Probabilités, modélisation et statistique

Chapitre 1 - Probabilité : Introduction et rappels

Raphaël Benerradi

Contenu pédagogique : Gwladys Toulemonde, Chloé Serre-Combe et Raphaël Benerradi

Polytech Montpellier - DevOps3 - Semestre 6

Année 2025-2026

Motivations du cours

Modéliser des phénomènes aléatoires ! Ça permet plein de chose, notamment pour l'étude du monde réel :

- tester des hypothèses sur un phénomène
  - *est-ce que la crème quantique fait rajeunir ma peau ?*
  - *existe-t-il des inégalités de genre dans l'accès aux disciplines scientifiques ?*
  - *le changement climatique est-il d'origine anthropique ?*
- extraire de l'information de phénomènes complexes
  - *quels facteurs influencent le vieillissement de ma peau ?*
  - *comment expliquer les inégalités de genre dans les disciplines scientifiques ?*
  - *qu'est-ce qui impacte la répartition géographique des forêts ?*
- prédire des événements pas encore observés
  - *à quel âge vont apparaître mes premiers signes de vieillissement de la peau ?*
  - *quel peut être l'effet d'une politique publique sur l'accès des femmes aux carrières scientifiques ?*
  - *quel sera la répartition des forêts sous tel scénario de changement climatique ?*

Motivations du cours

C'est relié à plein de domaines d'expertise scientifique !

- mathématiques : probabilités, statistiques, processus stochastiques, etc.
- informatique : apprentissage automatique, traitement du signal, etc.
- physique : mécanique quantique, physique statistique, thermodynamique, etc.

Ça s'applique pour plein de types de métiers !

- recherche académique ou industrielle
- fonction publique (ex : INSEE, ministères...)
- secteur privé (grandes entreprises, start-ups...)
- associations
- (voire même dans la vie quotidienne)

C'est utile (voir crucial) dans plein de domaines d'application !

- Mathématiques, informatique, physique
- Toutes les sciences expérimentales : biologie (santé, agronomie, écologie...), physique (mécanique, astronomie...), etc.
- En sciences humaines et sociales : psychologie, sociologie, économie, démographie, etc.
- (voire même dans la vie quotidienne)

Organisation du cours

- 9 séances de 3 heures (mix CM et TD)
- Notation
  - Contrôle continu (CC) : 3 interros 25 min en début de séance séances n°3 (../..), n°6 (../..), n°7 (../..)
  - Examen final (EF) : 1h30 (../..)
  - Note $= \max(0.3 \times \mathrm{CC} + 0.7 \times \mathrm{EF}, \mathrm{EF})$
  - Rattrapage : oral
- Contact : raphael.benerradi@umontpellier.fr
- Ressources : *https://raphbnrd.github.io/teaching/* section "Probabilités, modélisation et statistique"
- Autres ressources :
  - Livre de cours avec des exercices : **Jean-Pierre Lecoutre** (2023) *Statistique et probabilités - 7e éd.: Cours et exercices corrigés*
  - Livre de cours très complet : **Gilbert Saporta** (2011) *Probabilités, analyse des données et statistique*

Plan du cours

1. Probabilités : Introduction et rappels
   - Avoir des bases propres en probabilité
2. Variables aléatoires
   - Les outils pour observer une expérience aléatoire
3. Vecteurs aléatoires et indépendance
   - Observer le résultat de plusieurs expériences aléatoires
4. Convergence des variables aléatoires
   - Ce qui se passe quand on observe beaucoup d'expériences aléatoires
5. Statistique inférentielle
   - Tirer de l'information à partir d'observations (estimation, incertitude, confiance)

Vocabulaire de base (probabilités)

Définition : expérience aléatoire (épreuve)

Une **expérience aléatoire** est une expérience dont les résultats sont dus au hasard. Même si elle est répétée dans les mêmes conditions, elle ne donnera pas nécessairement les mêmes résultats.

Définition : éventualité

Une **éventualité** est un résultat possible d'une expérience aléatoire. On la notera en général ω.

