Chapitre XI

Probabilités discrètes

Table des matières

Dans ce chapitre, on utilisera la notation $\overline{A}$ pour désigner le complémentaire d'une sous-partie A d'un ensemble.

Partie A

Espaces probabilisés

1. Espaces probabilisables

a. Définitions

Définition 1. *Tribu*

Soit Ω un ensemble non vide et $\mathcal{T} \subset \mathcal{P}(\Omega)$. On dit que $\mathcal{T}$ est une **tribu** sur Ω si :
- $\Omega \subset \mathcal{T}$;
- $\mathcal{T}$ est stable par passage au complémentaire i.e. pour tout $A \in \mathcal{T}$, $\overline{A} \in \mathcal{T}$;
- $\mathcal{T}$ est stable par réunion dénombrable i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathcal{T}$, $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

Remarque 1.

On trouve parfois la terminologie σ-algèbre pour désigner une tribu et la notation $\mathcal{A}$ au lieu de $\mathcal{T}$.

Définition 2. *Espace probabilisable*

Soit Ω un ensemble non vide et $\mathcal{T}$ une tribu sur Ω. Le couple $(\Omega, \mathcal{T})$ est appelé **espace probabilisable**.

Exemple 1.

Soit Ω un ensemble non vide.
- $\mathcal{T} = \{\emptyset, \Omega\}$ est une tribu appelée **tribu grossière** ;
- $\mathcal{T} = \mathcal{P}(\Omega)$ est une tribu ;
- Si $A \subset \Omega$, $\mathcal{T} = \{\emptyset, A, \overline{A}, \Omega\}$ est une tribu

Exercice 1.

Soit Ω un ensemble non vide et $\mathcal{F} \subset \mathcal{P}(\Omega)$.

1. Montrer que l'intersection de toutes les tribus contenant $\mathcal{F}$ (en existe-t-il ?) est une tribu. On note cette tribu $\sigma(\mathcal{F})$ et on l'appelle **tribu engendrée par** $\mathcal{F}$.
2. Montrer que $\sigma(\mathcal{F})$ est la plus petite tribu contenant $\mathcal{F}$.

Remarque 2. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On utilise la terminologie suivante :
- Ω est l'**univers** ;
- les éléments A de $\mathcal{T}$ sont les **événements** ;

b. Les événements

Proposition 1.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On a les propriétés suivantes :
- $\emptyset \in \mathcal{T}$;
- $\mathcal{T}$ est stable par réunion et intersection finies ;
- $\mathcal{T}$ est stable par intersection dénombrable i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{T}$, alors $\bigcap\limits_{n\in\mathbb{N}} A_n \in \mathcal{T}$;
- pour tout $A, B \in \mathcal{T}$, $A \setminus B \in \mathcal{T}$.

Remarque 3. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $A, B \in \mathcal{T}$ des événements. On utilise la terminologie suivante concernant les événements :
- Ω est l'événement **certain** et $\emptyset$ l'événement **impossible** ;
- $\overline{A}$ est l'événement **contraire** de A ;
- si $A \cap B = \emptyset$, on dit que A et B sont des événements **incompatibles**.

Définition 3. *Système complet d'événements*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $(A_i)_{i\in I}$ une famille au plus dénombrable d'événements. On dit que la famille $(A_i)_{i\in I}$ est un **système complet d'événements** si elle forme une partition de Ω.

2. Espaces probabilisés

a. Définition

Définition 4. *Probabilité*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $P : \mathcal{T} \to [0,1]$ une application. On dit que P est une **probabilité** sur $(\Omega, \mathcal{T})$ si :

i) $P(\Omega) = 1$;

ii) pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux disjoints, la série $\sum P(A_n)$ converge et :

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} P(A_n). \qquad \boldsymbol{\sigma}\textbf{-additivité}$$

Remarque 4. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Pour $A \in \mathcal{T}$, on dit que $P(A)$ est **la probabilité de l'événement** A.

Définition 5. *Espace probabilisé*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Le triplet $(\Omega, \mathcal{T}, P)$ est appelé **espace probabilisé**.

Définition 6. *Événement négligeable/presque sûr*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A \in \mathcal{T}$. On dit que :

— l'événement A est **négligeable** si $P(A) = 0$.

— l'événement A est **presque sûr** si $P(A) = 1$.

b. Propriétés

Proposition 2.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A, B \in \mathcal{T}$. On a les propriétés suivantes :

i) $P(\emptyset) = 0$;

ii) si A et B sont incompatibles, $P(A \cup B) = P(A) + P(B)$;

iii) $P(\overline{A}) = 1 - P(A)$;

iv) si $A \subset B$, $P(B \setminus A) = P(B) - P(A)$;

v) si $A \subset B$, $P(A) \leq P(B)$;

vi) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Exercice 2.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Montrer que si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements, alors :

1. $\sum P(A_n)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n) = 1$.
2. pour tout $B \in \mathcal{T}$, $\sum P(A_n \cap B)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n \cap B) = P(B)$.

