Chapitre C1
Probabilités

I. Généralités

A. Univers et événements

Définition. On réalise une *expérience aléatoire* et on définit un ensemble Ω appelé univers, contenant les résultats possibles.

Remarque. Ainsi l'univers est juste un ensemble. On dit qu'il *modélise* l'expérience. Ces issues possibles sont aussi appelées les *observables*.

Exemple 1.

(i) On jette une pièce, elle tombe sur pile ou face.
Pour cette expérience aléatoire on choisit $\Omega = \{P, F\}$.

(ii) On lance un dé à six faces numérotées de 1 à 6. On choisit $\Omega = \{1, 2, 3, 4, 5, 6\}$.

(iii) On jette trois dés simultanément. On choisit pour univers :

(iv) On joue au tiercé, 21 chevaux sont au départ. On choisit pour univers :

Remarque. Le programme de première année est limité au cas où l'univers est fini. Dans ce cas on peut noter :

$$\Omega = \{\omega_1, \dots, \omega_n\} \qquad \text{avec} \qquad n \in \mathbb{N}.$$

Définitions. Un événement est un sous-ensemble de Ω.

Exemple 2. On lance trois dés. On définit les événements A : «les trois dés donnent le même numéro», B : «on obtient 421» et C : «la somme des dés est égale à 5». Alors :

Définitions.

- L'événement Ω est dit certain.
- L'événement $\varnothing$ est dit impossible.
- Si ω est un élément de Ω alors le singleton $\{\omega\}$ est appelé événement élémentaire.

Définitions. Si A et B sont deux événements alors :

- $A \cap B$ est l'événement «A et B», il a lieu si A et B ont lieu tous les deux.
- $A \cup B$ est l'événement «A ou B», il a lieu si A ou B a lieu.
- $\overline{A}$ est l'événement contraire de A, ou l'événement «non A», il a lieu si et seulement si A n'a pas lieu.
- On note $A - B = A \cap \overline{B}$, cet événement a lieu lorsque A a lieu mais pas B.
- A et B sont incompatibles s'ils ne peuvent avoir lieu ensemble, donc si $A \cap B = \varnothing$.
- A et B sont complémentaires si $A = \overline{B}$. Alors $A \cap B = \varnothing$ et $A \cup B = \Omega$.
- On dit que A implique B si $A \subseteq B$.

Définition. Un système complet d'événements est un ensemble $(A_i)_{1 \leqslant i \leqslant n}$ d'événements tel que :

- $\bigcup\limits_{i=1}^{n} A_i = \Omega$
- Pour tout $(i, j) \in \{1, \ldots, n\}^2$, si $i \neq j$ alors $A_i \cap A_j = \varnothing$.

Remarque. On dit aussi que l'ensemble des parties A_i est une partition de Ω.

Exemple. Pour tout événement A de Ω, la famille $\left(A, \overline{A}\right)$ est un système complet d'événements.

B. Probabilité

Définition. Une probabilité sur un univers fini Ω est une application

Définition. Un espace probabilisé fini est un couple (Ω, P) où Ω est un univers fini et P est une probabilité sur Ω.

Propositions. *Soit (Ω, P) un espace probabilisé. Alors pour tous événements A et B :*

(i)

(ii) $P(\overline{A}) =$

(iii) $P(\varnothing) =$

(iv) *Si* $A \subseteq B$ *alors*

(v) $P(A - B) =$

(vi) $P(A \cup B) =$

Démonstration.

(i) P est une application de $\mathcal{P}(\Omega)$ dans $[0,1]$ donc pour tout $A \in \mathcal{P}(\Omega)$: $P(A) \in [0,1]$

(ii) Les événements A et $\overline{A}$ sont incompatibles, et $A \cup \overline{A} = \Omega$.
Par additivité $P(A) + P(\overline{A}) = P(\Omega)$. Comme $P(\Omega) = 1$ alors $P(\overline{A}) = 1 - P(A)$.

(iii) Comme $\varnothing = \overline{\Omega}$ alors d'après le point précédent $P(\varnothing) = 1 - P(\Omega) = 0$.

(iv) On note $C = B - A$. Alors $B = A \cup C$ et $A \cap C = \varnothing$, *i.e.,* A et C sont incompatibles.
Par additivité $P(B) = P(A) + P(C)$. Comme $P(C) \geqslant 0$ d'après le point *(i)* alors $P(B) \geqslant P(A)$.

