

PROBABILITÉS

Contrairement au programme de première année, l'univers d'une expérience aléatoire n'est plus supposé fini.

1 Univers probabilisé

1.1 Notion de tribu

Définition 1.1 Tribu

On appelle **tribu** sur un ensemble Ω toute partie \mathcal{A} de $\mathcal{P}(\Omega)$ vérifiant les propriétés suivantes.

L'univers appartient à la tribu $\Omega \in \mathcal{A}$.

Stabilité par passage au complémentaire $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$.

Stabilité par union dénombrable $\forall (A_n) \in \mathcal{A}^{\mathbb{N}}, \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

REMARQUE. Une tribu est aussi parfois appelée une σ -algèbre.

Exemple 1.1

- $\mathcal{P}(\Omega)$ est une tribu sur Ω . On l'appelle la **tribu pleine**.
- $\{\emptyset, \Omega\}$ est une tribu sur Ω . On l'appelle la **tribu triviale**.

Définition 1.2 Événement

- On appelle **événement** tout élément de la tribu.
- Un événement est dit **élémentaire** si c'est un singleton.
- On appelle événement **contraire** d'un événement A le complémentaire \bar{A} de cet événement dans l'univers.
- L'ensemble vide est appelé événement **impossible**.
- Ω est appelé l'événement **certain**.
- On dit que deux événements A et B sont **incompatibles** si l'événement $A \cap B$ est impossible.
- On appelle **système complet d'événements** toute partition de l'univers formée d'événements.

Les propriétés suivantes des tribus découlent directement de la définition.

Proposition 1.1

Soit \mathcal{A} une tribu sur un ensemble Ω .

L'ensemble vide appartient à l'univers $\emptyset \in \mathcal{A}$.

Stabilité par intersection dénombrable $\forall (A_n) \in \mathcal{A}^{\mathbb{N}}, \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Stabilité par union ou intersection finie $\forall (A_1, \dots, A_n) \in \mathcal{A}^n, \bigcup_{i=1}^n A_i \in \mathcal{A}$ et $\bigcap_{i=1}^n A_i \in \mathcal{A}$.

Définition 1.3 Espace probabilisable

On appelle **espace probabilisable** tout couple (Ω, \mathcal{A}) où \mathcal{A} est une tribu sur Ω .

Tribu engendrée

Soit Ω un ensemble et \mathcal{C} une partie de $\mathcal{P}(\Omega)$. On appelle **tribu engendrée par \mathcal{C}** la plus petite tribu de Ω contenant \mathcal{C} . C'est l'intersection des tribus contenant \mathcal{C} .

Jeu de pile ou face infini

On considère l'expérience aléatoire consistant à lancer une infinité de fois une pièce. On associe face à 0 et pile à 1. On peut considérer l'univers $\Omega = \{0, 1\}^{\mathbb{N}^*}$ que l'on munit de la tribu $\mathcal{P}(\Omega)$. On note P_n l'événement «obtenir pile au lancer n » autrement dit

$$P_n = \{\omega \in \Omega, \omega_n = 1\}$$

L'univers Ω est généralement muni de la tribu engendrée par $\{P_n, n \in \mathbb{N}^*\}$.

Exemple 1.2

- L'événement «obtenir uniquement pile à partir du lancer n » est $\bigcap_{k \geq n} P_k$.
- L'événement «obtenir au moins un pile à partir du lancer n » est $\bigcup_{k \geq n} P_k$.
- L'événement «obtenir une infinité de fois pile» est $\bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} P_k$.

1.2 Espace probabilisé**Définition 1.4 Probabilité**

Soit (Ω, \mathcal{A}) un espace probabilisable. On appelle **probabilité** sur (Ω, \mathcal{A}) toute application $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ telle que

- $\mathbb{P}(\Omega) = 1$;
- pour toute suite d'événements **deux à deux disjoints**, $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.

Définition 1.5 Espace probabilisé

On appelle **espace probabilisé fini** tout triplet $(\Omega, \mathcal{A}, \mathbb{P})$ où (Ω, \mathcal{A}) est un espace probabilisable et \mathbb{P} une probabilité sur (Ω, \mathcal{A}) .