3. Propriétés élémentaires des probabilités

Théorème 1. *Continuité croissante/décroissante*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$.

- Si $(A_n)_{n\in\mathbb{N}}$ est croissante i.e. pour tout $n \in \mathbb{N}$, $A_n \subset A_{n+1}$, alors :
$$P(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$
- Si $(A_n)_{n\in\mathbb{N}}$ est décroissante i.e. pour tout $n \in \mathbb{N}$, $A_{n+1} \subset A_n$, alors :
$$P(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$

Proposition 3.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$.

- pour $n \in \mathbb{N}$, on a :
$$P(\bigcup_{i=0}^{n} A_n) \leq \sum_{i=0}^{n} P(A_i).$$
- On a :
$$P(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n=0}^{+\infty} P(A_i).$$

Corollaire 1.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Si, pour tout $n \in \mathbb{N}$, A_n est négligeable, alors $\displaystyle\bigcup_{n\in\mathbb{N}} A_n$ est négligeable.

4. Probabilité sur un univers au plus dénombrable

a. Caractérisation

Théorème 2.

Soit Ω un ensemble non vide au plus dénombrable.

i) Soit P est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ et on pose, pour $\omega \in \Omega$, $p_\omega = P(\{\omega\})$. Alors, pour tout $\omega \in \Omega$, $p_\omega \geq 0$ et :
$$\sum_{\omega \in \Omega} p_\omega = 1.$$

ii) Si $(p_\omega)_{\omega \in \Omega}$ est une suite de réels positifs telle que $\sum\limits_{\omega \in \Omega} p_\omega = 1$, alors il existe une unique probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que, pour tout $\omega \in \Omega$, $P(\{\omega\}) = p_\omega$. Elle est définie, pour $A \in \mathcal{P}(\Omega)$, par :
$$P(A) = \sum_{\omega \in A} p_\omega.$$

b. Exemples de probabilités sur $\mathbb{N}$

Exemple 2.

1. **Loi géométrique.** Soit $p \in]0,1[$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = (1-p)^n p$ pour $n \in \mathbb{N}$.
2. **Loi de Poisson.** Soit $\lambda \in \mathbb{R}_+^*$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = e^{-\lambda} \dfrac{\lambda^n}{n!}$ pour $n \in \mathbb{N}$.

Partie B

Probabilités conditionnelles

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

1. Conditionnement

Définition 7. *Probabilité conditionnelle*

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. Pour $A \in \mathcal{T}$, on appelle **probabilité (conditionnelle) de** A **sachant** B et on note $P_B(A)$ ou encore $P(A|B)$ la quantité :

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

L'application $P_B : \mathcal{T} \to [0, 1]$ définie par :

$$P_B : A \mapsto P_B(A)$$

est appelée **probabilité conditionnelle sachant** B.

Théorème 3.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. La probabilité conditionnelle $P_B : \mathcal{T} \to [0, 1]$ sachant B est une probabilité sur l'espace probabilisable $(\Omega, \mathcal{T})$.

2. Formules conditionnelles

a. Formule des probabilités composées

Théorème 4. *Formule des probabilités composées*

i) Soit $A, B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(A \cap B) = P(A|B)P(B).$$

ii) Soit $n \in \mathbb{N}^*$ et $A_1, \ldots, A_n \in \mathcal{T}$ tels que $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$. On a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \ldots \times P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right).$$

b. Formule des probabilités totales

Théorème 5. *Formule des probabilités totales*

Soit $(B_i)_{i\in I}$ un système complet d'événements. Pour tout événement $A \in \mathcal{T}$, on a :

$$P(A) = \sum_{i\in I} P(A|B_i)P(B_i).$$

avec la convention $P(A|B_i)P(B_i) = 0$ *si* $P(B_i) = 0$.

Démonstration.

Comme $(B_i)_{i\in I}$ est un système complet d'événements, on a :

$$\bigcup_{i\in I} B_i = \Omega \quad \text{et} \quad B_i \cap B_j = \emptyset \text{ si } i \neq j.$$

Par suite, on a :

$$A = \bigcup_{i\in I}(A \cap B_i) \text{ avec } (A \cap B_i) \cap (A \cap B_j) = \emptyset \text{ si } i \neq j.$$

En effet,

$$\bigcup_{i\in I}(A \cap B_i) = \bigcup_{i\in I} A \cap \bigcup_{i\in I} B_i = A \cap \Omega = A;$$

et

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset.$$

Par suite, par les propriétés de la fonction de probabilité P, on a :

$$P(A) = P\left(\bigcup_{i\in I}(A \cap B_i)\right) = \sum_{i\in I} P(A \cap B_i) = \sum_{i\in I} P(A|B_i)P(B_i).$$

□

Exercice 3.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$ et $P(\overline{B}) > 0$. Montrer que pour tout $A \in \mathcal{T}$, $P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$.

c. Formule de Bayes

Théorème 6. *Formule de Bayes*

Soit $A \in \mathcal{T}$ tel que $P(A) > 0$.

i) Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(B|A) = \frac{P(A|B)P(B)}{P(B)}.$$

ii) Soit $(B_i)_{i\in I}$ un système complet d'événements. On a, pour tout $j \in I$:

$$P(B_j) = \frac{P(A|B_j)P(B_j)}{\sum_{i\in I} P(A|B_i)P(B_i)}.$$

d. Exercice d'application des formules

Exercice 4.