(v) On sait que $A = (A - B) \cup (A \cap B)$ et $(A - B) \cap (A \cap B) = \varnothing$.
Par additivité $P(A) = P(A - B) + P(A \cap B)$, d'où le résultat.

(vi) On voit que $A \cup B = (A - B) \cup B$ et $(A - B) \cap B = \varnothing$.
Par additivité $P(A \cup B) = P(A - B) + P(B)$, puis grâce au point précédent on obtient le résultat. □

▷ **Exercice 1.**

Propositions. *Soit* (Ω, P) *un espace probabilisé,* $(A_1, \dots, A_n)$ *une famille d'événements.*

(i) (Additivité) Si les A_i *sont deux-à-deux incompatibles alors :*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

(ii) Dans tous les cas :

$$P(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} P(A_i).$$

<u>Démonstration</u>. On démontre ces propriétés simultanément par récurrence sur $n \in \mathbb{N}^*$.

<u>Initialisation.</u> Si $n = 1$ les propriétés sont évidentes.

<u>Hérédité.</u> Supposons que les propriétés sont acquises pour un $n > 0$ donné, et considérons une famille d'événements $A_1, \dots, A_{n+1}$. On note $A = A_1 \cup \dots \cup A_n$ et $B = A_{n+1}$.

(i) Supposons que les A_i sont deux-à-deux incompatibles. Alors A et B sont incompatibles. En effet par distributivité :

$$A \cap B = (A_1 \cup \dots \cup A_n) \cap A_{n+1} = (A_1 \cap A_{n+1}) \cup \dots \cup (A_1 \cap A_{n+1}) = \varnothing.$$

Ainsi l'additivité montre que $P(A \cup B) = P(A) + P(B)$, ce qui donne grâce à l'hypothèse de récurrence :

$$P(\bigcup_{i=1}^{n+1} A_i) = P(A \cup B) = P(A) + P(B) = \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) = \sum_{i=1}^{n+1} P(A_i).$$

(ii) Dans le cas général, on sait que :

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

Comme $P(A \cap B) \geqslant 0$ alors :

$$P(A \cup B) \leqslant P(A) + P(B).$$

Grâce à l'hypothèse de récurrence on obtient le résultat :

$$P(\bigcup_{i=1}^{n+1} A_i) = P(A \cup B) \leqslant P(A) + P(B) = P(\bigcup_{i=1}^{n} A_i) + P(A_{n+1}) \leqslant \sum_{i=1}^{n+1} P(A_i).$$

Les deux propriétés sont donc vraies au rang $n + 1$ si elles sont vraies au rang n.

<u>Conclusion.</u> Par récurrence les deux propriétés sont vraies pour tout $n \in \mathbb{N}^*$. □

Corollaire 1. *Soit* $(A_i)_{1 \leqslant i \leqslant n}$ *un système complet d'événements. Alors :*

$$\sum_{i=1}^{n} P(A_i) = 1.$$

Corollaire 2. *Si* Ω *est un univers fini alors :*

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

Proposition. *Une probabilité sur un univers fini est entièrement déterminée par les images des événements élémentaires.*
Plus précisément si $\Omega = \{\omega_1, \dots, \omega_n\}$ *et* $p_1, \dots, p_n$ *sont des réels tels que :*
(i) $\forall k = 1, \dots, n \quad p_k \in [0, 1]$
(ii) $\sum_{k=1}^{n} p_k = 1$
Alors il existe une et une seule probabilité P *sur* Ω *telle que :*

$$\forall k = 1, \dots, n \qquad P(\{\omega_k\}) = p_k.$$

Cette probabilité est définie par :

$$\forall A \in \mathcal{P}(\Omega) \qquad P(A) = \sum_{\omega_k \in A} p_k.$$

Remarque. Soit Ω un ensemble fini. Une famille $(p_\omega)_{\omega \in \Omega}$ de réels positifs de somme égale à 1 est appelée distribution de probabilité sur Ω.
La proposition ci-dessus signifie que les distributions de probabilités sur Ω correspondent bijectivement aux probabilités sur Ω.

