




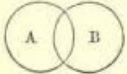
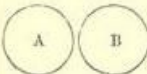
MAT4681 - Statistique pour les sciences

Arthur Charpentier

03 - Probabilités

été 2022

Ensembles / Sets

(i) Diagrammatic	(ii) Common Logic	(iii) Quantified	(iv) Symbolic
	$\left. \begin{array}{l} \text{All } A \text{ is } B \\ \text{All } B \text{ is } A \end{array} \right\}$	All A is all B	$\left. \begin{array}{l} A\bar{B} = 0 \\ \bar{A}B = 0 \end{array} \right\}$
	$\left. \begin{array}{l} \text{All } A \text{ is } B \\ \text{Some } B \text{ is not } A \end{array} \right\}$	All A is some B	$\left. \begin{array}{l} A\bar{B} = 0 \\ \bar{A}B = v \end{array} \right\}$
	$\left. \begin{array}{l} \text{All } B \text{ is } A \\ \text{Some } A \text{ is not } B \end{array} \right\}$	Some A is all B	$\left. \begin{array}{l} \bar{A}B = 0 \\ A\bar{B} = v \end{array} \right\}$
	$\left. \begin{array}{l} \text{Some } A \text{ is } B \\ \text{Some } A \text{ is not } B \\ \text{Some } B \text{ is not } A \end{array} \right\}$	Some A is some B	$\left. \begin{array}{l} AB = v \\ A\bar{B} = v \\ \bar{A}B = v \end{array} \right\}$
	No A is B	No A is any B	$AB = 0$

John Venn, *Symbolic Logic*, 1881.

Intersection & Union

Intersection

L'intersection de deux ensembles A et B , notée $A \cap B$, est l'ensemble des éléments de A qui sont aussi éléments de B

$$A \cap B = \{x : x \in A \text{ et } x \in B\}$$

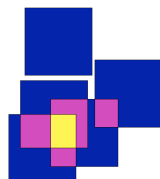
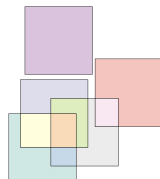
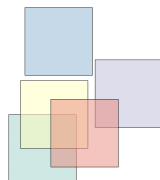
E.g. $\{1, 2, 3\} \cap \{2, 3, 4\} = \{2, 3\}$

Union

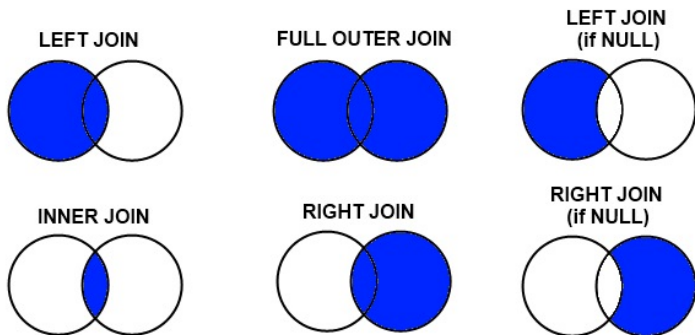
L'union de deux ensembles A et B , noté $A \cup B$, contient les éléments de A et les éléments de B

$$A \cup B = \{x : x \in A \text{ ou } x \in B\}$$

E.g. $\{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$



Fusionner et lier des données



via [Merge and Join DataFrames with Pandas](#)

Ensembles I

<https://aquaculture.ifremer.fr/Informations/Glossaire/Mois-en-R>

“Mois en ‘R’: *période des mois de septembre à avril dite favorable à la consommation des huîtres par opposition aux mois sans “R” de mai à août pendant laquelle on ne mangeait pas les coquillages autrefois car ils supportaient mal le transport.”*

Tous les mois de l'année :

$\Omega = \{ \text{Janvier, Février, Mars, Avril, Mai, Juin, Juillet, Août,} \\ \text{Septembre, Octobre, Novembre, Décembre} \}$

