

Probabilités

Binyze Mohamed

MP 2025-2026

Sommaire

1	Probabilités, espaces probabilisés	1
2	Probabilités conditionnelles et indépendance	3
3	Variables aléatoires : généralités	5
4	Variables aléatoires réelles discrètes	7
5	Variables aléatoires réelles à densité	10
6	Somme de variables aléatoires réelles	11
7	Espérance, variance et moments	12
8	Fonction génératrice d'une v.a.r. à valeurs dans $\mathbb{N}$	15
9	Inégalités, notions de convergence et théorèmes limites	17

1 Probabilités, espaces probabilisés

Tribus, espaces probabilisables et événements

Soit Ω un ensemble quelconque et $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω .

Définition 1.1.

tribu, espace probabilisable

Une partie $\mathcal{T}$ de $\mathcal{P}(\Omega)$ est appelée **tribu** (ou σ -**algèbre**) sur Ω si :

1. $\Omega \in \mathcal{T}$.
2. $\forall A \in \mathcal{T}, \overline{A} \stackrel{\text{d\'ef}}{=} \Omega \setminus A \in \mathcal{T}$.
3. $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$.

Le couple $(\Omega, \mathcal{T})$ est appelé **espace probabilisable**.

Dans la suite, le couple $(\Omega, \mathcal{T})$ désigne un espace probabilisable.

Exemple 1.1. ■ $\mathcal{T} = \{\emptyset, \Omega\}$ est une tribu sur Ω , dite tribu grossière.

■ $\mathcal{T} = \mathcal{P}(\Omega)$ est une tribu sur Ω , dite tribu discrète¹.

1. Ce choix est le plus usuel lorsque Ω est fini ou dénombrable.

Proposition 1.1.**propriétés**

1. $\emptyset \in \mathcal{T}$.
2. $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$.
3. $\forall A, B \in \mathcal{T}, A \cup B \in \mathcal{T}, A \cap B \in \mathcal{T}, A \setminus B \in \mathcal{T}$.

Définition 1.2.**événement**

Les parties A de Ω éléments de la tribu $\mathcal{T}$ sont appelées *événement* de Ω .

Probabilités, espaces probabilisés

Soit $(\Omega, \mathcal{T})$ un espace probabilisable.

Définition 1.3.**probabilité, espace probabilisé**

On appelle *probabilité* sur $(\Omega, \mathcal{T})$ toute application $\mathbb{P} : \mathcal{T} \rightarrow [0, 1]$ vérifiant :

1. $\mathbb{P}(\Omega) = 1$.
2. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles¹, la série $\sum \mathbb{P}(A_n)$ converge et

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n). \quad (\sigma\text{-additivité})$$

Le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ est appelé *espace probabilisé*.

1. $\forall (m, n) \in (\mathbb{N}^*)^2, m \neq n \implies A_m \cap A_n = \emptyset$.

Dans la suite, le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé.

Exemple 1.2. ■ Lorsque Ω est fini et $\mathcal{T} = \mathcal{P}(\Omega)$, on définit la *probabilité uniforme* sur Ω par $\mathbb{P}(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$.

Proposition 1.2.**propriétés élémentaires**

1. $\mathbb{P}(\emptyset) = 0$.
2. $\forall (A_0, \dots, A_n) \in \mathcal{T}^{n+1}$ deux à deux incompatibles,

$$\mathbb{P}\left(\bigcup_{k=0}^n A_k\right) = \sum_{k=0}^n \mathbb{P}(A_k).$$
3. Soit $A, B \in \mathcal{T}$.
 - a. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.
 - b. $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.
 - c. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
 - d. $A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $\forall (A_0, \dots, A_n) \in \mathcal{T}^{n+1}, \mathbb{P}\left(\bigcup_{k=0}^n A_k\right) \leq \sum_{k=0}^n \mathbb{P}(A_k)$.

Théorème 1.1.**continuité monotone**

1. Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$. Si $(A_n)_{n \in \mathbb{N}}$ est croissante, alors

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right). \quad (\text{continuité croissante})$$
2. Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$. Si $(A_n)_{n \in \mathbb{N}}$ est décroissante, alors

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right). \quad (\text{continuité décroissante})$$

Corollaire 1.1.**probabilité d'une union et d'une intersection dénombrable**

Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$.
$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{k=0}^n A_k\right) \quad \text{et} \quad \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{k=0}^n A_k\right).$$

Proposition 1.3.**sous-additivité**

Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$. $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$. (sous-additivité)

Théorème 1.2.**probabilité sur un univers au plus dénombrable**

Soit Ω un ensemble au plus dénombrable, noté $\Omega = \{\omega_n, n \in \mathbb{N}\}$. Soit $(p_n)_{n \in \mathbb{N}}$ une famille de réels positifs sommable de somme 1. Alors il existe une unique probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ telle que

$$\forall n \in \mathbb{N}, \mathbb{P}(\{\omega_n\}) = p_n \quad \text{et} \quad \forall A = \{\omega_i, i \in I \subset \mathbb{N}\} \in \mathcal{P}(\Omega), \mathbb{P}(A) = \sum_{i \in I} p_i.$$

Exemple 1.3. ■ Ω fini : $\Omega = \{\omega_1, \dots, \omega_n\}$. Une probabilité sur Ω est entièrement déterminée par le choix de $p_1, \dots, p_n \in \mathbb{R}^+$ avec $\sum_{i=1}^n p_i = 1$. En prenant $p_i = 1/n$, on définit l'équiprobabilité sur $[[1, n]]$.

Événements négligeables, événements presque sûrs**Définition 1.4.****événement négligeable, événement presque sûr**

1. On dit qu'un événement A est **négligeable** si $\mathbb{P}(A) = 0$.
2. On dit qu'un événement A est **presque sûr** (p.s) ou **quasi-certain** si $\mathbb{P}(A) = 1$.

Proposition 1.4.**propriétés quasi-certaines**

- | | |
|---|---|
| <ol style="list-style-type: none"> 1. Soit $A, B \in \mathcal{T}$ tel que $B \subset A$. <ol style="list-style-type: none"> a. A négligeable $\implies B$ négligeable. b. B p.s $\implies A$ p.s. | <ol style="list-style-type: none"> 2. Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$. <ol style="list-style-type: none"> a. $\forall n \in \mathbb{N}, A_n$ négligeable $\implies \bigcup_{n \in \mathbb{N}} A_n$ négligeable. b. $\bigcap_{n \in \mathbb{N}} A_n$ p.s $\implies \forall n \in \mathbb{N}, A_n$ p.s. |
|---|---|

2 Probabilités conditionnelles et indépendance

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé.

Probabilités conditionnelles**Définition 2.1.****probabilité conditionnelle**

Soit $B \in \mathcal{T}$ tel que $\mathbb{P}(B) \neq 0$. L'application $\mathbb{P}_B$ définie sur $\mathcal{T}$ par $\forall A \in \mathcal{T}, \mathbb{P}_B(A) \stackrel{\text{déf}}{=} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ est une probabilité sur $(\Omega, \mathcal{T})$ appelée la **probabilité de A sachant B** . On note aussi $\mathbb{P}(A | B)$ pour $\mathbb{P}_B(A)$.

Théorème 2.1.**formule des probabilités composées**

Soit $A_1, \dots, A_n \in \mathcal{T}$ tel que $\mathbb{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \neq 0$. $\mathbb{P}\left(\bigcap_{k=1}^n A_k\right) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \dots \mathbb{P}\left(A_n | \bigcap_{k=1}^{n-1} A_k\right)$.

