"Probabilités sur un univers dénombrable"

Introduction

Une expérience aléatoire est un procédé donnant des résultats dépendant du hasard ou pouvant s'y assimiler. Dans le secondaire, les probabilités ont été étudiées lorsque l'univers était un ensemble fini. Nous allons généraliser à un univers dénombrable. Les probabilités apparaissent avec les travaux de Galilée pour le duc de Toscane, se développent au XVII$^{\text{e}}$siècle avec Pascal et Fermat, au XVIII$^{\text{e}}$siècle avec les travaux de Bayes, et enfin le cadre mathématique rigoureux apparaît avec Kolmogorov en 1933.

1 Probabilités sur un univers dénombrable

1.1 Expérience aléatoire et univers

Définition 1.1. On appelle **expérience aléatoire** une expérience sur un système dont le résultat n'est pas connu d'avance et peut varier si on répète cette expérience.

Exemple 1.1. Jeter une pièce de monnaie, lancer des dés, prélever des boules dans une urne, lancer une fléchette en direction d'une cible, observer la position d'une particule dans un liquide, ...

Le résultat, par hypothèse unique, de la réalisation de l'expérience aléatoire est noté ω.

Définition 1.2. On appelle **univers** l'ensemble des résultats possibles. Il est noté Ω.

Exemple 1.2. On effectue deux jets successifs d'un dé : $\Omega = [1;6]^2$.

Exemple 1.3. On lance indéfiniment une pièce de monnaie : un résultat possible est alors une suite de l'ensemble $\{F, P\}$, donc $\Omega = \{F, P\}^{\mathbb{N}}$.

La difficulté vient du fait qu'il est possible, pour une même expérience aléatoire, de définir plusieurs univers, suivant ce que l'on entend par le terme « résultat possible ».

Par exemple, pour une expérience aléatoire consistant à prélever une boule dans une urne contenant 2 boules rouges et 3 boules noires, on peut considérer qu'un résultat possible est une couleur ($\Omega = \{N, R\}$) ou l'une des 5 boules ($\Omega = \{R_1, R_2, N_1, N_2, N_3\}$). De même, pour le lanceur d'une fléchette contre une cible, on peut considérer comme résultat possible le point d'impact ($\Omega = \mathbb{R}^2$, après avoir muni le plan d'un repère), ou la trajectoire suivie par la fléchette ($\Omega = \mathcal{C}([0;1];\mathbb{R}^3)$, ensemble des applications continues de $[0;1]$ dans $\mathbb{R}^3$).

Remarque 1.1. Si on répète la même expérience aléatoire d'univers Ω, on pourra choisir comme univers Ω^n dans le cas de n répétitions, et $\Omega^{\mathbb{N}}$ si on la répète indéfiniment.

1.2 Événements

C'est une propriété $\mathcal{E}$ énonçable (c'est-à-dire accessible à l'expérience), vérifiée ou non selon le résultat obtenu. On l'identifie à l'ensemble des résultats de l'expérience pour lesquels elle est vérifiée :

$$\mathcal{E} \text{ est identifiée à } \{\omega \in \Omega \mid \mathcal{E} \text{ est vraie lorsque } \omega \text{ est réalisé}\}.$$

En pratique, on ne s'intéresse souvent qu'à un sous-ensemble d'événements. C'est le cas par exemple lorsque $\Omega = \mathbb{R}$: on considère les parties boréliennes, et pas toutes les parties de $\mathbb{R}$.

Notation 1.1. On note $\mathcal{A}$ l'ensemble des événements relatifs à l'expérience aléatoire. On a donc par définition $\mathcal{A} \subset \mathcal{P}(\Omega)$.

1.2.1 Pourquoi ne pas prendre toutes les parties de Ω ?

Lorsque Ω est fini ou dénombrable, on peut prendre $\mathcal{A} = \mathcal{P}(\Omega)$ sans problème. Mais quand Ω est infini non dénombrable (comme $\mathbb{R}$), on ne peut pas définir une probabilité sur *toutes* les parties de $\mathbb{R}$ tout en gardant des propriétés naturelles comme la σ-additivité et le fait que la probabilité d'un intervalle soit sa longueur. Voici un exemple concret.

