Plan du chapitre 1

1 Mesures de probabilité

- Expérience aléatoire, univers, événements, σ-algèbre
- Mesures de probabilité
- Analyse combinatoire
- Mesures de probabilité conditionnelle
- Deux théorèmes et un pseudo-théorème importants
- Indépendance stochastique

Plan du chapitre 1

1 Mesures de probabilité

- Expérience aléatoire, univers, événements, σ-algèbre
- Mesures de probabilité
- Analyse combinatoire
- Mesures de probabilité conditionnelle
- Deux théorèmes et un pseudo-théorème importants
- Indépendance stochastique

Expérience aléatoire, univers, événements, σ-algèbre

Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ interroger un quidam sur ses intentions de vote

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \{\text{Macron}, \text{Le Pen}\}$

Un événement A :
Un sous-ensemble de Ω
Exemple : $A = \{\text{Macron}\}$

Remarques :
- Si $\omega \in A$, on dit que l'événement A se produit.
- Dans la suite, l'ensemble de tous les sous-ensembles de Ω sera noté $\mathcal{P}(\Omega)$.

Expérience aléatoire, univers, événements, σ-algèbre

Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ lancer d'un dé

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \{1, 2, 3, 4, 5, 6\}$

Un événement A :
Un sous-ensemble de Ω
Exemples : $A_1 = \{1\}$, $A_2 = \{2, 4, 6\}$, $A_3 = \{5, 6\}$

Remarques :
- Si A est un singleton, on dit que A est simple (e.g., A_1).
  Sinon, on dit que A est composé (e.g., A_2, A_3).
- Souvent, les événements composés s'expriment aisément sans mathématique (e.g., $A_2 =$ "obtenir un résultat pair").

Expérience aléatoire, univers, événements, σ-algèbre

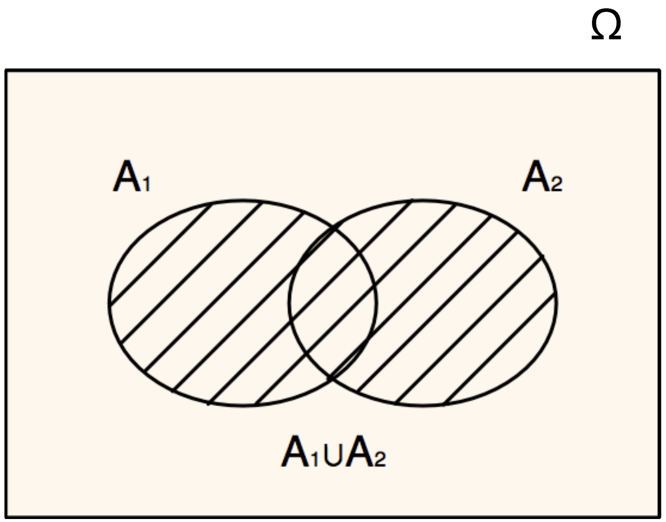
Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ lancer de deux dés (distinguables)

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,6)\}$

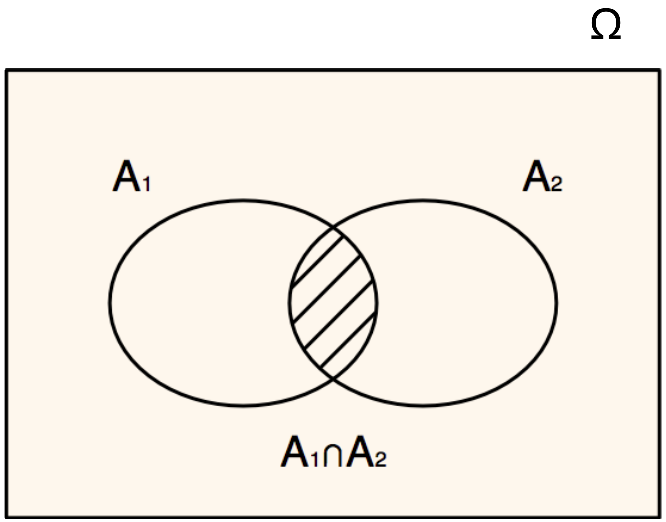
Un événement A :
Un sous-ensemble de Ω
Exemples : $A_1 = \{(1,4)\}$, $A_2 = \{(6,6)\}$, $A_3 = \{(5,6),(6,5)\}$

Remarque :
- On peut engendrer de nouveaux événements au moyen des opérations ensemblistes $\cup$, $\cap$, ^{c}, ..., associées aux opérations logiques "ou", "et", "non", ...
Exemple : $A_2 \cup A_3 = \{(6,6),(5,6),(6,5)\}$, i.e., "obtenir une somme égale à 12" ou "obtenir une somme égale à 11" = "obtenir une somme plus grande ou égale à 11".

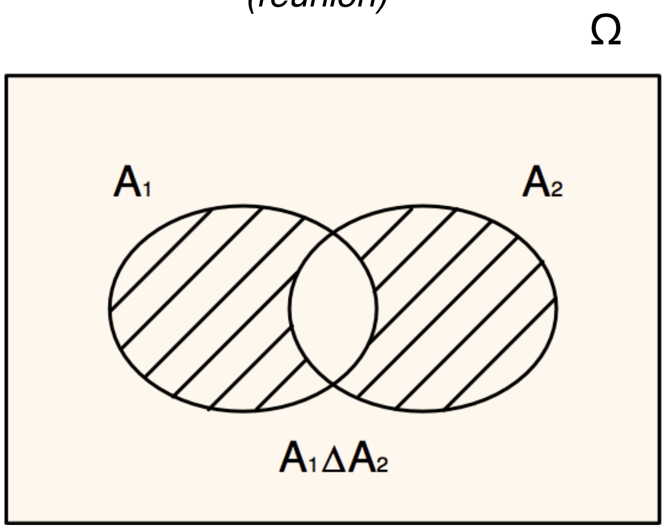
Expérience aléatoire, univers, événements, σ-algèbre



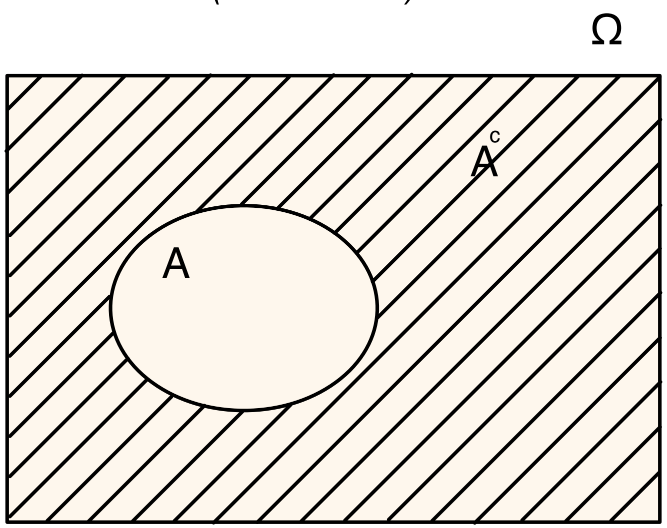
"ou" inclusif
(réunion)



"et"
(intersection)



"ou" exclusif
(différence symétrique)



"non"
(complémentaire)

Expérience aléatoire, univers, événements, σ-algèbre

Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ créer une start-up dans le but de percer dans les 5 ans

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \{\text{percer}, \text{ne pas percer}\}$

Un événement A :
Un sous-ensemble de Ω
Exemples : $A_1 = \{\text{percer}\}$, $A_2 = \emptyset$, $A_3 = \Omega$

Remarques :
- $\emptyset$ est dit événement impossible.
- Ω est dit événement certain.

Expérience aléatoire, univers, événements, σ-algèbre

Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ lancer une pièce de monnaie en l'air jusqu'à obtenir P (pile)

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \{P, (F, P), (F, F, P), (F, F, F, P), \ldots\}$

Un événement A :
Un sous-ensemble de Ω
Exemples : $A_1 = \{P, (F, P), (F, F, P)\}$, $A_2 = \{(F, F, F, P), (F, F, F, F, P), \ldots\}$

Remarques :
- A l'inverse des exemples précédents, cet Ω est infini.
- Ω est ici infini dénombrable (c'est-à-dire en bijection avec $\mathbb{N}$).

Expérience aléatoire, univers, événements, σ-algèbre

Expérience aléatoire E :
Une expérience dont on ne peut prédire le résultat avec certitude
Exemple : $E =$ mesurer le temps d'attente du bus 71 (en minutes)

L'univers $\Omega = \{\omega\}$:
L'ensemble de tous les résultats ω possibles de E
Exemple : $\Omega = \mathbb{R}_0^+ = (0, +\infty)$

Un événement A :
Un sous-ensemble de Ω
Exemples : $A_1 = [5, +\infty)$, $A_2 = (0, 15]$, ...

