

Chapitre 13

Variables aléatoires

Dans ce chapitre, Ω est un ensemble appelé **univers** et $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une tribu sur Ω . Les éléments de \mathcal{A} sont appelés **événements**.

On désigne aussi par \mathbb{P} une probabilité sur (Ω, \mathcal{A}) .

13.1 Variables aléatoires

13.1.1 Définition

Définition 1 (Variable aléatoire)

On appelle **variable aléatoire discrète** sur (Ω, \mathcal{A}) toute application X définie sur Ω et vérifiant :

- $X(\Omega)$ est **au plus dénombrable** (i.e. fini ou dénombrable)
- Pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\}) = \{\omega \in \Omega, X(\omega) = x\}$ est un **événement**, c'est-à-dire un élément de la tribu \mathcal{A} .

Le **support** de la variable aléatoire X est l'ensemble des valeurs prises par X :

$$X(\Omega) = \{X(\omega), \omega \in \Omega\} \subset E.$$

Si de plus $X(\Omega) \subset \mathbb{R}$, on dira qu'il s'agit d'une **variable aléatoire réelle discrète** (v.a.r.d.).

En général, l'univers Ω n'est pas explicité, on connaît seulement $X(\Omega)$.

Proposition 1 (Événements associés)

Soit X une variable aléatoire discrète sur (Ω, \mathcal{A}) et à valeurs dans un ensemble E .

Pour toute partie F de E , l'ensemble $X^{-1}(F) = \{\omega \in \Omega, X(\omega) \in F\}$ est un événement (i.e. un élément de \mathcal{A}).

On le note $(X \in F)$.

Preuve.

□

Proposition 2

Soit X une variable aléatoire discrète sur (Ω, \mathcal{A}) et à valeurs dans un ensemble E .

- Pour $x \in E$, on note $(X = x)$ (parfois $\{X = x\}$) l'événement $X^{-1}(\{x\})$.
- On a $(X \in \emptyset) = \emptyset$ et $(X \in E) = \Omega$.
- Pour toute partie F de E , on a $(X \in \overline{F}) = \overline{(X \in F)}$.
- Pour toute famille $(F_i)_{i \in I}$ de E , on a :

$$\left(X \in \bigcup_{i \in I} F_i \right) = \bigcup_{i \in I} (X \in F_i) \quad \text{et} \quad \left(X \in \bigcap_{i \in I} F_i \right) = \bigcap_{i \in I} (X \in F_i).$$

- Les événements $(X = x)_{x \in X(\Omega)}$ forment un système complet d'événements.

Preuve.

□

Notations si la variable aléatoire X est réelle :

$X^{-1}(\{1\})$	$(X = 1)$
$X^{-1}([a, b])$	$(a \leq X \leq b)$
$X^{-1}([a, +\infty[)$	$(a \leq X)$
$X^{-1}(] - \infty, a])$	$(X \leq a)$

Exemple 13.1. On lance deux dés à 6 faces une fois. La variable aléatoire S associée à un résultat la somme des deux valeurs obtenues. Le support de S est donc $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

On a par exemple (si les dés ne sont pas truqués) :

- L'événement $(S = 2)$ est aussi l'ensemble $\{(1, 1)\}$ et donc :

$$\mathbb{P}(S = 2) = \mathbb{P}(\{(1, 1)\}) = \frac{1}{36}$$

- L'événement $(S \leq 3)$ est aussi l'ensemble $\{(1, 1), (1, 2), (2, 1)\}$ et donc :

$$\mathbb{P}(S \leq 3) = \mathbb{P}(\{(1, 1), (1, 2), (2, 1)\}) = \frac{3}{36} = \frac{1}{12}$$

- L'événement $(S \leq 1)$ est aussi l'ensemble \emptyset et donc :

$$\mathbb{P}(S \leq 1) = \mathbb{P}(\emptyset) = 0$$

13.1.2 Loi d'une variable aléatoire

Définition 2 (Loi de probabilité)

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans un ensemble E .
On appelle **loi de probabilité** de X l'application :

$$\mathbb{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ F & \longmapsto \mathbb{P}(X \in F) \end{cases}$$

On peut démontrer qu'il s'agit d'une probabilité sur $X(\Omega)$ muni de la tribu $\mathcal{P}(X(\Omega))$.

Proposition 3

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans un ensemble E . La loi de probabilité de X est complètement déterminée par la donnée de :

- son support $X(\Omega)$ (au plus dénombrable),
- la valeur de $\mathbb{P}(X = x) = \mathbb{P}_X(\{x\})$ pour tout élément x de $X(\Omega)$.

Preuve.

□

Remarque : avec ces notations, la famille $(\mathbb{P}(X = x))_{x \in X(\Omega)}$ est une **distribution de probabilité** car :

$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x) \geq 0 \quad \text{et} \quad \sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \mathbb{P}\left(\bigcup_{x \in X(\Omega)} (X = x)\right) = \mathbb{P}(\Omega) = 1.$$

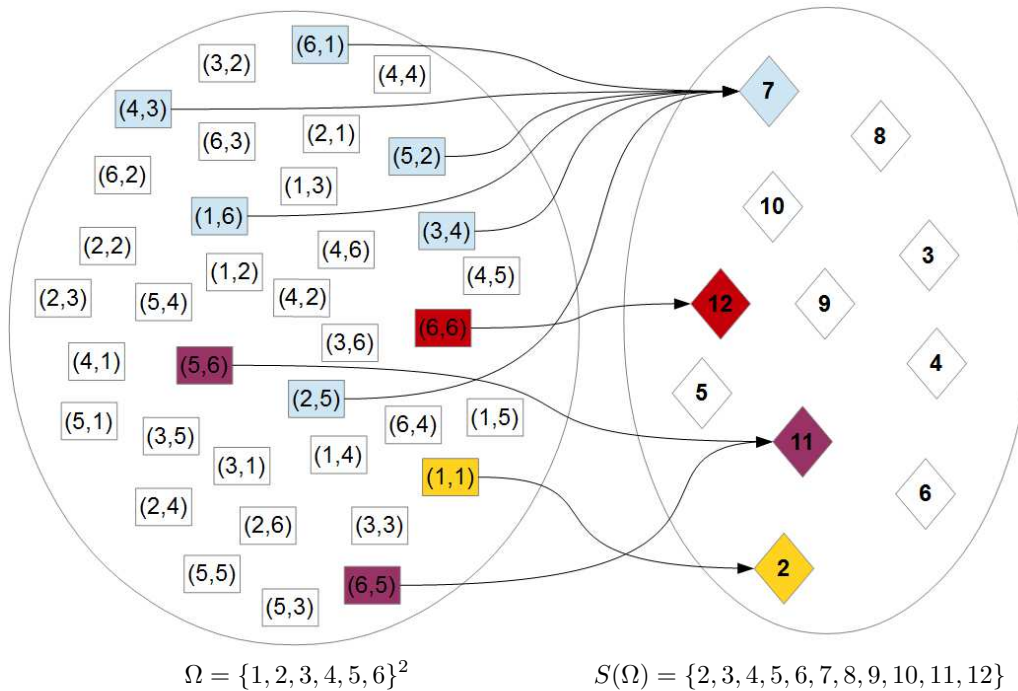
Exemple 13.2. On reprend l'exemple du lancer de deux dés, et de la variable aléatoire S .

On a déjà vu que $S(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

$$\begin{aligned} \mathbb{P}(S = 2) = \mathbb{P}_S(\{2\}) &= \mathbb{P}(\{(1, 1)\}) & \mathbb{P}(S = 3) = \mathbb{P}_S(\{3\}) &= \mathbb{P}(\{(1, 2), (2, 1)\}) \\ &= \frac{1}{36} & &= \frac{2}{36} = \frac{1}{18} \\ \mathbb{P}(S = 4) = \mathbb{P}_S(\{4\}) &= \mathbb{P}(\{(1, 3), (2, 2), (3, 1)\}) & \mathbb{P}(S = 5) = \mathbb{P}_S(\{5\}) &= \mathbb{P}(\{(1, 4), (2, 3), (3, 2), (4, 1)\}) \\ &= \frac{3}{36} = \frac{1}{12} & &= \frac{4}{36} = \frac{1}{9} \\ \mathbb{P}(S = 6) = \mathbb{P}_S(\{6\}) &= \mathbb{P}(\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}) & \mathbb{P}(S = 7) = \mathbb{P}_S(\{7\}) &= \mathbb{P}(\{(1, 6), (2, 5), (3, 4), (4, 3), \\ &= \frac{5}{36} & &= \frac{6}{36} = \frac{1}{6} \\ & & & \quad \quad \quad (5, 2), (6, 1)\} \\ \mathbb{P}(S = 8) = \mathbb{P}_S(\{8\}) &= \mathbb{P}(\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}) & \mathbb{P}(S = 9) = \mathbb{P}_S(\{9\}) &= \mathbb{P}(\{(3, 6), (4, 5), (5, 4), (6, 3)\}) \\ &= \frac{5}{36} & &= \frac{4}{36} = \frac{1}{9} \\ \mathbb{P}(S = 10) = \mathbb{P}_S(\{10\}) &= \mathbb{P}(\{(4, 6), (5, 5), (6, 4)\}) & \mathbb{P}(S = 11) = \mathbb{P}_S(\{11\}) &= \mathbb{P}(\{(5, 6), (6, 5)\}) \\ &= \frac{3}{36} = \frac{1}{12} & &= \frac{2}{36} = \frac{1}{18} \\ \mathbb{P}(S = 12) = \mathbb{P}_S(\{12\}) &= \mathbb{P}(\{(6, 6)\}) \\ &= \frac{1}{36} \end{aligned}$$

Interprétation : La variable aléatoire X permet de transférer la probabilité \mathbb{P} sur (Ω, \mathcal{A}) en une probabilité \mathbb{P}_X sur $(X(\Omega), \mathcal{P}(X(\Omega)))$.