Dans une population, on donne la probabilité définie, pour $n \in \mathbb{N}$ par p_n qu'une famille ait n enfants où, α étant un réel strictement positif :

$$p_n = \alpha \frac{2^n}{n!}.$$

On suppose qu'il est équiprobable qu'un enfant soit une fille ou un garçon.

1. Déterminer le nombre α.
2. Déterminer la probabilité qu'une famille ait au moins un garçon.
3. Déterminer la probabilité qu'une famille ait exactement 2 garçons et 3 filles.
4. On cherche à déterminer la probabilité qu'une famille ait exactement deux enfants sachant qu'elle a au moins deux filles.
   a) Déterminer la probabilité qu'une famille ait exactement deux filles.
   b) Déterminer la probabilité qu'une famille ait au moins deux filles.
   c) Conclure.

Correction.

1. Soit Ω l'ensemble des compositions (d'enfants) des familles de la population. La suite d'événements $(A_n)_{n\in\mathbb{N}}$ définie par :

   A_n="La famille possède exactement n enfants".

   étant un système complet d'événements, on a :

$$1 = P(\Omega) = P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} A_n = \sum_{n=0}^{+\infty} p_n.$$

   Par suite,

$$1 = \sum_{n=0}^{+\infty} p_n = \alpha \sum_{n=0}^{+\infty} \frac{2^n}{n!} = \alpha e^2.$$

D'où $\alpha = e^{-2}$.

2. On note B l'événement :

B="La famille possède uniquement des filles (ou aucun enfant)".

Alors $\overline{B}$="La famille possède au moins un garçon" est l'événement dont la probabilité est recherchée.

On a, d'après la formule des probabilités totales :

$$P(B) = \sum_{n=0}^{+\infty} P(B|A_n)P(A_n) = \sum_{n=0}^{+\infty} \frac{1}{2^n} p_n = e^{-2}.e = \frac{1}{e}.$$

3. On note C l'événement recherché, D l'événement "la famille possède 2 garçons et 3 filles (au moins)" et on reprend la suite (A_n) de la question précédente. Alors on a $C = D \cap A_2$ donc :

$$P(C) = P(D \cap A_5) = P(D|A_5)P(A_5) = \left(\binom{5}{2} \frac{1}{2^2} \frac{1}{2^3} \right) .e^{-2} \frac{2^5}{5!} = \frac{1}{12e^2}.$$

4. On note F l'événement "une famille a au moins deux filles" et on reprend la suite (A_n) de la question 1.

a) Avec les notations précédentes, ce qu'on doit calculer est :

$$P(F \cap A_2) = P(F|A_2)P(A_2) = \frac{1}{4}.e^{-2} \frac{4}{2} = \frac{1}{2e^2}.$$

b) D'après la formule des probabilités totales, on a :

$$P(F) = 1 - P(\overline{F}) = 1 - \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n).$$

et on a :

$$P(\overline{F}|A_n) = \binom{n}{0} \frac{1}{2^0} \frac{1}{2^n} + \binom{n}{1} \frac{1}{2^1} \frac{1}{2^{n-1}} = (n+1)\frac{1}{2^n}.$$

Ainsi, on a :

$$\begin{aligned}
P(\overline{F}) &= \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n) \\
&= \sum_{n=0}^{+\infty} (n+1) \frac{1}{2^n} p_n \\
&= e^{-2} \sum_{n=0}^{+\infty} \frac{n+1}{n!} \\
&= e^{-2} \left(\sum_{n=1}^{+\infty} \frac{n}{n!} + \sum_{n=0}^{+\infty} \frac{1}{n!} \right) \\
&= 2e^{-2} \sum_{n=0}^{+\infty} \frac{1}{n!} \\
&= \frac{2}{e}.
\end{aligned}$$

D'où $P(F) = 1 - \dfrac{2}{e}$.

c) On a :

$$P(A_2|F) = \frac{P(A_2 \cap F)}{P(F)} = \frac{1}{2e^2}.\frac{1}{1-\dfrac{2}{e}} = \frac{1}{2e(e-2)}.$$

3. Événements indépendants

a. Définition et premières propriétés

Définition 8. *Événements indépendants*

Soit $A, B \in \mathcal{T}$. On dit que A et B sont **indépendants** si

$$P(A \cap B) = P(A).P(B)$$

On justifie le choix de la terminologie d'"indépendance" par la proposition suivante :

Proposition 4.