Exemple (ensemble de Vitali). On considère $\Omega = [0, 1]$ et on aimerait avoir une probabilité P telle que $P([a, b]) = b - a$ pour tout intervalle $[a, b] \subset [0, 1]$ (c'est la loi uniforme). On peut montrer qu'il existe une partie $V \subset [0, 1]$ (appelée ensemble de Vitali) qui est « étrange » : si on prend toutes les translations de V par des nombres rationnels, on obtient une partition de $[0, 1]$ en une infinité dénombrable d'ensembles tous de même probabilité. Si on voulait assigner une probabilité à V, on aboutirait à une contradiction :

— Si $P(V) = 0$, alors par additivité dénombrable, $P([0, 1]) = \sum_{r\in\mathbb{Q}} P(V + r) = 0$, ce qui est impossible car $P([0, 1]) = 1$.

— Si $P(V) > 0$, alors $P([0, 1])$ serait infinie (somme d'une infinité de termes égaux à $P(V)$), ce qui est aussi impossible.

Ainsi, V ne peut pas être un événement auquel on attribue une probabilité. On est donc obligé de restreindre la collection des événements à une famille qui a de « bonnes » propriétés : une **tribu**.

1.2.2 Définition d'une tribu

Définition 1.3 (Tribu ou σ-algèbre)**.** Soit Ω un ensemble. Une famille $\mathcal{A} \subset \mathcal{P}(\Omega)$ est appelée **tribu** (ou σ-algèbre) si elle vérifie les trois propriétés suivantes :

1. $\Omega \in \mathcal{A}$.
2. Stabilité par complémentaire : pour tout $A \in \mathcal{A}$, on a $\overline{A} \in \mathcal{A}$.
3. Stabilité par réunion dénombrable : pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$, on a $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

À partir de ces axiomes, on déduit également que $\varnothing \in \mathcal{A}$ et que $\mathcal{A}$ est stable par intersection dénombrable.

On souhaite pouvoir définir certaines opérations sur les événements, et on utilise la correspondance suivante entre vocabulaires probabiliste et ensembliste :

Terminologie probabiliste	**Terminologie ensemble**	**Notation**
Événement certain	Ensemble entier	Ω
Événement impossible	Ensemble vide	$\varnothing$
Événement élémentaire	Singleton	$\{\omega\}$
Événement contraire de A	Complémentaire de A	$\overline{A}$
A et B	Réunion de A et B	$A \cup B$
A implique B	$A \subset B$	
A et B incompatibles	$A \cap B = \varnothing$	
ω réalise A	$\omega \in A$	

On exige que $\mathcal{A}$ contienne l'ensemble vide et l'univers, et qu'il soit stable par passage au complémentaire et par réunion dénombrable (c'est exactement la définition d'une tribu).

Exemple 1.4. 1. Si Ω est fini ou dénombrable, on s'intéresse naturellement aux événements élémentaires $\{\omega_i\}_{i\in I}$. La tribu engendrée est alors $\mathcal{P}(\Omega)$.

2. Si $\mathcal{F}$ est constitué d'un nombre fini ou dénombrable d'événements $(A_i)_{i\in I}$, qui forment une partition de Ω, la tribu engendrée est exactement l'ensemble des réunions quelconques d'événements A_i. Par exemple, en considérant une partition $\mathcal{F} = \{A, \overline{A}\}$ de Ω, on a $\mathcal{A} = \{\varnothing, \Omega, A, \overline{A}\}$.
3. Si $\Omega = \mathbb{R}$ ou $\Omega = \mathbb{R}^n$, on considère souvent comme tribu $\mathcal{A}$ la tribu des boréliens (engendrée par les ouverts, ou bien par les pavés $\prod_{i=1}^n [a_i; b_i]$, avec $a_i \leq b_i$, ou bien, dans le cas de $\mathbb{R}$, par les intervalles $]-\infty; a[$). On ne peut pas prendre $\mathcal{P}(\mathbb{R})$.