Dans l'exemple précédent, le couple de valeurs obtenues ne nous intéresse pas vraiment, on s'intéresse uniquement leur somme.



Convention : On peut étendre la notation $\mathbb{P}_X(\{x\})$ à **tous** les éléments de E :

- si $x \in X(\Omega)$ alors $\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x)$ (c'est la définition de \mathbb{P}_X),
- si $x \notin X(\Omega)$ alors $\mathbb{P}_X(\{x\}) = 0$.

On peut alors donner la définition suivante :

Définition 3

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans un même ensemble E .
 On dit que X et Y suivent la même loi si (avec la convention précédente) $\mathbb{P}_X = \mathbb{P}_Y$.
 On note $X \sim Y$.

Ainsi, $X \sim Y$ si pour tout $z \in E$, on a $\mathbb{P}(X = z) = \mathbb{P}(Y = z)$ si $z \in X(\Omega) \cap Y(\Omega)$, et $\mathbb{P}(X = z) = \mathbb{P}(Y = z) = 0$ sinon.

On admet les propositions suivantes.

Proposition 4 (Fonction de variable aléatoire)

Soit X une variable aléatoire discrète sur (Ω, \mathcal{A}) et à valeurs dans un ensemble E , et $f : E \rightarrow F$ une application. Alors l'application

$$f(X) : \begin{cases} \Omega & \rightarrow & F \\ \omega & \mapsto & f(X(\omega)) \end{cases}$$

est une variable aléatoire discrète sur (Ω, \mathcal{A}) et à valeur dans F .

On dit que $f(X)$ est une fonction de la variable aléatoire X .

Proposition 5

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans un même ensemble E , et $f : E \rightarrow F$ une application. On a :

$$X \sim Y \quad \implies \quad f(X) \sim f(Y)$$

Exemple 13.3. On lance un dé équilibré à 6 faces et on note X la variable aléatoire égale au résultat obtenu. On considère $f(x) = (x - 2)(x - 4)$. Déterminer la loi de $Y = f(X)$.

13.1.3 Exemples importants

Exemple 13.4. Loi uniforme sur E

Définition 4 (Loi uniforme)

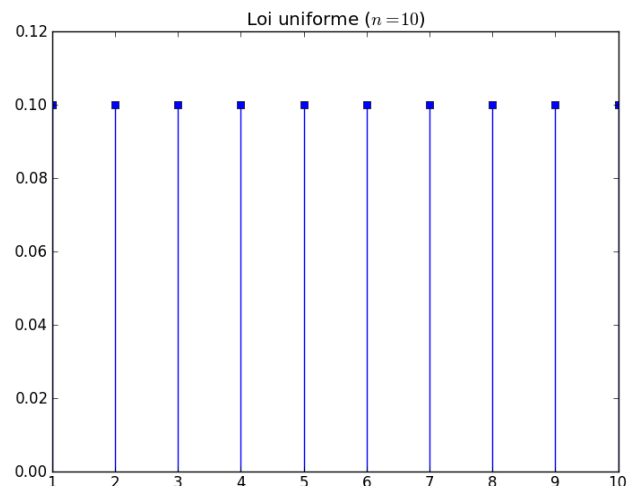
Soit $n \in \mathbb{N}^*$ et soit E un ensemble fini de cardinal n . On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ suit la **loi uniforme** sur E si :

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

On écrit alors $X \sim \mathcal{U}(E)$.

En particulier, quand $E = \llbracket 1, n \rrbracket$, et si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$ alors :

- $X(\Omega) = \llbracket 1, n \rrbracket$,
- $\forall k \in \llbracket 1, n \rrbracket, \quad \mathbb{P}(X = k) = \frac{1}{n}$.



Exemple 13.5. Loi de Bernoulli

Cette loi apparaît dans toute expérience « binaire », c'est-à-dire de type échec/succès comme le « pile ou face ».

On a alors deux résultats possibles (que l'on peut envoyer sur 0 et 1), le premier de probabilité $1 - p$, l'autre de probabilité p .

Définition 5 (Loi de Bernoulli)

Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit la **loi de Bernoulli de paramètre p** si

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

On écrit alors $X \sim \mathcal{B}(p)$.

Exemple 13.6. Loi binomiale

On reprend l'expérience de l'exemple précédent, et on la répète n fois (avec $n \in \mathbb{N}^*$). On observe donc une suite d'échecs (ou de 0) et de succès (ou de 1) **indépendants**.

On définit la variable aléatoire X comme étant le nombre de succès (ou de 1) observés.

On a $X(\Omega) = \llbracket 0, n \rrbracket$. Par exemple, calculons $\mathbb{P}(X = 1)$.

Calculons $\mathbb{P}(X = k)$ pour $k \in \llbracket 0, n \rrbracket$.

Le nombre de façons de positionner les k succès parmi les n expériences successives est $\binom{n}{k}$.

Pour chacun de ces résultats, la probabilité est $p^k(1-p)^{n-k}$. Et donc $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Définition 6

Soient $n \in \mathbb{N}^*$ et $p \in]0, 1[$.

On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit la **loi binomiale de paramètres n et p** si

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

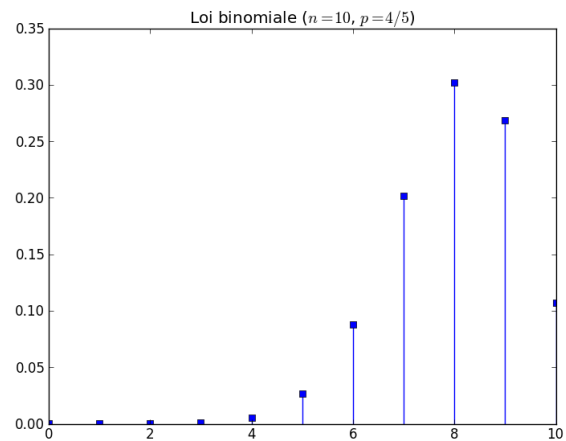
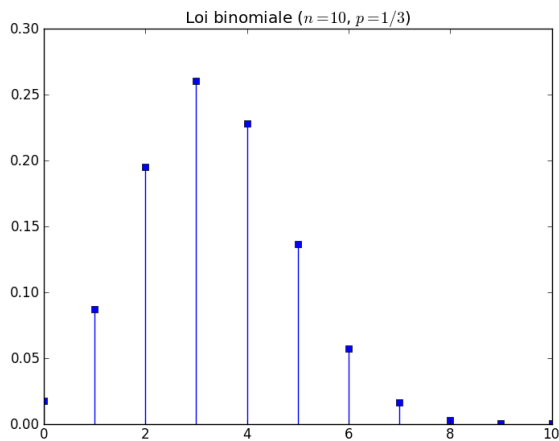
On écrit alors $X \sim \mathcal{B}(n, p)$.

Vérification : On définit bien une distribution de probabilité, puisque pour tout $k \in \llbracket 0, n \rrbracket$, on a $\mathbb{P}(X = k) \geq 0$ et :

$$\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

Exemples de situations concrètes :

- X = Nombre de fois où on obtient un boule blanche au cours de n tirages avec remise.
- X = Nombre de fois où on obtient 5 au cours de n lancers de dé.
- X = Nombre de bonnes réponses données dans un QCM de type Vrai/Faux (réponses aléatoires).



Exemple 13.7. Loi géométrique

On procède à une succession d'expériences identiques du type succès/échec jusqu'à obtenir un succès.

Les expériences sont indépendantes et on note $p \in]0, 1[$ la probabilité d'obtenir un succès.

La variable aléatoire est ici le **temps d'attente** noté T : c'est la variable qui à une répétition infinie de telles expériences, associe le rang où on observe le premier succès.

$$(T = k) = \{ \underbrace{(0, 0, \dots, 0)}_{k-1 \text{ échecs}}, 1 \}.$$

Et donc $T(\Omega) = \mathbb{N}^*$ et $\forall k \in \mathbb{N}_*, \mathbb{P}(T = k) = (1-p)^{k-1} p$.

Par convention, le résultat où l'on observe infiniment des échecs est noté ∞ . On a déjà vu, en utilisant la définition de probabilité, ou le théorème de continuité croissante/décroissante, que l'on a nécessairement $\mathbb{P}(T = \infty) = 0$.

Définition 7 (Loi géométrique)

Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit la **loi géométrique de paramètre p** si

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

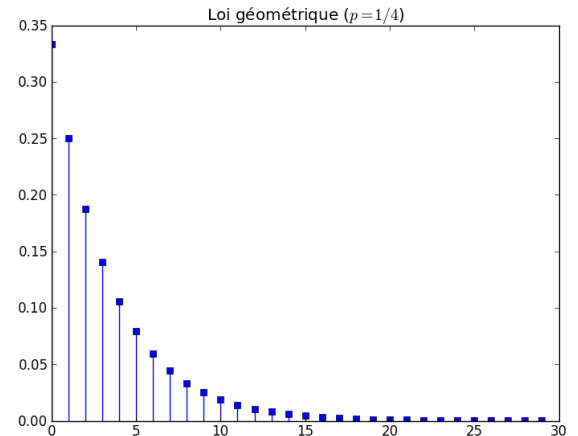
On écrit alors $X \sim \mathcal{G}(p)$.

Vérification : On définit bien une distribution de probabilité, puisque pour tout $k \in \mathbb{N}^*$, on a $\mathbb{P}(X = k) \geq 0$ et puisque :

$$\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \sum_{k=1}^{+\infty} (1 - p)^{k-1}p = \frac{p}{1 - (1 - p)} = 1.$$

Exemples de situations concrètes :

- X = nombre de lancers d'un dés jusqu'à ce qu'on obtienne 6.
- X = nombre de tirs de fléchettes nécessaires pour tirer dans le rond central (50 points).



Le résultat suivant est à retenir et à savoir démontrer rapidement.