Soit $A, B \in \mathcal{T}$ avec $P(B) > 0$. Les événements A et B sont indépendants si, et seulement si, $P(A|B) = P(A)$.

Proposition 5.

Soit $A, B \in \mathcal{T}$. Si A et B sont indépendants, alors

— $\overline{A}$ et $\overline{B}$ sont indépendants ;

— $\overline{A}$ et B sont indépendants (et A et $\overline{B}$ également).

b. Indépendance d'une famille d'événement

Définition 9. *Indépendance deux à deux/Indépendance mutuelle*

Soit $(A_i)_{i\in I}$ une famille à valeurs dans $\mathcal{T}$.

- On dit que les événements A_i, pour $i \in I$ sont **deux à deux indépendants** si, pour tous $i, j \in \mathbb{I}$ avec $i \neq j$, on a :

$$P(A_i \cap A_j) = P(A_i)P(A_j).$$

- On dit que les événements A_i, pour $i \in I$ sont **mutuellement indépendants** si, pour toute partie finie J de I, on a :

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{j\in J} P(A_i).$$

Remarque 5.

Il est clair que la mutuelle indépendance d'une famille d'événements implique l'indépendance deux à deux de ces événements mais **la réciproque est fausse !**

En effet, si on considère deux lancers de dé et les trois événements suivants sont deux à deux indépendants, mais pas mutuellement indépendants !

- A_1="La somme des deux lancers est pair" ;
- A_2="Le premier lancer est pair" ;
- A_3="Le second lancer est pair".

Partie C

Variables aléatoires discrètes

1. Variables aléatoires discrètes

a. Définition

Définition 10. *Variable aléatoire discrète*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$. On dit que la fonction X est une **variable aléatoire discrète** sur $(\Omega, \mathcal{T})$ si :

i) l'ensemble $X(\Omega)$ est au plus dénombrable ;

ii) pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement de $\mathcal{T}$, i.e.

$$X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{T}.$$

De plus, si $E \subset \mathbb{R}$, on dit que X est une variable aléatoire discrète **réelle**.

Proposition 6.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète. Pour tout $A \subset E$, $X^{-1}(A) \in \mathcal{T}$.

Pour des questions de lisibilité et de simplicité, on emploie les notations suivantes :

Notation 1.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète.

— Pour $A \subset E$, l'événement $X^{-1}(A)$ est noté $(X \in A)$.

— Pour $x \in E$, l'événement $X^{-1}(\{x\})$ est noté :

$$(X = x).$$

— Si $E \subset \mathbb{R}$, pour $x \in E$, les événements $X^{-1}(]-\infty, x[)$, $X^{-1}(]-\infty, x])$, $X^{-1}(]x, +\infty[)$ et $X^{-1}([x, +\infty[)$ sont notés respectivement :

$$(X < x), \qquad (X \leq x), \qquad (X > x), \qquad (X \geq x).$$

b. Loi de probabilité d'une variable aléatoire

Définition 11. *Loi de probabilité*

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. On appelle **loi de probabilité de** X l'application $P_X : \mathcal{P}(X(\Omega)) \to [0,1]$ définie par :

$$P_X : A \mapsto P(X \in A).$$

Proposition 7.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Le triplet $(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$ est un espace probabilisé.

Proposition 8.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Alors la famille $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements. En particulier, on a :

$$\sum_{x \in X(\Omega)} P(X = x) = 1.$$

Remarque 6.

— D'après ce qui précède, si X une variable aléatoire *discrète* sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, la loi de probabilité P_X est entièrement déterminée par les événements $(X = x)$ pour $x \in X(\Omega)$.

Ainsi, pour déterminer la loi de probabilité P_X de la variable aléatoire X, (il faut et) il suffit de déterminer les valeurs de $P(X = x)$ pour chaque $x \in X(\Omega)$.

— Si X est une variable aléatoire discrète à valeurs dans un ensemble au plus dénombrable E, $X(\Omega)$ n'est pas forcément égal à E. Dans la suite, on étendra la loi P_X de X à $\mathcal{P}(E)$ tout entier, en posant, pour $x \in E \setminus X(\Omega)$, $P_X(\{x\}) := 0$ ou encore $P(X = x) = 0$. Ainsi, P_X définit une loi de probabilité sur $(E, \mathcal{P}(E))$.

Définition 12. *Loi de probabilité discrète*

Soit E un ensemble au plus dénombrable et $(p_x)_{x \in E}$ une famille de réels positifs telle que $\sum_{x \in E} p_x = 1$. On appelle **loi (de probabilité) discrète sur** E la probabilité sur $(E, \mathcal{P}(E))$ définie par la famille $(p_x)_{x \in E}$. On notera souvent $\mathcal{L}$ une telle loi.