Exemple 2.1. ■ Une urne contient $r \geq 2$ boules rouges et $n \geq 2$ boules noires, indiscernables au toucher. On tire successivement quatre boules sans remise. Quelle est la probabilité que les deux premières soient rouges et les deux dernières noires ?

- Notons R_1 l'événement « la première boule tirée est rouge », R_2 l'événement « la deuxième boule tirée est rouge », N_3 l'événement « la troisième boule tirée est noire » et N_4 l'événement « la quatrième boule tirée est noire ». On cherche à calculer $\mathbb{P}(R_1 \cap R_2 \cap N_3 \cap N_4)$.
- On a par hypothèses :

$$\mathbb{P}(R_1) = \frac{r}{r+n}, \quad \mathbb{P}(R_2 | R_1) = \frac{r-1}{r+n-1}, \quad \mathbb{P}(N_3 | R_1 \cap R_2) = \frac{n}{r+n-2} \text{ et } \mathbb{P}(N_4 | R_1 \cap R_2 \cap N_3) = \frac{n-1}{r+n-3}.$$

D'après la formule des probabilités composées : $\mathbb{P}(R_1 \cap R_2 \cap N_3 \cap N_4) = \frac{r(r-1)n(n-1)}{(r+n) \dots (r+n-3)}.$

Définition 2.2.

système complet d'événements

Une famille $(A_i)_{i \in I}$ d'événements, où I est un ensemble au plus dénombrable, est appelée **systeme complet d'événements** si : $\Omega = \bigcup_{i \in I} A_i$ avec $A_i \cap A_j = \emptyset$ pour tout $i \neq j$.

Théorème 2.2.

formule des probabilités totales

Soit $(A_i)_{i \in I}$ un système complet d'événements tel que $\forall i \in I, \mathbb{P}(A_i) \neq 0$. $\forall B \in \mathcal{T}, \mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B | A_i) \mathbb{P}(A_i)$.

Théorème 2.3.

formule de Bayes

1. Soit $A, B \in \mathcal{T}$ tel que $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$. $\mathbb{P}(A | B) = \frac{\mathbb{P}(B | A) \mathbb{P}(A)}{\mathbb{P}(B)}.$
2. Soit $(A_i)_{i \in I}$ un système complet d'événements tel que $\forall i \in I, \mathbb{P}(A_i) \neq 0$.

$$\forall j \in I, \forall B \in \mathcal{T} \text{ tel que } \mathbb{P}(B) \neq 0, \quad \mathbb{P}(A_j | B) = \frac{\mathbb{P}(B | A_j) \mathbb{P}(A_j)}{\sum_{i \in I} \mathbb{P}(B | A_i) \mathbb{P}(A_i)}.$$

Indépendance

Définition 2.3.

indépendance de deux événements

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

Exemple 2.2. ■ On tire successivement et **sans remise** deux boules dans une urne contenant 5 boules blanches et 2 boules rouges. Les événements « la première boule tirée est blanche » et « la seconde boule tirée est blanche » ne sont pas indépendants.

En revanche, si l'on procède à un tirage **avec remise**, ces événements deviennent indépendants.

Proposition 2.1.

complémentaires de deux d'événements indépendants

Soit $A, B \in \mathcal{T}$. Les événements A et B sont indépendants si, et seulement si, A et $\overline{B}$ sont indépendants.

Définition 2.4.**famille mutuellement indépendants**

Soit $(A_i)_{i \in I}$ une famille d'événements, où I est un ensemble au plus dénombrable.

1. Les événements A_i , $i \in I$ sont **deux à deux indépendants** si $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ pour tout $i \neq j$.
2. Les événements A_i , $i \in I$ sont **mutuellement indépendants**¹ ou **indépendants** si

$$\forall p \in \mathbb{N}^*, \forall i_1, \dots, i_p \in I \text{ avec } i_1 < \dots < i_p, \mathbb{P}\left(\bigcap_{k=1}^p A_{i_k}\right) = \prod_{k=1}^p \mathbb{P}(A_{i_k}).$$

1. L'indépendance mutuelle entraîne l'indépendance des événements deux à deux, et la réciproque est fausse.

Remarque 2.1. ■ Une sous-famille d'événements mutuellement indépendants est aussi mutuellement indépendants.

Proposition 2.2.**complémentaires d'événements mutuellement indépendants**

Soit $(A_i)_{i \in I}$, où I est un ensemble au plus dénombrable, une famille d'événements mutuellement indépendants

et $(\varepsilon_i)_{i \in I}$ une famille de réels telle que $\varepsilon_i = 0$ ou 1. On pose $A_i^{\varepsilon_i} = \begin{cases} A_i & \text{si } \varepsilon_i = 0 \\ \overline{A_i} & \text{si } \varepsilon_i = 1 \end{cases}$.

Alors la famille $(A_i^{\varepsilon_i})_{i \in I}$ est constituée d'événements mutuellement indépendants.

3 Variables aléatoires : généralités

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé.

Généralités

Définition 3.1.**variable aléatoire réelle**

On appelle **variable aléatoire réelle** (v.a.r.) sur $(\Omega, \mathcal{T}, \mathbb{P})$ toute application $X : \Omega \longrightarrow \mathbb{R}$ vérifiant

$$\forall x \in \mathbb{R}, X^{-1}([-\infty, x]) \in \mathcal{T}.$$

Dans la suite, toutes les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

Notations : Soit X une v.a.r. et $A \subset \mathbb{R}$. On note $(X \in A) \stackrel{\text{déf}}{=} X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\}$.

En particulier, si $a \in \mathbb{R}$,

$$(X = a) = \{\omega \in \Omega, X(\omega) = a\}, (X \leq a) = \{\omega \in \Omega, X(\omega) \leq a\} \text{ et } (X > a) = \{\omega \in \Omega, X(\omega) > a\}.$$

Remarque 3.1. ■ Soit X une v.a.r. Pour tout intervalle I de $\mathbb{R}$, $(X \in I) \in \mathcal{T}$.

Proposition 3.1.**opérations sur les v.a.r.**

1. Soit $X_1, \dots, X_n$ des v.a.r. ($n \geq 1$) et $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ continue. L'application $\omega \in \Omega \longmapsto f(X_1(\omega), \dots, X_n(\omega))$ est une v.a.r. On la note $f(X_1, \dots, X_n)$.
2. Soit X une v.a.r. et $f : \mathbb{R} \longrightarrow \mathbb{R}$ monotone (ou monotone par morceaux). L'application composée $f \circ X$ est une v.a.r. On la note $f(X)$.
3. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. qui converge simplement vers une application $X : \Omega \longrightarrow \mathbb{R}$, alors X est une v.a.r.

Exemple 3.1. ■ Soit $X_1, \dots, X_n$ des v.a.r. Alors $\sum_{i=1}^n X_i$, $\prod_{i=1}^n X_i$, $\min_{1 \leq i \leq n} X_i$ et $\max_{1 \leq i \leq n} X_i$ sont des v.a.r.

■ Soit X une v.a.r. Alors $\lfloor X \rfloor$, X^2 , $|X|$ et $\cos X$ sont des v.a.r.

Loi d'une variable aléatoire réelle

$\mathcal{I}(\mathbb{R})$ désigne l'ensemble de tous les intervalles de $\mathbb{R}$.

