$ car $(X_k)_{k \geq 1}$ est identiquement distribuée d'où pour tout $k \geq 1$ $\mathbb{E}[X_1] = \mathbb{E}[X_k]$.

$$\begin{aligned} \Rightarrow \quad \mathbb{E}\left[(\bar{X}_n - \mathbb{E}[X_1])^2\right] &= \mathbb{E}\left[\frac{1}{n^2} \left(\sum_{k=1}^n X_k - \sum_{k=1}^n \mathbb{E}[X_k]\right)^2\right] \\ &= \frac{1}{n^2} \cdot \mathbb{E}\left[\left(\sum_{k=1}^n X_k - \mathbb{E}\left[\sum_{k=1}^n X_k\right]\right)^2\right] \quad \text{Espérance linéaire} \end{aligned}$$

2. Question de cours : énoncé et démonstration, penser à insister sur le fait que les variables sont IID et de carré intégrable.

d'où

$$\begin{aligned}
 \mathbb{E}\left[\left(\overline{X}_n - \mathbb{E}[X_1]\right)^2\right] &= \frac{1}{n^2} \cdot \text{Var}\left(\sum_{k=1}^n X_k\right) \\
 &= \frac{1}{n^2} \cdot \sum_{k=1}^n \text{Var}(X_k) \quad \text{car } (X_k)_{k \geq 1} \text{ sont indépendantes} \\
 &= \frac{1}{n^2} \cdot \sum_{k=1}^n \text{Var}(X_1) \quad \text{car } (X_k)_{k \geq 1} \text{ sont IID } \text{Var}(X_k) = \text{Var}(X_1) \\
 &= \frac{\text{Var}(X_1)}{n}
 \end{aligned}$$

$\text{Var}(X_1) < \infty$ car $\mathbb{E}[X_1^2] < \infty$ d'où le résultat :

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\left(\overline{X}_n - \mathbb{E}[X_1]\right)^2\right] = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var}(X_1) = 0.$$

□

4.2.2 Loi forte des grands nombres

Théorème 21. *La moyenne empirique $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ d'une suite $(X_k)_{k \geq 1}$ de variables aléatoires réelles IID intégrables converge presque sûrement et dans L^1 vers $\mathbb{E}[X_1]$ lorsque $n \rightarrow \infty$.*

Démonstration. (Dans le cas particulier où $\mathbb{E}[X_1^4] < \infty$.)

$$\begin{aligned}
 \mathbb{E}\left[\left(\overline{X}_n - \mathbb{E}[X_1]\right)^4\right] &= \frac{1}{n^4} \cdot \mathbb{E}\left[\left(\sum_{k=1}^n (X_k - \mathbb{E}[X_k])\right)^4\right] \\
 &= \frac{1}{n^4} \sum_{i_1, i_2, i_3, i_4 \in \mathbb{N}^4} \mathbb{E}[(X_{i_1} - \mathbb{E}[X_{i_1}])(X_{i_2} - \mathbb{E}[X_{i_2}])(X_{i_3} - \mathbb{E}[X_{i_3}])(X_{i_4} - \mathbb{E}[X_{i_4}])]
 \end{aligned}$$

En remarquant que $\mathbb{E}[X_j - \mathbb{E}[X_j]] = 0$, on peut dire que le terme correspond à un élément où un indice est différents des trois autres va être nul par indépendance. On va obtenir un terme non nul dans la situation où

- tous les indices sont égaux et on obtiendra alors $\mathbb{E}\left[(X_1 - \mathbb{E}[X_1])^4\right]$ car les X_i sont identiquement distribués;
- deux indices prennent une valeurs et les deux autres indices prennent une valeur (différente) (avec $i \neq j$) :

$$\begin{aligned}
 \mathbb{E}\left[(X - \mathbb{E}[X_i])^2(X_j - \mathbb{E}[X_j])^2\right] &= \mathbb{E}\left[(X - \mathbb{E}[X_i])^2\right] \mathbb{E}\left[(X_j - \mathbb{E}[X_j])^2\right] \quad (\text{par indep}) \\
 &= (\text{Var}(X_1))^2 \quad \text{car } (X_k)_{k \geq 1} \text{ sont identique distribuées}
 \end{aligned}$$