Notation 2.

Soit E un ensemble au plus dénombrable, $\mathcal{L}$ une loi discrète sur E définie par une famille $(p_x)_{x \in E}$ et X, Y des variables aléatoires discrètes à valeurs dans E.

— On dit que X **suit la loi discrète** $\mathcal{L}$ et on note $X \sim \mathcal{L}$ si $P_X = \mathcal{L}$ i.e. si pour tout

$x \in E$,

$$P(X = x) = p_x.$$

— On dit que X et Y **suivent la même loi** ou encore que X et Y sont **équidistribuées** et on note $X \sim Y$ si $P_X = P_Y$ i.e. si pour tout $x \in E$,

$$P(X = x) = P(Y = y).$$

c. Fonction d'une variable aléatoire discrète

Proposition-Notation 9.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, F un ensemble et $f : X(\Omega) \to F$ une fonction. On note $f(X)$ la variable aléatoire discrète $f \circ X$.

2. Lois usuelles sur un univers fini

a. Loi uniforme

Définition 13. *Loi uniforme*

Soit E un ensemble fini. On dit qu'une variable aléatoire X **suit une loi uniforme sur** E et on note $X \sim \mathcal{U}(E)$ si $X(\Omega) = E$ et, pour tout $x \in E$,

$$P(X = x) = \frac{1}{\#E}.$$

Si $E = [\![1, n]\!]$ pour un certain $n \in \mathbb{N}^*$, on notera simplement $\mathcal{U}(n)$ la loi uniforme sur $[\![1, n]\!]$.

b. Loi de Bernoulli

Définition 14. *Loi de Bernoulli*

Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X **suit une loi de Bernoulli de paramètre** p et on note $X \sim \mathcal{B}(p)$ si $X(\Omega) = \{0, 1\}$ et

$$P(X = 1) = p, \qquad P(X = 0) = 1 - p.$$

c. Loi binomiale

Définition 15. *Loi binomiale*

Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi binomiale de paramètres** n **et** p et on note $X \sim \mathcal{B}(n,p)$ si $X(\Omega) = [\![0,n]\!]$ et, pour tout $k \in [\![0,n]\!]$,

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

3. Lois usuelles sur un univers dénombrable

a. Loi géométrique

Définition 16. *Loi géométrique*

Soit $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi géométrique de paramètre** p et on note $X \sim \mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et, pour tout $k \in \mathbb{N}^*$,

$$P(X = k) = p(1-p)^{k-1}.$$

Remarque 7.

Une variable aléatoire suivant une loi géométrique de paramètre p modélise le temps d'attente du premier succès d'une suite d'épreuves de Bernoulli indépendantes de paramètre p.

Exercice 5.

Soit $p \in]0,1[$ et X une variable aléatoire qui suit une loi géométrique de paramètre p. Déterminer la valeur de $P(X > k)$ pour $k \in \mathbb{N}^*$.

Proposition 10.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Alors X suit une loi géométrique si, et seulement si, pour tout $k, l \in \mathbb{N}^*$,

$$P(X > k + l | X > k) = P(X > l).$$

b. Loi de Poisson

Définition 17. *Loi de Poisson*

Soit $\lambda \in \mathbb{R}_+^*$. On dit qu'une variable aléatoire X **suit une loi de Poisson de paramètre** λ et on note $X \sim \mathcal{P}(\lambda)$ si $X(\Omega) = \mathbb{N}$ et, pour tout $k \in \mathbb{N}$,

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Exercice 6.

Soit $\lambda \in \mathbb{R}_+^*$, $(p_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $]0,1[$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour chaque $n \in \mathbb{N}$, $X_n \sim \mathcal{B}(n, p_n)$.
Montrer que si $np_n \xrightarrow[n\to+\infty]{} \lambda$, alors, pour tout $k \in \mathbb{N}$,

$$P(X_n = k) \xrightarrow[n\to+\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Remarque 8.

Pour $n \in \mathbb{N}$ et $p \in]0,1[$ tels que $\lambda = np$ est proche de 0, une variable aléatoire qui suit une loi binomiale de paramètres n et p, suit approximativement une loi de Poisson de paramètre $\lambda = np$.

4. Couples de variables aléatoires

Dans ce paragraphe, $(\Omega, \mathcal{T}, P)$ désigne un espace probabilisé. Les valeurs aléatoires considérées sont définies sur cet espace.

a. Généralités

Définition 18. *Couple de variable aléatoires discrètes*

Soit X, Y des variables aléatoires discrètes sur un espace probabilisable. On appelle **couple de variables aléatoires discrètes sur** $(\Omega, \mathcal{T}, P)$ et on note (X, Y) l'application $\omega \mapsto (X(\omega), Y(\omega))$.