Définition 3.2.

loi d'une variable aléatoire réelle

Soit X une v.a.r. On appelle **loi de X** l'application

$$\begin{aligned} \mathbb{P}_X : \mathcal{I}(\mathbb{R}) &\longrightarrow [0, 1] \\ J &\longmapsto \mathbb{P}_X(J) = \mathbb{P}(X \in J) \end{aligned}.$$

Exemple 3.2. ■ Soit $\lambda \in \mathbb{R}$ et X la v.a.r. constante égale à λ . On a $\mathbb{P}_X(J) = \begin{cases} 1 & \text{si } \lambda \in J \\ 0 & \text{si } \lambda \notin J \end{cases} = 1_J(\lambda).$

Définition 3.3.

loi d'une famille de v.a.r.

Soit $X_1, \dots, X_n$ des v.a.r. On appelle **loi de la famille $(X_1, \dots, X_n)$** l'application

$$\begin{aligned} \mathbb{P}_{(X_1, \dots, X_n)} : \mathcal{I}(\mathbb{R})^n &\longrightarrow [0, 1] \\ \prod_{i=1}^n J_i &\longmapsto \mathbb{P}_{(X_1, \dots, X_n)}\left(\prod_{i=1}^n J_i\right) = \mathbb{P}\left(\bigcap_{i=1}^n (X_i \in J_i)\right) \end{aligned}.$$

Notation : On note $(X_1 \in J_1, \dots, X_n \in J_n)$ l'événement $\bigcap_{i=1}^n (X_i \in J_i)$.

Définition 3.4.

loi conditionnelle de X sachant A

Soit X une v.a.r. et $A \in \mathcal{T}$ tel que $\mathbb{P}(A) \neq 0$. On appelle **loi de X sachant A** l'application définie par

$$\forall J \in \mathcal{I}(\mathbb{R}), \quad \mathbb{P}_A(X \in J) = \frac{\mathbb{P}((X \in J) \cap A)}{\mathbb{P}(A)}.$$

On la note $(\mathbb{P}_A)_X$.

Fonction de répartition d'une variable aléatoire réelle

Définition 3.5.

fonction de répartition d'une v.a.r.

Soit X une v.a.r. sur $(\Omega, \mathcal{T}, \mathbb{P})$. On appelle **fonction de répartition de X** l'application

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow [0, 1] \\ t &\longmapsto F_X(t) = \mathbb{P}(X \leq t) \end{aligned}.$$

Proposition 3.2.

propriétés de F_X

1. F_X est croissante¹ sur $\mathbb{R}$.
 2. F_X est continue à droite en tout point de $\mathbb{R}$.
 3. $\lim_{t \rightarrow -\infty} F_X(t) = 0$.
 4. $\lim_{t \rightarrow +\infty} F_X(t) = 1$.
1. En particulier, F_X admet en tout point une limite à droite et une limite à gauche.

Proposition 3.3.

fonction de répartition de X vs la loi de X

Soient X une v.a.r. sur $(\Omega, \mathcal{T}, \mathbb{P})$ et $(a, b) \in \mathbb{R}^2$ tel que $a \leq b$.

1. $\mathbb{P}(X > a) = 1 - F_X(a)$.
2. $\mathbb{P}(X < a) = \lim_{t \rightarrow a^-} F_X(t)$.
3. $\mathbb{P}(X \geq a) = 1 - \lim_{t \rightarrow a^-} F_X(t)$.
4. $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.
5. $\mathbb{P}(a \leq X \leq b) = F_X(b) - \lim_{t \rightarrow a^-} F_X(t)$.
6. $\mathbb{P}(a \leq X < b) = \lim_{t \rightarrow b^-} F_X(t) - \lim_{t \rightarrow a^-} F_X(t)$.
7. $\mathbb{P}(a < X < b) = \lim_{t \rightarrow b^-} F_X(t) - F_X(a)$.
8. $\mathbb{P}(X = a) = F_X(a) - \lim_{t \rightarrow a^-} F_X(t)$.

Remarque 3.2. ■ La fonction de répartition caractérise la loi d'une variable aléatoire réelle : la connaissance de F_X permet de calculer $\mathbb{P}(X \in I)$ pour tout intervalle I de $\mathbb{R}$.

Corollaire 3.1.

continuité de F_X

Soit X une v.a.r. sur $(\Omega, \mathcal{T}, \mathbb{P})$ et $a \in \mathbb{R}$. F_X est continue en $a \iff \mathbb{P}(X = a) = 0$.

Définition 3.6.

fonction de répartition d'une famille de v.a.r.

Soit $X_1, \dots, X_n$ des v.a.r. sur $(\Omega, \mathcal{T}, \mathbb{P})$. On appelle **fonction de répartition de la famille** $(X_1, \dots, X_n)$ l'application¹

$$F_{(X_1, \dots, X_n)} : \begin{array}{ll} \mathbb{R}^n & \longrightarrow [0, 1] \\ (t_1, \dots, t_n) & \longmapsto F_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \mathbb{P}(X_1 \leq t_1, \dots, X_n \leq t_n) \end{array}.$$

1. La connaissance de $F_{(X_1, \dots, X_n)}$ permet de calculer $\mathbb{P}(X_1 \in I_1, \dots, X_n \in I_n)$ pour tous intervalles $I_1, \dots, I_n$ de $\mathbb{R}$.

Variables aléatoires réelles indépendantes

Définition 3.7.

famille de v.a.r. indépendante

Soit $(X_j)_{j \in J}$ une famille de v.a.r.

1. Les v.a.r. $X_j, j \in J$, sont dites **deux à deux indépendantes** si, pour toute famille $(I_j)_{j \in J}$ d'intervalles de $\mathbb{R}$, les événements $(X_j \in I_j), j \in J$, sont deux à deux indépendants.
2. Les v.a.r. $X_j, j \in J$, sont dites **indépendantes** (ou **mutuellement indépendantes**¹) si, pour toute famille $(I_j)_{j \in J}$ d'intervalles de $\mathbb{R}$, les événements $(X_j \in I_j), j \in J$, sont mutuellement indépendants.

1. L'indépendance mutuelle entraîne l'indépendance des v.a.r deux à deux. La réciproque est fausse.

Remarque 3.3. ■ Une sous-famille d'une famille de v.a.r. mutuellement indépendantes est aussi constituée de v.a.r. mutuellement indépendantes.

Théorème 3.1.

indépendance héritée (lemme des coalitions)

Soient $k \in \mathbb{N}^*$, $(n_0, \dots, n_k) \in \mathbb{N}^{k+1}$ tels que $0 = n_0 < n_1 < \dots < n_{k-1} < n_k = n$ et $f_i : \mathbb{R}^{n_i - n_{i-1}} \longrightarrow \mathbb{R}$, $1 \leq i \leq k$ une application continue. Si $(X_i)_{1 \leq i \leq n}$ est une famille de v.a.r. indépendantes alors la famille $(f_1(X_1, \dots, X_{n_1}), f_2(X_{n_1+1}, \dots, X_{n_2}), \dots, f_k(X_{n_{k-1}+1}, \dots, X_{n_k}))$ est indépendante.

Exemple 3.3. ■ Soit (X_1, X_2, X_3, X_4) une famille de v.a.r. indépendantes. La famille $(X_1^2, X_2 X_3, \cos(X_4))$ est indépendante.

4 Variables aléatoires réelles discrètes

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les variables aléatoires réelles seront définies sur cet espace.

Loi d'une variable aléatoire réelle discrète

Définition 4.1.

variable aléatoire réelle discrète

Une v.a.r. X est dite de **loi discrète** (v.a.r.d.) s'il existe $\Omega' \in \mathcal{T}$ tel que¹ $\mathbb{P}(\Omega') = 1$ et $D = X(\Omega')$ soit au plus dénombrable.