Remarques :
- Ω est ici infini non dénombrable.
- Les opérations ensemblistes sont encore permises.

Exemple : "attendre entre 5 et 15 minutes"$= [5, 15] = [5, +\infty) \cap (0, 15] = A_1 \cap A_2$
= "attendre au moins 5 minutes" et "attendre au plus 15 minutes".

Même pour des expériences aléatoires assez simples, la collection d'événements $\mathcal{P}(\Omega)$ peut être gigantesque.

Par exemple, lancer deux dés distinguables mène à 2^{36} *événements, soit plus de* 68 *milliards d'événements !*

Le plus souvent, on s'intéressera seulement à une petite collection $\mathcal{A}$ d'événements.

Par exemple, pour le lancer de deux dés distinguables, on peut être intéressé par la somme des résultats, ce qui mène à considérer

$$\mathcal{A} = \Big\{\emptyset, \{(1,1)\}, \{(1,2),(2,1)\}, \{(1,1),(1,2),(2,1)\}, \{(1,3),(2,2),(3,1)\}, \ldots, \Omega\Big\},$$

ou être intéressé par le fait de faire un double, ce qui mène à considérer

$$\mathcal{A} = \Big\{\emptyset, A = \{(1,1),(2,2),\ldots,(6,6)\}, A^c, \Omega\Big\}.$$

Expérience aléatoire, univers, événements, σ-algèbre

Dans tous les cas, cette collection d'événements devra être une σ-algèbre.

Définition

La collection $\mathcal{A}$ de sous-ensembles de Ω est une σ-algèbre si

- $\Omega \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow A_1 \cup A_2 \cup \ldots \in \mathcal{A}$

On peut vérifier que si on applique un nombre fini ou infini dénombrable d'opérations ensemblistes aux éléments d'un tel $\mathcal{A}$, on obtient encore un élément de $\mathcal{A}$:

- $\emptyset = \Omega^c \in \mathcal{A}$
- Si $A_1, A_2 \in \mathcal{A}$, alors $A_1 \cup A_2 = A_1 \cup A_2 \cup \emptyset \cup \emptyset \cup \ldots \in \mathcal{A}$
- Si $A_1, A_2 \in \mathcal{A}$, alors $A_1 \cap A_2 = (A_1^c \cup A_2^c)^c \in \mathcal{A}$
- $\ldots$

Expérience aléatoire, univers, événements, σ-algèbre

Pour une collection d'événements $\mathcal{A}$, on peut considérer la plus petite σ-algèbre, $\sigma(\mathcal{A})$, contenant $\mathcal{A}$. Par exemple, si $\mathcal{A} = \{A\}$, on obtient $\sigma(\mathcal{A}) = \big\{\emptyset, A, A^c, \Omega\big\}$.

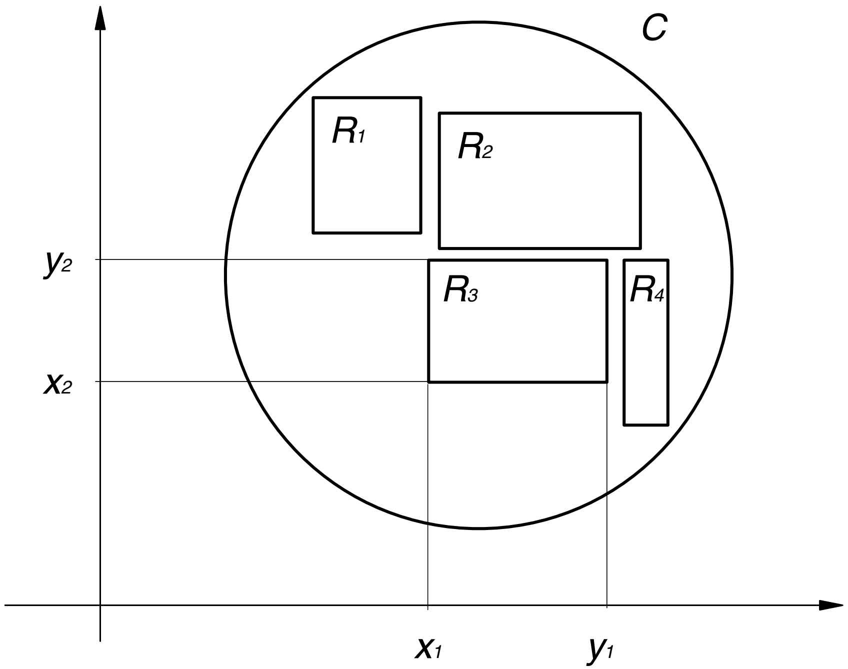
Si $\Omega = \mathbb{R}$, on considére souvent la σ-algèbre de Borel $\mathcal{B} = \sigma(\{(x, y] : x < y\})$.

Remarques :

- Les éléments $B \in \mathcal{B}$ sont appelés les boréliens.
- tous les intervalles de la forme $(x, y]$, (x, y), $[x, y]$, $[x, y)$, donc aussi (iii) les réunions finies de tels événements sont dans $\mathcal{B}$! (voir les TP).

Expérience aléatoire, univers, événements, σ-algèbre

Extension à $\mathbb{R}^+$: $\mathcal{B}^+ = \{B \cap \mathbb{R}^+ : B \in \mathcal{B}\}$
Extension à $\mathbb{R}_0^+$: $\mathcal{B}_0^+ = \{B \cap \mathbb{R}_0^+ : B \in \mathcal{B}\}$
$\ldots$
Extension à $\mathbb{R}^2$: $\mathcal{B}^2 = \sigma(\{(x_1, y_1] \times (x_2, y_2] : x_1 < y_1,\ x_2 < y_2\})$



Le disque C appartient à $\mathcal{B}^2$ car il se décompose en une union dénombrable de rectangles :
$C = R_1 \cup R_2 \cup R_3 \cup R_4 \cup \ldots$

Extension à $\mathbb{R}^k$: $\mathcal{B}^k = \sigma(\{(x_1, y_1] \times (x_2, y_2] \times \ldots \times (x_k, y_k] : x_i < y_i \ \forall i\})$

Plan du chapitre 1

1. Mesures de probabilité
   - Expérience aléatoire, univers, événements, σ-algèbre
   - Mesures de probabilité
   - Analyse combinatoire
   - Mesures de probabilité conditionnelle
   - Deux théorèmes et un pseudo-théorème importants
   - Indépendance stochastique

Nous désirons mesurer la probabilité, $P[A]$, de chaque événement A.

Soit E une expérience aléatoire reproductible.
Soit A un événement d'intérêt, dans une σ-algèbre $\mathcal{A}$ donnée.