Mois en ‘R’ : $R =$

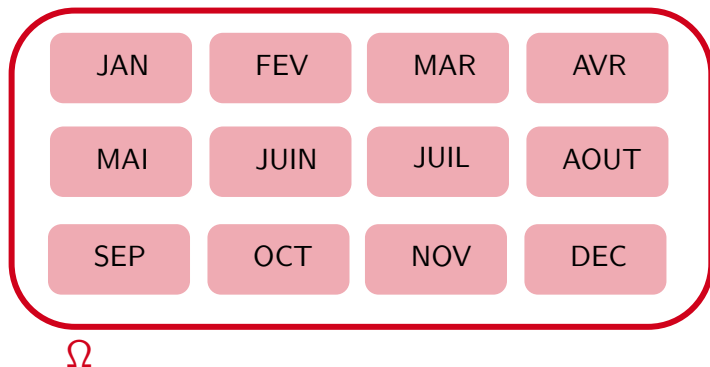
$\{ \text{Septembre, Octobre, Novembre, Décembre, Janvier, Février, Mars, Avril} \}$

Mois qui ont 31 jours :

$T = \{ \text{Janvier, Mars, Mai, Juillet, Août, Octobre, Décembre} \}$

Ensembles II

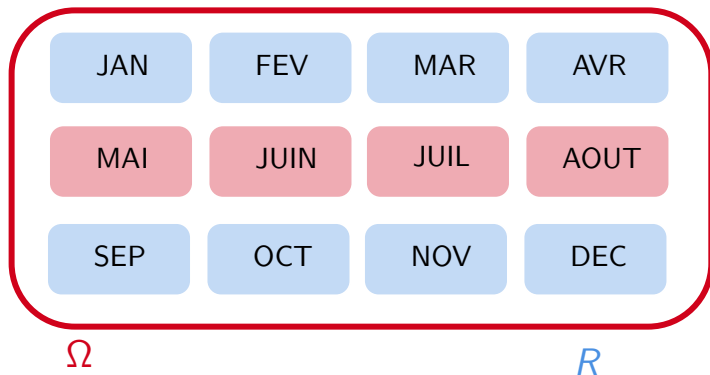
$\Omega = \{ \text{Janvier, Février, Mars, Avril, Mai, Juin, Juillet, Août,} \\ \text{Septembre, Octobre, Novembre, Décembre} \}$



Ensembles III

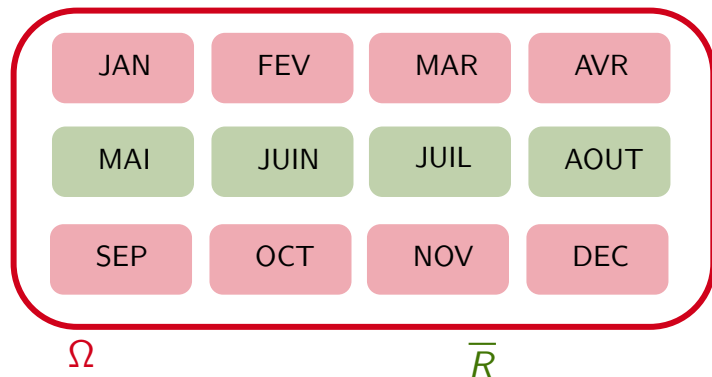
$R =$

$\{\text{Septembre, Octobre, Novembre, Décembre, Janvier, Février, Mars, Avril}\}$