1. Dans la plupart du temps, l'existence de l'ensemble Ω' est justifié par l'univers Ω .

Remarque 4.1. ■ Soit X une v.a.r.d. et $I \in \mathcal{I}(\mathbb{R})$.

- On a $(X \in I) = \bigcup_{x \in I \cap D} (X = x)$, donc si $\forall x \in D, (X = x) \in \mathcal{T}$ alors $(X \in I) \in \mathcal{T}$.
- $\mathbb{P}(X \in I) = \mathbb{P}(X \in I \cap D) = \mathbb{P}\left(\bigcup_{x \in I \cap D} (X = x)\right) = \sum_{x \in I \cap D} \mathbb{P}(X = x)$. On peut supprimer de D tous les éléments x tels que $\mathbb{P}(X = x) = 0$, les x restants sont appelés les **valeurs possibles** de la v.a.r.d. X .

Proposition 4.1.**loi d'une v.a.r.d.**

La loi d'une v.a.r.d. X est caractérisée par la donnée de¹ D et de l'application $x \in D \mapsto \mathbb{P}(X = x)$.

1. Dans la plupart du temps, l'ensemble D est $X(\Omega)$.

Remarque 4.2. ■ Soit X une v.a.r.d. Pour tout $A \subset D$, $\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)$.

Théorème 4.1.**existence d'une suite infinie de v.a.r. indépendantes**

Si $\mathcal{L}$ désigne une loi d'une v.a.r. discrète, il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ sur lequel il existe une suite $(X_n)_{n \in \mathbb{N}}$ de v.a.r. mutuellement indépendantes et suivant chacune la loi $\mathcal{L}$.

Définition 4.2.**v.a.r.d. de loi usuelle**

Soit X une v.a.r.d. On dit que la loi de X est **discrète usuelle** s'il existe un intervalle J de $\mathbb{Z}$ et une bijection

$$\begin{aligned} \varphi: J &\longrightarrow D \\ k &\longmapsto x_k \end{aligned}$$

Exemples classiques de lois discrètes usuelles

Soit X une v.a.r.d.

Définition 4.3.**loi uniforme**

On dit que X suit la **loi uniforme** sur un ensemble fini E et, on note $X \sim \mathcal{U}(E)$, si :

$$X(\Omega) = E \quad \text{et} \quad \forall x \in X(\Omega), \quad \mathbb{P}(X = x) = \frac{1}{\text{Card}(E)}.$$

Situations-types :

- Si X est la valeur du lancer d'un dé équilibré, alors $X \sim \mathcal{U}([1, 6])$.
- On tire « au hasard » un entier entre 1 et n . Soit X le chiffre obtenu, alors $X \sim \mathcal{U}([1, n])$.

Définition 4.4.**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 :

$$X(\Omega) = \{0, 1\} \quad \text{et} \quad \mathbb{P}(X = 0) = 1 - p, \quad \mathbb{P}(X = 1) = p.$$

Remarque 4.3. ■ L'événement $(X = 1)$ est appelé « succès » et l'événement $(X = 0)$ est appelé « échec ».

■ Les lois de Bernoulli servent à modéliser les épreuves à deux issues : succès ou échec.

Situations-types :

- On lance une pièce et soit X égale à 1 si l'on obtient « pile » et 0 si on l'obtient « face ». Alors $X \sim \mathcal{B}(p)$ où p est la probabilité d'obtenir « pile ».
- On lance un dé équilibré, dont les 6 faces sont numérotées de 1 à 6 et soit X égale à 1 si l'on obtient un chiffre multiple de 3 et 0 sinon. Alors $X \sim \mathcal{B}(p)$ où $p = \mathbb{P}(X = 1) = \mathbb{P}(\{3, 6\}) = \frac{2}{6} = \frac{1}{3}$.

Remarque 4.4. ■ Soit $A \in \mathcal{T}$. La v.a.r.d. 1_A suit la loi de Bernoulli de paramètre $p = \mathbb{P}(A)$.

Définition 4.5.**loi binomiale**

On dit que X suit la **loi binomiale** de paramètres $n \in \mathbb{N}^*$ et $p \in]0, 1[$ et, on note $X \sim \mathcal{B}(n, p)$, si :

$$X(\Omega) = \llbracket 0, n \rrbracket \quad \text{et} \quad \forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Remarque 4.5. ■ Les lois binomiales servent à modéliser le nombre de succès lors de la répétition d'épreuves de Bernoulli indépendantes.

■ La loi $\mathcal{B}(1, p)$ est simplement la loi $\mathcal{B}(p)$.

Situations-types :

■ On lance une pièce n fois et soit X le nombre de piles obtenus, alors $X \sim \mathcal{B}(n, p)$ où p est la probabilité d'obtenir « pile ».

■ Une urne contient des boules blanches et des boules rouges. On tire n boules **avec remise** dans cette urne et soit X le nombre de boules blanches obtenues. Alors $X \sim \mathcal{B}(n, p)$ où p est la probabilité d'obtenir une boules blanche.

Définition 4.6.**loi géométrique**

On dit que X suit la **loi géométrique** de paramètre $p \in]0, 1[$ et, on note $X \sim \mathcal{G}(p)$, si :

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

Remarque 4.6. ■ Une loi géométrique permet de modéliser le **temps d'attente** du premier succès dans la répétition d'épreuves de Bernoulli indépendantes et de même paramètre p .

Situations-types :

■ On lance une pièce une infinité de fois et soit X le rang du premier « pile » obtenu, alors $X \sim \mathcal{G}(p)$ où p est la probabilité d'obtenir « pile ».

■ On lance indéfiniment un dé et soit T le temps d'attendre du premier 6, alors $T \sim \mathcal{G}(p)$ où $p = \frac{1}{6}$.

Définition 4.7.**loi de Poisson**

On dit que X suit la **loi de Poisson** de paramètre $\lambda > 0$ et, on note $X \sim \mathcal{P}(\lambda)$, si :

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

Remarque 4.7. ■ Les lois de Poisson servent à modéliser le nombre de succès lors d'une grande répétition d'épreuves ayant une faible probabilité de réussite, on parle d'**événements rares**.

Situations-types :

■ Il n'y a pas de situation type où l'on peut reconnaître la loi de Poisson. La loi de Poisson peut être vue comme une « loi limite ».

Loi conjointe et lois marginales

X et Y désignent deux v.a.r.d.

Définition 4.8.**loi conjointe, loi marginale**

1. La loi de la v.a.d. $Z: \Omega \longrightarrow \mathbb{R}^2$ est appelée **loi conjointe** de X et Y .
 $\omega \longmapsto (X(\omega), Y(\omega))$

2. Les lois de X et Y sont appelées **lois marginales** de $Z = (X, Y)$.

Proposition 4.2.**la loi conjointe détermine les lois marginales**

Soient D_1 , D_2 l'ensemble des valeurs possibles respectivement de X et Y . On a :

$$\forall x \in D_1, \mathbb{P}(X = x) = \sum_{y \in D_2} \mathbb{P}(X = x, Y = y) \quad \text{et} \quad \forall y \in D_2, \mathbb{P}(Y = y) = \sum_{x \in D_1} \mathbb{P}(X = x, Y = y).$$

5 Variables aléatoires réelles à densité

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les variables aléatoires réelles seront définies sur cet espace.