La définition fréquentielle de $P[A]$ est

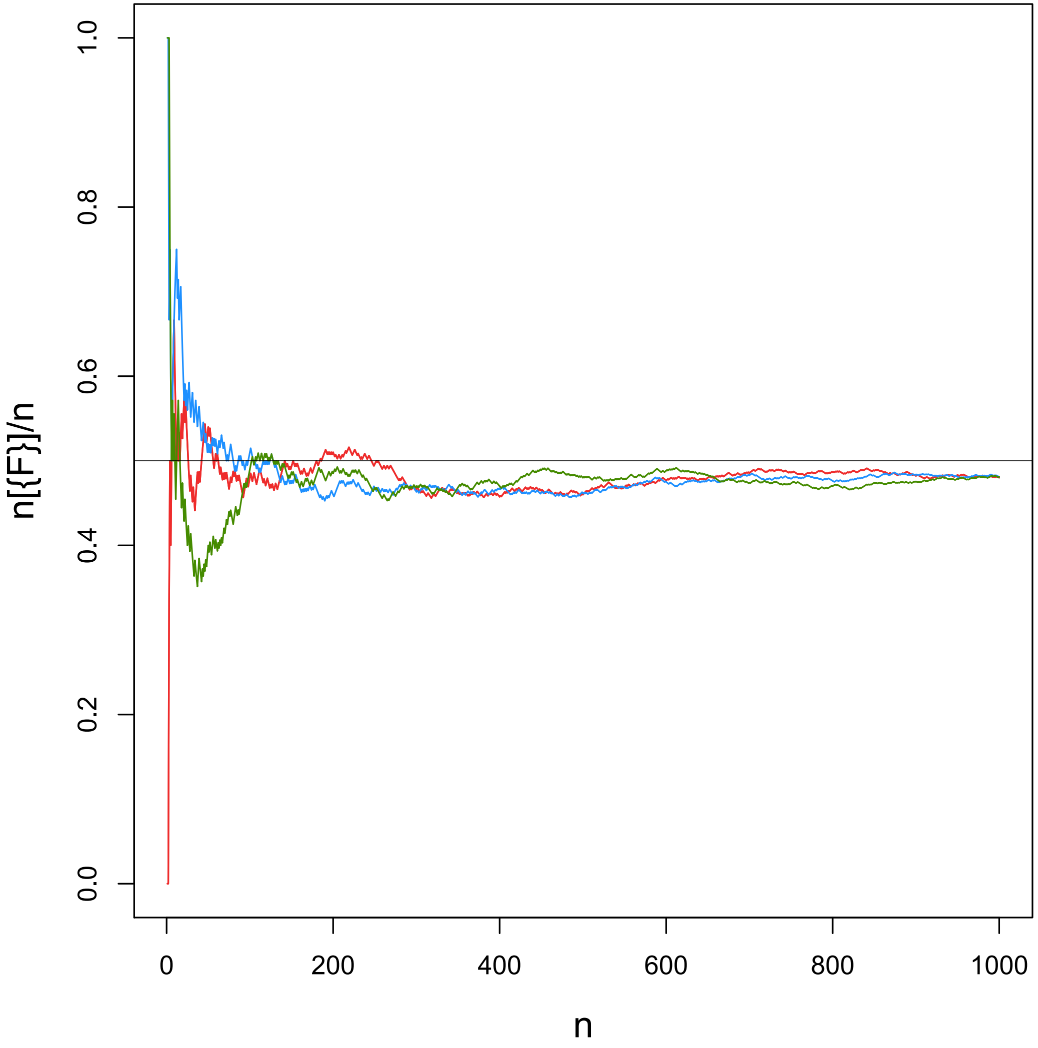
$$P[A] = \lim_{n\to\infty} \frac{n[A]}{n},$$

où $n[A]$ est le nombre de réalisations de A dans n répétitions indépendantes de E.

Au 18ème siècle, Buffon obtint $n(\{F\}) = 2048$ résultats "Face" en lançant $n = 4040$ fois une pièce équilibrée $\rightsquigarrow P[\{F\}] \approx \frac{n(\{F\})}{n} = 0.5069$

Au début du 20ème siècle, Karl Pearson obtient $P[\{F\}] \approx \frac{n(\{F\})}{n} = \frac{12012}{24000} = 0.5005$

Aujourd'hui, on utilise un ordinateur



Exercice : vérifier qu'il découle directement de la définition fréquentielle que

(1) $P[A] \geq 0$ pour tout $A \in \mathcal{A}$.

(2) $P[\Omega] = 1$.

(3) $P[A_1 \cup A_2] = P[A_1] + P[A_2]$, pour tout $A_1, A_2 \in \mathcal{A}$ tels que $A_1 \cap A_2 = \emptyset$,

ou, plus généralement,

(3') $P[A_1 \cup A_2 \cup \ldots] = P[A_1] + P[A_2] + \ldots$, pour tout $A_1, A_2, \ldots \in \mathcal{A}$ tels que $A_i \cap A_j = \emptyset$ si $i \neq j$.

Mesures de probabilité

Certaines expériences ne sont pas reproductibles (mesurer la température à la Bastoche le 1/9/2043 à 12h30), ce qui mène à adopter une définition plus générale.

Définition

Soit $\mathcal{A}$ une σ-algèbre. La fonction d'ensemble

$$\begin{aligned} P : \mathcal{A} &\to \mathbb{R} \\ A &\mapsto P[A] \end{aligned}$$

est une mesure de probabilité sur $\mathcal{A}$ si

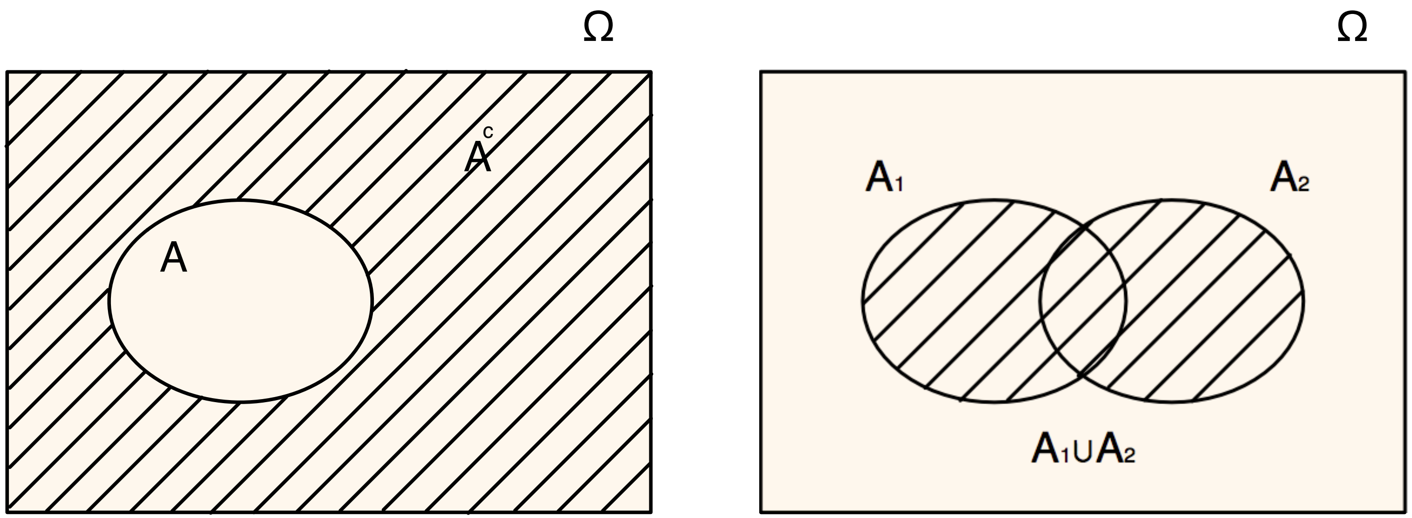
- $P[A] \geq 0$ *pour tout* $A \in \mathcal{A}$
- $P[\Omega] = 1$
- $P[A_1 \cup A_2 \cup \ldots] = P[A_1] + P[A_2] + \ldots$, *pour tout* $A_1, A_2, \ldots \in \mathcal{A}$ *tels que* $A_i \cap A_j = \emptyset$ *si* $i \neq j$.

C'est ce qu'on appelle la définition axiomatique de probabilité
(on dit que le triple $(\Omega, \mathcal{A}, P)$ est un espace probabilisé).

Mesures de probabilité

Il découle des axiomes (voir les TP) que toute mesure de probabilité vérifie

- $P[A^c] = 1 - P[A]$
- $0 \leq P[A] \leq 1$
- $P[\emptyset] = 0$
- $P[A_1 \setminus A_2] = P[A_1] - P[A_1 \cap A_2]$
- Si $A_1 \subset A_2$, alors $P[A_1] \leq P[A_2]$
- Si $A_1 \cap A_2 = \emptyset$, alors $P[A_1 \cup A_2] = P[A_1] + P[A_2]$
- $P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2]$



Si $\#\Omega < \infty$, définir une mesure de probabilité est strictement équivalent à associer à chaque élément ω_i de Ω un nombre $p_i (\geq 0)$, qui sera la valeur de $P[\{\omega_i\}]$.
Seule restriction : $\sum_{i=1}^{\#\Omega} p_i = 1$.

résultats possibles	ω_1	ω_2	$\dots$	$\omega_{\#\Omega}$
probabilités	p_1	p_2	$\dots$	$p_{\#\Omega}$