Proposition 11.

Soit X, Y des variables aléatoires discrètes à valeurs dans E, F respectivement. Alors le couple de variables aléatoires (X, Y) est une variable aléatoire discrète à valeurs dans $E \times F$.

Proposition 12.

Soit (X, Y) un couple de variables aléatoires discrètes. Alors la famille d'événements :

$$((X = x) \cap (Y = y))_{(x,y)\in X(\Omega)\cap Y(\Omega)}$$

est un système complet d'événements appelé **système complet d'événements associé au couple** (X, Y).

b. Loi conjointe

Définition 19. *Loi conjointe*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **loi conjointe de** (X, Y) et on note $P_{(X,Y)}$ la loi de probabilité du couple (X, Y).

Proposition 13.

Soit (X, Y) un couple de variables aléatoires discrètes. La loi $P_{(X,Y)}$ est déterminée par la donnée des $P((X = x) \cap (Y = y))$ pour chaque couple $(x, y) \in X(\Omega) \cap Y(\Omega)$.

c. Lois marginales

Définition 20. *Loi marginales*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **première loi marginale du couple** la loi de probabilité de X et **deuxième loi marginale du couple** la loi de probabilité de Y.

Théorème 7.

Soit (X, Y) un couple de variables aléatoires discrètes. On a :

— pour tout $x \in X(\Omega)$, $P(X = x) = \displaystyle\sum_{y \in Y(\Omega)} P((X = x) \cap (Y = y))$.

— pour tout $y \in Y(\Omega)$, $P(Y = y) = \displaystyle\sum_{x \in X(\Omega)} P((X = x) \cap (Y = y))$.

d. Lois conditionnelles

Définition 21. *loi conditionnelle*

Soit (X, Y) un couple de variables aléatoires discrètes et $x \in X(\Omega)$ tel que $P(X = x) = 0$. On appelle **loi conditionnelle de** Y **sachant** $X = x$ la probabilité notée $P_{(X=x)}$ sur $(\Omega, \mathcal{T})$ déterminée, pour $y \in Y(\Omega)$, par :

$$P_{(X=x)}(Y = y) = P(Y = y | X = x).$$

Exercice 7.

Soit (X, Y) un couple de variables aléatoires discrètes tel que, pour tout $x \in X(\Omega)$, $P(X = x) = 0$.

1. Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. Montrer que

$$P((X = x) \cap (Y = y)) = P(X = x) P(Y = y | X = x).$$

2. Soit $y \in Y(\Omega)$. Montrer que

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

Remarque 9.

On généralise toutes les notions précédentes aux cas d'un n-uplet de variables aléatoires.

5. Indépendance

Proposition 14.

Soit (X,Y) un couple de variables aléatoires discrètes. Les variables aléatoires discrètes X et Y sont indépendantes si, et seulement si, pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$,

$$P((X=x) \cap (Y=y)) = P(X=x)P(Y=y).$$

Proposition 15.

Soit (X,Y) un couple de variables aléatoires discrètes et f,g des fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement. Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 16. *Loi d'une somme de variables aléatoires discrètes*

Soit (X,Y) un couple de variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $n \in \mathbb{N}$,

$$P(X+Y=n) = \sum_{k=0}^{n} P(X=k)P(Y=n-k).$$

Exercice 8.

Soit $\lambda, \mu \in \mathbb{R}_+^*$ et (X,Y) un couple de variables aléatoires discrètes tel que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$. Montrer que si X et Y sont indépendantes, alors $X+Y \sim \mathcal{P}(\lambda+\mu)$.

Partie D

Espérance, variance

Sauf mention contraire, les variables aléatoires discrètes considérées dans cette partie définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et sont à valeurs réelles.

1. Espérance

a. Définitions et exemples

Définition 22. *Espérance*

Soit X une variable aléatoire discrète .

- Si X est à valeurs à valeurs dans $\mathbb{R}_+$, on appelle **espérance de** X et on note $E(X)$ la quantité (potentiellement infinie) :

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

- Si X est à valeurs dans $\mathbb{R}$, on dit que X **admet une espérance finie** si la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, on appelle **espérance de** X et on note $E(X)$ la quantité :

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

Exemple 3.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $E(X) = \dfrac{1}{p}$, Si $X \sim \mathcal{G}(p)$, alors pour tout $k \in \mathbb{N}^*$, $P(X = k) = p(1-p)^{k-1}$. D'après le critère de D'Alembert, $u_k = kp(1-p)^{k-1}$ est le terme général d'une série convergente car $0 < 1 - p < 1$. Par suite, X est d'espérance finie et

$$E(X) = \sum_{k=1}^{+\infty} kp(1-p)^{k-1}.$$

Calculons la valeur de cette somme. La série entière $\sum_{k \geq 0} x^k$ a pour rayon de convergence 1, donc sa somme $f : x \mapsto \frac{1}{1-x}$ est C^1 sur $]-1, 1[$ et on a, pour tout $x \in]-1, 1[$:

$$\begin{aligned} \frac{1}{(1-x)^2} = f'(x) &= \frac{d}{dx} \sum_{k=0}^{+\infty} x^k \\ &= \sum_{k=0}^{+\infty} \frac{d}{dx} x^k \\ &= \sum_{k=0}^{+\infty} kx^{k-1} \end{aligned}$$