Vocabulaire de base (probabilités)

Définition : univers (ensemble fondamental)

L'**univers** est l'ensemble de toutes les éventualités (tous les résultats possibles) d'une expérience aléatoire. On le notera en général Ω.

Définition : événement

Un **événement** est un sous-ensemble de l'univers Ω.

Exemples

1. On jette un dé une fois et on regarde le résultat.
   - Les éventualités sont 1, 2, 3, 4, 5, 6.
   - $\Omega = \{1, 2, 3, 4, 5, 6\}$.
   - "Le résultat est un nombre pair" est un événement : $\{2, 4, 6\} \subset \Omega$.

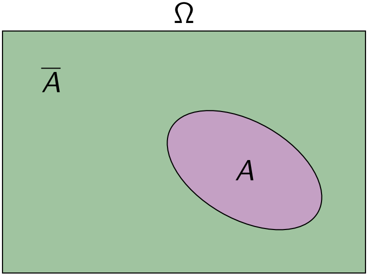
2. On considère une population animale composée à part égale d'individus des deux sexes (F : Femelle, M : Mâle).
   On s'intéresse à l'épreuve suivante : extraire trois individus de cette population.
   - FMF est un exemple d'éventualité.
   - $\Omega = \{FFF, FFM, FMF, MFF, MMF, MFM, FMM, MMM\}$.
   - "Extraire au moins deux individus femelles" est un événement : $\{FFF, FFM, FMF, MFF\} \subset \Omega$.

Le complémentaire

Définition : événement complémentaire (contraire)

Soit A un événement. L'événement **complémentaire** (ou **contraire**) de A, noté $\overline{A}$ (ou A^c), est l'événement constitué des éventualités de Ω qui ne sont pas dans A.

$$\overline{A} = \{\omega \in \Omega \mid \omega \notin A\}$$

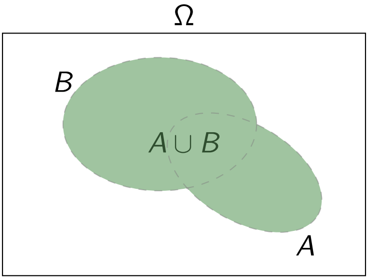


L'union

Définition : union d'événements

Soit A et B deux événements. L'**union** de A et B, notée $A \cup B$, est l'événement constitué des éventualités qui sont dans A, dans B, ou dans les deux.

$$A \cup B = \{\omega \in \Omega \mid \omega \in A \text{ ou } \omega \in B\}$$

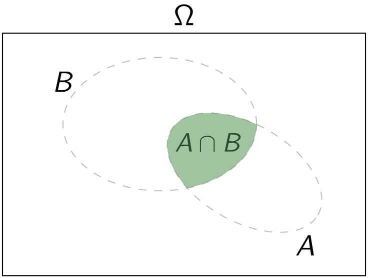


L'intersection

Définition : intersection d'événements

Soit A et B deux événements. L'**intersection** de A et B, notée $A \cap B$, est l'événement constitué des éventualités qui sont à la fois dans A et dans B.

$$A \cap B = \{\omega \in \Omega \mid \omega \in A \text{ et } \omega \in B\}$$

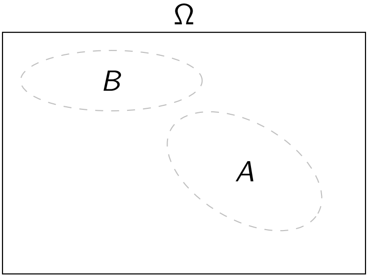


Evénements disjoints

Définition : événements disjoints

Deux événements A et B sont dits **disjoints** s'ils n'ont aucun élément en commun, c'est-à-dire si leur intersection est vide :

$$A \cap B = \emptyset$$



Propriétés sur l'union et l'intersection

Soient A, B et C trois événements de l'univers Ω.