$$\begin{aligned}
& \mathbb{E}\left[(\overline{X}_n - \mathbb{E}[X_n])^4\right] = \frac{1}{n^4} \mathbb{E}\left[(X_1 - \mathbb{E}[X_1])^4\right] + \frac{3n \cdot (n-1)}{n^3} \cdot (\text{Var}(X_1))^2 \\
\Rightarrow & \sum_{n=1}^{\infty} \mathbb{E}\left[(\overline{X}_n - \mathbb{E}[X_n])^4\right] < \infty \\
\Rightarrow & \mathbb{E}\left[\sum_{n=1}^{\infty} (\overline{X}_n - \mathbb{E}[X_n])^4\right] < \infty \\
\Rightarrow & \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} (\overline{X}_n - \mathbb{E}[X_n])^4 < \infty\right\}\right) = 1
\end{aligned}$$

Pour presque tout $\omega \in \Omega$

$$\sum_{n=1}^{\infty} (\overline{X}_n - \mathbb{E}[X_1])^4 < \infty \Rightarrow \lim_{n \rightarrow \infty} (\overline{X}_n - \mathbb{E}[X_1])^4 = 0$$

□

4.3 Fonction caractéristique et convergence en loi

4.3.1 Fonction caractéristique

Définition 15. Soit X un vecteur aléatoire à valeurs dans \mathbb{R}^d . On appelle fonction caractéristique de X la fonction $\phi_X: \mathbb{R}^d \rightarrow \mathbb{C}$. $\forall u \in \mathbb{R}^d$, $\phi_X(u) = \mathbb{E}[\exp(iu \cdot X)]$ où $u \cdot X = \sum_{i=1}^n u_i X_i$ (produit scalaire).

Si X suit la loi :

- de Bernoulli ($p \in [0, 1]$) : $\phi_X(u) = (1-p) + p e^{iu}$ $u \in \mathbb{R}$;
- Binomiale ($n \in \mathbb{N}^*$, $p \in [0, 1]$) : $\phi_X(u) = ((1-p) + p e^{iu})^n$ $u \in \mathbb{R}$;
- Géométrique ($p \in]0, 1]$) : $\phi_X(u) = \frac{p e^{iu}}{1 - (1-p) e^{iu}}$ $u \in \mathbb{R}$;
- de Poisson ($\lambda > 0$) : $\phi_X(u) = \exp(\lambda(e^{iu} - 1))$ $u \in \mathbb{R}$;
- Uniforme sur $[a, b]$: $\phi_X(u) = \frac{\sin\left(\frac{(b-a)u}{n}\right)}{\frac{(b-a)u}{n}} e^{iu \frac{a+b}{2}}$ $u \in \mathbb{R}$;
- Exponentielle ($\lambda > 0$) : $\phi_X(u) = \frac{\lambda}{\lambda - iu}$ $u \in \mathbb{R}$;
- Gaussienne ($\mu \in \mathbb{R}$, $\sigma^2 > 0$) : $\phi_X(u) = \exp\left(iu\mu - \frac{\sigma^2 u^2}{2}\right)$ $u \in \mathbb{R}$.

Proposition 22. La fonction caractéristique d'une variable aléatoire caractérise sa loi si :

$$X \text{ et } Y \text{ ont même loi } (X \stackrel{\mathcal{L}}{=} Y) \Leftrightarrow \forall u \in \mathbb{R}^d, \phi_X(u) = \phi_Y(u)$$

Propriété de la fonction caractéristique

Lemme 23. Soit X une variable aléatoire à valeurs réelles de carré intégrable : $\mathbb{E}[X^2] < \infty$. Alors sa fonction caractéristique est de classe C^2 et admet un développement limité en 0.

$$\boxed{\phi_X(u) = 1 + iu\mathbb{E}[X] - \frac{u^2}{2}\mathbb{E}[X^2] + o(u^2)}$$

4.3.2 Convergence en loi (Théorème limite centrale)

Définition 16. On dit qu'une suite X_n de variables aléatoires $(X_n)_{n \geq 1}$ à valeurs dans \mathbb{R}^d converge en loi vers X à valeurs dans \mathbb{R}^d et on note $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ si $\forall f: \mathbb{R}^d \rightarrow \mathbb{R}$ continue et bornée $\lim_{n \rightarrow +\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(x)]$.