▷ **Exercice 2.**

Définition. Soit Ω un univers fini de cardinal $n \in \mathbb{N}^*$. On définit alors une probabilité P sur Ω en posant :

$$\forall \omega \in \Omega \qquad P(\{\omega\}) = \frac{1}{n}.$$

Cette probabilité est appelée probabilité uniforme sur Ω.
On dit également qu'il y a équiprobabilité sur Ω.

Proposition. *Si les événements élémentaires de* Ω *sont équiprobables, alors pour tout événement* A *:*

$$P(A) = \frac{\operatorname{Card} A}{\operatorname{Card} \Omega}.$$

Exemple 3. On possède une urne contenant 10 boules numérotées de 1 à 10.
Trois fois de suite on pioche une boule, on note le numéro obtenu et on la remet dans l'urne.

a. Donner un espace probabilisé modélisant cette expérience.
b. Calculer la probabilité des événements :
   - A : «On obtient uniquement des 10.»
   - B : «On n'obtient jamais le 10.»
   - C : «On obtient au moins une fois le 10.»
   - D : «On obtient exactement une fois le 10.»

▷ **Exercices 3, 4.**

II. Probabilités conditionnelles

A. Définition

Définition. Soit (Ω, P) un espace probabilisé et A un événement de probabilité non-nulle. Pour tout $B \subseteq \Omega$ on note :

Ce réel est appelé probabilité de B sachant A.

Proposition. *P_A est une probabilité sur Ω.*

Définition. On dit que P_A est une probabilité conditionnelle.
Plus précisément c'est la probabilité conditionnée par A.

Remarque. Les propriétés d'une probabilité donnent des formules comme :

- $P_A(\overline{B}) = 1 - P_A(B)$
- Si $C \subseteq B$ alors : $P_A(C) \leqslant P_A(B)$
- Si $(B_1, \ldots, B_n)$ est un système complet d'événements alors : $\displaystyle\sum_{i=1}^{n} P_A(B_i) = 1$

Démonstration.

▷ **Exercice 5.**

B. Propriétés

Proposition. *Soit A et B deux événements de probabilités non-nulles. Alors :*

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

Démonstration. Il suffit d'utiliser la définition des probabilités conditionnelles :

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \qquad \text{et} \qquad P_B(A) = \frac{P(A \cap B)}{P(B)}$$ □

Théorème : Formule des probabilités composées. *Soit n un entier naturel non-nul, et $(A_i)_{1 \leqslant i \leqslant n}$ une famille d'événements telle que $P\left(\bigcap_{i=1}^{n-1} A_i\right)$ est non-nulle. Alors :*

Exemple. Si $A_1 = A$ et $A_2 = B$ alors

Démonstration.

▷ **Exercice 6.**

Remarque. Lien avec les arbres de probabilités.

Théorème : Formule des probabilités totales. *Soit* $(A_i)_{1\leqslant i\leqslant n}$ *un système complet d'événements de probabilités non-nulles. Alors pour tout événement* B *:*

Exemple. Soit A un événement ni impossible, ni certain. Alors pour tout événement B :

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

<u>Démonstration de la formule des probabilité totales.</u>

Exemple 4. On possède un dé et une pièce. On jette le dé. Si on obtient le chiffre k alors on jette k fois la pièce. Quelle est la probabilité de n'obtenir que des Piles ?

Théorème : inversion cause-effet. *Soit A et B deux événements de probabilités non-nulles. Alors :*

Démonstration.

Remarque : Formule de Bayes. En combinant les deux formules précédentes, on obtient la proposition suivante.

Soit $(A_i)_{1\leqslant i\leqslant n}$ un système complet d'événements de probabilités non-nulles et B un événement de probabilité non-nulle. Alors pour tout $k \in \{1, \ldots, n\}$:

Exemple 4 (suite). On n'a obtenu que des Piles. Quelle est la probabilité que le dé ait donné k ?

▷ **Exercice 7.**

C. Événements indépendants

Définition. Deux événements A et B sont dits indépendants pour la probabilité P si :

Proposition. *On suppose que A est de probabilité non-nulle.*
Dans ce cas A et B sont indépendants si et seulement si :

▷ **Exercice 8.**

Exemple 5. Une urne contient $n \geqslant 2$ boules dont a sont noires et les autres sont blanches. On pioche deux boules de suite dans cette urne, sans remise. On note A l'événement «la première boule est noire» et B l'événement «la seconde boule est noire».

(i) Calculer $P(A)$, $P(B)$, puis $P(A \cap B)$.