Proposition 1.2 Propriétés des probabilités

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

- $\mathbb{P}(\emptyset) = 0$.
- Soit $(A, B) \in \mathcal{A}^2$. Alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- Soit $A \in \mathcal{A}$. Alors $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.
- Soit $(A, B) \in \mathcal{A}^2$. Si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- Soit \mathcal{S} un système complet d'événements. Alors $\sum_{A \in \mathcal{S}} \mathbb{P}(A) = 1$.

Proposition 1.3

Soient Ω un univers **fini** ou **dénombrable**.

- Soit \mathbb{P} une probabilité sur $(\Omega, \mathcal{P}(\Omega))$. Alors la famille $(\mathbb{P}(\{\omega\}))_{\omega \in \Omega}$ est sommable de somme 1.
- Soit (p_ω) une famille sommable de réels positifs de somme 1. Alors il existe une unique probabilité sur $(\Omega, \mathcal{P}(\Omega))$ telle que $\mathbb{P}(\{\omega\}) = p_\omega$ pour tout $\omega \in \Omega$.

Pile ou face infini

On considère un jeu de pile ou face modélisé par l'univers $\Omega = \{0, 1\}^{\mathbb{N}^*}$ que l'on munit de la tribu \mathcal{A} engendrée par les événements «obtenir pile au $n^{\text{ème}}$ lancer» pour $n \in \mathbb{N}$.

Un théorème de Kolmogorov assure l'existence d'une probabilité sur (Ω, \mathcal{A}) telle que pour tout $n \in \mathbb{N}^*$ et tout $\tilde{\omega} \in \{0, 1\}^n$

$$\mathbb{P}(\{\omega \in \Omega, \forall i \in \llbracket 1, n \rrbracket, \omega_i = \tilde{\omega}_i\}) = \frac{1}{2^n}$$

Autrement dit, il existe une probabilité telle que la probabilité d'obtenir une séquence **finie** donnée au cours des n premiers lancers est $\frac{1}{2^n}$.

On peut alors prouver sans peine que les événements «obtenir pile au $n^{\text{ème}}$ lancer» ou «obtenir face au $n^{\text{ème}}$ lancer» sont de probabilité $\frac{1}{2}$.

1.3 Propriétés des probabilités

Proposition 1.4 Continuité croissante et décroissante

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un univers probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Continuité croissante Si (A_n) est **croissante** pour l'inclusion, alors $\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right)$.

Continuité décroissante Si (A_n) est **décroissante** pour l'inclusion, alors $\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right)$.

Corollaire 1.1

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un univers probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$$

REMARQUE. On peut donner un sens à l'inégalité précédente même si la série $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ diverge, puisqu'étant à termes positifs, elle diverge dans ce cas vers $+\infty$.

Définition 1.6 Événement négligeable/presque sûr

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

Événement négligeable On dit que A est **négligeable** si $\mathbb{P}(A) = 0$.

Événement presque sûr On dit que A est **presque sûr** si $\mathbb{P}(A) = 1$.

**ATTENTION !**

- A peut être un événement négligeable sans que $A = \emptyset$.
- A peut être un événement presque sûr sans que $A = \Omega$.

Proposition 1.5

- Une réunion **finie ou dénombrable** d'événements négligeables est un événement négligeable.
- Une intersection **finie ou dénombrable** d'événements presque sûrs est un événement presque sûr.

Exercice 1.1**Lemme de Borel-Cantelli et loi du zéro-un de Borel**

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On pose $B_n = \bigcup_{k \geq n} A_k$ et $A = \bigcap_{n \in \mathbb{N}} B_n$.

1. On suppose que la série $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ converge.

(a) Montrer que $\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}(B_n)$.

(b) En déduire que $\mathbb{P}(A) = 0$.

2. On suppose que les A_n sont mutuellement indépendants et que la série $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ diverge.