Théorème 24. La suite $(X_n)_{n \geq 1}$ de variables aléatoires à valeurs dans \mathbb{R}^d convergence en loi vers une variable aléatoire X à valeurs dans \mathbb{R}^d ssi

$$\phi_{X_n} \xrightarrow[n \rightarrow \infty]{} \phi_X(u) \quad \forall u \in \mathbb{R}^d.$$

Proposition 25.

- Si $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ où X_n et X sont des variables aléatoires à valeurs dans \mathbb{R}^d et $g: \mathbb{R}^d \rightarrow \mathbb{R}$ continue alors $g(X_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} g(X)$.
- Si $(X_n)_{n \geq 1}$ converge en probabilité vers X alors $(X_n)_{n \geq 1}$ converge en loi vers X .

4.4 Le théorème de la limite centrale

4.4.1 Théorèmes

Théorème 26. Soit $(X_j)_{j \geq 1}$ une suite de variables aléatoires à valeurs réelles IID telle que $\mathbb{E}[X_1^2] < \infty$. On note $\sigma = \sqrt{\text{Var}(X_1)} > 0$ et $\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$ (moyenne empirique). Alors

$$\underbrace{\frac{\sqrt{n}}{\sigma}}_{(1)} \underbrace{(\bar{X}_n - \mathbb{E}[X_i])}_{(2)} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y \sim \mathcal{N}(0, 1)$$

où (1) $\xrightarrow[n \rightarrow \infty]{} \infty$ et (2) $\xrightarrow[n \rightarrow \infty]{} \infty$ (loi des grands nombres).

Démonstration. On va utiliser les fonctions caractéristiques.

$$\underbrace{\phi_{\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mathbb{E}[X_i])}}_{\text{calcul?}}(u) \xrightarrow[n \rightarrow \infty]{} \phi_Y(u) \quad \text{où } Y \sim \mathcal{N}(0, 1).$$

$$\phi_{\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mathbb{E}[X_i])}(u) = \mathbb{E} \left[\exp \left(iu \frac{\sqrt{n}}{\sigma} (\bar{X}_n - \mathbb{E}[X_i]) \right) \right]$$

mais

$$\bar{X}_n = \mathbb{E}[X_i] = \frac{1}{n} \sum_{j=1}^n (X_j - \mathbb{E}[X_j])$$

d'où

$$\begin{aligned}
\phi_{\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mathbb{E}[X_i])}(u) &= \mathbb{E} \left[\exp \left(\frac{i u}{\sigma \sqrt{n}} \sum_{j=1}^n (X_j - \mathbb{E}[X_j]) \right) \right] \\
&= \mathbb{E} \left[\prod_{j=1}^n \exp \left(\frac{i u}{\sigma \sqrt{n}} (X_j - \mathbb{E}[X_j]) \right) \right] \quad \text{mais } (X_j)_{j \geq 1} \text{ indépendants} \\
&= \prod_{j=1}^n \mathbb{E} \left[\exp \left(\frac{i u}{\sigma \sqrt{n}} (X_j - \mathbb{E}[X_j]) \right) \right] \quad (X_j)_{j \geq 1} \text{ IID} \\
&= \left(\underbrace{\mathbb{E} \left[\exp \left(\frac{i u}{\sigma \sqrt{n}} (X_1 - \mathbb{E}[X_1]) \right) \right]}_{=\phi_{X_1 - \mathbb{E}[X_1]} \left(\frac{u}{\sigma \sqrt{n}} \right)} \right)^n \quad \text{On a } \phi_Z(v) = \mathbb{E}[e^{ivZ}]. \\
&= \left(\phi_{X_1 - \mathbb{E}[X_1]} \left(\frac{u}{\sigma \sqrt{n}} \right) \right)^n
\end{aligned}$$