Loi d'une variable aléatoire réelle à densité

Définition 5.1.**variable aléatoire réelle à densité**

1. Une v.a.r. X est dite de **loi à densité** (ou de **loi continue**) si sa fonction de répartition F_X est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé d'un sous-ensemble F fini ou vide.

2. On appelle alors **densité** de X la fonction f_X définie sur $\mathbb{R}$ par :

$$\forall t \in \mathbb{R}, f_X(t) = \begin{cases} F'_X(t) & \text{si } t \notin F \\ 0 & \text{si } t \in F \end{cases}.$$
Proposition 5.1.**caractérisation d'une densité**

La densité f_X est une fonction positive, continue sur $\mathbb{R} \setminus F$, d'intégrale convergente et valant 1 sur $\mathbb{R}$.

Remarque 5.1. ■ Si X est une v.a.r. de loi à densité, alors $\forall a \in \mathbb{R}, \mathbb{P}(X = a) = 0$.

Proposition 5.2.**fonction de répartition de X vs la loi de X**

Soit X une v.a.r. de loi à densité et $I \in \mathcal{I}(\mathbb{R})$ avec $a = \inf I$ et $b = \sup I$.

$$\mathbb{P}(X \in I) = F_X(b) - F_X(a) = \int_a^b f_X(t) dt.$$

En particulier, $\forall x \in \mathbb{R}, F_X(x) = \int_{-\infty}^x f_X(t) dt$.

Exemples classiques de lois à densité usuelles

Soit X une v.a.r. à densité.

Définition 5.2.**loi uniforme**

Soit $(a, b) \in \mathbb{R}^2$ tel que $a < b$. On dit que X suit la **loi uniforme** sur $[a, b]$ et, on note $X \sim \mathcal{U}([a, b])$, si :

$$\forall t \in \mathbb{R}, f_X(t) = \frac{1}{b-a} 1_{[a,b]}(t) = \begin{cases} \frac{1}{b-a} & \text{si } t \in [a, b] \\ 0 & \text{si } t \notin [a, b] \end{cases}.$$

Définition 5.3.**loi exponentielle**

On dit que X suit la **loi exponentielle** de paramètre $\lambda > 0$ et, on note $X \sim \mathcal{E}(\lambda)$, si :

$$\forall t \in \mathbb{R}, f_X(t) = \lambda e^{-\lambda t} 1_{[0, +\infty[}(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}$$

Définition 5.4.**loi Gamma**

On dit que X suit la **loi Gamma** de paramètre (α, λ) avec $\alpha, \lambda > 0$ et, on note $X \sim \Gamma(\alpha, \lambda)$, si :

$$\forall t \in \mathbb{R}, f_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} 1_{]0, +\infty[}(t) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}.$$

où $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$.

Définition 5.5.**loi gaussienne**

On dit que X suit la **loi gaussienne** (ou **normale**) de paramètre (μ, σ^2) avec $\mu \in \mathbb{R}$, $\sigma > 0$ et, on note $X \sim \mathcal{N}(\mu, \sigma^2)$, si :

$$\forall t \in \mathbb{R}, f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right).$$

Lorsque $X \sim \mathcal{N}(0, 1)$, on dit X suit la **loi normale centrée réduite**.

6 Somme de variables aléatoires réelles

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les variables aléatoires réelles seront définies sur cet espace.

Proposition 6.1.**somme de v.a.r.d. indépendantes**

Soient X et Y deux v.a.r.d. indépendantes dont l'ensemble des valeurs possibles est respectivement D_1 et D_2 . La v.a.r. $S = X + Y$ est discrète dont l'ensemble des valeurs possibles est $D = \{x + y, (x, y) \in D_1 \times D_2\}$ de loi est donnée par ¹ :

$$\forall s \in D, \mathbb{P}(S = s) = \sum_{x \in D_1} \mathbb{P}(X = x) \mathbb{P}(Y = s - x) = \sum_{y \in D_2} \mathbb{P}(Y = y) \mathbb{P}(X = s - y).$$

1. Cette formule est appelée la **convolution discrète** des lois de X et Y .

Exemple 6.1 (Stabilité de la loi de Poisson).

■ Soient X et Y deux v.a.r. indépendantes telles que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$. Alors $S \sim \mathcal{P}(\lambda + \mu)$.

- $D = S(\Omega) = \mathbb{N}$. Pour $n \in S(\Omega)$, on a

$$\mathbb{P}(S = n) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k) \mathbb{P}(Y = n - k) = \sum_{k=0}^n e^{-\lambda} \frac{\lambda^k}{k!} e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} = e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \lambda^k \mu^{n-k} = e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^n}{n!}.$$

- D'où $\forall n \in \mathbb{N}, \mathbb{P}(S = n) = e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^n}{n!}$ i.e. $S \sim \mathcal{P}(\lambda + \mu)$.

Proposition 6.2.**somme d'une v.a.r.d. et d'une v.a.r. à densité indépendantes**

Soient X une v.a.r.d. dont l'ensemble des valeurs possibles est D et Y une v.a.r. à densité de densité f_Y telles que X et Y sont indépendantes.

La v.a.r. $S = X + Y$ est à densité de densité f_S est donnée par :

$$\forall t \in \mathbb{R}, f_S(t) = \sum_{x \in D} \mathbb{P}(X = x) f_Y(t - x).$$
Proposition 6.3.**somme de v.a.r. à densité indépendantes**

Soient X et Y deux v.a.r. à densité indépendantes de densités respectives f_X et f_Y .

La v.a.r. $S = X + Y$ est à densité de densité f_S est donnée par ¹ :

$$\forall t \in \mathbb{R}, f_S(t) = \int_{-\infty}^{+\infty} f_X(x) f_Y(t - x) dx.$$

1. Cette formule est appelée la **convolution des densités** f_X et f_Y .

7 Espérance, variance et moments

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les variables aléatoires réelles seront définies sur cet espace.

Espérance

Définition 7.1.

espérance d'une v.a.r.

1. Soit X une v.a.r. de loi discrète dont l'ensemble des valeurs possibles est D . On dit que X admet une **espérance** si la famille $(x\mathbb{P}(X = x))_{x \in D}$ est sommable. Dans ce cas, on appelle **espérance de X** le réel

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

2. Soit X une v.a.r. de loi à densité de densité f_X . On dit que X admet une **espérance** si l'application $t \mapsto tf_X(t)$ est intégrable sur $\mathbb{R}$. Dans ce cas, on appelle **espérance de X** le réel

$$\mathbb{E}(X) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} tf_X(t)dt.$$

Exemple 7.1 (Espérance des variables aléatoires discrètes usuelles).

- $X \sim \mathcal{B}(p)$. On a $X(\Omega) = \{0, 1\}$ est fini donc X admet une espérance et

$$\mathbb{E}(X) = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) = p.$$

- $X \sim \mathcal{B}(n, p)$. On a $X(\Omega) = [0, n]$ est fini donc X admet une espérance et

$$\mathbb{E}(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = np \underbrace{\sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}}_{=1} = np.$$

- $X \sim \mathcal{G}(p)$. On a $X(\Omega) = \mathbb{N}^*$. La série $\sum_{k \geq 1} k\mathbb{P}(X = k) = p \sum_{k \geq 1} k(1-p)^{k-1}$ est une série dérivée géométrique de raison $1-p$ convergente car $|1-p| < 1$ donc X admet une espérance et

$$\mathbb{E}(X) = p \sum_{k=1}^{+\infty} k(1-p)^{k-1} = p \underbrace{\left(\frac{1}{1-x} \right)' \Big|_{x=1-p}}_{=1/p^2} = \frac{1}{p}.$$

- $X \sim \mathcal{P}(\lambda)$. On a $X(\Omega) = \mathbb{N}$. La série $\sum_{k \geq 0} k\mathbb{P}(X = k) = e^{-\lambda} \sum_{k \geq 0} k \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k \geq 0} \frac{\lambda^k}{k!}$ est une série exponentielle convergente donc X admet une espérance et

$$\mathbb{E}(X) = e^{-\lambda} \sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \underbrace{\sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{=e^{\lambda}} = \lambda.$$

Exemple 7.2 (Espérance des variables aléatoires à densité usuelles).