Définition 1.4. On appelle **événement** tout élément de la tribu $\mathcal{A}$.

Définition 1.5. On appelle **événement élémentaire** tout singleton (partie constituée d'un seul élément) de la tribu $\mathcal{A}$.

Exemple 1.5. On lance deux fois un dé à six faces. On pose $\Omega = [1; 6]^2$ et $\mathcal{A} = \mathcal{P}(\Omega)$. On considère A : « la somme obtenue est supérieure ou égale à 11 ». On a $A = \{(5; 6), (6; 5), (6; 6)\}$ et A est un événement. On considère B : « la somme obtenue est divisible par 3 et par 4 ». On a $B = \{(6; 6)\}$ et B est un événement élémentaire.

Remarque 1.2. Ne pas confondre résultat possible ($\omega \in \Omega$) et événement élémentaire ($\{\omega\} \in \mathcal{P}(\Omega)$). En particulier, on verra, lorsqu'on aura déjà une probabilité P, que l'écriture $P(\omega)$ n'a aucun sens, et on prendra bien garde à écrire $P(\{\omega\})$.

Définition 1.6. Le couple $(\Omega; \mathcal{A})$, où Ω est l'univers et $\mathcal{A}$ une tribu de parties de Ω, est appelé **espace probabilisable**.

1.3 Probabilité

Il s'agit d'affecter à chaque événement A un poids $P(A)$ indiquant la « chance » d'être réalisé si l'on effectue l'expérience aléatoire.

Des considérations relatives à l'expérience peuvent conduire, dans le cas où Ω est fini, à affecter des poids de probabilité identiques à chaque événement élémentaire : ce sera l'hypothèse d'équiprobabilité, par exemple lorsqu'on jette un dé équilibré, que l'on tire des cartes dans un jeu bien battu, que l'on prélève des boules indiscernables...

Une autre approche est l'approche « fréquentiste » : on se donne un événement fixé A, on répète n fois l'expérience aléatoire et on note n_A le nombre de fois où l'événement A a été réalisé. La fréquence de réalisation de A au cours de ces n répétitions est donc $\frac{n_A}{n}$. On démontre que, lorsque n tend vers l'infini, $\frac{n_A}{n}$ fluctue de moins en moins autour d'une valeur limite f_A (c'est la loi faible des grands nombres), appelée fréquence de réalisation de A.

La limite principale de cette approche est qu'elle ne concerne que des expériences aléatoires aisément répétables, et dans des conditions identiques. Cependant, l'application f possède les propriétés élémentaires suivantes :

1. $\forall A \in \mathcal{A},\ f_A \geq 0$ (positivité)
2. $f_\Omega = 1$ (totalité)
3. $\forall (A, B) \in \mathcal{A}^2,\ A \cap B = \varnothing \Rightarrow f_{A \cup B} = f_A + f_B$ (additivité)

Par analogie avec ces trois propriétés de la fréquence, on définit ce qu'est une probabilité de manière purement axiomatique (présentation due à Kolmogorov), sans référence à une quelconque observation de la réalisation de l'expérience aléatoire :

Définition 1.7. Étant donné un espace probabilisable $(\Omega; \mathcal{A})$, on appelle **probabilité** sur $(\Omega; \mathcal{A})$ toute application $P : \mathcal{A} \to \mathbb{R}$ satisfaisant aux trois conditions :

1. $\forall A \in \mathcal{A},\ P(A) \geq 0$ (positivité)
2. $P(\Omega) = 1$ (totalité)
3. Pour toute suite (A_n) d'éléments deux à deux disjoints,

$$P\left(\bigcup_n A_n\right) = \sum_{n=1}^{\infty} P(A_n) \quad (\sigma\text{-additivité})$$

Remarque 1.3. Avec le vocabulaire de la théorie de la mesure, P est donc une mesure positive finie, de masse totale égale à 1.