Mais si $\mathbb{E}[Z^2] < \infty$ alors :

$$\begin{aligned}
\phi_Z(v) &= 1 + iv\mathbb{E}[Z] - \frac{v^2}{2}\mathbb{E}[Z^2] + o(v^2) \\
&= \left(1 + \frac{i u}{\sigma \sqrt{n}} \times 0 - \frac{u^2}{2\sigma^2 n} \underbrace{\mathbb{E}[(X_1 - \mathbb{E}[X_1])^2]}_{\sigma^2} + o\left(\frac{1}{n}\right) \right)^n \\
&= \left(1 - \frac{u^2}{2n} + o\left(\frac{1}{n}\right) \right)^n \\
&= \exp \left(n \log \left(1 - \frac{u^2}{2n} + o\left(\frac{1}{n}\right) \right) \right) \xrightarrow{n \rightarrow \infty} e^{-\frac{u^2}{2}} = \phi_Y(u) \quad \text{où } Y \sim \mathcal{N}(0, 1).
\end{aligned}$$

□

Théorème 27. Si $(X_n)_{n \geq 1}$ converge en loi vers $a \in \mathbb{R}^{d_1}$ où X_n est à valeurs dans \mathbb{R}^{d_1} , $(Y_n)_{n \geq 1}$ converge en loi vers Y où Y et Y_n sont à valeurs dans \mathbb{R}^{d_2} . Alors :

$$(X_n, Y_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} (a, Y).$$

Une conséquence est que dans le cas où $d_1 = d_2 = 1$, $X_n Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} aY$.

4.4.2 Intervalle de confiance et méthode de Monte-Carlo

Le principe de la méthode de Monte Carlo repose sur la loi forte de grands nombres. Son principe consiste à générer la réalisation d'une suite IID de variables aléatoires $(X_j)_{j \geq 1}$ pour évaluer numériquement $\mathbb{E}[X_i] < \infty$.

La moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j \xrightarrow[n \rightarrow +\infty]{\text{P.S.}} \mathbb{E}[X_i]$. Si X_1 est de carré intégrable, le théorème de la limite centrale permet d'estimer la vitesse de convergence de cette approximation.

Si $\mathbb{E}[X_i] < \infty$, le théorème de la limite centrale donne :

$$\underbrace{\frac{\sqrt{n}}{\sigma}}_{\xrightarrow[n \rightarrow \infty]{} \infty} \underbrace{(\overline{X}_n - \mathbb{E}[X_i])}_{\xrightarrow[n \rightarrow \infty]{} \infty} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y \sim \mathcal{N}(0, 1)$$

converge en loi de $(Z_n)_{n \geq 1}$ vers Z . $\lim_{n \rightarrow +\infty} \mathbb{E}[f(Z_n)] = \mathbb{E}[f(Z)]$ pour toute fonction continue et bornée. On va choisir

$$f(x) = I_{\{|x| \leq a\}} = \begin{cases} 1 & \text{si } |x| \leq a \\ 0 & \text{sinon} \end{cases} .$$

$$\begin{aligned} \mathbb{E} \left[f \left(\frac{\sqrt{n}}{\sigma} (\overline{X}_n - \mathbb{E}[X_1]) \right) \right] &= \mathbb{P} \left(\left\{ \left| \frac{\sqrt{n}}{\sigma} (\overline{X}_n - \mathbb{E}[X_1]) \right| \leq a \right\} \right) \\ &= \mathbb{P} \left(\left\{ \overline{X}_n - \frac{a\sqrt{n}}{\sigma} \leq \mathbb{E}[X_1] \leq \overline{X}_n + \frac{a\sqrt{n}}{\sigma} \right\} \right) \end{aligned}$$

On a envie de conclure que

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left\{ \overline{X}_n - \frac{a\sqrt{n}}{\sigma} \leq \mathbb{E}[X_1] \leq \overline{X}_n + \frac{a\sqrt{n}}{\sigma} \right\} \right) = \mathbb{E}[f(Y)] \quad (4.1)$$

mais si $(Z_n)_{n \geq 1}$ converge en loi vers Z alors on a le résultat suivant : $\lim_{n \rightarrow \infty} \mathbb{E}[f(Z_n)] = \mathbb{E}[f(Z)]$ pour toute fonction f bornée et telle que $\mathbb{P}(\{z \in D_f\}) = 0$ où D_f est l'ensemble des points de discontinuités de f .

Dans notre cas $D_f = \{-a; a\}$, $\mathbb{P}(Y \in \{-a, a\}) = \mathbb{P}(\{Y = a\}) + \mathbb{P}(\{Y = -a\}) = 0 + 0 = 0$ car Y à densité, ce qui implique $\mathbb{P}(\{Y \in D_f\}) = 0$, d'où l'égalité 4.1.