Exercice de colle (E1)

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ suivant la loi géométrique de paramètre $p \in]0, 1[$. Démontrer que pour tout $n \in \mathbb{N}$, on a :

$$\mathbb{P}(X > n) = (1 - p)^n.$$

Exemple 13.8. On suppose que $X \sim \mathcal{G}(p)$ avec $p \in]0, 1[$. On rappelle que pour tout $n \in \mathbb{N}$, $\mathbb{P}(X > n) = (1 - p)^n$. Montrons que X est une loi sans mémoire, c'est-à-dire que :

$$\forall (k, n) \in (\mathbb{N}^*)^2, \quad \mathbb{P}_{(X > k)}(X > k + n) = \mathbb{P}(X > n)$$

Exercice de colle (E2)

Réciproquement, soit X une variable aléatoire discrète telle que $X(\Omega) \subset \mathbb{N}^*$ et telle que :

$$\forall (k, n) \in (\mathbb{N}^*)^2, \quad \mathbb{P}_{(X > k)}(X > k + n) = \mathbb{P}(X > n)$$

Montrer que X suit une loi géométrique dont on précisera le paramètre.

Exemple 13.9. Loi de Poisson

Définition 8

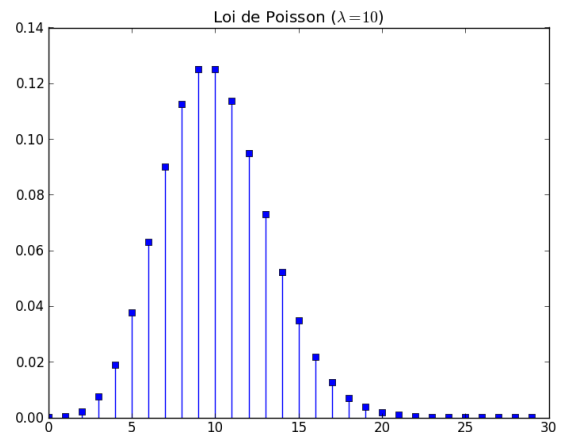
Soit $\lambda \in]0, +\infty[$. On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit la **loi de Poisson de paramètre λ** si

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

On écrit alors $X \sim \mathcal{P}(\lambda)$.

Vérification : On définit bien une distribution de probabilité, puisque pour tout $k \in \mathbb{N}$, on a $\mathbb{P}(X = k) \geq 0$ et :

$$\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda} e^{-\lambda} = 1.$$



Interprétation en termes d'événements rares : Contrairement aux lois précédemment définies, cette loi n'est, a priori, pas liée à une expérience type. Denis Poisson (1781-1840) l'introduit dans son ouvrage *Recherches sur la probabilité des jugements* en 1837. Il l'obtient comme « limite » de la loi binomiale $X_n \sim \mathcal{B}(n, p_n)$ avec $\lim_{n \rightarrow +\infty} np_n = \lambda$ strictement positif.

Exercice de colle (E2)

Pour $n \in \mathbb{N}$, on se donne $X_n \sim \mathcal{B}(n, p_n)$ avec $\lim_{n \rightarrow +\infty} np_n = \lambda > 0$. Démontrer que :

$$\forall k \in \mathbb{N}, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

En pratique : Soit X une variable aléatoire suivant une loi $\mathcal{B}(n, p)$.

Si p est petit et si n est grand, les valeurs numériques des probabilités $\mathbb{P}(X = k)$ sont très voisines de $e^{-np} \frac{(np)^k}{k!}$. Cela signifie que X se comporte approximativement comme si elle suivait une loi de Poisson de paramètre np .

Exemple 13.10. On lance 100 fois deux dés non truqués et on appelle X la variable aléatoire donnant le nombre de double six obtenus. La loi exacte de X est $\mathcal{B}(100, 1/36)$ qui conduit à des calculs compliqués.

On pose $\lambda = np = \frac{100}{36}$, on a le tableau suivant :

k	0	1	2	3	4	5	10
$\mathbb{P}(X = k)$ avec $X \sim \mathcal{B}(100, 1/36)$	0.05978	0.1708	0.24156	0.22546	0.15621	0.08579	0.00038
$\mathbb{P}(X = k)$ avec $X \sim \mathcal{P}(100/36)$	0.06218	0.17271	0.23988	0.22211	0.15424	0.08569	0.00047

La différence n'est pas très grande, surtout lorsque k est proche de la valeur moyenne

$$E(X) = np = \frac{100}{36} \approx 2,78.$$

En pratique : On considère qu'une très bonne approximation de la loi binomiale $\mathcal{B}(n, p)$ est la loi de Poisson $\mathcal{P}(np)$ si (les conditions varient selon les auteurs) :

- $n \geq 30, p \leq 0,1$ et $np < 15$
- $n \geq 20$ et $p \leq 0,05$
- $n \geq 20, p \leq 0,1$ et $np < 5$
- $n \geq 100$ et $np < 10$

D'une manière générale, il faut des assez grandes valeurs de n et des petites valeurs de p .

Exemples de situations concrètes :

- X = Nombre de défauts sur une pièce usinée.
- X = Nombre de clients entrant dans un magasin en une journée.
- X = Nombre d'erreurs de frappe sur une page d'un livre.
- X = Nombre de particules alpha émises par un matériau radioactif en une minute.

13.1.4 Loi conditionnelle de X sachant un événement A

Définition 9

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $A \in \mathcal{A}$.

On appelle loi probabilité conditionnelle de X sachant A , la loi de probabilité associée à X pour la probabilité conditionnelle \mathbb{P}_A (probabilité sachant A).

$$\forall F \subset X(\Omega), \quad (\mathbb{P}_A)_X(F) = \mathbb{P}_A(X \in F) = \mathbb{P}((X \in F)|A) = \frac{\mathbb{P}((X \in F) \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(X^{-1}(F) \cap A)}{\mathbb{P}(A)}.$$

13.1.5 Couples de variables aléatoires

Dans tout ce paragraphe, on considère des variables aléatoires discrètes sur (Ω, \mathcal{A}) .

$$X : \Omega \longrightarrow X(\Omega) \subset E \quad \text{et} \quad Y : \Omega \longrightarrow Y(\Omega) \subset F.$$

Définition 10

Si X et Y sont deux variables aléatoires discrètes sur (Ω, \mathcal{A}) à valeurs dans E et F respectivement, l'application :

$$(X, Y) : \begin{cases} \Omega & \longrightarrow & X(\Omega) \times Y(\Omega) \subset E \times F \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{cases}$$

est une variable aléatoire discrète appelée couple de variables aléatoires.

Définition 11

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, la loi de probabilité de (X, Y) est appelée **loi conjointe** de (X, Y) . Elle est complètement déterminée par la donnée de la valeur de :

$$\mathbb{P}\left((X, Y) = (x, y)\right) = \mathbb{P}(X = x, Y = y) \text{ pour tout élément } (x, y) \text{ de } X(\Omega) \times Y(\Omega).$$

Les lois de X et de Y sont appelées **lois marginales** du couple (X, Y) .

Exemple 13.11. Dans une urne contenant 6 boules blanches, 4 boules rouges et 2 boules noires, on tire 3 boules. On note X la variable aléatoire donnant le nombre de boules blanches et Y celle donnant le nombre de boules noires. $X(\Omega) = \llbracket 0, 3 \rrbracket$ et $Y(\Omega) = \llbracket 0, 2 \rrbracket$. Il y a $\binom{12}{3} = 220$ tirages possibles et :

$$\forall (i, j) \in \llbracket 0, 3 \rrbracket \times \llbracket 0, 2 \rrbracket, \quad \mathbb{P}(X = i, Y = j) = \frac{\binom{6}{i} \binom{2}{j} \binom{4}{3-i-j}}{220}.$$

On convient dans cette égalité que si $3 - i - j \notin \llbracket 0, 4 \rrbracket$ alors $\binom{4}{3-i-j} = 0$.

Par le calcul, on obtient la loi conjointe de (X, Y) .

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	Loi de Y
$Y = 0$		36/220	60/220	20/220	
$Y = 1$		48/220	30/220	0	
$Y = 2$			0	0	
Loi de X					1

Proposition 6

Avec les notations précédentes, si l'on connaît la loi conjointe de (X, Y) alors on connaît ses lois marginales, mais la réciproque est fautive.

Preuve.

□

Exemple 13.12. On construit deux couples de variables (X, Y) et (X', Y') à valeurs dans $\llbracket 0, 1 \rrbracket^2$ ayant mêmes lois marginales mais des lois conjointes distinctes.

- La première suit une loi uniforme :

$$\forall (i, j) \in \llbracket 0, 1 \rrbracket^2, \mathbb{P}((X, Y) = (i, j)) = \frac{1}{4}.$$

On a alors :

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2} \text{ et } \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \frac{1}{2}.$$

	$X = 0$	$X = 1$	Loi de Y
$Y = 0$	1/4	1/4	
$Y = 1$	1/4	1/4	
Loi de X			1

- La seconde est donnée par :

$$\mathbb{P}((X', Y') = (0, 0)) = \frac{1}{6}, \mathbb{P}((X', Y') = (0, 1)) = \frac{1}{3},$$

$$\mathbb{P}((X', Y') = (1, 0)) = \frac{1}{3} \text{ et } \mathbb{P}((X', Y') = (1, 1)) = \frac{1}{6}.$$

On a encore :

$$\mathbb{P}(X' = 0) = \mathbb{P}(X' = 1) = \frac{1}{2} \text{ et } \mathbb{P}(Y' = 0) = \mathbb{P}(Y' = 1) = \frac{1}{2}.$$

	$X' = 0$	$X' = 1$	Loi de Y'
$Y' = 0$	1/6	1/3	
$Y' = 1$	1/3	1/6	
Loi de X'			1

Par conséquent, les lois marginales d'un couple de variables aléatoires ne suffisent pas pour définir sa loi conjointe.

Définition 12 (Lois conditionnelles)

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.

• Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$ on appelle loi conditionnelle de X sachant $(Y = y)$ la loi de probabilité définie sur $(X(\Omega), \mathcal{P}(X(\Omega)))$ par :

$$\begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ \{x\} & \longmapsto \mathbb{P}_{(Y=y)}(X = x) = \mathbb{P}(X = x | Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \end{cases}$$

• Pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$ on appelle loi conditionnelle de Y sachant $(X = x)$ la loi de probabilité définie sur $(Y(\Omega), \mathcal{P}(Y(\Omega)))$ par

$$\begin{cases} \mathcal{P}(Y(\Omega)) & \longrightarrow [0, 1] \\ \{y\} & \longmapsto \mathbb{P}_{(X=x)}(Y = y) = \mathbb{P}(Y = y | X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)} \end{cases}$$

Exemple 13.13. On reprend l'exemple de l'urne contenant 6 boules blanches, 4 boules rouges et 2 boules noires, dans laquelle on tire 3 boules.

Déterminer la loi conditionnelle de X sachant $(Y = 1)$.

Généralisation au cas de n variables aléatoires.

Définition 13

Soient X_1, \dots, X_n sont des variables aléatoires discrètes sur (Ω, \mathcal{A}) , à valeurs dans E_1, \dots, E_n respectivement. Alors l'application :

$$(X_1, \dots, X_n) : \begin{cases} \Omega & \longrightarrow X_1(\Omega) \times \dots \times X_n(\Omega) \subset E_1 \times \dots \times E_n \\ \omega & \longmapsto (X_1(\omega), \dots, X_n(\omega)) \end{cases}$$

est une variable aléatoire discrète appelée vecteur de variables aléatoires.

Définition 14

Si X_1, \dots, X_n sont des variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, la loi de probabilité de (X_1, \dots, X_n) est appelée **loi conjointe** de (X_1, \dots, X_n) . Elle est complètement déterminée par la donnée de la valeur de :

$$\mathbb{P}\left((X_1, \dots, X_n) = (x_1, \dots, x_n)\right) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \text{ pour tout élément } (x_1, \dots, x_n) \text{ de } X_1(\Omega) \times \dots \times X_n(\Omega).$$

Les lois de X_1, \dots, X_n sont appelées **lois marginales** du vecteur (X_1, \dots, X_n) .

13.1.6 Variables aléatoires indépendantes

Définition 15 (Indépendance de deux variables aléatoires)

Deux variables aléatoires discrètes X et Y définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ sont dites **indépendantes** si, pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants, ce qui s'écrit :

$$\forall (A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), \quad \mathbb{P}\left((X \in A) \cap (Y \in B)\right) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

On note dans ce cas $X \perp\!\!\!\perp Y$.

On admet la proposition suivante.

Proposition 7

Soient X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors :

$$X \perp\!\!\!\perp Y \iff \left(\forall (x, y) \in X(\Omega) \times Y(\Omega), \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)P(Y = y) \right).$$

Exercice de colle (E2)

Soit (X, Y) un couple de variables aléatoires à valeurs dans \mathbb{N}^2 tel que :

$$\forall (p, q) \in \mathbb{N}^2, \mathbb{P}((X, Y) = (p, q)) = \mathbb{P}(X = p, Y = q) = \lambda \frac{p+q}{p!q!2^{p+q}}.$$

Déterminer λ , calculer les lois marginales puis montrer que les variables aléatoires X et Y sont indépendantes.

Exemple 13.14. Soient X, Y deux variables aléatoires indépendantes tels que $X \sim \mathcal{U}(\{-1, 1\})$ et $Y \sim \mathcal{U}(\{-1, 1\})$. On pose $U = X + Y$ et $V = X - Y$. Les variables aléatoires U et V sont-elles indépendantes ?

On généralise la définition d'indépendance à plusieurs variables aléatoires. On pourrait démontrer les propositions par récurrence.

Définition 16 (Indépendance de n variables aléatoires)

Soit X_1, \dots, X_n sont n variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, on dit que (X_1, \dots, X_n) est une famille de variables indépendantes si :

$$\forall (A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \dots \times \mathcal{P}(X_n(\Omega)), \quad \mathbb{P}\left(\bigcap_{i=1}^n (X_i \in A_i)\right) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i).$$

Proposition 8

Si X_1, \dots, X_n sont n variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors (X_1, \dots, X_n) est une famille de variables indépendantes si et seulement si :

$$\forall (x_1, \dots, x_n) \in \prod_{i=1}^n X_i(\Omega), \quad \mathbb{P}((X_1, \dots, X_n) = (x_1, \dots, x_n)) = \prod_{i=1}^n \mathbb{P}(X_i = x_i).$$

Il n'y a pas de notation officielle au programme pour écrire que (X_1, \dots, X_n) est une famille de variables indépendantes.

Définition 17 (Indépendance d'une suite de variables aléatoires)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que ces variables aléatoires sont indépendantes, si toute sous-famille finie de termes de $(X_n)_{n \in \mathbb{N}}$ est une famille de variables aléatoires indépendantes. Pour cela, il faut et il suffit de vérifier que pour tout $n \in \mathbb{N}$, les variables aléatoires X_0, \dots, X_n sont indépendantes.

Définition 18 (Suite i.i.d.)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que ces variables aléatoires forment une suite **indépendante et identiquement distribuée** (on note i.i.d.) si l'on a :

- Pour tout $(i, j) \in \mathbb{N}^2$, on a $X_i \sim X_j$ (les X_n suivent toutes la même loi),
- $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes.

Exemple 13.15. Par exemple, la répétition indépendante d'une même expérience de type échec/succès comme le jeu de pile ou face infini, donne une suite de variables aléatoires i.i.d. de Bernoulli.

Fonction de variables aléatoires :

On rappelle que si $X : \Omega \rightarrow X(\Omega) \subset E$ est une variable aléatoire discrète et si $f : E \rightarrow F$ est une application, la variable aléatoire $f(X)$ est :

$$f(X) : \begin{cases} \Omega & \rightarrow & F \\ \omega & \mapsto & f(X(\omega)) \end{cases}$$

On admet alors les proposition suivante.

Proposition 9

- Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans E et F , et f, g deux application définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement. On a l'implication :

$$X \perp\!\!\!\perp Y \quad \implies \quad f(X) \perp\!\!\!\perp g(Y).$$

- Ce résultat s'étend à n variables aléatoires :

$$(X_1, \dots, X_n) \text{ indépendantes} \quad \implies \quad (f_1(X_1), \dots, f_n(X_n)) \text{ indépendantes,}$$

où f_1, \dots, f_n sont des fonctions définies sur $X_1(\Omega), \dots, X_n(\Omega)$ respectivement.

Proposition 10 (Lemme des coalitions)

Si X_1, \dots, X_n sont n variables aléatoires discrètes indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans E , alors pour toutes fonctions f et g définies sur $\prod_{i=1}^p X_i(\Omega)$ et $\prod_{i=p+1}^n X_i(\Omega)$ respectivement, on a :

$$f(X_1, \dots, X_p) \perp\!\!\!\perp g(X_{p+1}, \dots, X_n)$$

Remarque : On a ici deux coalitions (deux fonctions f et g), on admettra aussi que le résultat est encore valable pour plus de deux coalitions.

13.1.7 Somme de variables aléatoires indépendantes

L'énoncé suivant n'est pas au programme. Il faudra systématiquement justifier les calculs.

Soit X_1 et X_2 deux variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ et **indépendantes**. Alors $X = X_1 + X_2$ est une variable aléatoire réelle discrète et

$$\forall x \in X(\Omega), \quad \mathbb{P}(X_1 + X_2 = x) = \sum_{x_1 \in X_1(\Omega)} \mathbb{P}(X_1 = x_1) \mathbb{P}(X_2 = x - x_1).$$

En effet : Si pose $X = X_1 + X_2$, on a évidemment $X(\Omega) \subset \mathbb{R}$. On a :

$$X(\Omega) = \{x_1 + x_2, x_1 \in X_1(\Omega) \text{ et } x_2 \in X_2(\Omega)\} \subset \mathbb{R}.$$

Par la proposition 4, l'application $X = f(X_1, X_2) = X_1 + X_2 : \Omega \rightarrow X(\Omega)$ est une variable aléatoire (réelle) discrète.

Déterminons sa loi : pour $x \in X(\Omega)$, on calcule $\mathbb{P}(X = x)$.

Proposition 11 (Somme de lois binomiales)

Soient X_1, \dots, X_q des variables aléatoires (mutuellement) indépendantes suivant des lois binomiales telles que :

$$X_1 \sim \mathcal{B}(n_1, p), X_2 \sim \mathcal{B}(n_2, p), \dots, X_q \sim \mathcal{B}(n_q, p).$$

Alors la variable aléatoire $X_1 + \dots + X_q$ suit une loi binomiale : $\sum_{i=1}^q X_i \sim \mathcal{B}(n_1 + \dots + n_q, p)$.

Preuve. (E2) On le montre dans le cas où $q = 2$. Le cas général s'obtiendrait ensuite par récurrence sur q .

Soient $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$ deux variables indépendantes. On a $X_1(\Omega) = \llbracket 0, n_1 \rrbracket$ et $X_2(\Omega) = \llbracket 0, n_2 \rrbracket$, donc $(X_1 + X_2)(\Omega) = \llbracket 0, n_1 + n_2 \rrbracket$.