(a) Soit $(n, p) \in \mathbb{N}^2$. Montrer que

$$\mathbb{P}\left(\bigcap_{k=n}^{n+p} \overline{A_k}\right) \leq \exp\left(-\sum_{k=n}^{n+p} \mathbb{P}(A_k)\right)$$

(b) En déduire que $\mathbb{P}(A) = 1$.

1.4 Probabilité conditionnelle**Définition 1.7 Probabilité conditionnelle**

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A, B) \in \mathcal{A}^2$ tel que $\mathbb{P}(B) \neq 0$. On appelle **probabilité de A sachant B** notée $\mathbb{P}(A | B)$ ou $\mathbb{P}_B(A)$ le quotient $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Proposition 1.6

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $B \in \mathcal{A}$ tel que $\mathbb{P}(B) \neq 0$. Alors $\mathbb{P}_B : A \in \mathcal{A} \mapsto \mathbb{P}_B(A)$ est une probabilité sur (Ω, \mathcal{A}) .

Proposition 1.7 Formule des probabilités composées

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_1, \dots, A_n) \in \mathcal{A}^n$ tel que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \mathbb{P}(A_2 | A_1) \mathbb{P}(A_3 | A_1 \cap A_2) \dots \mathbb{P}(A_n | A_1 \cap \dots \cap A_{n-1})$$

Proposition 1.8 Formule des probabilités totales

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, \mathcal{S} un système complet d'événements et B un événement. Alors

$$\mathbb{P}(B) = \sum_{A \in \mathcal{S}} \mathbb{P}(B | A) \mathbb{P}(A)$$

REMARQUE. La propriété reste vraie si \mathcal{S} est un système-quasi complet d'événements.

1.5 Événements indépendants

Définition 1.8 Événements indépendants

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On dit que deux événements A et B sont **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Proposition 1.9

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient A et B deux événements avec $\mathbb{P}(B) > 0$. Alors A et B sont indépendants si et seulement si $\mathbb{P}(A | B) = \mathbb{P}(A)$



ATTENTION ! Dire que deux événements sont indépendants ne signifie pas qu'ils sont incompatibles.

Au contraire, deux événements incompatibles ne sont généralement pas indépendants à moins que l'un d'entre eux soit de probabilité nulle.

Définition 1.9 Famille d'événements indépendants

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_i)_{i \in I}$ une famille d'événements. On dit que $(A_i)_{i \in I}$ est une famille d'événements **mutuellement indépendants** si, pour toute partie **finie** J de I

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

REMARQUE. Il découle directement de la définition que toute sous-famille d'une famille d'événements mutuellement indépendants est encore une famille d'événements mutuellement indépendants.



ATTENTION ! Si $(A_i)_{i \in I}$ est une famille finie d'événements mutuellement indépendants alors les A_i sont indépendants deux à deux. Cependant, la réciproque est fausse.

Exemple 1.3 Pile ou face infini

Si l'on considère le jeu de pile ou face infini muni de la tribu et de la probabilité vues précédemment, la famille $(P_n)_{n \in \mathbb{N}^*}$ où P_n est l'événement «obtenir pile au $n^{\text{ème}}$ lancer» est une famille d'événements mutuellement indépendants.

Exercice 1.2

Soit $(A_i)_{i \in I}$ une famille d'événements mutuellement indépendants d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On considère une famille $(B_i)_{i \in I}$ telle que pour tout $i \in I$, $B_i = A_i$ ou $B_i = \overline{A_i}$. Montrer que $(B_i)_{i \in I}$ est également une famille d'événements mutuellement indépendants.

2 Variables aléatoires discrètes

2.1 Définitions

Définition 2.1 Variable aléatoire discrète

Soit (Ω, \mathcal{A}) un espace probabilisable. On appelle **variable aléatoire discrète** sur (Ω, \mathcal{A}) toute application X définie sur Ω telle que

- $X(\Omega)$ est fini ou dénombrable ;
- pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{A}$.

REMARQUE. Si X est une variable aléatoire discrète sur un espace probabilisable (Ω, \mathcal{A}) , alors $X(\mathcal{A}) = \{X(A), A \in \mathcal{A}\} = \mathcal{P}(X(\Omega))$.