Pour n grand :

$$\begin{aligned} \mathbb{P} \left(\left\{ \overline{X}_n - \frac{a\sqrt{n}}{\sigma} \leq \mathbb{E}[X_1] \leq \overline{X}_n + \frac{a\sqrt{n}}{\sigma} \right\} \right) &\simeq \frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{-\frac{y^2}{2}} dy \\ &= \mathbb{E}[f(Y)] = \mathbb{P}(\{Y \in [-a, a]\}) \end{aligned}$$

Soit I_n l'intervalle de confiance :

$$I_n = \left[\overline{X}_n - \frac{a\sqrt{n}}{\sigma}, \overline{X}_n + \frac{a\sqrt{n}}{\sigma} \right] \quad \mathbb{P}(\{\mathbb{E}[X_1] \in I_n\}) = \frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{-\frac{y^2}{2}} dy .$$

On règle

1. a tel que :

$$\frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{-\frac{y^2}{2}} dy = \begin{cases} 0,95 & a = 1,96 \\ 0,99 & a = 2,58 \end{cases} ;$$

2. n pour avoir I_n aussi petit que possible suivant la précision demandée.

$\sigma^2 = \mathbb{E}[X_1^2] - (\mathbb{E}[X_1])^2$ le problème est que dans la pratique σ est inconnue. il faut donc estimer σ . On a approximé $\mathbb{E}[X_1]$ par $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ et on peut approximer $\mathbb{E}[X_1^2]$ par $\frac{1}{n} \sum_{k=1}^n X_k^2$. De plus, la loi forte des grands nombres nous dit que $\frac{1}{n} \sum_{k=1}^n X_k^2 \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}[X_1^2]$ d'où

$$V_n = \frac{1}{n} \sum_{k=1}^n X_k^2 - \left(\frac{1}{n} \sum_{k=1}^n X_k \right)^2 \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \sigma^2$$

Donc

$$\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - \mathbb{E}[X_1]) \underset{n \rightarrow +\infty}{\overset{\mathcal{L}}{\rightarrow}} Y \sim \mathcal{N}(0, 1)$$

en faisant le même calcul, on a que

$$\mathbb{P}(\{\mathbb{E}[X_1] \in I'_n\}) = \frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{-\frac{y^2}{2}} dy \quad \text{où} \quad I'_n = \left[\bar{X}_n - \frac{a\sqrt{V_n}}{\sqrt{n}}, \bar{X}_n + \frac{a\sqrt{V_n}}{\sqrt{n}} \right].$$

Mais

$$\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - \mathbb{E}[X_1]) = \underbrace{\frac{\sigma}{\sqrt{V_n}}}_{\xrightarrow{\mathcal{L}} 1} \times \underbrace{\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mathbb{E}[X_1])}_{\xrightarrow{\mathcal{L}} Y} \xrightarrow{\mathcal{L}} Y.$$

Annexes

Questions de cours

Fonctions génératrices

Définition 10. Soit X une variable aléatoire $X: \Omega \rightarrow \mathbb{N}$. On appelle fonction génératrice de X la fonction $G_X: [-1] \rightarrow \mathbb{R}$ et elle est définie par

$$G_x(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbb{P}(\{X = k\}).$$

Proposition 8. Si $X: \Omega \rightarrow \mathbb{N}$ et $Y: \Omega \rightarrow \mathbb{N}$, X et Y sont indépendantes alors

$$G_{X+Y}(s) = G_X(s)G_Y(s)$$

Démonstration. $G_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y]$ par indépendance on obtient $G_{X+Y}(s) = \mathbb{E}[f(x)f(y)] = \mathbb{E}[f(x)]\mathbb{E}[f(y)] = G_X(s)G_Y(s)$ où $f(x) = s^X$ \square

Fonctions génératrices à savoir les calculer.