On écrit la formule des probabilités totales avec le système complet d'événements, $(X_1 = i)_{i \in \{0, \dots, n_1\}}$. On obtient pour $k \in \llbracket 0, n_1 + n_2 \rrbracket$, on a :

$$\begin{aligned} \mathbb{P}(X_1 + X_2 = k) &= \sum_{i=0}^{n_1} \mathbb{P}\left((X_1 + X_2 = k) \cap (X_1 = i)\right) = \sum_{i=0}^{n_1} \mathbb{P}\left((X_2 = k - i) \cap (X_1 = i)\right) \\ &\hspace{15em} X_1, X_2 \text{ sont indépendantes} \\ &= \sum_{i=0}^{n_1} \mathbb{P}(X_1 = i) \underbrace{\mathbb{P}(X_2 = k - i)}_{=0 \text{ si } k-i < 0} = \sum_{i=0}^k \mathbb{P}(X_1 = i) \mathbb{P}(X_2 = k - i) \\ &= \sum_{i=0}^k \binom{n_1}{i} p^i (1-p)^{n_1-i} \binom{n_2}{k-i} p^{k-i} (1-p)^{n_2-(k-i)} \\ &= \sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k-i} p^k (1-p)^{n_1+n_2-k} \\ &= p^k (1-p)^{n_1+n_2-k} \sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k-i} \end{aligned}$$

On rappelle l'identité de Vandermonde, démontrée dans le chapitre « Probabilités ».

Lemme (Formule de Van der Monde)

Pour tout $n, p \in \mathbb{N}$ et tout $k \in \llbracket 0, n + p \rrbracket$ entier naturel, avec la convention $\binom{N}{q} = 0$ si $q \notin \llbracket 0, N \rrbracket$, on a :

$$\sum_{i=0}^k \binom{n}{i} \binom{p}{k-i} = \binom{n+p}{k},$$

En reportant dans les égalités précédentes avec $n = n_1$ et $p = n_2$, on obtient bien que

$$\mathbb{P}(X_1 + X_2 = k) = p^k (1-p)^{n_1+n_2-k} \binom{n_1 + n_2}{k}$$

et donc que $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$. □

Exercice de colle (E1 - Somme de lois de Poisson)

Soient X_1, \dots, X_q des variables aléatoires (mutuellement) indépendantes suivant des lois de Poissons telles que :

$$X_1 \sim \mathcal{P}(\lambda_1), X_2 \sim \mathcal{P}(\lambda_2), \dots, X_q \sim \mathcal{P}(\lambda_q).$$

Alors la variable aléatoire $X_1 + \dots + X_q$ suit une loi de Poisson : $\sum_{i=1}^q X_i \sim \mathcal{P}(\lambda_1 + \dots + \lambda_q)$.

On le montre dans le cas où $q = 2$. Le cas général s'obtiendrait ensuite par récurrence sur q .

13.2 Moments d'une variable aléatoire réelle

13.2.1 Espérance

On rappelle la définition vue en première année (cas fini).

Définition 19

Soit X une variable aléatoire réelle sur (Ω, \mathbb{P}) un espace probabilisé fini. On note $X(\Omega) = \{x_1, \dots, x_N\}$.
On appelle **espérance** de X la somme

$$\mathbb{E}(X) = \sum_{n=1}^N x_n \mathbb{P}(X = x_n).$$

Interprétation : L'espérance est une « moyenne probabiliste ».

Prenons par exemple, l'ensemble des notes d'un devoir d'une classe. Supposons que

2 élèves ont 10/20	3 élèves ont 7/20	2 élèves ont 6/20
4 élèves ont 12/20	1 élève a 16/20	3 élèves ont 9/20

La moyenne de classe est

$$\begin{aligned} m &= \frac{1}{15}(2 \times 10 + 3 \times 7 + 4 \times 12 + 1 \times 16 + 2 \times 6 + 3 \times 9) \\ &= \frac{2}{15} \times 10 + \frac{3}{15} \times 7 + \frac{4}{15} \times 12 + \frac{1}{15} \times 16 + \frac{2}{15} \times 6 + \frac{3}{15} \times 9 \end{aligned}$$

Les fractions correspondent aux proportions d'élèves ayant une note donnée. Dans le calcul d'une espérance, cette proportion est remplacée par la probabilité qu'un élève ait cette note. L'espérance est donc la valeur moyenne d'une variable aléatoire que l'on peut espérer obtenir.

On en donnera aussi en fin de chapitre une interprétation naturelle grâce à la loi des grands nombres.

On généralise cette notion aux variables aléatoires réelles discrètes. On utilise pour cela, la notion de famille sommable.

Cas d'une variable aléatoire positive

Définition 20

Soit X une variable aléatoire réelle discrète **positive** sur (Ω, \mathcal{A}, P) . L'ensemble $X(\Omega)$ est contenu dans \mathbb{R}^+ et la famille $\{x\mathbb{P}(X = x), x \in X(\Omega)\}$ est une famille au plus dénombrable d'éléments positifs.

On appelle **espérance** de X : $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \in [0, +\infty]$.

On dit que X est d'espérance finie si $\mathbb{E}(X) < +\infty$, c'est-à-dire si la famille $\{x\mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable.

Remarque : on peut étendre cette définition aux variables aléatoires discrètes à valeurs dans $[0, +\infty]$. On conviendra alors que dans ce cas :

$$\mathbb{P}(X = +\infty) = 0 \quad \text{et} \quad \ll +\infty\mathbb{P}(X = +\infty) \gg = 0.$$

En utilisant les notions introduites sur les familles sommables, on peut manipuler les sommes (finies ou dénombrables) de nombres positifs, la finitude de la somme prouvera la sommabilité.

Exemple 13.16. Soit $A \in \mathcal{A}$ un événement. On appelle fonction indicatrice de A l'application suivante :

$$\mathbf{1}_A : \begin{cases} \Omega & \longrightarrow & \{0, 1\} \\ \omega \in A & \longmapsto & 1 \\ \omega \notin A & \longmapsto & 0 \end{cases}$$

On a les égalités $(\mathbf{1}_A = 0) = \bar{A}$ et $(\mathbf{1}_A = 1) = A$. Ainsi, $\mathbf{1}_A$ est donc une variable aléatoire de Bernoulli de paramètre $p = \mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(A)$. La variable aléatoire $\mathbf{1}_A$ est à support fini $\{0, 1\}$, elle admet donc une espérance :

$$\mathbb{E}(\mathbf{1}_A) = 0 \cdot \mathbb{P}(\mathbf{1}_A = 0) + 1 \cdot \mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(A).$$

• **Espérance des lois usuelles :**

Proposition 12

- Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$ alors $\mathbb{E}(X) = \frac{n+1}{2}$
- Si $X \sim \mathcal{G}(p)$ alors $\mathbb{E}(X) = \frac{1}{p}$
- Si $X \sim \mathcal{B}(n, p)$ alors $\mathbb{E}(X) = np$
- Si $X \sim \mathcal{P}(\lambda)$ alors $\mathbb{E}(X) = \lambda$

Preuve.(D1)

□

Autre expression lorsque $X(\Omega) \subset \mathbb{N}$:

Proposition 13

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} .

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n) \in [0, +\infty].$$

Et en particulier, on a l'équivalence : $\mathbb{E}(X) < +\infty \iff \sum \mathbb{P}(X \geq n)$ converge

Preuve. Il s'agit du théorème de Fubini pour les sommes de familles à termes positifs.

□

On retiendra néanmoins le raisonnement suivant.

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} . Dans ce cas :

$$(X \geq n) = (n \leq X < n + 1) \cup (X \geq n + 1) = (X = n) \cup (X \geq n + 1).$$

Et comme l'union est disjointe, par σ -additivité :

$$\mathbb{P}(X = n) = \mathbb{P}(X \geq n) - \mathbb{P}(X \geq n + 1)$$

Sous les mêmes hypothèses, on aurait pu démontrer que :

- $\mathbb{P}(X = n) = \mathbb{P}(X > n - 1) - \mathbb{P}(X > n)$,
- $\mathbb{P}(X = n) = \mathbb{P}(X \leq n) - \mathbb{P}(X \leq n - 1)$,
- $\mathbb{P}(X = n) = \mathbb{P}(X < n + 1) - \mathbb{P}(X < n)$.

Exercice de colle (E3)

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} . On suppose que X est d'espérance finie. Démontrer que $\lim_{N \rightarrow +\infty} (N + 1)\mathbb{P}(X \geq N + 1) = 0$, et retrouver la deuxième expression de l'espérance de X .

Cas d'une variable aléatoire à valeurs réelles ou complexes

Définition 21

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs \mathbb{R} ou \mathbb{C} .

On dit que X est d'espérance finie si la famille $\{x\mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable. Dans ce cas, appelle espérance de X :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x).$$

Interprétation en distinguant les cas fini et dénombrable :

- Si $X(\Omega)$ est fini, alors X est d'espérance finie.
- Si $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ est dénombrable alors X est d'espérance finie si $\sum x_n\mathbb{P}(X = x_n)$ converge **absolument**.

Remarque : Si X est **bornée**, alors elle est d'espérance finie. Dans le cas où X est à support fini, c'est évident.

Supposons $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ est dénombrable. Comme X est bornée, il existe $M > 0$ tel que pour tout $n \in \mathbb{N}$, on ait $|x_n| \leq M$ et donc :

$$|x_n\mathbb{P}(X = x_n)| \leq M\mathbb{P}(X = x_n)$$

et par définition, la série $\sum \mathbb{P}(X = x_n)$ converge. Donc par les théorèmes de comparaison, $\sum |x_n\mathbb{P}(X = x_n)|$ converge également.

On peut regrouper les deux cas en utilisant la notion de famille sommable : Pour tout $x \in X(\Omega)$, on a :

$$|x\mathbb{P}(X = x)| \leq M\mathbb{P}(X = x).$$

Et comme la famille $\{\mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable par définition de probabilité, par comparaison, la famille $\{x\mathbb{P}(X = x), x \in X(\Omega)\}$ l'est aussi, donc X est d'espérance finie.

Définition 22

Une variable aléatoire réelle discrète X est dite **centrée** si elle est d'espérance finie et si $\mathbb{E}(X) = 0$.