(ii) A et B sont-ils indépendants ?

▷ **Exercice 9.**

Définition. Des événements $A_1, \ldots, A_n$ sont mutuellement indépendants si pour tous entiers $i_1 \ldots i_m$ tels que $1 \leqslant i_1 < \cdots < i_m \leqslant n$:

$$P(A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_{i_1}) \times \cdots \times P(A_{i_m}).$$

Remarque (autre écriture). Soit $I = \{1, \ldots, n\}$. Les A_i sont mutuellement indépendants si

$$\forall J \subseteq I \qquad P\left(\bigcap_{j \in J} A_j \right) = \prod_{j \in J} P(A_j).$$

Remarque. L'indépendance mutuelle a lieu dès que l'on répète des expériences identiques et indépendantes.

Exemple. On jette n fois un dé, et on note A_i l'événement : le tirage i donne un 6. Alors les événements $A_1, \ldots, A_n$ sont mutuellement indépendants.

Remarque. L'indépendance des A_i deux-à-deux n'implique pas leur indépendance mutuelle.

▷ **Exercice 10.**

III. Variables aléatoires

A. Définitions

Définition. Soit Ω un univers fini et E un ensemble.

Une variable aléatoire sur Ω à valeurs dans E est une fonction de Ω dans E.

Si X est une variable aléatoire alors $X(\Omega)$ est l'ensemble des valeurs qu'elle peut prendre.

Une variable aléatoire réelle est une variable aléatoire $X : \Omega \to \mathbb{R}$.

Une variable aléatoire réelle est entière si $X(\Omega) \subseteq \mathbb{Z}$.

Notations. Soit X une variable aléatoire réelle. On note

- $X = x$ l'événement $X^{-1}(\{x\})$ où x est un réel
- $X \leqslant x$ l'événement $X^{-1}(]-\infty, x])$
- $a \leqslant X \leqslant b$ l'événement $X^{-1}([a, b])$ où a et b sont deux réels
- $X \in A$ l'événement $X^{-1}(A)$ où A est une partie de $\mathbb{R}$
- etc.

Proposition. *L'ensemble* $\{(X = x) \mid x \in X(\Omega)\}$ *est un système complet d'événements.*

Démonstration. Tout élément ω de Ω admet une image $x = X(\omega)$ dans $X(\Omega)$, donc appartient à l'événement $(X = x)$. Ceci montre que $\Omega = \bigcup_{x \in X(\Omega)} (X = x)$.

Si x et y sont deux éléments distincts de $X(\Omega)$, alors $(X = x) \cap (X = y) = \varnothing$ car si un élément ω était dans cette intersection alors on aurait $X(\omega) = x$ et $X(\omega) = y$. $\square$

Remarques.

(i) L'événement $(X = x)$ est l'ensemble des éléments ω de Ω tels que $X(\omega) = x$, donc il existe plusieurs façons d'écrire cette partition :

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

(ii) Plus généralement, si $f : E \to F$ est une application surjective alors la famille $\{f^{-1}(\{y\}) \mid y \in F\}$ est une partition de E.

Définition. Une loi de probabilité de X est une application $P_X : X(\Omega) \to \mathbb{R}$ vérifiant :

Remarques.

(i) Si $P : \mathcal{P}(\Omega) \to \mathbb{R}$ est une probabilité sur Ω alors l'application $x \mapsto P(X = x)$ est une loi de probabilité de X.

On notera généralement $P(X = x)$ au lieu de $P_X(x)$.

(ii) Si $p : X(\Omega) \to \mathbb{R}$ est une loi de probabilité de X alors :

$$\forall x \in E \qquad 0 \leqslant p(x) \leqslant 1.$$

Ceci car tous les $p(x)$ sont positifs et leur somme est égale à 1, donc ils sont tous inférieurs ou égaux à 1 :

$$1 - p(x) = \sum_{\substack{y \in X(\Omega) \\ y \neq x}} p(y) \geqslant 0.$$

(iii) Notons $X(\Omega) = \{x_1, \dots, x_n\}$. Alors une loi de probabilité de X est une application $x_i \mapsto p_i$ où les p_i sont des éléments de $[0, 1]$ dont la somme est égale à 1.

Définition. La fonction $x \mapsto P(X \leqslant x)$ est appelée fonction de répartition de X.