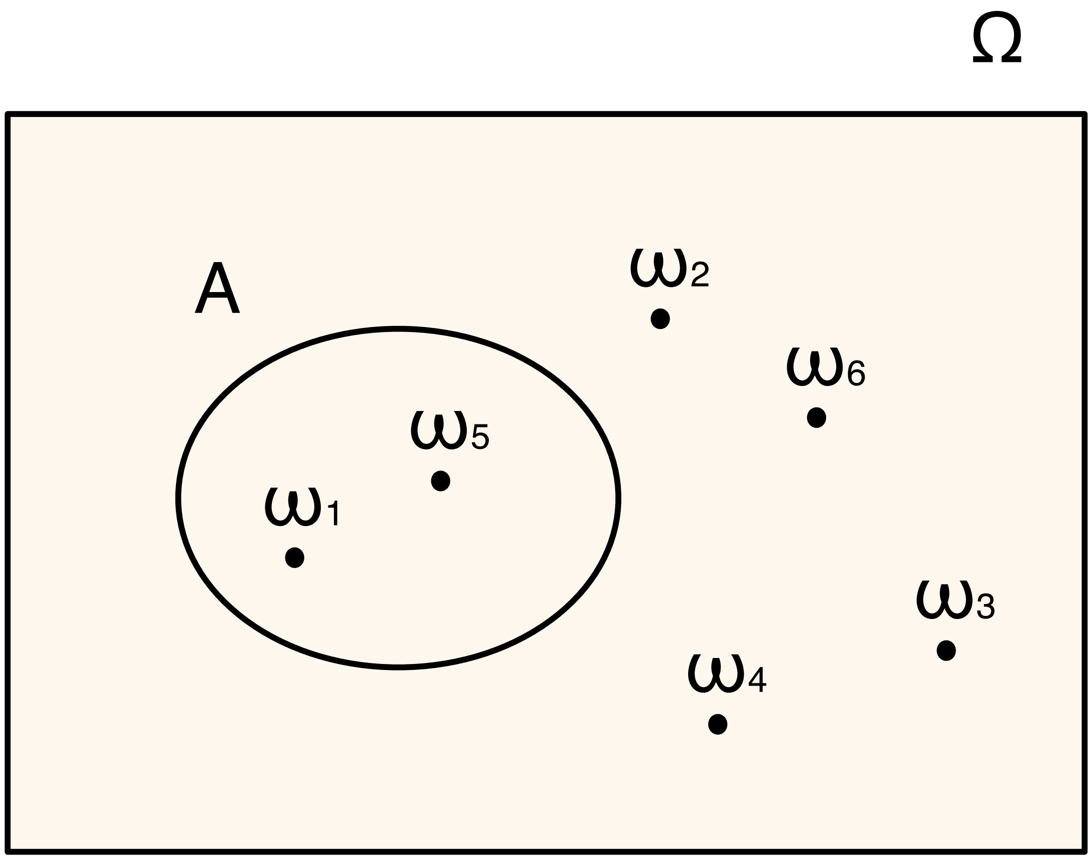
La mesure de probabilité qui en résulte est alors donnée par

$$\begin{aligned} P : \mathcal{P}(\Omega) &\to \mathbb{R} \\ A &\mapsto \sum_{i : \omega_i \in A} P[\{\omega_i\}] \end{aligned}$$

avec, par convention, $P[\emptyset] = 0$.

Remarque : si $\#\Omega < \infty$, on peut toujours travailler avec $\mathcal{A} = \mathcal{P}(\Omega)$.

Mesures de probabilité



$$P[A] = \sum_{i:\omega_i \in A} P[\{\omega_i\}] = P[\{\omega_1\}] + P[\{\omega_5\}]$$

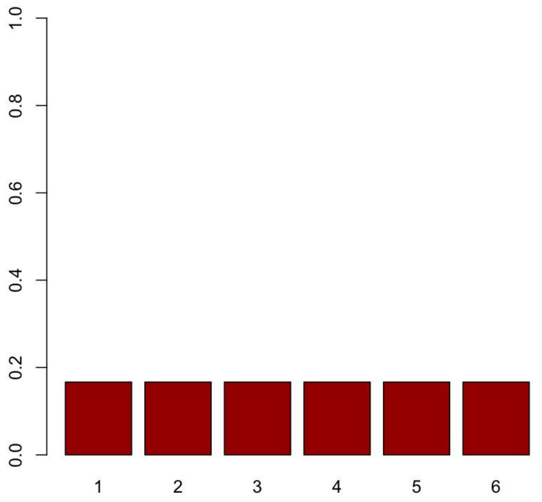
Exemple : pour le lancer d'un dé ($\rightsquigarrow \Omega = \{1, 2, \ldots, 6\}$), on peut prendre

$$(p_1, p_2, p_3, p_4, p_5, p_6) = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$$

(hypothèse d'un dé équilibré).

Ceci mène à

$$\begin{aligned} P[\text{obtenir un résultat pair}] &= P[\{2,4,6\}] \\ &= P[\{2\}] + P[\{4\}] + P[\{6\}] \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= \frac{1}{2} \end{aligned}$$

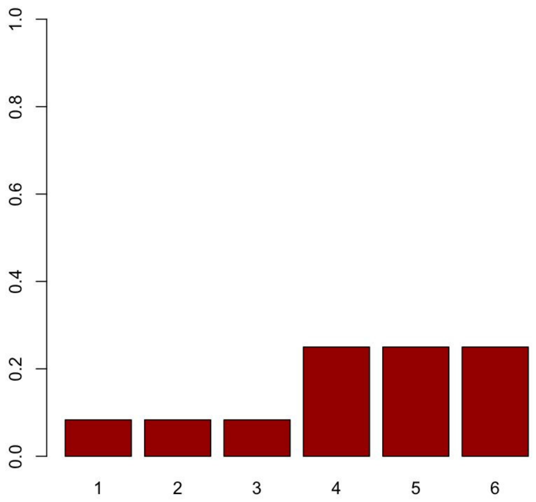


Mais peut-être le dé est-il plutôt caractérisé par

$$(p_1, p_2, p_3, p_4, p_5, p_6) = \left(\frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right).$$

Dans ce cas, on obtient

$$\begin{aligned} P[\text{obtenir un résultat pair}] &= P[\{2,4,6\}] \\ &= P[\{2\}] + P[\{4\}] + P[\{6\}] \\ &= \frac{1}{12} + \frac{1}{4} + \frac{1}{4} \\ &= \frac{7}{12} \left(> \frac{1}{2}\right) \end{aligned}$$



Comment choisir la mesure de probabilité P ?

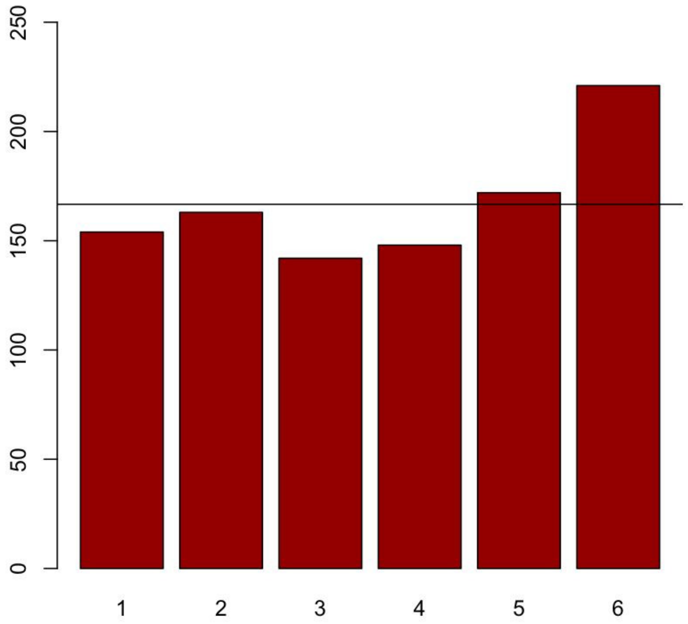
- Le choix est parfois guidé par les hypothèses naturelles que l'on fait (e.g., dé équilibré).
- Dans la suite de la partie "probabilités" du cours, on supposera toujours que P est connue (ou que l'on fait des hypothèses nous permettant de déterminer P).

Mais si on ne connaît pas P, on peut recourir à l'inférence statistique

- pour "estimer" P,
- pour confronter notre "croyance" sur P à des données empiriques...

Imaginons que 1000 lancers du même dé aient livré les données suivantes :

valeurs possibles	fréquences observées
1	154
2	163
3	142
4	148
5	172
6	221



Est-ce que ceci permet (ou non) d'infirmer l'hypothèse que le dé est équilibré ?

Etudier la validité d'une hypothèse (ou d'une théorie scientifique) en la confrontant à des données empiriques est l'un des usages principaux de l'inférence statistique.

Un cas particulier important : l'équiprobabilité

Si $\#\Omega < \infty$, on peut considérer le cas où $p_1 = p_2 = \ldots = p_{\#\Omega}$.