Proposition 14 (Théorème du transfert)

Soit X une variable aléatoire réelle (ou complexe) discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $f : X(\Omega) \rightarrow \mathbb{R}$ ou \mathbb{C} une application. La variable aléatoire $f(X)$ est d'espérance finie si et seulement si la famille $\{f(x)\mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable. Dans ce cas, :

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

Preuve. Soit X une variable aléatoire réelle (ou complexe) discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $f : X(\Omega) \rightarrow \mathbb{R}$ ou \mathbb{C} une application, on note $Y = f(X)$.

$$\Omega \xrightarrow{X} X(\Omega) \xrightarrow{f} Y(\Omega)$$

• Pour tout $y \in Y(\Omega)$, on note $A_y = f^{-1}(\{y\}) = \{x \in X(\Omega), f(x) = y\}$.

Alors $X(\Omega) = \bigcup_{y \in Y(\Omega)} A_y$, et il s'agit d'une union disjointe ($f(x)$ ne peut pas valoir à la fois y et y' si $y \neq y'$) et dénombrable.

• D'autre part, pour $y \in Y(\Omega)$, on décrit l'événement $(Y = y)$ à l'aide A_y et de X :

$$\begin{aligned} \omega \in (Y = y) &\iff Y(\omega) = y \\ &\iff f(X(\omega)) = y \\ &\iff X(\omega) \in A_y \\ &\iff \omega \in (X \in A_y) \end{aligned}$$

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

• On travaille d'abord avec des sommes de termes positifs. Les calculs se font donc dans $[0, +\infty]$.

$$\begin{aligned} \sum_{x \in X(\Omega)} |f(x)|\mathbb{P}(X = x) &= \sum_{y \in Y(\Omega)} \sum_{x \in A_y} |f(x)|\mathbb{P}(X = x) \quad (\text{sommation par paquets}) \\ &= \sum_{y \in Y(\Omega)} \sum_{x \in A_y} |y|\mathbb{P}(X = x) \quad (\text{car pour } x \in A_y, \text{ on a } f(x) = y) \\ &= \sum_{y \in Y(\Omega)} |y| \sum_{x \in A_y} \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} |y|\mathbb{P}(Y = y) \quad (\text{d'après ce qui précède}) \end{aligned}$$

Ainsi, la famille $\{f(x)\mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable si et seulement si la famille $\{y\mathbb{P}(Y = y), y \in Y(\Omega)\}$ est sommable, c'est-à-dire si et seulement si Y est d'espérance finie. Et dans ce cas, les calculs précédents sont encore valables sans les valeurs absolues, ce qui donne :

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

□

Interprétation en distinguant les cas fini et dénombrable :

- Si $X(\Omega)$ est fini, alors $f(X)$ est d'espérance finie.
- Si $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ est dénombrable alors $f(X)$ est d'espérance finie si et seulement si $\sum f(x_n)\mathbb{P}(X = x_n)$ converge absolument.

Ce théorème permet de calculer l'espérance de $f(X)$ sans avoir à déterminer sa loi de probabilité, en utilisant uniquement celle de X .

Exemple 13.17. On suppose que X suit une loi uniforme sur $A = \{-2, -1, 0, 1, 2\}$. Calculer l'espérance de X^2 :

- en utilisant le théorème du transfert :

- en revenant à la définition :

Proposition 15 (Premières propriétés de l'espérance)

Soient X et Y deux variables aléatoires réelles (ou complexes) discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. On a les propriétés suivantes.

1. Linéarité :

- (a) Si $X = b$ est une variable aléatoire constante ($X(\Omega) = \{b\}$), alors elle est d'espérance finie et $\mathbb{E}(b) = \mathbb{E}(X) = b$.
- (b) Si X est d'espérance finie, alors pour tout $a, b \in \mathbb{R}$, $aX + b$ l'est aussi et

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

- (c) Si X et Y sont d'espérance finie, alors pour tout $a, b \in \mathbb{R}$, $aX + bY$ l'est aussi et

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

2. Comparaison : Si $|X| \leq Y$ et si Y est d'espérance finie, alors X l'est aussi.

3. Positivité : Si X est d'espérance finie et à valeurs positives, alors $\mathbb{E}(X) \geq 0$.

4. Croissance : Si X et Y sont d'espérance finie et si $X \leq Y$ ($\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$) alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

5. Inégalité triangulaire : Si X est d'espérance finie alors $|X|$ l'est aussi et on a $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

Preuve.

1(a) :

1(b) :

1(c) et 2 : admis.

3 :

4 :

5 :

□

Le point 1 de cette proposition a pour conséquence la proposition suivante.

Proposition 16

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un univers probabilisé.

L'ensemble $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ des variables aléatoires discrètes réelles ou complexes d'espérance finie est un espace vectoriel et l'espérance est une forme linéaire sur $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ (notation pas explicitement au programme).

Preuve.

□

Exemple 13.18. Si X est une variable aléatoire discrète réelle d'espérance finie, alors la variable aléatoire $X - E(X)$ est centrée.

Proposition 17 (Définie positivité)

Soient X est une variable aléatoire réelle discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$. On a l'implication suivante.

$$\left. \begin{array}{l} X \geq 0 \\ \text{et} \\ E(X) = 0 \end{array} \right\} \implies \mathbb{P}(X = 0) = 1 \quad (\text{l'événement } (X = 0) \text{ est presque sûr})$$

Preuve.

□

On admet enfin la proposition suivante.

Proposition 18 (Espérance du produit de deux variables indépendantes)

Si X et Y sont d'espérance finie et si elles sont **indépendantes**, alors la variable aléatoire XY est d'espérance finie et on a :

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

13.2.2 Variance

Pour commencer ce paragraphe, on définit une notion qui n'est pas explicitement au programme mais qui est souvent introduite en épreuve.

Définition 23

Soit X une variable aléatoire discrète réelle (ou complexe) sur $(\Omega, \mathcal{A}, \mathbb{P})$ et p un entier naturel non nul. On dit que X admet un moment d'ordre p si la variable aléatoire X^p est d'espérance finie, ou encore (par le théorème de transfert) si la famille $\{x^p \mathbb{P}(X = x), x \in X(\Omega)\}$ est sommable. Dans ce cas, on définit le moment d'ordre p de X par :

$$m_p(X) = \mathbb{E}(X^p) \stackrel{\text{Th. de transfert}}{=} \sum_{x \in X(\Omega)} x^p \mathbb{P}(X = x).$$

Exercice de colle (E2)

Soit X une variable aléatoire discrète réelle (ou complexe) sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Montrer que si X admet un moment d'ordre $r \geq 2$, alors elle admet un moment d'ordre s pour tout $s \in \{1, \dots, r\}$.

On vérifie d'abord que pour tout $x \in \mathbb{R}$, on a $|x|^{r-1} \leq 1 + |x|^r$.

Proposition 19

Soient X et Y deux variables aléatoires discrètes réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$.

- Si la variable aléatoire X^2 est d'espérance finie, alors X l'est aussi.
- **Inégalité de Cauchy-Schwarz** : Si X^2 et Y^2 sont d'espérance finie, alors XY l'est aussi et on a :

$$\left(\mathbb{E}(XY)\right)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2),$$

avec égalité si il existe $\lambda \in \mathbb{R}$ tel que $X = \lambda Y$ presque sûrement ou $Y = \lambda X$ presque sûrement.

Preuve. • Le premier point est l'exercice de colle précédent avec $r = 2$.

- Montrons le second point (**D3**).

□

La définition de variance est donnée par la proposition suivante.

Proposition 20 (Formule d'Huyghens-König)

Soit X une variable aléatoire discrète réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$. Si la variable aléatoire X^2 est d'espérance finie alors X l'est aussi et on appelle variance de X :

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0.$$

L'écart-type est donné par $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Preuve. Le cas fini a déjà été vu en première année. On s'intéresse ici au cas dénombrable, on note donc $X(\Omega) = \{x_n, n \in \mathbb{N}\}$. On remarque tout d'abord que pour tout $n \in \mathbb{N}$, on a $2|x_n| \leq 1 + x_n^2$ et donc

$$|x_n|P(X = x_n) \leq \frac{1}{2}P(X = x_n) + \frac{1}{2}x_n^2P(X = x_n).$$

Or par définition de probabilité, la série $\sum P(X = x_n)$ converge (elle a pour somme 1). De plus, puisque X^2 est d'espérance finie, par le théorème de transfert, la série $\sum_{n \in \mathbb{N}} x_n^2 P(X = x_n)$ converge (absolument, et elle a pour somme $\mathbb{E}(X^2)$).

Par comparaison, la série $\sum |x_n|P(X = x_n)$ converge, et donc X est d'espérance finie.

Puisque $(X - \mathbb{E}(X))^2 = X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2$, par linéarité, $(X - \mathbb{E}(X))^2$ est d'espérance finie.

On peut donc poser $V(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$, puisque cette espérance existe.

Il reste à vérifier l'égalité $\mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Par linéarité de l'espérance, on a ($\mathbb{E}(X) = b = \text{cst}$) :

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

Enfin, puisque, $(X - \mathbb{E}(X))^2 \geq 0$ on a $V(X) \geq 0$ (positivité de l'espérance) et donc on peut poser : $\sigma(X) = \sqrt{V(X)}$.

□

Interprétation : La variable aléatoire $(X - \mathbb{E}(X))^2$ mesure la distance au carré de X à sa valeur moyenne. L'espérance de cette variable aléatoire donne la valeur moyenne de cette distance au carré. Ainsi l'écart-type mesure la répartition de X autour de sa valeur moyenne. S'il est grand, X prend « souvent » des valeurs éloignées de $\mathbb{E}(X)$. S'il est petit, les valeurs de X sont davantage concentrées autour de $\mathbb{E}(X)$.

Pour calculer la variance de X , d'après la formule d'Huyghens-König, il suffit de déterminer ses **moments** d'ordre 1 et 2, c'est-à-dire $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$.

Définition 24

Avec les notations précédentes, on dit que X est réduite si $\sigma(X) = 1$.

Proposition 21 (Propriétés de la variance)

Soit X une variable aléatoire discrète réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ avec X^2 d'espérance finie. Soient a et b des réels. Alors $X + b$ et aX admettent une variance et

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

Preuve.