Notation 2.1

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

Pour tout $A \in \mathcal{P}(X(\Omega))$, $X^{-1}(A)$ est un élément de \mathcal{A} donc un événement. L'événement $X^{-1}(A)$ se note plutôt $X \in A$.

L'événement $X^{-1}(\{x\})$ se note plutôt $X = x$.

On peut alors parler des probabilités $\mathbb{P}(X \in A)$ et $\mathbb{P}(X = x)$.

Si X est une variable aléatoire réelle et $x \in \mathbb{R}$, alors $\mathbb{P}(X \in]-\infty, x])$ et $\mathbb{P}(X \in [x, +\infty[)$ se note alors plutôt $\mathbb{P}(X \leq x)$ et $\mathbb{P}(X \geq x)$ respectivement.

Définition 2.2 Loi d'une variable aléatoire

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle **loi** de X l'application \mathbb{P}_X :

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

Notation 2.2

Si X et Y sont des variables aléatoires de même loi, on note $X \sim Y$.

Lorsque X suit une loi \mathcal{L} donnée, on notera également $X \sim \mathcal{L}$.

Proposition 2.1

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors la loi \mathbb{P}_X de X est une probabilité sur l'espace probabilisable $(X(\Omega), \mathcal{P}(X(\Omega)))$.

Corollaire 2.1

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors \mathbb{P}_X est entièrement déterminée par la donnée des $\mathbb{P}(X = x)$ pour $x \in X(\Omega)$.

REMARQUE. C'est pour cela qu'en pratique, lorsque l'on demande la loi d'une variable aléatoire X sur Ω , on ne demande pas $\mathbb{P}(X \in A)$ pour tout $A \in \mathcal{P}(X(\Omega))$ mais $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

Définition 2.3 Loi conditionnelle

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et $B \in \mathcal{A}$ tel que $P(B) > 0$. On appelle **loi de X conditionnée par l'événement B** l'application $P_{X|B} : \begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ A & \longmapsto P(X \in A | B) \end{cases}$.

REMARQUE. A nouveau, quand on demande en pratique la loi conditionnelle de X conditionnée par l'événement B , on se contente de donner $P(X = x | B)$ pour tout $x \in X(\Omega)$ plutôt que $P(X \in A | B)$ pour tout $A \in \mathcal{P}(X(\Omega))$.

2.2 Lois usuelles**Définition 2.4 Loi géométrique**

Soient X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N}^* et $p \in [0, 1]$. On dit que X suit la **loi géométrique de paramètre p** si

$$\forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = (1 - p)^{n-1} p$$

Pour abréger, on dit que X suit la loi $\mathcal{G}(p)$ ou encore $X \sim \mathcal{G}(p)$.

Interprétation de la loi géométrique

On considère une succession d'épreuves de Bernoulli indépendantes de même paramètre p . Le rang d'apparition du premier succès suit une loi géométrique de paramètre p .

Proposition 2.2 Loi sans mémoire

Soit X une variable aléatoire discrète à valeurs dans \mathbb{N}^* . X suit une loi géométrique si et seulement si

$$\forall (m, n) \in \mathbb{N}, \mathbb{P}(X > m + n | X > m) = \mathbb{P}(X = n)$$

Définition 2.5 Loi de Poisson

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N}^* et $\lambda \in \mathbb{R}_+^*$. On dit que X suit la **loi de Poisson de paramètre λ** si

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

Pour abréger, on dit que X suit la loi $\mathcal{P}(\lambda)$ ou encore $X \sim \mathcal{P}(\lambda)$.

Proposition 2.3 Approximation d'une loi binomiale par une loi de Poisson

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires telle que $X_n \sim \mathcal{P}(\lambda_n)$ pour tout $n \in \mathbb{N}$. On suppose que $n p_n \xrightarrow{n \rightarrow +\infty} \lambda > 0$. Alors pour tout $k \in \mathbb{N}$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

REMARQUE. On dit que la suite (X_n) converge en loi vers une variable aléatoire de loi $\mathcal{P}(\lambda)$.