- $X \sim \mathcal{U}([a, b])$. La fonction $t \mapsto tf_X(t) = \frac{t}{b-a} 1_{[a, b]}(t)$ est intégrable sur $\mathbb{R}$ donc X admet une espérance et

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \frac{t}{b-a} 1_{[a, b]}(t) dt = \int_a^b \frac{t}{b-a} dt = \frac{a+b}{2}.$$

- $X \sim \mathcal{E}(\lambda)$. La fonction $t \mapsto tf_X(t) = \lambda t e^{-\lambda t} 1_{[0, +\infty[}(t)$ est intégrable sur $\mathbb{R}$ (car $t^2 tf_X(t) \xrightarrow[t \rightarrow +\infty]{} 0$) donc X admet une espérance et

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \lambda t e^{-\lambda t} 1_{[0,+\infty[}(t) dt = \int_0^{+\infty} \lambda t e^{-\lambda t} dt \underset{s=\lambda t}{=} \underbrace{\frac{1}{\lambda} \int_0^{+\infty} s e^{-s} ds}_{=\Gamma(2)=1} = \frac{1}{\lambda}.$$

■ $X \sim \Gamma(\alpha, \lambda)$. La fonction $t \mapsto t f_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^\alpha e^{-\lambda t} 1_{]0,+\infty[}(t)$ est intégrable sur $\mathbb{R}$ (car $t^2 \cdot t f_X(t) \xrightarrow[t \rightarrow +\infty]{} 0$) donc X admet une espérance et

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^\alpha e^{-\lambda t} 1_{]0,+\infty[}(t) dt = \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^\alpha e^{-\lambda t} dt \underset{s=\lambda t}{=} \frac{1}{\lambda \Gamma(\alpha)} \underbrace{\int_0^{+\infty} s^\alpha e^{-s} ds}_{=\Gamma(\alpha+1)} = \frac{\Gamma(\alpha+1)}{\lambda \Gamma(\alpha)} = \frac{\alpha}{\lambda}.$$

■ $X \sim \mathcal{N}(\mu, \sigma^2)$. La fonction $t \mapsto t f_X(t) = \frac{t}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$ est intégrable sur $\mathbb{R}$ (car $t^2 \cdot t f_X(t) \xrightarrow[t \rightarrow \pm\infty]{} 0$) donc X admet une espérance et

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \frac{t}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \underset{s=(t-\mu)/\sqrt{2}\sigma}{=} \frac{1}{\sqrt{\pi}} \left(\underbrace{\mu \int_{-\infty}^{+\infty} e^{-s^2} ds}_{=\sqrt{\pi}} + \underbrace{\sqrt{2}\sigma \int_{-\infty}^{+\infty} s e^{-s^2} ds}_{=0} \right) = \mu.$$

Théorème 7.1.**théorème de transfert**

Soit X une v.a.r. et $g: \mathbb{R} \rightarrow \mathbb{R}$ une application de sorte que $Y = g(X)$ soit une v.a.r.

1. Supposons X est de loi discrète dont l'ensemble des valeurs possibles est D . On a équivalence entre :

- (i) Y admet une espérance ;
- (ii) $(g(x)\mathbb{P}(X=x))_{x \in D}$ sommable.

Dans ce cas, $\mathbb{E}(Y) = \sum_{x \in D} g(x)\mathbb{P}(X=x).$

2. Supposons X est de loi à densité de densité f_X . On a équivalence entre :

- (i) Y admet une espérance ;
- (ii) $t \mapsto g(t)f_X(t)$ intégrable sur $\mathbb{R}$.

Dans ce cas, $\mathbb{E}(Y) = \int_{-\infty}^{+\infty} g(t)f_X(t)dt.$

Théorème 7.2.**théorème de transfert d'un couple de v.a.r.d.**

Soient X et Y deux v.a.r.d. dont l'ensemble des valeurs possibles respectivement D_1 et D_2 . Soit $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ une application de sorte que $Z = g(X, Y)$ soit une v.a.r. Alors Z admet une espérance si, et seulement si, la famille $(g(x, y)\mathbb{P}(X=x, Y=y))_{(x,y) \in D_1 \times D_2}$ est sommable. Dans ce cas, $\mathbb{E}(Z) = \sum_{(x,y) \in D_1 \times D_2} g(x, y)\mathbb{P}(X=x, Y=y).$

Proposition 7.1.**propriétés de l'espérance**

Soient X et Y deux v.a.r.

1. Linéarité : si X et Y admettent une espérance, alors pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda X + \mu Y$ aussi. De plus

$$\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$$

2. Supposons X admet une espérance.

a. Positivité : $X \geq 0 \implies \mathbb{E}(X) \geq 0.$

b. Si $X \geq 0$, alors $\mathbb{E}(X) = 0 \implies X = 0$ p.s.

3. Croissance : si X et Y admettent une espérance, alors $X \leq Y \implies \mathbb{E}(X) \leq \mathbb{E}(Y).$

4. Inégalité triangulaire : X admet une espérance si, et seulement si $|X|$ aussi. De plus $|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$

5. Domination : si $|X| \leq Y$ et si Y admet une espérance, alors X aussi.

Remarque 7.1. ■ En particulier, si $|X|$ est bornée, alors X admet une espérance.

Théorème 7.3.**espérance de variables aléatoires indépendantes**

Si X et Y sont deux v.a.r. indépendantes admettent une espérance, alors ¹ XY aussi et $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

1. La réciproque est fausse.

Définition 7.2.**v.a.r. centrée**

Une v.a.r. X qui admet une espérance est dite **centrée** si $\mathbb{E}(X) = 0$.

Remarque 7.2. ■ Si X admet une espérance, alors $X - \mathbb{E}(X)$ est centrée.

Moments et variance**Définition 7.3.****moments d'ordre k**

Soit X une v.a.r. et $k \in \mathbb{N}^*$. On dit que X admet un **moment d'ordre k** si X^k admet une espérance. Dans ce cas, $\mathbb{E}(X^k)$ s'appelle le **moment d'ordre k** de X . Ainsi :

$$\mathbb{E}(X^k) = \begin{cases} \sum_{x \in D} x^k \mathbb{P}(X = x) & \text{si } X \text{ est de loi discrète dont l'ensemble des valeurs possibles est } D. \\ \int_{-\infty}^{+\infty} t^k f_X(t) dt & \text{si } X \text{ est de loi à densité de densité } f_X. \end{cases}$$

Proposition 7.2.

Soit X une v.a.r. admet un moment d'ordre $k \in \mathbb{N}^*$.

1. X admet un moment d'ordre j pour tout $j \in \llbracket 1, k \rrbracket$. 2. $X + \alpha$ admet un moment d'ordre k pour tout $\alpha \in \mathbb{R}$.