Par suite, on a, comme $0 < 1 - p < 1$:

$$\begin{aligned} E(X) &= \sum_{k=1}^{+\infty} kp(1-p)^{k-1} \\ &= p \sum_{k=0}^{+\infty} k(1-p)^{k-1} \\ &= pf'(1-p) = \frac{p}{(1-(1-p))^2} \\ E(X) &= \frac{1}{p} \end{aligned}$$

— Si $X \sim \mathcal{P}(\lambda)$ alors $E(X) = \lambda$.

Si $X \sim \mathcal{P}(\lambda)$, alors pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$. D'après le critère de D'Alembert, $u_k = k\dfrac{\lambda^k}{k!}$ est le terme général d'une série convergente. Par suite, X est d'espérance finie et

$$E(X) = \sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!}.$$

Calculons la valeur de cette somme. On a :

$$\begin{aligned}
E(X) &= \sum_{k=0}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!}\\
&= p\sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!}\\
&= \lambda e^{-\lambda}\textstyle\sum_{k=1}^{+\infty}\dfrac{\lambda^{k-1}}{(k-1)!}\\
&= \lambda e^{-\lambda}\textstyle\sum_{k=0}^{+\infty}\dfrac{\lambda^{k}}{(k)!}\\
&= \lambda e^{-\lambda}e^{\lambda}\\
E(X) &= \lambda
\end{aligned}$$

Exercice 9.

1. Soit $c \in \mathbb{R}$ et X une variable aléatoire discrète presque sûrement égale à c. Calculer l'espérance de X.
2. Déterminer les espérances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Exercice 10.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ telle que, pour $n \in \mathbb{N}^*$, $P(X = n) = \dfrac{\alpha}{n^2}$ pour un certain $\alpha > 0$.

1. Déterminer α.
2. Quelle est l'espérance de X ?

b. Propriétés de l'espérance

Théorème 8. *Théorème de transfert*

Soit X une variable aléatoire discrète et $f : X(\Omega) \to \mathbb{R}$ une fonction. La variable aléatoire $f(X)$ est d'espérance finie si, et seulement si, la famille $(f(x)P(X = x))_{x\in X(\Omega)}$ est sommable. Dans ce cas, on a :

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

Théorème 9. *Linéarité de l'espérance*

Soit X, Y des variables aléatoires discrètes et $\lambda, \mu \in \mathbb{R}$. Si X et Y sont d'espérance finie, alors la variable aléatoire discrète $\lambda X + \mu Y$ est d'espérance finie et on a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Proposition 17. *Positivité/Croissance*

Soit X, Y des variables aléatoires discrètes.

— Si X est à valeurs positives, alors $E(X) \geq 0$ et $E(X) = 0$ implique $X = 0$ presque sûrement.

— Si X, Y sont d'espérance finie et $X \leq Y$, alors $E(X) \leq E(Y)$.

Proposition 18.

Soit X une variable aléatoire discrète. Alors X est d'espérance finie si, et seulement si, $|X|$ est d'espérance finie. Dans ce cas on a $|E(X)| \leq E(|X|)$.

Théorème 10. *Espérance d'un produit de variables indépendantes*

Soit X, Y des variables aléatoires discrètes. Si X et Y sont **indépendantes** et d'espérance finie, alors la variable aléatoire XY est d'espérance finie et on a :

$$E(XY) = E(X)E(Y).$$

Théorème 11. *Inégalité de Markov*

Soit X une variable aléatoire discrète à valeurs positives. Si X est d'espérance finie, alors, pour tout $a > 0$, on a :

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

2. Variance

a. Moments

Définition 23. *Moments*

Soit X une variable aléatoire discrète et $r \in \mathbb{N}$. On dit que X **admet un moment d'ordre** r si la variable aléatoire discrète X^r admet une espérance finie et dans ce cas, on appelle **moment d'ordre** r la quantité $E(X^r)$.

Proposition 19.

Soit X une variable aléatoire discrète. Si X admet un moment d'ordre 2, alors X est d'espérance finie.