2.3 Couples et uplets de variables aléatoires

Remarquons tout d'abord que si X et Y sont des variables aléatoires discrètes sur un espace probabilisable $(\Omega, \mathcal{A}, \mathbb{P})$, il en est de même du couple (X, Y) , c'est-à-dire de l'application $\omega \in \Omega \mapsto (X(\omega), Y(\omega))$.

Définition 2.6 Loi conjointe, lois marginales

Soient X et Y des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
La loi du couple (X, Y) est appelée **loi conjointe** des variables aléatoires X et Y .
Les lois de X et Y sont appelées les **lois marginales** du couple (X, Y) .

Proposition 2.4

Soient X et Y des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
La loi conjointe de X et Y permet de retrouver les lois marginales. Plus précisément

$$\begin{aligned}\forall x \in X(\Omega), \mathbb{P}(X = x) &= \sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y)) \\ \forall y \in Y(\Omega), \mathbb{P}(Y = y) &= \sum_{x \in X(\Omega)} \mathbb{P}((X, Y) = (x, y))\end{aligned}$$



ATTENTION ! Les lois marginales ne permettent pas de retrouver la loi conjointe (à moins que les variables aléatoires soient indépendantes).

Définition 2.7 Vecteur aléatoire

De même, si X_1, \dots, X_n sont des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors (X_1, \dots, X_n) est une variable aléatoire appelée **vecteur aléatoire**. Les lois de X_1, \dots, X_n sont les lois marginales de ce vecteur aléatoire.

2.4 Variables aléatoires indépendantes

Définition 2.8 Couple de variables aléatoires indépendantes

Soient X et Y des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Les variables aléatoires X et Y sont dites **indépendantes** si

$$\forall (A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), \mathbb{P}((X, Y) \in A \times B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Proposition 2.5

Soient X et Y des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Les variables aléatoires X et Y sont indépendantes si et seulement si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Définition 2.9 Famille de variables aléatoires mutuellement indépendantes

Soit $(X_i)_{i \in I}$ une famille **finie** de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
On dit que $(X_i)_{i \in I}$ est une famille de variables aléatoires **mutuellement indépendantes** si

$$\forall (A_i)_{i \in I} \in \prod_{i \in I} \mathcal{P}(X_i(\Omega)), \quad \mathbb{P}\left(\bigcap_{i \in I} X_i \in A_i\right) = \prod_{i \in I} \mathbb{P}(X_i \in A_i)$$

Une famille **infinie** de variables aléatoires est une famille de variables aléatoires mutuellement indépendantes si toute sous-famille finie l'est.

Proposition 2.6

Soit $(X_i)_{i \in I}$ une famille **finie** de variables aléatoires sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
 $(X_i)_{i \in I}$ est une famille de variables aléatoires mutuellement indépendantes si et seulement si

$$\forall (x_i)_{i \in I} \in \prod_{i \in I} X_i(\Omega), \quad \mathbb{P}\left(\bigcap_{i \in I} X_i = x_i\right) = \prod_{i \in I} \mathbb{P}(X_i = x_i)$$

REMARQUE. Toute sous-famille d'une famille finie de variables mutuellement indépendantes est une famille de variables aléatoires mutuellement indépendantes.



ATTENTION ! Si des variables aléatoires sont mutuellement indépendantes, alors ces variables aléatoires sont indépendantes deux à deux mais la réciproque est fausse.

Exemple 2.1

Soit X et Y deux variables aléatoires indépendantes de loi uniforme sur $\{-1, 1\}$. On pose $Z = XY$. Alors X, Y, Z sont deux à deux indépendantes mais pas mutuellement indépendantes.

Exercice 2.1

Montrer que $(A_i)_{i \in I}$ est une famille d'événements mutuellement indépendants si et seulement si $(\mathbb{1}_{A_i})_{i \in I}$ est une famille de variables aléatoires mutuellement indépendantes.

Proposition 2.7 Lemme des coalitions

Soient X_1, \dots, X_n des variables aléatoires discrètes mutuellement indépendantes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Soient $m \in \llbracket 1, n-1 \rrbracket$, f une application définie sur $\prod_{k=1}^m X_k(\Omega)$ et g une application définie sur $\prod_{k=m+1}^n X_k(\Omega)$. Alors, $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont des variables aléatoires indépendantes.