Théorème 7.4.**inégalité de Cauchy-Schwartz**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2. Alors XY admet une espérance et

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

avec égalité si, et seulement si $X = 0$ ou $\exists \lambda \in \mathbb{R}, Y = \lambda X$ p.s.

Définition 7.4.**variance, écart-type**

Soit X une v.a.r. admet un moment d'ordre 2.

1. On appelle **variance** de X le réel positif $\mathbb{V}(X) \stackrel{\text{déf}}{=} \mathbb{E}((X - \mathbb{E}(X))^2)$.

2. On appelle **écart-type** de X le réel $\sigma(X) \stackrel{\text{déf}}{=} \sqrt{\mathbb{V}(X)}$.

Proposition 7.3.**propriétés de variance**

Soit X une v.a.r. admet une variance et $(a, b) \in \mathbb{R}^2$.

1. $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. (König-Huygens)

2. $\mathbb{V}(X) = 0 \iff X$ est constante p.s.

3. $aX + b$ admet une variance et

$$\mathbb{V}(aX + b) = a^2 \mathbb{V}(X) \text{ et } \sigma(aX + b) = |a| \sigma(X).$$

Définition 7.5.**v.a.r. centrée réduite**

Soit X une v.a.r. admet une variance.

Si $\mathbb{V}(X) = 1$, on dit que X est **réduite**. Si $\mathbb{E}(X) = 0$ et $\mathbb{V}(X) = 1$, on dit que X est **centrée réduite**.

Remarque 7.3. ■ Si X admet une variance > 0 alors $\frac{X}{\sigma(X)}$ est réduite et $X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

Définition 7.6.**covariance**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2. On définit la **covariance** du couple (X, Y) par

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

Proposition 7.4.**propriétés de covariance**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2.

1. $\mathbb{C}(X, Y) = \mathbb{C}(Y, X)$ et $\mathbb{C}(X, X) = \mathbb{V}(X)$.
2. $\mathbb{C}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
3. Si X et Y sont indépendantes, alors ¹ $\mathbb{C}(X, Y) = 0$.
4. $|\mathbb{C}(X, Y)| \leq \sigma(X)\sigma(Y)$.

1. La réciproque est fausse.

Proposition 7.5.**variance de la somme**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2.

1. $X + Y$ admet une variance et $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathbb{C}(X, Y)$.
2. Si X et Y sont indépendantes, alors $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

Définition 7.7.**coefficient de corrélation linéaire**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2 telles que $\mathbb{V}(X) > 0$ et $\mathbb{V}(Y) > 0$. On définit le **coefficient de corrélation linéaire** du couple (X, Y) par

$$\rho(X, Y) \stackrel{\text{déf}}{=} \frac{\mathbb{C}(X, Y)}{\sigma(X)\sigma(Y)}.$$
Proposition 7.6.**propriétés du coefficient de corrélation linéaire**

Soient X et Y deux v.a.r. admettent un moment d'ordre 2 telles que $\mathbb{V}(X) > 0$ et $\mathbb{V}(Y) > 0$.

1. $\rho(X, Y) \in [-1, 1]$.
2. $\rho(X, Y) = 1 \iff \exists \alpha > 0, \beta \in \mathbb{R}, Y = \alpha X + \beta \text{ p.s.}$
3. $\rho(X, Y) = -1 \iff \exists \alpha < 0, \beta \in \mathbb{R}, Y = \alpha X + \beta \text{ p.s.}$
4. Si X et Y sont indépendantes, alors ¹ $\rho(X, Y) = 0$.

1. La réciproque est fausse.

8 Fonction génératrice d'une v.a.r. à valeurs dans $\mathbb{N}$

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les v.a.r. seront définies sur cet espace à valeurs dans $\mathbb{N}$.

Définition 8.1.**fonction génératrice**

On appelle **fonction génératrice** de X la fonction

$$G_X(t) \stackrel{\text{déf}}{=} \mathbb{E}(t^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n.$$

Proposition 8.1.**propriétés de G_X**

1. G_X est au moins définie sur $[-1, 1]$
2. $G_X(1) = 1$ et $\forall t \in [-1, 1], |G_X(t)| \leq 1$.
3. La série entière $\sum_{n \geq 0} \mathbb{P}(X = n)t^n$ CN sur $[-1, 1]$.
4. G_X est continue sur $[-1, 1]$.
5. G_X est de classe $\mathcal{C}^\infty$ sur $] -1, 1[$.
6. $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}$.

Remarque 8.1. ■ G_X caractérise la loi de probabilité. Autrement dit, X et Y ont même loi $\iff G_X = G_Y$.

Exemple 8.1 (Fonction génératrice des variables aléatoires discrètes usuelles).

- $X \sim \mathcal{B}(p)$. On a $X(\Omega) = \{0, 1\}$ est fini donc la série $\sum_{k \geq 0} \mathbb{P}(X = k)t^k$ converge pour tout $t \in \mathbb{R}$ et

$$\forall t \in \mathbb{R}, G_X(t) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)t^k = \mathbb{P}(X = 0)t^0 + \mathbb{P}(X = 1)t^1 = (1 - p) + pt.$$

- $X \sim \mathcal{B}(n, p)$. On a $X(\Omega) = [0, n]$ est fini donc la série $\sum_{k \geq 0} \mathbb{P}(X = k)t^k$ converge pour tout $t \in \mathbb{R}$ et

$$\forall t \in \mathbb{R}, G_X(t) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)t^k = \sum_{k=0}^n \binom{n}{k} (tp)^k (1-p)^{n-k} = ((1-p) + pt)^n.$$

- $X \sim \mathcal{G}(p)$. La série $\sum_{k \geq 1} \mathbb{P}(X = k)t^k = \frac{p}{1-p} \sum_{k \geq 1} ((1-p)t)^k$ converge si, et seulement si, $|(1-p)t| < 1$ (série géométrique) et

$$\forall t \in \left] \frac{-1}{1-p}, \frac{1}{1-p} \right[, G_X(t) = \sum_{k=1}^{+\infty} \mathbb{P}(X = k)t^k = \frac{p}{1-p} \sum_{k=1}^{+\infty} ((1-p)t)^k = \frac{p}{1-p} \frac{(1-p)t}{1 - (1-p)t} = \frac{pt}{1 - (1-p)t}.$$

- $X \sim \mathcal{P}(\lambda)$. La série $\sum_{k \geq 0} \mathbb{P}(X = k)t^k = e^{-\lambda} \sum_{k \geq 0} \frac{(\lambda t)^k}{k!}$ est une série exponentielle convergente pour tout $t \in \mathbb{R}$ et

$$\forall t \in \mathbb{R}, G_X(t) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)t^k = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}.$$

Proposition 8.2.**fonction génératrice de la somme**

Si $X_1, \dots, X_k$ sont indépendantes, alors $G_{X_1 + \dots + X_k} = \prod_{i=1}^k G_{X_i}$.

Remarque 8.2 (Somme de Bernoulli indépendantes).

- Soient $X_1, \dots, X_n$ des v.a.r. indépendantes suivant la même loi de Bernoulli de paramètre $p \in [0, 1]$. Alors $S = X_1 + \dots + X_n$ suit la loi binomiale de paramètres n et p . En effet :

$$\forall t \in \mathbb{R}, G_S(t) = \prod_{k=1}^n G_{X_k}(t) = (G_{X_1}(t))^n = ((1-p) + pt)^n.$$

Donc $S \sim \mathcal{B}(n, p)$.