On doit alors avoir

$$p_i = \frac{1}{\#\Omega} \quad \forall i.$$

Donc, pour tout $A \in \mathcal{P}(\Omega)$,

$$\begin{aligned} P[A] &= \sum_{i:\omega_i \in A} P[\{\omega_i\}] \\ &= \frac{1}{\#\Omega} + \frac{1}{\#\Omega} + \ldots + \frac{1}{\#\Omega} \quad (\#A \text{ fois}) \\ &= \frac{\#A}{\#\Omega} = \frac{\text{nombre de cas "favorables"}}{\text{nombre de cas possibles}}. \end{aligned}$$

Mesures de probabilité

Si Ω est infini dénombrable, on définit encore une mesure de probabilité en associant à chaque élément ω_i de Ω un nombre $p_i (\geq 0)$ qui sera la valeur de $P[\{\omega_i\}]$. La restriction devient $\sum_{i=1}^{\infty} p_i = 1$.

résultats possibles	ω_1	ω_2	ω_3	$\dots$
probabilités	p_1	p_2	p_3	$\dots$

On calcule encore la probabilité d'un événement A au moyen de la règle

$$\begin{aligned} P : \mathcal{P}(\Omega) &\rightarrow \mathbb{R} \\ A &\mapsto \sum_{i:\omega_i \in A} P[\{\omega_i\}], \end{aligned}$$

avec $P[\emptyset] = 0$. Notons que l'équiprobabilité est impossible ici (pourquoi ?)

Remarque : si Ω est infini dénombrable, on peut toujours travailler avec $\mathcal{A} = \mathcal{P}(\Omega)$.

Exemple : si on lance une pièce jusqu'à obtenir pile pour la première fois ($\Omega = \{P, (F, P), (F, F, P), (F, F, F, P), \ldots\}$), il est naturel de prendre

$$(p_1, p_2, p_3, p_4, \ldots) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots\right),$$

qui livre bien $\sum_{i=1}^{\infty} p_i = 1$.

On a alors

$$\begin{aligned} P[\text{il faut au plus trois lancers}] &= P[\{P, (F, P), (F, F, P)\}] \\ &= P[\{P\}] + P[\{(F, P)\}] + P[\{(F, F, P)\}] \\ &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} \end{aligned}$$

et

$$\begin{aligned} P[\text{il faut au moins trois lancers}] &= P[\{(F, F, P), (F, F, F, P), \ldots\}] \\ &= P[\{(F, F, P)\}] + P[\{(F, F, F, F, P)\}] + \ldots \\ &= \frac{1}{8} + \frac{1}{16} + \ldots = \frac{1}{4}. \end{aligned}$$

Mesures de probabilité

Le cas où Ω est infini non dénombrable est plus complexe.

Dans l'exemple où on observe le temps d'attente (en minutes) du bus 71 ($\Omega = \mathbb{R}_0^+$), et si on suppose qu'un 71 passe exactement toutes les 10 minutes, on peut prendre

$$\begin{aligned} P : \{B \cap [0,10] : B \in \mathcal{B}\} &\to \mathbb{R} \\ A &\mapsto P[A] = \frac{|A|}{|[0,10]|} = \frac{|A|}{10}, \end{aligned}$$

où $|C|$ représente la mesure de l'ensemble C (la longueur pour un intervalle).

On vérifie que P est bien une mesure de probabilité et on calcule par exemple

$$P[\text{attendre entre 5 et 10 minutes}] = P[[5,10]] = \frac{|[5,10]|}{10} = \frac{5}{10} = \frac{1}{2}.$$

On peut montrer qu'on ne peut pas travailler ici avec $\mathcal{A} = \mathcal{P}(\Omega)$.

Plan du chapitre 1

1. Mesures de probabilité
   - Expérience aléatoire, univers, événements, σ-algèbre
   - Mesures de probabilité
   - Analyse combinatoire
   - Mesures de probabilité conditionnelle
   - Deux théorèmes et un pseudo-théorème importants
   - Indépendance stochastique

Revenons sur la situation d'équiprobabilité, pour laquelle

$$P[A] = \frac{\#A}{\#\Omega} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Pour évaluer $P[A]$, il faudra pouvoir compter le nombre d'éléments d'un ensemble.

C'est l'analyse combinatoire, à travers la règle du produit et les concepts de

- permutations,
- arrangements,
- combinaisons,

qui permettra ce comptage dans des situations complexes.

Analyse combinatoire : règle du produit

Le produit cartésien de deux ensembles A et B est défini par

$$A \times B = \{(a, b) : a \in A,\ b \in B\}.$$

La règle du produit dit que $A \times B$ est de cardinalité $(\#A) \times (\#B)$.

Plus généralement,

$$A \times B \times \ldots \times F = \{(a, b, \ldots, f) : a \in A,\ b \in B, \ldots,\ f \in F\}$$

est de cardinalité $(\#A) \times (\#B) \times \ldots \times (\#F)$.

Exemples :

- Menu avec 3 entrées, 2 plats, 2 desserts $\rightsquigarrow 3 \times 2 \times 2 = 12$ repas possibles
- Si on doit pronostiquer les résultats de 8 matchs de tennis, on a $2^8 = 256$ pronostics possibles

Analyse combinatoire : permutations

Soit A un ensemble de n éléments.

Définition

Une permutation est une suite ordonnée $(a_1, a_2, \ldots, a_n)$ des n éléments de A.

Il y a $n! = n \times (n-1) \times \ldots \times 2 \times 1$ permutations possibles*.

$$
\begin{aligned}
2! &= 2 \\
3! &= 6 \\
4! &= 24 \\
5! &= 120 \\
6! &= 720 \\
7! &= 5\,040 \\
8! &= 40\,320
\end{aligned}
$$

La formule de Stirling

$$\lim_{n \to +\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1$$

montre que le nombre de permutations augmente plus vite que a^n pour tout $a > 0$.

Au 100m olympique, il y a $8! = 40\,320$ manières d'attribuer les couloirs

*. Par convention, on pose $0! = 1$

Définition

Un arrangement de longueur $k(< n)$ est une suite ordonnée $(a_1, a_2, \ldots, a_k)$ de k éléments distincts de A.

Il y a

$$A_n^k = n \times (n-1) \times \ldots \times (n-k+1) = \frac{n!}{(n-k)!}$$

tels arrangements.

Au 100m olympique, il y a $A_8^3 = 336$ podiums possibles

Analyse combinatoire : combinaisons

Définition

Une combinaison de longueur $k (\leq n)$ est un ensemble $\{a_1, a_2, \ldots, a_k\}$ contenant k éléments distincts de A.

L'ordre n'importe pas (>< arrangements), et le nombre de telles combinaisons est

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!} \quad \left(= \binom{n}{k}\right)$$

Au 100m olympique, il y a $C_8^3 = \frac{336}{3!} = 56$ façons de choisir qui montera sur le podium

---

Les coefficients binomiaux $C_n^k = \binom{n}{k}$, qui apparaissent dans le binôme de Newton $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$, vérifient

- $\binom{n}{0} = 1$, $\binom{n}{1} = n$, $\binom{n}{2} = \frac{n(n-1)}{2}, \ldots$
- $\binom{n}{k} = \binom{n}{n-k}$
- $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