1. $A \cup \overline{A} = \Omega$
2. $A \setminus B = A \cap \overline{B}$ avec $A \setminus B = \{\omega \in A | \omega \notin B\}$
3. **Associativité** : $(A \cup B) \cup C = A \cup (B \cup C)$ et $(A \cap B) \cap C = A \cap (B \cap C)$
4. **Distributivité** : $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ et $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
5. $\overline{A} \cap \overline{B} = \overline{A \cup B}$ et $\overline{A} \cup \overline{B} = \overline{A \cap B}$

Exemples: On considère l'univers $\Omega = \{1, 2, 3, 4, 5, 6\}$, et les événements $A = \{1, 2, 3\}, B = \{5, 6\}$ et $C = \Omega$.

- $\overline{A} = \{4, 5, 6\}$
- $\overline{C} = \emptyset$
- $A \cup B = \{1, 2, 3, 5, 6\}$
- $A \cap B = \emptyset$

Système complet d'événements

Définition : système complet d'événements

Un **système complet d'événements** est une suite $E_1, E_2, ..., E_m$ de sous-ensembles non vides de Ω vérifiant

- $\forall i \neq j\,,\ E_i \cap E_j = \emptyset$
- $E_1 \cup E_2 \cup ... \cup E_m = \Omega$

Autrement dit, $E_1, E_2, ..., E_m$ est une partition de Ω.

Remarque: Si $E_1, E_2, ..., E_m$ est un système complet de Ω et A est un événement de Ω, alors $E_1 \cap A$, $E_2 \cap A$, ... et $E_m \cap A$ constituent une partition de A.

Exemple: Soit $\Omega = \{1, 2, 3, 4, 5, 6\}$ les résultats d'un test scientifique. Les événements $\{i\}$ avec $i \in [\![1, 6]\!]$ constituent un système complet d'événements de Ω.

Tribu et espace de probabilisable

Définition : tribu

Un ensemble d'événements $\mathcal{A}$ (i.e. un ensemble de sous-ensembles de Ω) est une **tribu** sur un ensemble Ω si :

1. $\mathcal{A}$ n'est pas vide,
2. $\mathcal{A}$ est stable par complémentaire : si $A \in \mathcal{A}$, alors $\overline{A} \in \mathcal{A}$,
3. $\mathcal{A}$ est stable par réunion dénombrable : si $A_1, A_2, \ldots \in \mathcal{A}$, alors $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

On dit que alors que $(\Omega, \mathcal{A})$ est un **espace mesurable**, ou dans le cadre des probabilités, un **espace probabilisable**.

Exemples :

- La tribu grossière : $\mathcal{A} = \{\emptyset, \Omega\}$,
- La tribu discrète : $\mathcal{A} = \mathcal{P}(\Omega)$ (tous les sous-ensembles de Ω).

Pour $\Omega = \{1, 2, 3, 4, 5, 6\}$:

- $\mathcal{A} = \{\emptyset, \{1, 2\}, \{3, 4, 5, 6\}, \Omega\}$ est une tribu,
- $\mathcal{A} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \Omega\}$ n'est pas une tribu.

Remarque : Dans la suite, on considérera par défaut la tribu discrète $\mathcal{A} = \mathcal{P}(\Omega)$ quand Ω est au plus dénombrable, et la tribu borélienne $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$ quand $\Omega = \mathbb{R}^n$ (voir *Wikipedia*).

Mesure de probabilité

Définition : mesure de probabilité

Une **mesure de probabilité** est une application $\mathbb{P} : \mathcal{A} \to [0,1]$ telle que :

1. $\mathbb{P}(\Omega) = 1$,
2. Pour toute suite dénombrable d'événements $E_1, E_2, \ldots$ disjoints deux à deux dans $\mathcal{A}$:
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$$

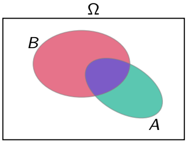
On dit que $(\Omega, \mathcal{A}, \mathbb{P})$ est un **espace probabilisé**.