Théorème 8.1.**espérance et variance**

1. X admet une espérance si, et seulement si, G_X est dérivable en 1. Dans ce cas $\mathbb{E}(X) = G'_X(1)$.
2. X admet une variance si, et seulement si, G_X est deux fois dérivable en 1. Dans ce cas

$$G''_X(1) = \mathbb{E}(X^2) - \mathbb{E}(X) \quad \text{et} \quad \mathbb{V}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2.$$

9 Inégalités, notions de convergence et théorèmes limites

$(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé. Toutes les variables aléatoires réelles seront définies sur cet espace.

Inégalités classiques

Théorème 9.1.

Inégalité de Markov

Soit X une v.a.r. **positive** admettant une espérance. $\forall \alpha > 0, \mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}(X)}{\alpha}$.

Théorème 9.2.

Inégalité de Bienyamé-Tchebychev

Soit X une v.a.r. admettant un moment d'ordre 2. $\forall \varepsilon > 0, \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$.

Théorème 9.3.

Inégalité de Jensen

Soit X est une v.a.r. admettant une espérance et $f: \mathbb{R} \rightarrow \mathbb{R}$ convexe. Si $f(X)$ admet une espérance, alors

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

Convergence en loi, convergence en probabilité

Définition 9.1.

convergence en probabilité

Soient $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. et X une v.a.r. On dit que $(X_n)_{n \in \mathbb{N}}$ **converge en probabilité** vers X si

$$\forall \varepsilon > 0, \mathbb{P}(|X_n - X| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0.$$

On écrit alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X$.

Exemple 9.1. ■ Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. indépendantes telles que $\mathbb{P}(X_n = -1) = \mathbb{P}(X_n = 1) = \frac{1}{2}$. Alors

$$\frac{1}{n} \sum_{j=1}^n X_j \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

- Posons $S_n = \frac{1}{n} \sum_{j=1}^n X_j$. On a $\mathbb{E}(S_n) = 0$ et $\mathbb{V}(S_n) = \frac{1}{n} \mathbb{V}(X_1) = \frac{1}{n} \underbrace{\mathbb{E}(X_1^2)}_{=1} - \frac{1}{n} \underbrace{\mathbb{E}(X_1)^2}_{=0} = \frac{1}{n}$.
- D'après l'inégalité de Bienyamé-Tchebychev, $\forall \varepsilon > 0, \mathbb{P}(|S_n| \geq \varepsilon) \leq \frac{1}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$.

D'où $S_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0$.

Proposition 9.1.

Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions qui converge simplement vers $f: \mathbb{R} \rightarrow \mathbb{R}$ et X une v.a.r. de sorte que $f_n(X)$ et $f(X)$ reste des v.a.r. Alors $f_n(X) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} f(X)$.

Définition 9.2.**convergence en loi**

Soient $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. et X une v.a.r. On dit que $(X_n)_{n \in \mathbb{N}}$ **converge en loi** vers X si ¹

$$\forall t \in \mathbb{R} \setminus D_X, \quad F_{X_n}(t) \xrightarrow{n \rightarrow +\infty} F_X(t).$$

où D_X est l'ensemble des points de discontinuité de F_X . On écrit alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$.

1. Les v.a.r. X_n et X peuvent être définies sur des espaces distincts, puisque seules leurs lois sont en cause.

Proposition 9.2.**convergence en loi pour une v.a.r. discrète**

Soient X et $(X_n)_{n \in \mathbb{N}}$ des v.a.r. à valeurs dans $\mathbb{N}$. On a équivalence entre :

$$(i) \quad X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X; \quad (ii) \quad \forall k \in \mathbb{N}, \quad \mathbb{P}(X_n = k) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(X = k).$$

Application 9.1 (Approximation de la loi binomiale par la loi de Poisson).

Soit $(p_n)_{n \in \mathbb{N}}$ une suite de réels de $[0, 1]$ telle que $np_n \xrightarrow{n \rightarrow +\infty} \lambda > 0$ et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. suivant toutes la loi binomiale de paramètre (n, p_n) .

La suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une v.a.r. qui suit la loi de poisson de paramètre λ .

Soit $k \in \llbracket 0, n \rrbracket$. On a

$$\mathbb{P}(X_n = k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{n(n-1) \dots (n-k+1)}{k!} p_n^k (1 - p_n)^{n-k} \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!} p_n^k (1 - p_n)^{n-k} \underset{n \rightarrow +\infty}{\sim} \frac{\lambda^k}{k!} (1 - p_n)^{n-k}.$$

$$\text{Or } (1 - p_n)^{n-k} = \exp((n-k) \ln(1 - p_n)) \xrightarrow{n \rightarrow +\infty} e^{-\lambda}. \text{ D'où } \mathbb{P}(X_n = k) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Proposition 9.3.**la cv en probabilité implique la cv en loi**

Soient X et $(X_n)_{n \in \mathbb{N}}$ des v.a.r. $X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$. (la réciproque est fausse)

Théorèmes limites**Théorème 9.4.****loi faible des grands nombres (LFGN)**

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. indépendantes et de même loi, admettant un moment d'ordre 2.

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

Théorème 9.5.**théorème de la limite centrée**

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. indépendantes et de même loi ¹, admettant un moment d'ordre 2.

On pose $\mu = \mathbb{E}(X_1)$, $\sigma = \sigma(X_1)$. On a $\frac{1}{\sigma\sqrt{n}} \left(\sum_{k=1}^n X_k - n\mu \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ où $X \sim \mathcal{N}(0, 1)$.

1. On n'a pas besoin de connaître la loi des X_n .

Lois discrètes usuelles				
Variable aléatoire	Loi	Espérance	Variance	Fonction génératrice
$X \sim \mathcal{B}(p)$	$X(\Omega) = \{0, 1\}$ $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$	p	$p(1 - p)$	$\forall t \in \mathbb{R}, G_X(t) = (1 - p) + pt$
$X \sim \mathcal{B}(n, p)$	$X(\Omega) = \llbracket 0, n \rrbracket$ $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$	$\forall t \in \mathbb{R}, G_X(t) = ((1 - p) + pt)^n$
$X \sim \mathcal{G}(p)$	$X(\Omega) = \mathbb{N}^*$ $\mathbb{P}(X = k) = p(1 - p)^{k-1}$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$	$\forall t \in \left] \frac{-1}{1 - p}, \frac{1}{1 - p} \right[, G_X(t) = \frac{pt}{1 - (1 - p)t}$
$X \sim \mathcal{P}(\lambda)$	$X(\Omega) = \mathbb{N}$ $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ	$\forall t \in \mathbb{R}, G_X(t) = e^{\lambda(t-1)}$
Lois à densité usuelles				
Variable aléatoire	Densité	Espérance	Variance	Fonction de répartition
$X \sim \mathcal{U}([a, b])$	$f_X(t) = \begin{cases} \frac{1}{b-a} & \text{si } t \in [a, b] \\ 0 & \text{si } t \notin [a, b] \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$F_X(t) = \begin{cases} 0 & \text{si } t \leq a \\ \frac{t-a}{b-a} & \text{si } t \in [a, b] \\ 1 & \text{si } t \geq b \end{cases}$
$X \sim \mathcal{E}(\lambda)$	$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$F_X(t) = \begin{cases} 1 - e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}$
$X \sim \Gamma(\alpha, \lambda)$	$f_X(t) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	pas d'expression explicite seule la définition !
$X \sim \mathcal{N}(\mu, \sigma^2)$	$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$	μ	σ^2	pas d'expression explicite seule la définition !