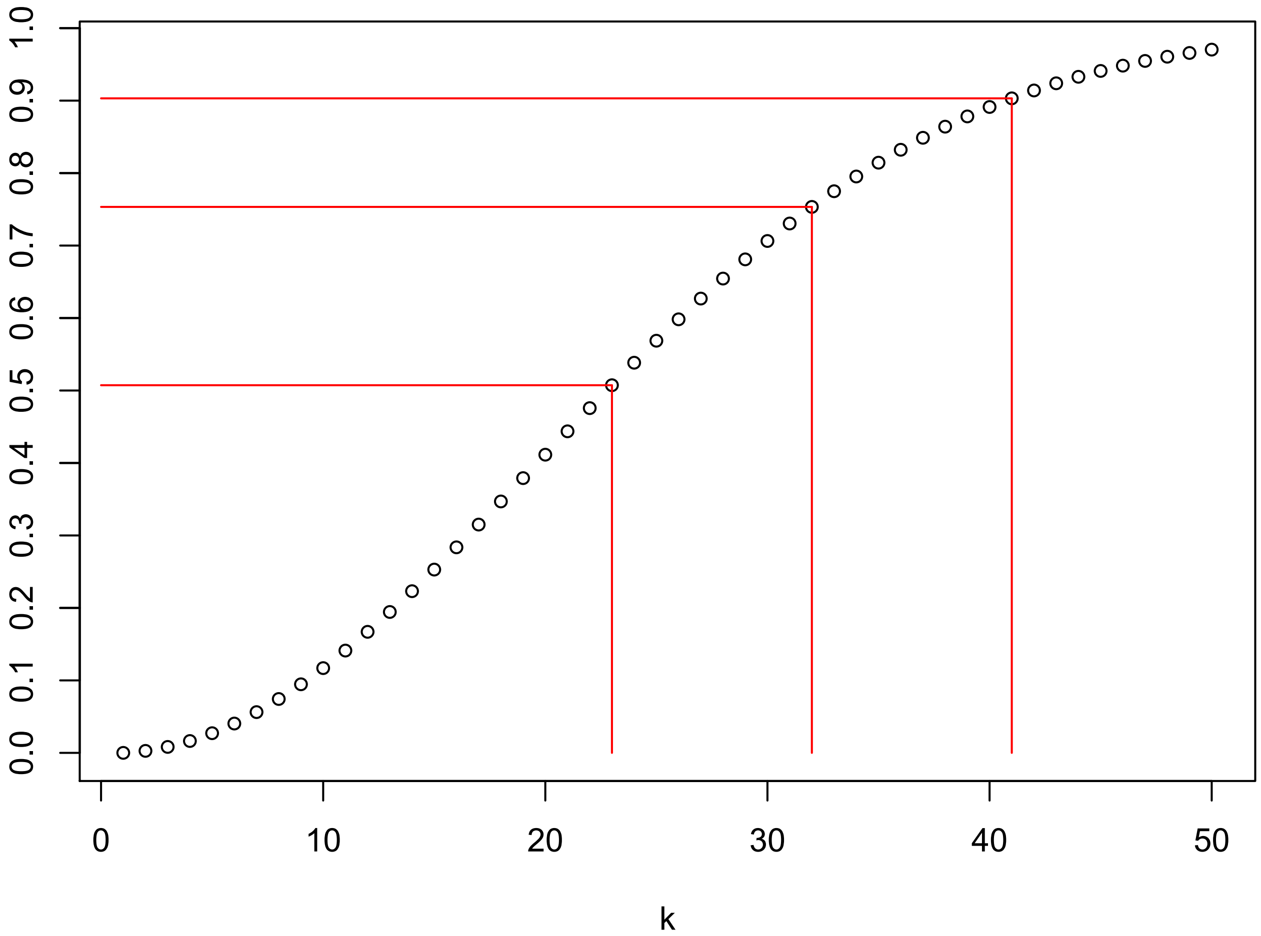
Si on choisit au hasard $k(\leq 365)$ personnes, quelle est la probabilité que deux au moins aient la même date d'anniversaire ? [*]

$\Omega = \{(a_1, \ldots, a_k) : a_i \in \{1, 2, \ldots, 365\}\} \qquad (\rightsquigarrow \#\Omega = 365^k)$
$\mathcal{A} = \mathcal{P}(\Omega)$ (on peut car $\#\Omega < \infty$)
P : équiprobabilité

B : le sous-ensemble de Ω tel que deux composantes au moins coïncident
B^c : le sous-ensemble de Ω tel que les composantes sont deux à deux différentes
$\#B^c = A^k_{365} = \frac{365!}{(365-k)!}$

$$\Rightarrow P[\text{deux anniversaires au moins coïncident}] = P[B] = 1 - P[B^c] = 1 - \frac{\#B^c}{\#\Omega} = 1 - \frac{A^k_{365}}{365^k}$$

*. On négligera les années bissextiles



Probabilité qu'un moins deux anniversaires coïncident parmi k, en fonction de k

Au lotto (où on choisit 6 numéros parmi 45), quelle est la probabilité d'avoir exactement k bons numéros ($k = 0, 1, \ldots, 6$) ?

$\Omega =$ l'ensemble de tous les tirages possibles
$\mathcal{A} = \mathcal{P}(\Omega)$ (on peut car $\#\Omega < \infty$)
P : équiprobabilité

B_k : le sous-ensemble de Ω contenant exactement k numéros du bulletin du joueur, donc k numéros parmi les 6 du bulletin et $6 - k$ numéros parmi les $45 - 6$ autres

$$\Rightarrow P[\text{exactement } k \text{ bons numéros}] = P[B_k] = \frac{\#B_k}{\#\Omega} = \frac{\binom{6}{k} \times \binom{45-6}{6-k}}{\binom{45}{6}}$$

k	0	1	2	3	4	5	6
$P[B_k]$	0.40	0.42	0.15	0.022	0.0014	0.000029	0.00000012

Plan du chapitre 1

1. Mesures de probabilité
   - Expérience aléatoire, univers, événements, σ-algèbre
   - Mesures de probabilité
   - Analyse combinatoire
   - Mesures de probabilité conditionnelle
   - Deux théorèmes et un pseudo-théorème importants
   - Indépendance stochastique

Mesures de probabilité conditionnelle

Souvent, on dispose d'une information qui permet de mettre à jour la probabilité d'un événement A. La forme la plus simple que peut prendre cette information est la connaissance qu'un événement B s'est produit.

Considérons un espace probabilisé $(\Omega, \mathcal{A}, P)$ et des événements A, B.

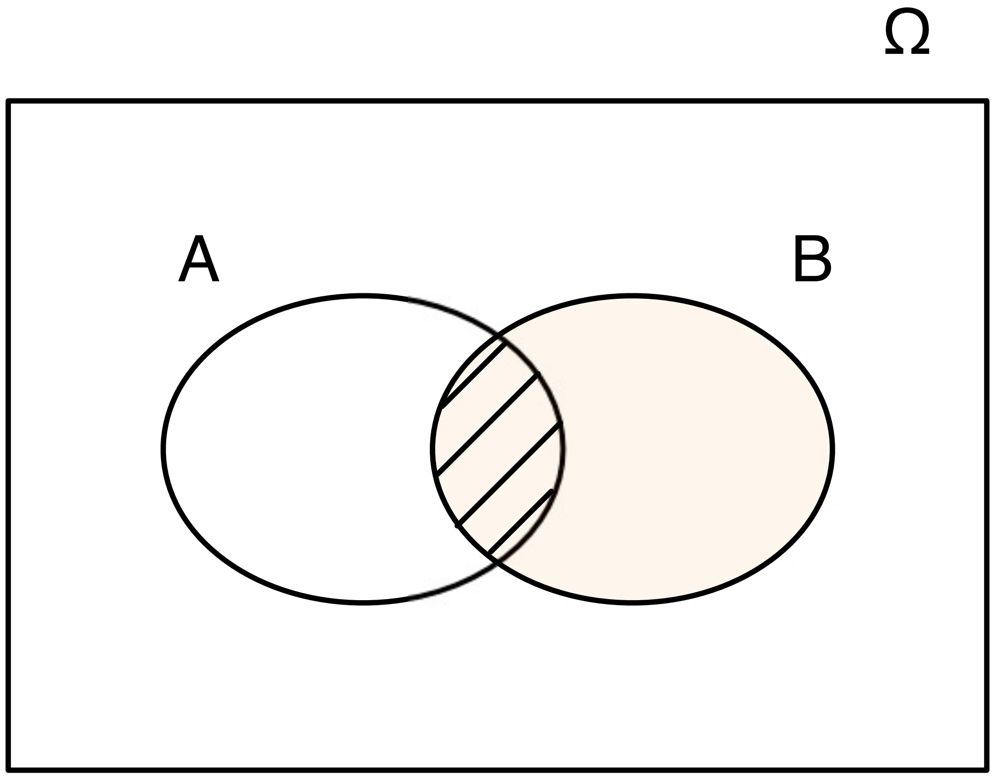
Définition

La probabilité conditionnelle de A sachant que B se produit est

$$P[A\,|\,B\,] = \frac{P[A \cap B]}{P[B]}$$

(*on suppose donc que* $P[B] > 0$).

$P[A|\Omega] = P[A]$ (l'information que Ω se produit est inutile !)



Si on a l'information que B se produit, les cas possibles sont associés à B et les cas favorables sont associés à $A \cap B \rightsquigarrow P[A|B] = P[A \cap B]/P[B]$.