Définition 1.8. Soit $A \in \mathcal{A}$.

— Si $P(A) = 0$, on dit que A est un événement **presque impossible**.

— Si $P(A) = 1$, on dit que A est un événement **presque certain**.

Remarque 1.4 (Cas d'un univers fini)**.** On se place dans le cas d'un univers fini $\Omega = \{\omega_1, \dots, \omega_n\}$, et on s'intéresse aux événements élémentaires en posant comme tribu $\mathcal{A} = \mathcal{P}(\Omega)$. La donnée de n nombres réels positifs $p_1, \dots, p_n$ de somme égale à 1 permet de définir une probabilité P sur $\mathcal{A}$ en posant :

— $\forall i \in [1; n],\ P(\{\omega_i\}) = p_i$,

— $\forall A \in \mathcal{P}(\Omega),\ P(A) = \sum_{i:\omega_i \in A} p_i$ (car P doit être additive).

Exemple 1.6. Jet d'un dé, $\Omega = [1; 6]$. La donnée de $p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{10}$ et $p_6 = \frac{5}{10}$ définit une probabilité P. Si A est l'événement « le résultat est pair », alors

$$P(A) = p_2 + p_4 + p_6 = \frac{7}{10}.$$

On procède de la même façon lorsque Ω est dénombrable, en se donnant une suite de réels positifs $(p_i)_{i \in \mathbb{N}}$ tels que la série $\sum_{i \geq 0} p_i$ converge vers 1.

Remarque 1.5 (Cas particulier important : l'équiprobabilité)**.** Ω fini, de cardinal n. Des considérations relatives à l'expérience (dé équilibré, boules indiscernables, ...) peuvent conduire à définir une probabilité P sur $\mathcal{P}(\Omega)$ telle que

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\}).$$

On a alors nécessairement $P(\{\omega_i\}) = \frac{1}{n}$ pour tout i, et pour tout $A \in \mathcal{P}(\Omega)$,

$$P(A) = \frac{\mathrm{card}(A)}{n}.$$

1.4 Propriétés d'une probabilité

Propriété 1.1. Soit P une probabilité sur $(\Omega; \mathcal{A})$. Alors :

1. $P(\varnothing) = 0$.
2. $\forall (A, B) \in \mathcal{A}^2,\ A \subset B \Rightarrow P(A) \leq P(B)$.
3. $\forall A \in \mathcal{A},\ P(A) \in [0; 1]$ et $P(A) = 1 - P(\overline{A})$.
4. **Formule du crible de Poincaré** : $\forall (A_1, \dots, A_n) \in \mathcal{A}^n$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) + \cdots + (-1)^{n-1} P(A_1 \cap \cdots \cap A_n).$$

5. Si $(A_i)_{i\in\mathbb{N}}$ est une suite croissante d'événements (au sens de l'inclusion), alors la suite $(P(A_i))_{i\in\mathbb{N}}$ est croissante et

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcup_{i=0}^{+\infty} A_i\right).$$

6. Si $(A_i)_{i\in\mathbb{N}}$ est une suite décroissante d'événements, alors $(P(A_i))_{i\in\mathbb{N}}$ est décroissante et

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcap_{i=0}^{+\infty} A_i\right).$$

Remarque 1.6. Pour $n = 2$, la formule de Poincaré s'écrit $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$. Pour $n = 3$:

$$P(A_1\cup A_2\cup A_3) = P(A_1)+P(A_2)+P(A_3)-P(A_1\cap A_2)-P(A_1\cap A_3)-P(A_2\cap A_3)+P(A_1\cap A_2\cap A_3).$$

Définition 1.9. Le triplet $(\Omega; \mathcal{A}; P)$ est appelé **espace probabilisé**.

1.5 Probabilité conditionnelle

On se donne un événement B qui est réalisé et, disposant de cette information, on souhaite modifier la probabilité affectée aux événements A de la tribu $\mathcal{A}$ (ceux qui ont une intersection vide avec B auront alors naturellement une probabilité nulle, et B aura bien sûr une probabilité égale à 1). On ne change donc pas d'espace probabilisable (c'est toujours le couple $(\Omega, \mathcal{A})$), on modifie seulement l'application P.