Exemple 2.2

Si X, Y et Z sont trois variables aléatoires discrètes réelles mutuellement indépendantes, alors $\cos(X^2 + Y^2)$ et e^Z sont indépendantes.

Proposition 2.8

Si l'on se donne une suite de lois de variables aléatoires discrètes, il existe un espace probabilisé tel qu'il existe une suite de variables aléatoires discrètes mutuellement indépendantes suivant ces lois.

REMARQUE. Peu d'intérêt pratique. Cette proposition sert essentiellement à légitimer certains énoncés d'exercices.

Jeu de pile ou face infini

La proposition précédente permet de donner un cadre théorique au jeu de pile ou face infini. En effet, on peut affirmer l'existence d'un espace probabilisé pour lequel il existe une suite de variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre donné ($1/2$ si la pièce lancée est équilibrée).

3 Espérance, variance et covariance

Dans tout ce paragraphe, les variables aléatoires considérées sont des variables aléatoires réelles (à valeurs dans \mathbb{R}).

3.1 Espérance

Définition 3.1 Espérance d'une variable aléatoire discrète réelle positive

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que $X(\Omega) \subset \mathbb{R}_+$. L'**espérance** de X est la somme, dans $\mathbb{R}_+ \cup \{+\infty\}$, de la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$. On la note $\mathbb{E}(X)$.

Définition 3.2 Espérance d'une variable aléatoire discrète réelle

Soit X une variable aléatoire discrète réelle sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. 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, la somme de cette famille est l'**espérance** de X . On la note $\mathbb{E}(X)$.

REMARQUE. Il est primordial de remarquer que l'espérance d'une variable aléatoire ne dépend que de sa loi.

REMARQUE. Une variable aléatoire d'espérance finie est dite centrée si $\mathbb{E}(X) = 0$.

Proposition 3.1 Espérance des lois usuelles

- Si $X \sim \mathcal{G}(p)$, alors $\mathbb{E}(X) = \frac{1}{p}$.
- Si $X \sim \mathcal{P}(\lambda)$, alors $\mathbb{E}(X) = \lambda$.

Proposition 3.2 Propriétés de l'espérance

Linéarité Soient X et Y deux variables aléatoires discrètes réelles d'espérances finies. Soit $(\lambda, \mu) \in \mathbb{R}^2$. Alors $\lambda X + \mu Y$ est d'espérance finie et $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.

Positivité Soit X une variable aléatoire discrète positive. Alors $\mathbb{E}(X) \geq 0$.

Croissance Soient X et Y deux variables aléatoires discrètes réelles d'espérances finies. Si $X \leq Y$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

Inégalité triangulaire Soit X une variable aléatoire d'espérance finie. Alors $|X|$ est d'espérance finie et $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

REMARQUE. Si $|X| \leq Y$ et Y est d'espérance finie, alors X est d'espérance finie et $|\mathbb{E}(X)| \leq \mathbb{E}(|X|) \leq \mathbb{E}(Y)$.

Proposition 3.3 Formule de transfert

Soient X une variable aléatoire discrète et $f : X(\Omega) \rightarrow \mathbb{R}$. Alors $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable et, dans ce cas,

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

REMARQUE. L'intérêt de la formule de transfert est qu'elle permet de calculer l'espérance de $f(X)$ sans connaître la loi de $f(X)$ mais uniquement la loi de X .

Exercice 3.1

Soit X une variable suivant la loi géométrique de paramètre p . Déterminer l'espérance de $1/X$.

Proposition 3.4 Espérance et indépendance

Soient X et Y deux variables aléatoires discrètes réelles **indépendantes** d'espérances finies. Alors XY est d'espérance finie et

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

3.2 Variance

Définition 3.3 Moment d'une variable aléatoire

Soit X une variable aléatoire discrète réelle. On dit que X admet un moment d'ordre r si X^r est d'espérance finie. Dans ce cas, le moment d'ordre r de X est $\mathbb{E}(X^r)$.