Théorème 12. *Inégalité de Cauchy-Schwarz*

Soit X, Y des variables aléatoires discrètes. Si X et Y admettent des moments d'ordre 2, alors XY est d'espérance finie et on a :

$$E(XY)^2 \leq E(X^2)E(Y^2).$$

b. Variance, écart-type

Définition 24. *Variance*

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On appelle **variance de** X et on note $V(X)$ la quantité :

$$V(X) = E\left((X - E(X))^2\right).$$

Théorème 13. *Formule de Kœnig-Huygens*

Soit X une variable aléatoire admettant un moment d'ordre 2. Alors on a :

$$V(X) = E(X^2) - E(X)^2.$$

Exemple 4.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $V(X) = \dfrac{1-p}{p^2}$,

— Si $X \sim \mathcal{P}(\lambda)$ alors $V(X) = \lambda$.

Exercice 11.

1. Montrer les résultats de l'exemple précédent.
2. Déterminer les variances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Proposition 20.

Soit $\lambda, \mu \in \mathbb{R}$ et X une variable aléatoire discrète admettant un moment d'ordre 2. On a :

$$V(\lambda X + \mu) = \lambda^2 V(X).$$

Définition 25. *Écart-type*

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On appelle **écart-type de** X et on note $\sigma(X)$ la quantité :

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

Définition 26. *Variable centrée réduite*

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On dit que la variable aléatoire X est **centrée réduite** si $E(X) = 0$ et $\sigma(X) = 1$.

Proposition 21.

Soit X une variable aléatoire discrète admettant un moment d'ordre 2 telle que $V(X) \neq 0$. Alors la variable aléatoire $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est une variable aléatoire centrée réduite.

Théorème 14. *Inégalité de Bienaymé-Tchebychev*

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On a, pour tout $\varepsilon > 0$:

$$P\left(|X - E(X)| \geq \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2}.$$

3. Covariance

a. Généralités

Définition 27. *Covariance*

Soit (X, Y) un couple de variables aléatoires discrètes tel que X et Y admettent des moments d'ordre 2. On appelle **covariance du couple** (X, Y) et on note $\mathrm{Cov}(X, Y)$ la quantité :

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

Théorème 15. *Formule de Kœnig-Huygens*

Soit (X, Y) un couple de variables aléatoires discrètes tel que X et Y admettent des moments d'ordre 2. On a :

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

Proposition 22.

Soit (X, Y) un couple de variables aléatoires discrètes tel que X et Y admettent des moments d'ordre 2. Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0$.

b. Variance d'une somme

Théorème 16.

Soit (X, Y) deux variables aléatoires discrètes admettant un moment d'ordre 2. On a :

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y).$$

Soit $n \in \mathbb{N}^*$. On a, plus généralement, si $(X_1, ..., X_n)$ est un n-uplet de variables aléatoires discrètes, on a :

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

Corollaire 2.

Soit (X, Y) deux variables aléatoires discrètes admettant un moment d'ordre 2. Si X et Y sont indépendantes, on a :

$$V(X + Y) = V(X) + V(Y).$$

4. Loi faible des grands nombres

Théorème 17. *Loi faible des grands nombres*

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes, deux à deux indépendantes, de même loi et admettant un moment d'ordre 2. On pose $m = E(X_1)$ et, pour $n \in \mathbb{N}^*$, $S_n = X_1 + \ldots X_n$. Alors on a, pour tout $\varepsilon > 0$:

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow[n \to +\infty]{} 0.$$

5. Fonctions génératrices

a. Définition et premières propriétés

Définition 28. *Fonction génératrice*

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. On appelle **fonction génératrice de** X et on note G_X la fonction :

$$G_X : t \mapsto E(t^X).$$

Théorème 18.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice G_X de X est définie sur $[-1,1]$ et on a, pour tout $t \in [-1,1]$:

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

Proposition 23.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice de X est de classe C^∞ sur $]-1,1[$ et on a, pour tout $k \in \mathbb{N}$,

$$G^{(k)}(0) = k!P(X=k).$$

Corollaire 3.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$.

— La loi de probabilité P_X de X est entièrement déterminée par sa fonction génératrice G_X.

— les variables X et Y ont la même loi si, et seulement si, $G_X = G_Y$.

Proposition 24.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice.

$$G'(1) = E(X) \qquad G''(1) = E(X(X-1))$$

Exercice 12.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. Exprimer $E(X^2)$ et $V(X)$.

Exemple 5.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$,

— Si $X \sim \mathcal{P}(\lambda)$ alors $G_X(t) = e^{\lambda(1-t)}$.

Exercice 13.

1. Montrer les résultats de l'exemple précédent.
2. Déterminer les fonctions génératrices des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Théorème 19.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $t \in [-1, 1]$:

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

Exercice 14.

Soit $k, m, n \in \mathbb{N}^*$ tels que $k = nm$ et X, Y des variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes. On pose $Z = X + Y$ et on suppose que $X \sim \mathcal{U}([\![0, n-1]\!])$ et $Z \sim \mathcal{U}([\![0, k-1]\!])$.

1. Déterminer la fonction génératrice de Y.
2. En déduire la loi de Y.