Définition 1.10. Soit B un événement de probabilité non nulle. On appelle **probabilité conditionnelle à** B, ou probabilité sachant B, associée à P l'application

$$P_B : \mathcal{A} \to [0; 1], \qquad P_B(A) = \frac{P(A\cap B)}{P(B)}.$$

On note aussi $P(A \mid B)$ pour $P_B(A)$. Il faut alors être attentif au fait que c'est une simple notation, et que $A \mid B$ n'est pas un événement.

Propriété 1.2. L'application P_B est une probabilité sur $(\Omega; \mathcal{A})$.

Remarque 1.7 (Cas particulier de l'équiprobabilité)**.** Ω fini, $\mathcal{A} = \mathcal{P}(\Omega)$, P uniforme. Soit B de probabilité non nulle. Alors

$$P_B(A) = \frac{\mathrm{card}(A\cap B)}{\mathrm{card}(B)}.$$

Ainsi, pour tout $\omega \in \overline{B}$, $P_B(\{\omega\}) = 0$, et pour tout $\omega \in B$, $P_B(\{\omega\}) = \frac{1}{\mathrm{card}(B)}$.

Propriété 1.3. Soit B un événement de probabilité non nulle. Alors $\forall A \in \mathcal{A},\ P(A\cap B) = P(A \mid B)P(B)$.

Par récurrence, on obtient la formule des probabilités en chaîne :

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1}),$$

sous réserve que toutes les probabilités conditionnelles soient bien définies.

1.6 Formule des probabilités totales

Définition 1.11. On appelle **système complet d'événements** toute famille d'événements $(A_i)_{i \in I}$ (I étant une partie non vide de $\mathbb{N}$) telle que :

1. $\forall i \in I,\ A_i \neq \varnothing$,
2. $\forall (i,j) \in I^2,\ i \neq j \Rightarrow A_i \cap A_j = \varnothing$,
3. $\bigcup_{i \in I} A_i = \Omega$.

Propriété 1.4. Soit $(A_i)_{i \in I}$ un système complet d'événements. Alors on a la formule des probabilités totales :

$$\forall A \in \mathcal{A}, \qquad P(A) = \sum_{i \in I} P(A \mid A_i)\, P(A_i),$$

sous réserve que pour tout i, $P(A_i) \neq 0$ (afin que les probabilités conditionnelles soient bien définies).

En particulier, pour le système complet $\{B, \overline{B}\}$:

$$P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B}).$$

1.7 Indépendance de deux événements

On traduit le fait que la réalisation de l'un n'a pas d'influence sur la réalisation de l'autre.

Définition 1.12. A est dit **indépendant** de B si $P(A \mid B) = P(A)$ ou si $P(B) = 0$.

Propriété 1.5. A est indépendant de B si et seulement si $P(A \cap B) = P(A)P(B)$.

La propriété précédente montre que la relation est symétrique, et on dit alors que A et B sont indépendants.

Exemple 1.7. On jette un dé équilibré. $\Omega = [1;6]$, $\mathcal{A} = \mathcal{P}(\Omega)$, P uniforme. Soient $A = \{2,4,6\}$, $B = \{5,6\}$ et $C = \{5\}$. On a $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$, $P(C) = \frac{1}{6}$, $P(A \cap B) = \frac{1}{6}$, $P(A \cap C) = 0$, $P(B \cap C) = \frac{1}{6}$. Ainsi A et B sont indépendants, mais A et C ainsi que B et C ne le sont pas.

Remarque 1.8. Ne pas confondre incompatibilité ($A \cap B = \varnothing$) et indépendance ($P(A \cap B) = P(A)P(B)$). La notion d'indépendance dépend de la probabilité P.

Propriété 1.6. A et B indépendants $\iff$ A et $\overline{B}$ indépendants $\iff$ $\overline{A}$ et B indépendants $\iff$ $\overline{A}$ et $\overline{B}$ indépendants.