REMARQUE. Un moment d'une variable aléatoire ne dépend que de sa loi.

Exercice 3.2

Soit $r \in \mathbb{N}$. Montrer que l'ensemble des variables aléatoires discrètes réelles admettant un moment d'ordre r forment un \mathbb{R} -espace vectoriel.

Exercice 3.3

Montrer que si une variable aléatoire réelle discrète admet un moment d'ordre r alors elle admet un moment d'ordre k pour tout $k \in \llbracket 0, r \rrbracket$.

Proposition 3.5

L'ensemble des variables aléatoires discrètes réelles admettant un moment d'ordre 2 forment un \mathbb{R} -espace vectoriel.

Exercice 3.4

Soit E l'ensemble des variables aléatoires discrètes réelles admettant un moment d'ordre 2. Montrer que l'application $(X, Y) \in E^2 \mapsto \mathbb{E}(XY)$ est bien définie, bilinéaire, symétrique et positive.

Proposition 3.6

Si une variable aléatoire réelle discrète admet un moment d'ordre 2, elle est d'espérance finie.

Proposition 3.7 Inégalité de Cauchy-Schwarz

Soient X et Y deux variables aléatoires discrètes réelles admettant chacune un moment d'ordre 2. Alors XY est d'espérance finie et

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

Définition 3.4 Variance et écart-type

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2.

- On appelle **variance** de X le réel positif

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

- On appelle **écart-type** de X le réel

$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

REMARQUE. La variance et l'écart-type d'une variable aléatoire ne dépendent que de sa loi.

REMARQUE. On dit qu'une variable aléatoire X est **réduite** si $\mathbb{V}(X) = 1$.

Exercice 3.5

Soit X une variable aléatoire discrète réelle de variance nulle. Montrer que X est presque sûrement constante égale à $\mathbb{E}(X)$, autrement dit que $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

Proposition 3.8 Théorème de König-Huygens

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2 (et donc également une espérance finie). Alors

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

Proposition 3.9 Variance d'une fonction affine d'une variable aléatoire

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Soit $(a, b) \in \mathbb{R}^2$. Alors $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.

REMARQUE. Si X est une variable aléatoire admettant une variance finie non nulle, alors $\frac{X - \mathbb{E}(X)}{\sigma(X)}$ est **centrée réduite**.

Proposition 3.10 Variance des lois usuelles

- Si $X \sim \mathcal{G}(p)$, alors $\mathbb{V}(X) = \frac{1-p}{p^2}$.
- Si $X \sim \mathcal{P}(\lambda)$, alors $\mathbb{V}(X) = \lambda$.

3.3 Covariance**Définition 3.5 Covariance de deux variables aléatoires**

Soient X et Y deux variables aléatoires discrètes réelles admettant un moment d'ordre 2. On appelle **covariance de X et Y** le réel

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

REMARQUE. $\mathbb{V}(X) = \text{Cov}(X, X)$

REMARQUE. D'après l'inégalité de Cauchy-Schwarz :

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$$

REMARQUE. La covariance est «presque» un produit scalaire. C'est une forme bilinéaire symétrique positive mais pas nécessairement définie.

Proposition 3.11

Soient X et Y deux variables aléatoires discrètes réelles admettant un moment d'ordre 2. Alors

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

Proposition 3.12 Covariance de variables aléatoires

Si X et Y sont deux variables aléatoires discrètes réelles **indépendantes** admettant un moment d'ordre 2, alors $\text{Cov}(X, Y) = 0$.



ATTENTION ! La réciproque est fautive. Il suffit par exemple de prendre $X \sim \mathcal{U}(\{-1, 1\})$ et $Y = -X$. On a bien $\text{Cov}(X, Y) = 0$ mais X et Y ne sont évidemment pas indépendantes.