□

Exemple 13.19. Montrer que si X est une variable aléatoire telle que X^2 est d'espérance finie et telle que $\mathbb{V}(X) \neq 0$, alors $Y = \frac{X - E(X)}{\sigma(X)}$ est centrée et réduite.

Proposition 22 (Variance des lois usuelles)

- Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$ alors $\mathbb{V}(X) = \frac{n^2 - 1}{12}$
- Si $X \sim \mathcal{B}(n, p)$ alors $\mathbb{V}(X) = np(1 - p)$
- Si $X \sim \mathcal{G}(p)$ alors $\mathbb{V}(X) = \frac{1 - p}{p^2}$
- Si $X \sim \mathcal{P}(\lambda)$ alors $\mathbb{V}(X) = \lambda$

Preuve. (D1)

□

On sait que si X^2 et Y^2 sont d'espérance finie, alors XY l'est aussi. On définit alors la covariance de ces deux variables.

Définition 25 (Covariance)

Soient X, Y deux variables aléatoires discrètes réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ avec X^2 et Y^2 d'espérance finie. Alors $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$ admet une espérance finie, on l'appelle **covariance** de X et de Y .

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

Preuve.

□

Proposition 23

Soient X, Y et Z des variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ avec X^2, Y^2 et Z^2 d'espérance finie. On a les propriétés suivantes.

1. $\forall (a, b) \in \mathbb{R}^2, \quad \mathbb{V}(aX + bY) = a^2\mathbb{V}(X) + b^2\mathbb{V}(Y) + 2ab\text{Cov}(X, Y).$
2. $\text{Cov}(X, X) = \mathbb{V}(X).$
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ (*symétrie*)
4. $\forall (a, b) \in \mathbb{R}^2, \quad \text{Cov}(aX + bY, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z)$ (*linéarité à gauche*)
5. Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0$ (mais la réciproque est fausse).

Preuve.

□

Proposition 24 (Variance d'une somme finie de variables aléatoires)

Soit (X_1, \dots, X_n) des variables aléatoires discrètes réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ admettant un moment d'ordre 2 (X_i^2 d'espérance finie). On a :

- $$\mathbb{V}(X_1 + \dots + X_n) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j).$$

- Si X_1, \dots, X_n sont **deux-à-deux indépendantes**, c'est-à-dire que pour tous $i, j \in \llbracket 1, n \rrbracket$ distincts, X_i et X_j sont indépendantes, alors $\mathbb{V}(X_1 + \dots + X_n) = \mathbb{V}(X_1) + \dots + \mathbb{V}(X_n)$.

Preuve. On montre le premier point par récurrence sur n .

Initialisation : on a vu que $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{Cov}(X, Y)$ et donc le résultat est vrai pour $n = 2$.

Hérédité : Soit $n \geq 2$ un entier pour lequel la propriété est vérifiée. On se donne $(X_1, \dots, X_n, X_{n+1})$ des variables aléatoires admettant un moment d'ordre 2 et on pose $Y = X_1 + \dots + X_n$. On a

$$\mathbb{V}(X_1 + \dots + X_n + X_{n+1}) = \mathbb{V}(X_{n+1} + Y) =$$

Si X_1, \dots, X_n sont deux-à-deux indépendantes, alors pour tout i, j distincts on a $\text{Cov}(X_i, X_j) = 0$ d'où le second point. □

13.3 Convergence et approximations

13.3.1 Inégalités de Markov

Proposition 25 (Inégalités de Markov)

Soit X une variable aléatoire discrète réelle et **positive** sur $(\Omega, \mathcal{A}, \mathbb{P})$. On a les deux majorations suivantes.

- Première inégalité de Markov : Si X est d'espérance finie, alors $\forall t > 0, \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t},$
- Seconde inégalité de Markov : Si X^2 est d'espérance finie, alors $\forall t > 0, \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X^2)}{t^2}.$
- Cas général : Soit $p \in \mathbb{N}^*$. Si X^p est d'espérance finie, alors $\forall t > 0, \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X^p)}{t^p}.$

Remarque : Extrait du sujet de mathématiques CCP PSI 2018.

« **Q32.** Énoncer et démontrer l'inégalité de Markov pour une variable aléatoire finie Y sur $(\Omega, \mathcal{A}, \mathbb{P})$. Montrer que ce résultat est encore vrai lorsque Y est une variable aléatoire discrète non nécessairement finie. »

Preuve.

• (E1) Si X est d'espérance finie, alors :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} xP(X = x) =$$

Autre rédaction possible (E3) : en utilisant une fonction indicatrice. Soit $t > 0$, on note A l'événement $(X \geq t)$. On appelle fonction indicatrice de A l'application suivante :

$$\mathbf{1}_A : \begin{cases} \Omega & \longrightarrow \{0, 1\} \\ \omega \in A & \longmapsto 1 \\ \omega \notin A & \longmapsto 0 \end{cases}$$

On a donc en particulier $\mathbb{E}(\mathbf{1}_A) = 0 \cdot \mathbb{P}(\mathbf{1}_A = 0) + 1 \cdot \mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(A) = \mathbb{P}(X \geq t)$.

On remarque alors que : $t\mathbf{1}_A \leq X$ (inégalité de fonctions).

En effet :

Ainsi, par croissance de l'espérance, on obtient $\mathbb{E}(t\mathbf{1}_A) \leq \mathbb{E}(X)$.

Et par linéarité de l'espérance : $t\mathbb{E}(\mathbf{1}_A) = t\mathbb{P}(A) \leq \mathbb{E}(X)$.

En divisant par $t > 0$, on obtient bien

$$\mathbb{P}(A) = \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Remarque : Cette rédaction n'est pas fondamentalement différente de la précédente. Mais elle a l'avantage d'être valable également dans le cas de variables aléatoires continues, fréquemment étudiées en école d'ingénieur.

La notion de fonction indicatrice peut aussi être introduite dans certains sujets de concours.

• Si X^2 est d'espérance finie, en appliquant la première inégalité avec X^2 et t^2 , on obtient :

• Si X^p est d'espérance finie, en appliquant la première inégalité avec X^p et t^p , on obtient :

□

Remarque : Les inégalités de Markov donnent des majorations de $\mathbb{P}(X \geq t)$ avec des hypothèses assez faibles. Cependant, ces majorations ne sont pas très fines. Prenons par exemple, $X \sim \mathcal{G}(p)$. On a donc $X(\Omega) = \mathbb{N}^*$ et $\mathbb{P}(X = n) = p(1-p)^{n-1}$. On sait que X et X^2 sont d'espérance finie et que :

$$\bullet \mathbb{E}(X) = \frac{1}{p} = C_1(p) \qquad \bullet \mathbb{E}(X^2) = \mathbb{V}(X) + \mathbb{E}(X)^2 = \frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2} = C_2(p).$$

Les inégalités de Markov donnent (avec $t = n \in \mathbb{N}^*$) :

$$\forall n \in \mathbb{N}^*, \quad \mathbb{P}(X \geq n) \leq \frac{C_1(p)}{n} \quad \text{et} \quad \mathbb{P}(X \geq n) \leq \frac{C_2(p)}{n^2}.$$

Calculons $\mathbb{P}(X \geq n)$ pour $n \in \mathbb{N}$.

$$\mathbb{P}(X \geq n) = \sum_{k=n}^{+\infty} \mathbb{P}(X = k) = \sum_{k=n}^{+\infty} p(1-p)^{k-1} = p(1-p)^{n-1} \frac{1}{1-(1-p)} = (1-p)^{n-1} = e^{-a(n-1)}$$

avec $-a = \ln(1-p) < 0$.

Ainsi les majorations données par les inégalités de Markov ne sont pas optimales.

13.3.2 Inégalité de Bienaymé-Tchebychev

Proposition 26 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire discrète réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ avec X^2 d'espérance finie alors on a :

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$$

ou en passant à l'événement contraire : $\forall \varepsilon > 0, \quad 0 \leq 1 - \mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$

Preuve. (D2) On utilise la seconde inégalité de Markov appliquée à la variable aléatoire $|X - \mathbb{E}(X)|$ discrète et positive.

□

Interprétation : $\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon)$ est la probabilité que X soit « proche » de sa valeur moyenne, c'est-à-dire de son espérance. La deuxième majoration montre que lorsque ε est grand et/ou la variance $V(X)$ est petite, $\mathbb{P}(|X - E(X)| < \varepsilon)$ est proche de 1. Avec certains choix de ε (en général pas trop grand), on peut avoir $\mathbb{P}(|X - E(X)| < \varepsilon)$ aussi proche de 1 que l'on veut. Là encore, les hypothèses sont peu restrictives, mais la majoration obtenue n'est pas optimale.

13.3.3 Loi faible des grands nombres

On donne ici une nouvelle interprétation de l'espérance d'une variable aléatoire. L'idée est de répéter un grand nombre de fois une expérience aléatoire définissant une variable aléatoire X .

La loi faible des grand nombre traduit le fait que la moyenne de tous les résultats obtenus est « proche » de $\mathbb{E}(X)$.

Soient X_1, \dots, X_n soient des variables aléatoires deux-à-deux indépendantes obtenues en répétant n fois la même expérience, elles suivent donc toutes une même loi. On note $S_n = X_1 + \dots + X_n$, ainsi $\frac{S_n}{n}$ est la **moyenne** des résultats obtenus. Si

$\mathbb{E}(X_i) = m$, alors $\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_n)) = m$. Le résultat suivant est à savoir retrouver rapidement.