1.8 Indépendance de n événements

Définition 1.13. On dit que n événements $A_1, \ldots, A_n$ sont **mutuellement indépendants** si pour toute partie non vide $I \subset \{1, \ldots, n\}$,

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

La définition s'étend au cas d'une suite d'événements (A_n), en considérant l'indépendance mutuelle des événements de toute sous-famille finie.

Remarque 1.9. L'indépendance mutuelle entraîne l'indépendance deux à deux, mais la réciproque est fausse.

Exemple 1.8. On lance deux fois une pièce de monnaie équilibrée. $\Omega = \{P, F\}^2$, P uniforme. Soient $A =$ « Pile au premier lancer », $B =$ « Pile au second lancer », $C =$ « les deux jets ont donné le même résultat ». On a $P(A) = P(B) = P(C) = \frac{1}{2}$, $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, mais $P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8}$. Les événements A, B, C sont indépendants deux à deux mais pas mutuellement indépendants.

2 Variables aléatoires et lois usuelles

2.1 Variable aléatoire

Une variable aléatoire sur un univers Ω est une application définie sur Ω à valeurs dans un ensemble E.

Remarque 2.1. Une variable aléatoire n'est donc ni une variable (car c'est une application) ni aléatoire.

Une variable aléatoire est dite **réelle** si elle est à valeurs dans $\mathbb{R}$ et **complexe** si elle est à valeurs dans $\mathbb{C}$.

2.2 Loi d'une variable aléatoire

On utilise l'allégement $P(X = x)$ à la place de $P(\{X = x\})$.

Théorème 2.1. Si X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, alors l'application

$$P_X : \mathcal{P}(X(\Omega)) \to [0; 1], \qquad A \mapsto P(X \in A)$$

est une probabilité sur $X(\Omega)$ appelée **loi de** X et notée P_X.

Propriété 2.1. La loi d'une variable aléatoire X est déterminée de manière unique par la distribution de probabilités $(P(X = x))_{x \in X(\Omega)}$. Plus précisément,

$$\forall A \subset X(\Omega), \quad P_X(A) = \sum_{x \in A} P(X = x).$$

Notation 2.1. On note $X \sim Y$ lorsque $P_X = P_Y$, ce qui équivaut à $X(\Omega) = Y(\Omega)$ et $\forall x \in X(\Omega),\ P(X = x) = P(Y = x)$.

2.3 Lois usuelles

2.3.1 Loi uniforme

Soit E un ensemble fini non vide. On dit que X suit la **loi uniforme** sur E si X est à valeurs dans E et

$$\forall x \in E, \quad P(X = x) = \frac{1}{\mathrm{card}(E)}.$$

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

Exemple 2.1. Une urne contient n boules numérotées de 1 à n ; on en prend une au hasard et on note X le numéro tiré. Alors $X \sim \mathcal{U}([1, n])$.

2.3.2 Loi de Bernoulli

Soit $p \in]0, 1[$. On dit que X suit la **loi de Bernoulli** de paramètre p si elle est à valeurs dans $\{0, 1\}$ et $P(X = 1) = p$. On note $X \sim \mathcal{B}(p)$.

Exemple 2.2. Toute épreuve à deux issues peut être représentée par une loi de Bernoulli, à condition de noter 0 et 1 les deux issues possibles.

2.3.3 Loi binomiale

Soit $p \in]0, 1[$ et $n \in \mathbb{N}^*$. On dit que X suit la **loi binomiale** de paramètres (n, p) si X est à valeurs dans $[0, n]$ et

$$\forall k \in [0, n], \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

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

Exemple 2.3. On lance n fois une pièce équilibrée et on considère X égal au nombre de piles obtenu. Alors $X \sim \mathcal{B}(n, \frac{1}{2})$.