Proposition 3.13 Variance d'une somme et théorème de Pythagore

Soient X_1, \dots, X_n des variables aléatoires réelles discrètes admettant des moments d'ordre 2. Alors

$$\mathbb{V}\left(\sum_{k=1}^n X_k\right) = \sum_{1 \leq i, j \leq n} \text{Cov}(X_i, X_j)$$

Si les X_k sont **indépendantes deux à deux**,

$$\mathbb{V}\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n \mathbb{V}(X_k)$$

3.4 Inégalités classiques

Proposition 3.14 Inégalité de Markov

Soient X une variable aléatoire discrète réelle **positive** d'espérance finie et $a \in \mathbb{R}_+^*$. Alors

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

Si on ne suppose plus X positive, on peut affirmer que

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

Corollaire 3.1 Inégalité de Bienaymé-Tchebychev

Soient X une variable aléatoire discrète réelle admettant un moment d'ordre 2 et $\varepsilon \in \mathbb{R}_+^*$. Alors

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2} = \frac{\sigma(X)^2}{\varepsilon^2}$$

Corollaire 3.2 Loi faible des grands nombres

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires deux à deux indépendantes, de même loi et admettant des moments d'ordre

2. Alors, en posant $S_n = \sum_{k=1}^n X_k$ et en notant m l'espérance commune des X_k ,

$$\forall \varepsilon \in \mathbb{R}_+^*, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0$$

REMARQUE. On peut en fait prouver à l'aide de l'inégalité de Bienaymé-Tchebychev qu'en notant σ^2 la variance commune des X_k ,

$$\forall \varepsilon \in \mathbb{R}_+^*, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

La loi faible des grands nombres s'en déduit immédiatement.

4 Fonctions génératrices

Définition 4.1 Fonction génératrice

Soit X une variable aléatoire à valeurs dans \mathbb{N} . On appelle fonction génératrice de X la fonction G_X définie par

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

REMARQUE. A nouveau, la fonction génératrice d'une variable aléatoire ne dépend que de sa **loi**.

REMARQUE. Puisque la série $\sum_{k \in \mathbb{N}} \mathbb{P}(X = k)$ converge (vers 1),

- la série entière $\sum_{k \in \mathbb{N}} \mathbb{P}(X = k)t^k$ a un rayon de convergence supérieur ou égal à 1 ;

- elle converge normalement sur le disque fermé de centre 0 et de rayon 1 ;
- sa somme G_X est continue sur ce même disque ;
- G_X est de classe \mathcal{C}^∞ sur $] -1, 1[$.

Fonctions génératrices des lois usuelles

- Si $X \sim \mathcal{B}(p)$, $G_X(t) = 1 - p + pt$.
- Si $X \sim \mathcal{B}(n, p)$, $G_X(t) = (1 - p + pt)^n$.
- Si $X \sim \mathcal{G}(p)$, $G_X(t) = \frac{pt}{1 - (1 - p)t}$.
- Si $X \sim \mathcal{P}(\lambda)$, $G_X(t) = e^{\lambda(t-1)}$.

Proposition 4.1 Fonction génératrice et loi

Soit X une variable aléatoire à valeurs dans \mathbb{N} . Alors

$$\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}$$

Proposition 4.2

Deux variables aléatoires à valeurs dans \mathbb{N} sont égales si et seulement si elles ont la même fonction génératrice.

Proposition 4.3 Fonction génératrice et espérance

Soit X une variable aléatoire à valeurs dans \mathbb{N} . Alors X admet une espérance finie si et seulement si G_X est dérivable en 1 et, dans ce cas, $\mathbb{E}(X) = G'_X(1)$.

Proposition 4.4 Fonction génératrice et variance

Soit X une variable aléatoire à valeurs dans \mathbb{N} . Alors X admet un moment d'ordre 2 si et seulement si G_X est deux fois dérivable en 1 et, dans ce cas, $\mathbb{V}(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.

Proposition 4.5 Fonction génératrice d'une somme finie de variables aléatoires indépendantes

Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes à valeurs dans \mathbb{N} . Alors

$$G_{\sum_{k=1}^n X_k} = \prod_{k=1}^n G_{X_k}$$

Exercice 4.1

Soient X_1, \dots, X_n des variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs $\lambda_1, \dots, \lambda_n$.

Déterminer la loi de $S = \sum_{k=1}^n X_k$.