Remarque. On suppose que X est entière. Alors pour tout $k \in \mathbb{Z}$:

$$P(X = k) =$$

Ceci permet de calculer certaines lois de probabilité.

▷ **Exercice 11.**

B. Espérance, variance, écart-type

Définitions. Soit X une variable aléatoire réelle sur un univers fini.

- L'espérance mathématique de X est le réel :

- La variance de X est l'espérance de la variable aléatoire $(X - E(X))^2$:

- L'écart-type de X est la racine carrée de la variance :

Remarques.

(i) L'espérance mathématique correspond à la moyenne d'une série statistique.
C'est la valeur moyenne de X sur un grand nombre d'expériences.

(ii) La variance et l'écart type sont des paramètres de dispersion. Plus ils sont grands, plus la variable aléatoire X peut être éloignée de son espérance. On constate en général que
- Dans 60% des cas X est dans l'intervalle $[E(X) - \sigma(X), E(X) + \sigma(X)]$.
- Dans 95% des cas X est dans l'intervalle $[E(X) - 2\sigma(X), E(X) + 2\sigma(X)]$.

(iii) Si X a une unité, alors $E(X)$ et $\sigma(X)$ ont également cette unité.

▷ **Exercice 12.**

Proposition (définition alternative de l'espérance). *Soit X une variable aléatoire sur un univers fini. Alors :*

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

Démonstration. L'écriture $(X = x) = \{\omega \in \Omega \mid X(\omega) = x\}$ donne :

$$E(X) = \sum_{x \in X(\Omega)} x P(X = x) = \sum_{x \in X(\Omega)} \left(x \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P(\{\omega\}) \right).$$

On en déduit :

$$E(X) = \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} x P(\{\omega\}) = \sum_{x \in X(\Omega)} \sum_{\omega \in X^{-1}(\{x\})} X(\omega) P(\{\omega\}).$$

L'ensemble des $X^{-1}(\{x\})$ pour x parcourant $X(\Omega)$ est une partition de Ω donc :

$$E(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}).$$ □

C. Propriétés

Théorème de transfert. *Soit X une variable aléatoire définie sur un univers fini Ω. Soit f une fonction de $X(\Omega)$ dans $\mathbb{R}$. Alors :*

Remarque. Si $f : X(\Omega) \to \mathbb{R}$ est une fonction alors $Y = f(X)$ est une variable aléatoire réelle sur Ω.

De plus $Y(\Omega) = f(X(\Omega))$, et si P_X est une loi de probabilité de X alors on obtient une loi de probabilité P_Y de Y en posant pour tout $y \in Y(\Omega)$:

$$P_Y(y) = P_X(f^{-1}(\{y\})) = \sum_{\substack{x \in X(\Omega) \\ f(x)=y}} P_X(x).$$

On écrit plutôt : $\quad P(Y = y) = P(X \in f^{-1}(\{y\})) = \sum_{\substack{x \in X(\Omega) \\ f(x)=y}} P(X = x)$

Exemple. Dans le cas où $f(x) = x^2$: $\quad P(X^2 = 4) = P(X = 2) + P(X = -2)$

Démonstration. En notant $Y = f(X)$:

$$E(Y) = \sum_{y \in Y(\Omega)} yP(Y = y) = \sum_{y \in Y(\Omega)} \left(\sum_{\substack{x \in X(\Omega) \\ f(x)=y}} yP(X = x) \right) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

Il s'agit bien du résultat que l'on voulait démontrer. □

Corollaire. *Soit a et b deux réels. Alors :*

Démonstration pour l'espérance.

▷ **Exercice 13.**

Exemple. La variable aléatoire $Y = X - E(X)$ est dite *centrée*, elle vérifie :

$E(Y) =$ $V(Y) =$

La variable aléatoire $T = \frac{X-E(X)}{\sigma(X)}$ est dite *centrée réduite*, elle vérifie :

$E(T) =$ $\sigma(T) =$

Proposition (Formule de König-Huyghens). Johann Samuel König (ou Kœnig), Allemagne, 1712 – 1757, Christian Huygens (ou Huyghens), Hollande, 1629 – 1695.
Pour toute variable aléatoire X sur un univers fini Ω :

Remarque. On utilise le théorème de transfert pour calculer $E(X^2)$:

Démonstration.

▷ **Exercices 14, 15.**