Exercice de colle (E2)

Soit (X_1, \dots, X_n) des variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ deux-à-deux indépendantes et suivant une même loi. On suppose que les X_i^2 sont d'espérance finie et on note

$$S_n = X_1 + \dots + X_n, \quad m = \mathbb{E}(X_1) \quad \text{et} \quad \sigma = \sigma(X_1)$$

Montrer que :

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Proposition 27 (Loi faible des grands nombres)

Soit $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires discrètes indépendantes et identiquement distribuées (i.i.d.) sur (Ω, \mathcal{A}, P) admettant un moment d'ordre 2. Alors, en notant

$$S_n = X_1 + \dots + X_n, \quad m = E(X_1) \text{ et } \sigma = \sigma(X_1)$$

on a la limite suivante :

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Preuve. Avec les hypothèses de l'énoncé, on peut utiliser l'exercice de colle précédent :

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Ainsi, par le théorème d'encadrement, on a bien $\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) = 0.$ □

Remarque : Le résultat est encore valable si l'on suppose que les X_i sont deux-à-deux indépendantes seulement.

Exemple 13.20. Une urne contient des boules blanches avec la proportion p mais on ne connaît pas cette proportion. Lors d'une expérience, on effectue n tirages avec remise et on obtient k boules blanches. Intuitivement, si n est grand, $\frac{k}{n}$ est proche de p . La loi faible des grands nombres traduit cette intuition.

On note $X_i \sim \mathcal{B}(p)$ donnant le résultat au i -ème tirage.

On a $E(X_i) = p = m$ et $V(X_i) = p(1-p)$ donc $\sigma = \sqrt{V(X)} = \sqrt{p(1-p)}$. Et donc, d'après l'inégalité encadrée ci-dessus :

$$\mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}.$$

On suppose ne pas connaître p , on voudrait savoir avec quelle erreur probabiliste on peut approximer p par $\frac{k}{n}$. On ne connaît pas $p(1-p)$, mais par une étude rapide de $x \mapsto x(1-x)$ (parabole tournée vers les $y < 0$), on montrerait que pour tout $p \in]0, 1[$, on a $p(1-p) \leq \frac{1}{4}$.

Prenons par exemple, $n = 1000$ et $\varepsilon = 0,1$. On a donc $\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| \geq 0,1 \right) \leq \frac{1}{4n\varepsilon^2} = \frac{0,1}{4} = 0,025$.

Ou encore, en passant à l'événement contraire $\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| < 0,1 \right) \geq 0,975$.

Ainsi si $n = 1000$ et si $S_{1000} = k$, la probabilité que p appartienne à l'**intervalle de confiance** $\left] \frac{k}{1000} - 0,1; \frac{k}{1000} + 0,1 \right[$ est d'au moins 0,975. Si par exemple, on a tiré 875 boules blanches sur 1000, alors avec 97,5% de chance, la proportion de boules blanches est comprise entre 0,775 et 0,975.

13.4 Fonction génératrice d'une variable aléatoire à valeurs dans \mathbb{N}

13.4.1 Définitions et premières propriétés

Définition 26

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} . On appelle fonction génératrice de X la fonction G_X définie par

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

Proposition 28

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} et G_X sa fonction génératrice. On a

- G_X est une série entière dont le rayon de convergence R est au moins égal à 1, il y a convergence normale de cette série sur $[-1, 1]$ et $G_X(1) = 1$.
- La fonction G_X est continue sur $[-1, 1]$ et de classe \mathcal{C}^∞ sur $] - 1, 1[$.
- Pour tout $n \in \mathbb{N}$, on a $\mathbb{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}$.

Et donc la loi de X est **complètement déterminée** par sa fonction génératrice G_X .

Preuve.

□

13.4.2 Fonctions génératrices des lois usuelles

Il faut savoir retrouver rapidement les fonction génératrices suivantes (à connaître!).

Exercice de colle (E1)

- Si $X \sim \mathcal{U}([1, n])$ alors $G_X(t) = \frac{1}{n}(t + t^2 + \dots + t^n)$ avec $R = +\infty$.
- Si $X \sim \mathcal{B}(n, p)$ alors $G_X(t) = (1 - p + pt)^n$ avec $R = +\infty$.
- Si $X \sim \mathcal{G}(p)$ alors $G_X(t) = \frac{pt}{1 - (1 - p)t}$ avec $R = \frac{1}{1 - p} > 1$.
- Si $X \sim \mathcal{P}(\lambda)$ alors $G_X(t) = e^{\lambda(t-1)}$ avec $R = +\infty$.

Preuve.

• Si $X \sim \mathcal{U}([1, n])$:

• Si $X \sim \mathcal{B}(n, p)$:

• Si $X \sim \mathcal{G}(p)$:

• Si $X \sim \mathcal{P}(\lambda)$:

□

13.4.3 Lien entre fonction génératrice et $\mathbb{E}(X)$ ou $\mathbb{V}(X)$

Soit $R \geq 1$ le rayon de convergence de la série entière G_X . On sait que G_X est de classe \mathcal{C}^∞ sur l'intervalle ouvert de convergence $] - R, R[$ et en tant que série entière on peut la dériver terme à terme sur $] - R, R[$. Si X et $X(X - 1)$ sont d'espérance finie, comparons les expressions suivantes.

$\forall t \in] - R, R[, \quad G'_X(t) = \sum_{n=0}^{+\infty} nt^{n-1}\mathbb{P}(X = n)$	$\mathbb{E}(X) = \sum_{n=0}^{+\infty} n\mathbb{P}(X = n)$
$G''_X(t) = \sum_{n=0}^{+\infty} n(n-1)t^{n-2}\mathbb{P}(X = n)$	$\mathbb{E}(X(X-1)) = \sum_{n=0}^{+\infty} n(n-1)\mathbb{P}(X = n)$

Si $R > 1$, on peut choisir $t = 1$ et obtenir des égalités. On admet la proposition suivante, qui étend ce résultat au cas général ($R \geq 1$).

Proposition 29

Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} et G_X sa fonction génératrice. Alors, la variable aléatoire X admet une espérance $\mathbb{E}(X)$ **si et seulement si** G_X est dérivable (à gauche) en 1 et dans ce cas, on a

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} n\mathbb{P}(X = n) = G'_X(1).$$

Preuve. (D3) Pour $n \in \mathbb{N}$, on note $f_n : t \mapsto t^n\mathbb{P}(X = n)$. Alors f_n est de classe \mathcal{C}^1 sur $[0, 1]$ et pour tout $t \in [0, 1]$, on a :

$$f'_n(t) = nt^{n-1}\mathbb{P}(X = n).$$

• Supposons que $\mathbb{E}(X) < +\infty$, montrons que G_X est dérivable sur $[0, 1]$ (sens facile).

Alors la série $\sum_{n \in \mathbb{N}} n\mathbb{P}(X = n)$ est convergente.

• Supposons que $\mathbb{E}(X) = +\infty$ montrons que G_X n'est pas dérivable en 1 (sens difficile).

□

On peut, de la même manière, en dérivant une seconde fois, obtenir un lien entre $G_X(t)$ et $\mathbb{V}(X)$.
On pourra remarquer que si X admet une variance, alors X et $X(X-1)$ aussi et on a :

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \mathbb{E}(X(X-1) + X) - \mathbb{E}(X)^2 \\ &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2\end{aligned}$$

Exemple 13.21. Soit X une variable aléatoire à valeurs dans \mathbb{N} dont la fonction génératrice est donnée par :

$$\forall t \in \mathbb{R}, \quad G_X(t) = ae^{1+t^2}.$$

- Déterminer a .
- Donner la loi de X et calculer $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

13.4.4 Fonction génératrice de la somme de variables aléatoires indépendantes

Proposition 30

• Soient X et Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} et G_X, G_Y leurs fonctions génératrices respectives. Si X et Y sont **indépendantes**, alors

$$G_{X+Y} = G_X G_Y.$$

• Soient X_1, \dots, X_n des variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} et G_{X_1}, \dots, G_{X_n} leurs fonctions génératrices respectives. Si X_1, \dots, X_n sont **indépendantes**, alors

$$G_{X_1+\dots+X_n} = G_{X_1} \times \dots \times G_{X_n}.$$

Preuve.

Application 1 :

Retrouver que si $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$ sont indépendantes, alors $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$.

□

Application 2 :

Retrouver que si $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$ sont indépendantes, alors $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

13.5 Annexe

13.5.1 Lois, espérances, variances

Nom	Support	Loi	Espérance	Variance	Fonct. Gén.
$X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$ avec $n \in \mathbb{N}^*$	$X(\Omega) = \llbracket 1, n \rrbracket$	$\mathbb{P}(X = k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$G_X(t) = \frac{t + \dots + t^n}{n}$
$X \sim \mathcal{B}(p)$ avec $p \in]0, 1[$	$X(\Omega) = \{0, 1\}$	$\mathbb{P}(X = 1) = p$ $\mathbb{P}(X = 0) = 1 - p$	p	$p(1-p)$	$G_X(t) = 1 - p + pt$
$X \sim \mathcal{B}(n, p)$ avec $p \in]0, 1[$, $n \in \mathbb{N}^*$	$X(\Omega) = \llbracket 0, n \rrbracket$	$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$G_X(t) = (1 - p + pt)^n$
$X \sim \mathcal{G}(p)$ avec $p \in]0, 1[$	$X(\Omega) = \mathbb{N}^*$	$\mathbb{P}(X = k) = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$G_X(t) = \frac{pt}{1 - (1-p)t}$
$X \sim \mathcal{P}(\lambda)$ avec $\lambda \in]0, +\infty[$	$X(\Omega) = \mathbb{N}$	$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ	$G_X(t) = e^{\lambda(t-1)}$

13.5.2 Quelques sommes de séries à connaître

Proposition 31 (Série géométrique et dérivées)

$$\forall p \in]-1, 1[, \quad \sum_{n=0}^{+\infty} p^n = \frac{1}{1-p}, \quad \sum_{n=0}^{+\infty} np^{n-1} = \frac{1}{(1-p)^2} \quad \text{et} \quad \sum_{n=0}^{+\infty} n(n-1)p^{n-2} = \frac{2}{(1-p)^3}$$

Proposition 32 (Série exponentielle)

$$\forall x \in \mathbb{R}, \quad \sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x.$$