Les parents du roi ont deux enfants.
Quelle est la probabilité que le roi ait une soeur ? *

$\Omega = \{(G, G), (G, F), (F, G), (F, F)\}$
$\mathcal{A} = \mathcal{P}(\Omega)$
P : équiprobabilité

$B = \{(G, G), (G, F), (F, G)\}$ "l'un des enfants au moins est un garçon"

Alors, avec $A = \{(G, F), (F, G), (F, F)\}$ "l'un des enfants est une fille", on obtient

$$\begin{aligned} P[\text{le roi a une soeur}] &= P[A \,|\, B] = \frac{P[A \cap B]}{P[B]} \\ &= \frac{P[\{(G, F), (F, G)\}]}{P[\{(G, G), (G, F), (F, G)\}]} = \frac{2/4}{3/4} = \frac{2}{3}. \end{aligned}$$

*. On supposera que seul le premier héritier mâle accède au trône

Mesures de probabilité conditionnelle

Les probabilités conditionnelles sont d'application courante en assurance :

- En assurance non-vie, les primes dépendent du lieu d'habitation car, par exemple,

$$P[\text{subir un sinistre} \,|\, \text{vivre à Charleroi}] > P[\text{subir un sinistre} \,|\, \text{vivre à Lasnes}].$$

- En assurance vie, les primes sont plus élevées quand on avance en âge car

$$P[\text{décéder sous peu} \,|\, \text{être vieux}] > P[\text{décéder sous peu} \,|\, \text{être jeune}].$$

  Les "tables de mortalité" reprennent $P[\text{décéder avant } n+1 \text{ ans} \,|\, \text{on a atteint } n \text{ ans}]$ pour différents n ;

Les applications s'étendent à de nombreux domaines, dont celui de l'économie où les probabilités conditionnelles jouent un rôle important en théorie des jeux.

Plan du chapitre 1

1 Mesures de probabilité

- Expérience aléatoire, univers, événements, σ-algèbre
- Mesures de probabilité
- Analyse combinatoire
- Mesures de probabilité conditionnelle
- Deux théorèmes et un pseudo-théorème importants
- Indépendance stochastique

Deux théorèmes et un pseudo-théorème importants

Théorème ("Théorème de multiplication")

Pour tout $A, B \in \mathcal{A}$ *avec* $P[A] > 0$*, on a* $P[A \cap B] = P[A]P[B|A]$.

Ce résultat n'est pas vraiment un théorème, mais seulement un "pseudo-théorème", puisqu'il ne fait que reformuler la définition de $P[B|A]$!

Exemple :

Au hasard, on choisit séquentiellement deux personnes dans un groupe contenant 4 hommes et 6 femmes. Quelle est la probabilité que deux femmes soient choisies ?

Si F_i est l'événement "la *i*ème personne choisie est une femme", alors le théorème de multiplication fournit directement $P[F_1 \cap F_2] = P[F_1]P[F_2|F_1] = \frac{6}{10} \times \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$.

Deux théorèmes et un pseudo-théorème importants

Ce résultat s'étend à plus de deux événements.

Par exemple, on vérifie facilement que

$$P[A \cap B \cap C] = P[A]P[B|A]P[C|A \cap B]$$

(exercice)

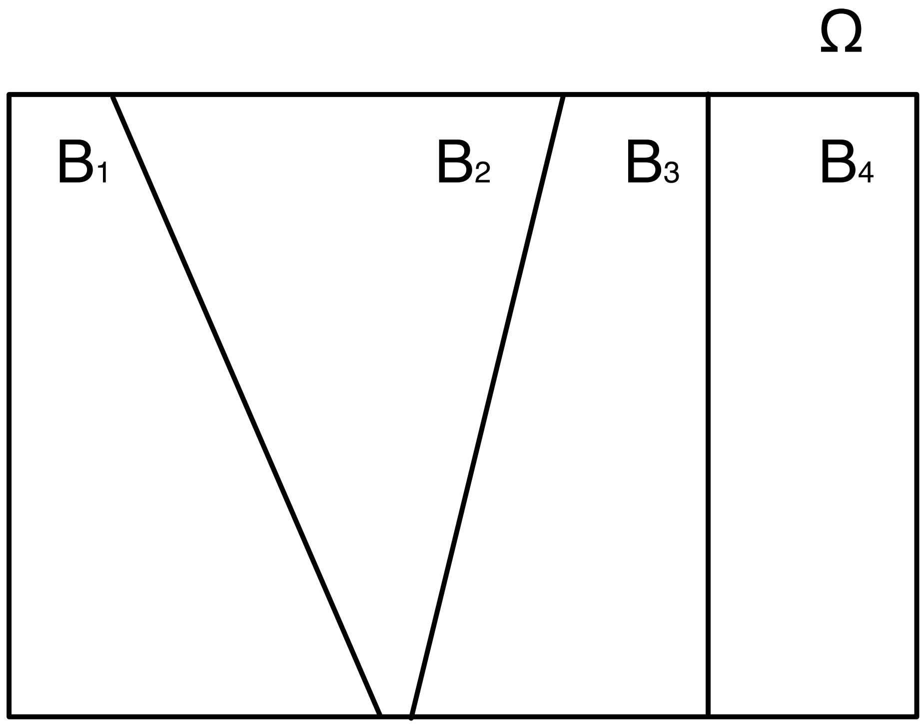
Exemple :

Au hasard, on choisit séquentiellement trois personnes dans un groupe contenant 4 hommes et 6 femmes. Quelle est la probabilité que trois femmes soient choisies ?

Si F_i est l'événement "la $i$ème personne choisie est une femme", alors on obtient $P[F_1 \cap F_2 \cap F_3] = P[F_1]P[F_2|F_1]P[F_3|F_1 \cap F_2] = \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} = \ldots = \frac{1}{6}$.

Deux théorèmes et un pseudo-théorème importants

Les deux vrais théorèmes impliquent une partition de Ω en événements $B_1, \ldots, B_k \in \mathcal{A}$ (ce qui signifie que $B_1 \cup \ldots \cup B_k = \Omega$ et $B_i \cap B_j = \emptyset$ pour $i \neq j$).



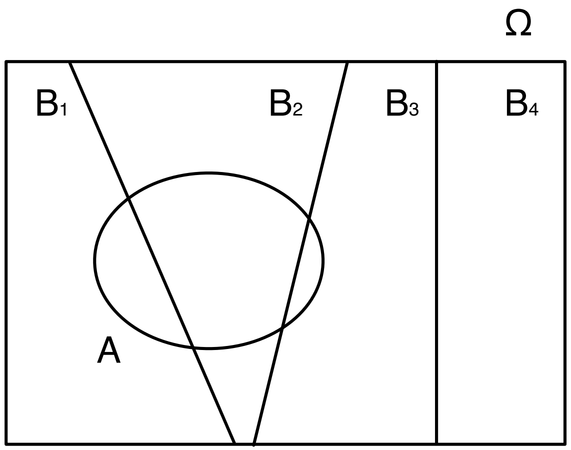
Partition de Ω en B_1, B_2, B_3, B_4 : exactement un des B_i se produit.

Théorème (Formule des probabilités totales)

Supposons que $B_1, \ldots, B_k \in \mathcal{A}$ forment une partition de Ω et que $P[B_i] > 0 \ \forall i$.

Alors, $\forall A \in \mathcal{A}$, on a $P[A] = \sum_{i=1}^{k} P[A|B_i]P[B_i]$.

Preuve :



$$
\begin{aligned}
P[A] &= P[(A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)] \\
&= P[A \cap B_1] + P[A \cap B_2] + \ldots + P[A \cap B_k] \\
&= P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + \ldots + P[A|B_k]P[B_k]
\end{aligned}
$$

$\square$

La probabilité $P[A]$ est donc une moyenne pondérée des probabilités $P[A|B_1], \ldots, P[A|B_k]$, avec poids $P[B_1], \ldots, P[B_k]$ (notons en effet que $\sum_{i=1}^{k} P[B_i] = 1$)

Deux théorèmes et un pseudo-théorème importants

Théorème (Formule des probabilités totales)

Supposons que $B_1, \ldots, B_k \in \mathcal{A}$ forment une partition de Ω et que $P[B_i] > 0 \ \forall i$.

Alors, $\forall A \in \mathcal{A}$, on a $P[A] = \sum_{i=1}^{k} P[A|B_i]P[B_i]$.

Exemple :

$E =$ au tennis, jouer un point où on sert
$A =$ gagner le point

$B_1 =$ réussir son 1er service ($P[B_1] = .70$)
$B_2 =$ rater son 1er service, mais réussir le 2nd ($P[B_2] = .25$)
$B_3 =$ faire une double faute ($\rightsquigarrow P[B_3] = 1 - P[B_1] - P[B_2] = .05$)

Si $P[A|B_1] = .9$ et $P[A|B_2] = .6$, on obtient

$$\begin{aligned} P[A] &= P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + P[A|B_3]P[B_3] \\ &= .9 \times .70 + .6 \times .25 + 0 \times .05 = .78 \end{aligned}$$

Deux théorèmes et un pseudo-théorème importants

Théorème (Formule de Bayes)

Supposons que $B_1, \ldots, B_k \in \mathcal{A}$ *forment une partition de* Ω *et que* $P[B_i] > 0 \ \forall i$.