2.3.4 Loi géométrique

Soit $p \in]0,1[$. La variable aléatoire X suit la **loi géométrique** de paramètre p lorsque $P(X=0)=0$ et

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

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

2.3.5 Loi de Poisson

Soit $\lambda > 0$. Une variable aléatoire X suit la **loi de Poisson** de paramètre λ lorsque

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

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

3 Espérance et variance

3.1 Espérance d'une variable aléatoire discrète

Définition 3.1. La variable aléatoire réelle discrète X à valeurs dans un ensemble fini ou dénombrable $\{x_n,\ n \in I\}$ est dite **d'espérance finie** lorsque la famille $(x_n P(X=x_n))_{n\in I}$ est sommable. Si tel est le cas, on appelle **espérance** de X le réel

$$E(X) = \sum_{n\in I} x_n P(X=x_n).$$

On dit que X est **centrée** si $E(X)=0$.

Théorème 3.1. Si X est à valeurs dans $\mathbb{N}$, alors

$$E(X) = \sum_{n=1}^{\infty} P(X \geq n).$$

Théorème 3.2 (Théorème de transfert)**.** Si X est une variable aléatoire et f une application à valeurs réelles définie sur $X(\Omega)$, alors $f(X)$ est d'espérance finie si et seulement si la série $\sum_{n\geq 0} P(X=x_n)\, f(x_n)$ converge absolument. Dans ce cas,

$$E(f(X)) = \sum_{n\geq 0} P(X=x_n)\, f(x_n).$$

Théorème 3.3. Soient X, Y deux variables aléatoires réelles discrètes. On a :

— $\forall \alpha \in \mathbb{R},\ E(\alpha X + Y) = \alpha E(X) + E(Y)$.

— Si $X \geq 0$, alors $E(X) \geq 0$.

Théorème 3.4. Si X et Y sont deux variables aléatoires discrètes indépendantes, alors $E(XY) = E(X)E(Y)$.

3.2 Variance

Théorème 3.5 (Théorème de Koenig-Huyghens)**.** Si X^2 est d'espérance finie, alors X est elle-même d'espérance finie et

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

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

Théorème 3.6. $\forall (a, b) \in \mathbb{R}^2,\ V(aX + b) = a^2 V(X)$.

Remarque 3.1. Une variable aléatoire admettant une variance est dite **centrée réduite** lorsque $E(X) = 0$ et $V(X) = 1$. Par exemple, si $V(X) \neq 0$, alors $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Exercice corrigé : Espérance et variance des lois géométrique et de Poisson

Exercice 3.1. Soit $p \in]0, 1[$ et $X \sim \mathcal{G}(p)$ (loi géométrique : $P(X = k) = p(1 - p)^{k-1}$ pour $k \geq 1$). Soit $\lambda > 0$ et $Y \sim \mathcal{P}(\lambda)$ ($P(Y = n) = e^{-\lambda} \frac{\lambda^n}{n!}$, $n \geq 0$). Calculer $E(X)$, $V(X)$, $E(Y)$ et $V(Y)$.

Corrigé 3.1. 1. Loi géométrique. On rappelle $\sum_{k\geq 1} kq^{k-1} = 1/(1-q)^2$ et $\sum_{k\geq 1} k^2 q^{k-1} = (1+q)/(1-q)^3$ pour $|q| < 1$. $E(X) = p \sum_{k\geq 1} k(1-p)^{k-1} = p \cdot \frac{1}{p^2} = \frac{1}{p}$. $E(X^2) = p \sum_{k\geq 1} k^2 (1-p)^{k-1} = p \cdot \frac{2-p}{p^3} = \frac{2-p}{p^2}$. D'où $V(X) = E(X^2) - E(X)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

2. Loi de Poisson. $E(Y) = \sum_{n\geq 0} n e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \lambda e^{\lambda} = \lambda$. $E(Y^2) = E(Y(Y-1)) + E(Y) = \lambda^2 + \lambda$. Donc $V(Y) = E(Y^2) - \lambda^2 = \lambda$.

$$\boxed{E(X) = \frac{1}{p},\ V(X) = \frac{1-p}{p^2}, \qquad E(Y) = \lambda,\ V(Y) = \lambda.}$$