Propriétés :

1. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
2. $\mathbb{P}(\emptyset) = 0$
3. Si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Preuve des propriétés

1. Le premier point de la définition donne $\mathbb{P}(\overline{A} \cup A) = \mathbb{P}(\Omega) = 1$
   Le second point donne $\mathbb{P}(\overline{A} \cup A) = \mathbb{P}(\overline{A}) + \mathbb{P}(A)$
   Ainsi, $\mathbb{P}(\overline{A}) + \mathbb{P}(A) = 1$
   D'où $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
2. $1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) = 1 + \mathbb{P}(\emptyset) \quad$ (propriété 1)
   Donc $\mathbb{P}(\emptyset) = 0$
3. Soient $A \subset B$. $\mathbb{P}(B) = \mathbb{P}\Big((B \setminus A) \cup A\Big) = \mathbb{P}(B \setminus A) + \mathbb{P}(A)$
   Or $\mathbb{P}(B \setminus A) \geq 0$ car $\mathbb{P}$ est à valeur positives par définition.
   Donc $\mathbb{P}(A) \leq \mathbb{P}(B)$.



4. $$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}\Big((A \cap B) \cup (A \setminus (A \cap B)) \cup (B \setminus (A \cap B))\Big)\\ &= \mathbb{P}(A \cap B) + \mathbb{P}\big(A \setminus (A \cap B)\big) + \mathbb{P}\big(B \setminus (A \cap B)\big)\\ &= \mathbb{P}(A \cap B) + \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\end{aligned}$$

Principe de comptage dans le cas équiprobable

Soit $\Omega = \{\omega_1, \ldots, \omega_n\}$ un univers fini. Si Chaque élément de l'univers a la même probabilité (**équiprobabilité**), alors pour toute éventualité ω_i de Ω :

$$\mathbb{P}(\{\omega_i\}) = \frac{1}{|\Omega|} = \frac{1}{n} \qquad \text{avec } |\Omega| \text{ le cardinal de } \Omega.$$

Dans ce cas, la probabilité de tout événement A est : $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$

Exemples : On considère un dé à 6 faces non pipé : $\Omega = \{1, 2, 3, 4, 5, 6\}$, et A l'événement "on obtient une face paire" $A = \{2, 4, 6\}$.
Alors $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$.

Arrangements

La question est : *"Combien y a-t-il de façons de ranger k objets pris au hasard parmi n objets distincts ?"*

Les arrangements correspondent à un tirage avec ordre et sans répétition (sans remise).

$$\text{il y a } A_n^k = n \times (n-1) \times \ldots \times (n-k+1) = \frac{n!}{(n-k)!} \text{ possibilités.}$$

Exemples : Lors de la finale olympique du 100m il y a huit finalistes. Combien y a-t-il de possibilités de répartition des médailles d'or, d'argent et de bronze ?
On a $k = 3$, $n = 8$, donc $A_8^3 = \frac{8!}{5!} = 8 \times 7 \times 6 = 336$ possibilités.

Combinaisons

Les combinaisons correspondent à un tirage sans ordre et sans répétition (sans remise).

$$\text{il y a } C_n^k = \frac{n \times (n-1) \times \ldots \times (n-k+1)}{k!} = \frac{n!}{(n-k)!k!} \text{ possibilités.}$$

Remarque : $C_n^k = \frac{n!}{(n-k)!k!}$ est appelé coefficient binomial. En général, on le note $\binom{n}{k}$, ce qui se lit "k parmi n".

Exemples : Le nombre de tirages de 4 cartes dans un paquet de 32 cartes est

$$C_{32}^4 = \frac{32!}{4!28!} = \frac{32 \times 31 \times 30 \times 29}{4 \times 3 \times 2} = 35960.$$

Probabilités conditionnelles : un exemple

On dispose de 2 dés parfaits, un noir et un blanc. On lance les 2 dés et on s'intéresse à l'événement $A = \{\text{la somme des dés vaut } 5\}$. Il y a quatre couples de solutions possibles, donc $\mathbb{P}(A) = \frac{4}{36} = \frac{1}{9}$.

Supposons maintenant que nous disposions d'une information supplémentaire :

a) L'événement $B = \{\text{le dé blanc vaut } 5\}$ est réalisé. Il est alors clair que l'événement A ne se réalisera pas, quelle que soit la valeur du dé noir. On dit que la probabilité que A se réalise sachant B est nulle : $\mathbb{P}(A|B) = 0$.
Remarque : $\mathbb{P}(A|B)$ se lit "P de A sachant B".

b) L'événement $C = \{\text{le dé blanc vaut } 1\}$ est réalisé. Dans ce cas, la seule situation qui permette d'obtenir 5 est que le dé noir prenne la valeur 4, ce qui a une chance sur 6 de se produire : $\mathbb{P}(A|C) = \frac{1}{6}$.