Alors, $\forall A \in \mathcal{A}$ *avec* $P[A] > 0$*, on a*

$$P[B_j|A] = \frac{P[A|B_j]P[B_j]}{\sum_{i=1}^{k} P[A|B_i]P[B_i]},$$

pour tout $j = 1, \ldots, k$.

Preuve :

$$P[B_j|A] = \frac{P[B_j \cap A]}{P[A]} = \frac{P[A|B_j]P[B_j]}{P[A]} = \frac{P[A|B_j]P[B_j]}{\sum_{i=1}^{k} P[A|B_i]P[B_i]},$$

par le théorème de multiplication et la formule des probabilités totales. $\square$

Deux théorèmes et un pseudo-théorème importants

Exemple :

$E =$ faire un test covid
$A =$ le test est positif

$B_1 =$ la personne a le covid ($P[B_1] = .003$)
$B_2 =$ la personne est saine ($\rightsquigarrow P[B_2] = P[B_1^c] = 1 - P[B_1] = .997$)

Supposons que $P[A|B_1] = .95$ et $P[A^c|B_2] = .95$.

Alors $P[A|B_2] = 1 - P[A^c|B_2] = 1 - .95 = .05$, ce qui fournit

$$\begin{aligned} P[B_1|A] &= \frac{P[A|B_1]P[B_1]}{P[A|B_1]P[B_1] + P[A|B_2]P[B_2]} \\ &= \frac{.95 \times .003}{.95 \times .003 + .05 \times .997} \\ &\approx .05(!) \end{aligned}$$

1 Mesures de probabilité

- Expérience aléatoire, univers, événements, σ-algèbre
- Mesures de probabilité
- Analyse combinatoire
- Mesures de probabilité conditionnelle
- Deux théorèmes et un pseudo-théorème importants
- Indépendance stochastique

Indépendance stochastique

En général, le fait que B se produise influence la probabilité que A se produise :

$$P[A|B] \neq P[A]$$

Si ce n'est pas le cas, on dit que A et B sont indépendants.

Définition

Les événements A et B (de probabilités non nulles) sont indépendants $\Leftrightarrow$
$P[A|B] = P[A] \Leftrightarrow P[A \cap B] = P[A]P[B] \Leftrightarrow P[B|A] = P[B]$

La définition "A et B sont indépendants $\Leftrightarrow P[A \cap B] = P[A]P[B]$" est plus générale car elle tolère que A ou B soient de probabilité nulle.

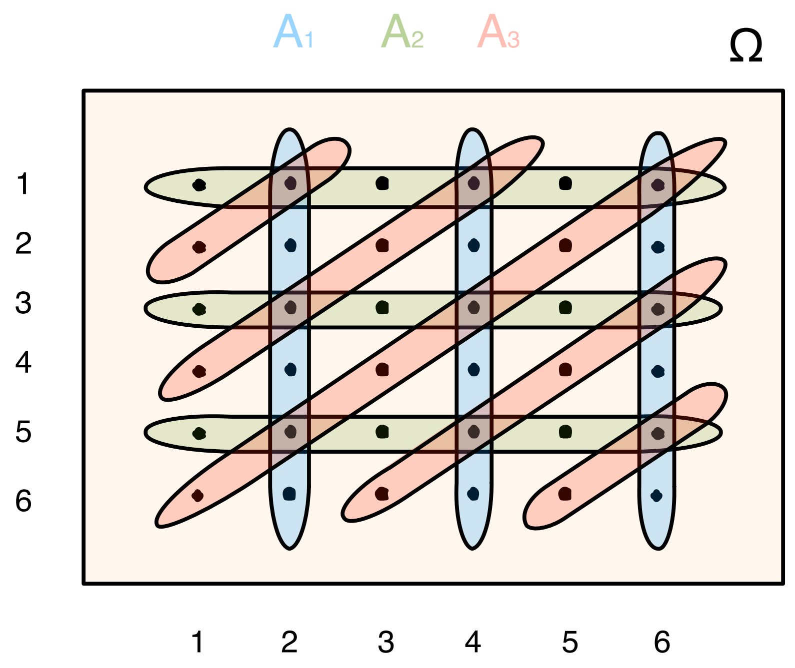
Mais les autres définitions sont préférables pour l'intuition.

Exemple :

$E =$ lancer de deux dés (distinguables)
$\Omega = \{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,6)\}$
$\mathcal{A} = \mathcal{P}(\Omega)$
P : équiprobabilité

$A_1 =$ le résultat du 1er dé est pair
$A_2 =$ le résultat du 2nd dé est impair
$A_3 =$ la somme des résultats est impaire

$P[A_1 \cap A_2] = \frac{9}{36} = \frac{18}{36} \times \frac{18}{36} = P[A_1]P[A_2]$
$P[A_1 \cap A_3] = \frac{9}{36} = \frac{18}{36} \times \frac{18}{36} = P[A_1]P[A_3]$
$P[A_2 \cap A_3] = \frac{9}{36} = \frac{18}{36} \times \frac{18}{36} = P[A_2]P[A_3]$
$\Rightarrow$ ces événements sont "2 à 2 indépendants"



Indépendance stochastique

Extension à un nombre fini $n \geq 3$ événements :

Définition

Les événements $A_1, A_2, \ldots, A_n$ sont mutuellement indépendants $\Leftrightarrow$
$\forall k \in \{2, 3, \ldots, n\}, \forall 1 \leq i_1 < i_2 < \ldots < i_k \leq n,$
$P[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] = P[A_{i_1}]P[A_{i_2}] \ldots P[A_{i_k}].$

Extension à un nombre infini (dénombrable) d'événements :

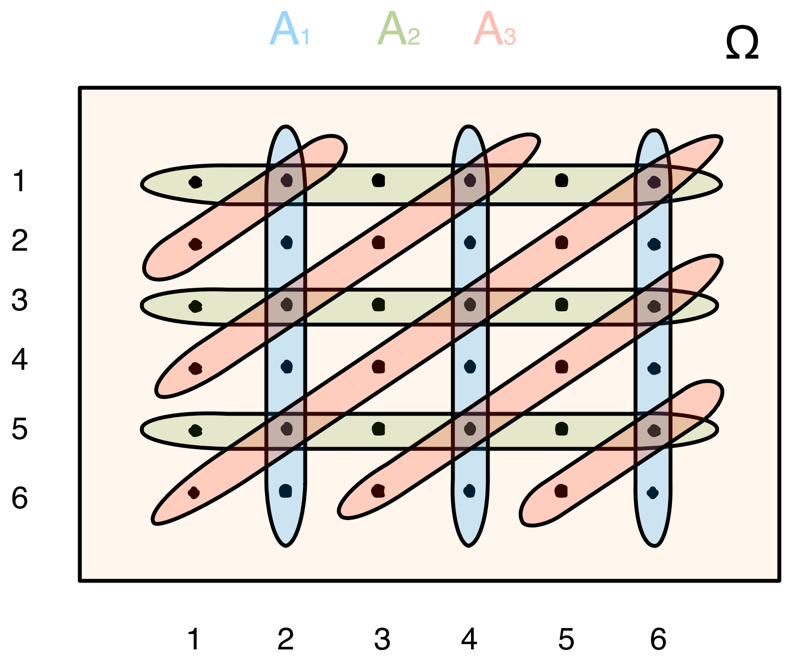
Définition

Les événements $A_1, A_2, \ldots$ sont mutuellement indépendants $\Leftrightarrow$
$\forall k \in \{2, 3, \ldots\}, \forall 1 \leq i_1 < i_2 < \ldots < i_k,$
$P[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] = P[A_{i_1}]P[A_{i_2}] \ldots P[A_{i_k}].$

Revenons à l'exemple précédent.

A_1 = le résultat du 1er dé est pair
A_2 = le résultat du 2nd dé est impair
A_3 = la somme des résultats est impaire

$P[A_1 \cap A_2] = P[A_1]P[A_2]$
$P[A_1 \cap A_3] = P[A_1]P[A_3]$
$P[A_2 \cap A_3] = P[A_2]P[A_3]$
$\Rightarrow$ ces événements sont 2 à 2 indépendants



Mais ces événements ne sont pas mutuellement indépendants car

$$P[A_1 \cap A_2 \cap A_3] = \frac{9}{36} \neq \frac{18}{36} \times \frac{18}{36} \times \frac{18}{36} = P[A_1]P[A_2]P[A_3].$$

Pour l'indépendance mutuelle de $n \geq 3$ événements, il est donc important d'aussi considérer $k \geq 3$ dans les définitions de la page précédente.