Si X suit une loi binomiale de paramètre $n \in \mathbb{N}^*$ $p \in [0, 1]$: $G_X(s) = (sp + (1 - p))^n$.
 $X = \sum_{i=1}^n X_i$ où X_i suit une loi de Bernoulli de paramètre p pour $i = 1, \dots, n$, $(X_i)_{i=1, \dots, n}$ IID. On a alors $G_X(s) = \prod_{i=1}^n G_{X_i}(s)$ où

$$G_{X_i}(s) = \mathbb{E}[s^{X_i}] = s^0 \mathbb{P}(X_i = 0) + s \mathbb{P}(X_i = 1) = (1 - p) + ps$$

donc on a $G_X(s) = \prod_{i=1}^n [(1 - p) + ps]$.

Si N suit une loi de Poisson de paramètre $\lambda > 0$.

$$G_N(s) = \sum_{k=0}^{\infty} \mathbb{P}(N = k) \cdot s^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda + \lambda s} = e^{\lambda(s-1)}.$$

Convergence

Définition 14. Pour $n \rightarrow \infty$, on dit qu'une suite $(X_n)_{n \geq 1}$ de variables à valeurs dans \mathbb{R}^d converge vers la variable aléatoire X à valeurs dans \mathbb{R}^d .

1. *Presque sûrement* si

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

pour presque tous les points ω (excepté pour un ensemble de probabilité égale à 0) la suite de réels $X_n(\omega)$ converge vers X_n .

2. En probabilité si

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

3. Dans L^1 si les variables $(X_n)_{n \geq 1}$ et X sont intégrables et

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

4. Dans L^2 si les variables $(X_n)_{n \geq 1}$ et X sont de carré intégrable et

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

Loi faible des grands nombres

Penser à insister sur le fait que la suite de variables est IID et de carré intégrable.

Proposition 20. La moyenne empirique $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ d'une suite $(X_k)_{k \geq 1}$ de variables aléatoires réelles IID et de carré intégrable converge dans L^2 (et donc dans L^1 et en probabilité) vers $\mathbb{E}[X_1]$.

Démonstration. On suppose que $(X_j)_{j \geq 1}$ vérifie $\mathbb{E}[X^2] < \infty$ pour tout $j \geq 1$. On veut montrer que

$$\lim_{n \rightarrow \infty} \mathbb{E}[(\overline{X}_n - \mathbb{E}[X_1])^2] = 0.$$

$$\mathbb{E}[(\overline{X}_n - \mathbb{E}[X_1])^2] = \mathbb{E}\left[\left(\frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}[X_1]\right)^2\right]$$

Mais $\mathbb{E}[X_1] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}[X_1] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}[X_k]$ car $(X_k)_{k \geq 1}$ est identiquement distribuée d'où pour tout $k \geq 1$ $\mathbb{E}[X_1] = \mathbb{E}[X_k]$.

$$\begin{aligned} \Rightarrow \mathbb{E}[(\overline{X}_n - \mathbb{E}[X_1])^2] &= \mathbb{E}\left[\frac{1}{n^2} \left(\sum_{k=1}^n X_k - \sum_{k=1}^n \mathbb{E}[X_k]\right)^2\right] \\ &= \frac{1}{n^2} \cdot \mathbb{E}\left[\left(\sum_{k=1}^n X_k - \mathbb{E}\left[\sum_{k=1}^n X_k\right]\right)^2\right] \quad \text{Espérance linéaire} \end{aligned}$$

d'où

$$\begin{aligned} \mathbb{E}[(\overline{X}_n - \mathbb{E}[X_1])^2] &= \frac{1}{n^2} \cdot \text{Var}\left(\sum_{k=1}^n X_k\right) \\ &= \frac{1}{n^2} \cdot \sum_{k=1}^n \text{Var}(X_k) \quad \text{car } (X_k)_{k \geq 1} \text{ sont indépendantes} \\ &= \frac{1}{n^2} \cdot \sum_{k=1}^n \text{Var}(X_1) \quad \text{car } (X_k)_{k \geq 1} \text{ sont IID } \text{Var}(X_k) = \text{Var}(X_1) \\ &= \frac{\text{Var}(X_1)}{n} \end{aligned}$$

$\text{Var}(X_1) < \infty$ car $\mathbb{E}[X_1^2] < \infty$ d'où le résultat :

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[(\overline{X}_n - \mathbb{E}[X_1])^2\right] = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var}(X_1) = 0.$$

□