Probabilités conditionnelles

Définition : probabilité conditionnelle

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient E_1 et E_2 deux événements tels que $\mathbb{P}(E_1) \neq 0$.
La **probabilité conditionnelle** de E_2 étant donné E_1, $\mathbb{P}(E_2|E_1)$, indique la probabilité que E_2 se produise sachant que E_1 s'est déjà produit :

$$\mathbb{P}(E_2|E_1) = \frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_1)}$$

Théorème de Bayes

Soient A et B deux événements de probabilité non nulle. Alors :

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

Remarque : Si la réalisation ou la non réalisation de E_1 n'affecte pas E_2, alors $\mathbb{P}(E_2|E_1) = \mathbb{P}(E_2)$.

Exemple : Probabilité conditionnelle

On considère les résultats d'un test médical donnés dans le tableau suivant.

Genre / Test	positif	négatif
homme	12	30
femme	18	40

- Quelle est la probabilité d'avoir un test positif ?
- Quelle est la probabilité d'avoir un test positif chez les hommes ?

Formule des probabilités totales

Formule des probabilités totales (cas fini)

Soit $\{B_1, B_2, \ldots, B_n\}$ une partition de Ω et A un événement. On a :

- $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i)$
- Si de plus, $\forall i$, $\mathbb{P}(B_i) \neq 0$, alors $\mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)$

Exemple :
Un laboratoire a mis au point un alcootest. 2% des personnes contrôlées par la police sont réellement en état d'ébriété. Les premiers essais ont conduit aux résultats suivants :

- Lorsqu'une personne est réellement en état d'ébriété, 95 fois sur 100 l'alcootest se révèle positif.
- Lorsqu'une personne n'est pas en état d'ébriété, 96 fois sur 100 l'alcootest se révèle négatif.

Quelle est la probabilité pour qu'une personne soit réellement en état d'ébriété lorsque l'alcootest est positif ?

Bien définir les événements avant de se lancer dans les calculs !

Indépendance

L'indépendance est un concept fondamental de la théorie des probabilités. Elle permet de formaliser que deux événements n'interagissent pas l'un sur l'autre.

Définition : événements indépendants

Deux événements E_1 et E_2 sont indépendants si

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1) \times \mathbb{P}(E_2)$$

Plus généralement, $E_1, \ldots, E_k$ sont indépendants si

$$\mathbb{P}(E_1 \cap E_2 \cap \ldots \cap E_k) = \mathbb{P}(E_1) \times \mathbb{P}(E_2) \times \ldots \times \mathbb{P}(E_k)$$

- L'indépendance de E_1 et E_2 est équivalente à $\mathbb{P}(E_1|E_2) = \mathbb{P}(E_1)$ et/ou $\mathbb{P}(E_2|E_1) = \mathbb{P}(E_2)$.
- L'indépendance entre E_1 et E_2 se note $E_1 \perp\!\!\!\perp E_2$.

Exemple : Indépendance

On considère un dé à 6 faces non pipé et les événements $A = \{2, 4, 6\}$ et $B = \{3, 6\}$.

- $\mathbb{P}(A) = \frac{3}{6}$
- $\mathbb{P}(B) = \frac{2}{6}$
- $\mathbb{P}(A \cap B) = \frac{1}{6}$

On vérifie donc que $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B) = \frac{3}{6} \times \frac{2}{6} = \frac{1}{6}$. A et B sont donc indépendants.

Remarques sur l'indépendance

- Un événement de probabilité nulle est toujours indépendant de tous les autres.
- Indépendance et incompatibilité sont des notions différentes !

Propriété

Si A et B sont indépendants, alors les paires $\{A, \overline{B}\}$, $\{\overline{A}, B\}$ et $\{\overline{A}, \overline{B}\}$ sont aussi constituées d'événements indépendants.