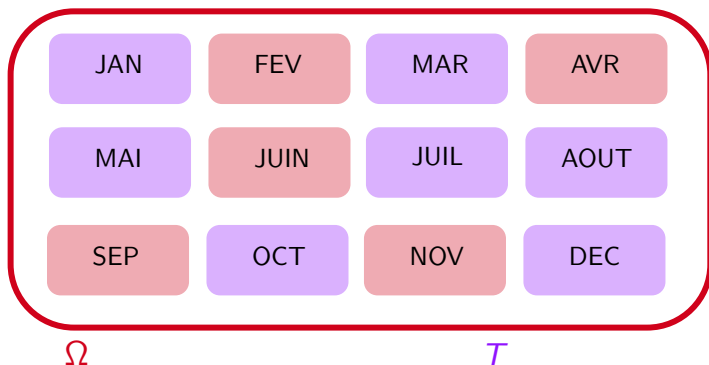
Ensembles IV

$$\overline{R} = \Omega \setminus R = \{Mai, Juin, Juillet, Août\}$$



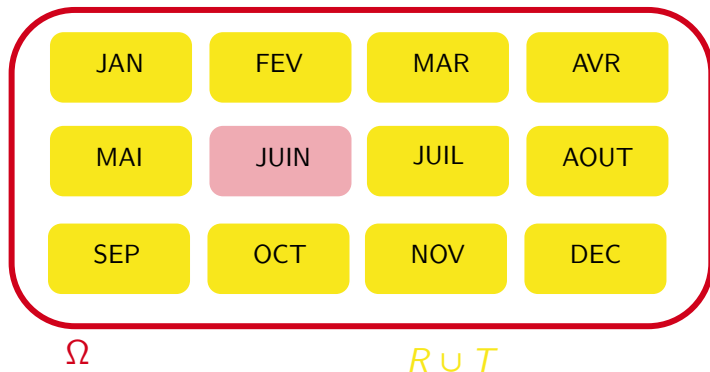
Ensembles V

$$T = \{\text{Janvier, Mars, Mai, Juillet, Août, Octobre, Décembre}\}$$



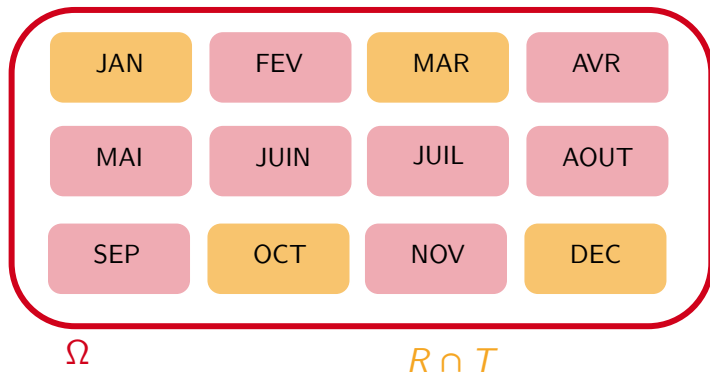
Ensembles VI

$$R \cup T = \Omega \setminus \{Juin\}$$



Ensembles VII

$$R \cap T = \{\text{Janvier}, \text{Mars}, \text{Octobre}, \text{Décembre}\}$$



Produit I

L'ADN est formé de suites de 4 nucléotides $N = \{A, C, G, T\}$

Combien existe-t-il de séquences de 3 nucléotides différentes ?

Réponse: $4 \times 4 \times 4 = 4^3 = 64$

$$\begin{Bmatrix} A \\ C \\ G \\ T \end{Bmatrix} \begin{Bmatrix} A \\ C \\ G \\ T \end{Bmatrix} \begin{Bmatrix} A \\ C \\ G \\ T \end{Bmatrix}$$

Produit II

AAA, AAC, AAG, AAT, ACA, ACC, ACG, ACT, AGA, AGC,
AGG, AGT, ATA, ATC, ATG, ATT, CAA, CAC, CAG, CAT,
CCA, CCC, CCG, CCT, CGA, CGC, CGG, CGT, CTA, CTC,
CTG, CTT, GAA, GAC, GAG, GAT, GCA, GCC, GCG, GCT
GGA, GGC, GGG, GGT, GTA, GTC, GTG, GTT, TAA, TAC,
TAG, TAT, TCA, TCC, TCG, TCT, TGA, TGC, TGG, TGT,
TTA, TTC, TTG, TTT

Combien existe-t-il de séquences de 3 nucléotides sans doublon ?

Réponse: $4 \times 3 \times 2 = 24$

$$\left\{ \begin{array}{c} A \\ C \\ G \\ T \end{array} \right\} \left\{ \begin{array}{c} \square \\ \square \\ \square \end{array} \right\} \left\{ \begin{array}{c} \square \\ \square \end{array} \right\}$$

Produit III

ACG, ACT, AGC, AGT, ATC, ATG, CAG, CAT, CGA, CGT, CTA, CTG, GAC, GAT, GCA, GCT, GTA, GTC, TAC, TAG, TCA, TCG, TGA, TGC

On cherche ici le nombre de permutations de 3 éléments parmi 4

$$4 \times 3 \times 2 = 24 = \frac{4!}{(4-3)!}$$

Combinaisons et permutation I

abc et *cba* sont deux permutations de $\{a, b, c\}$

Donnez toutes les permutations de 3 éléments de $\{a, b, c, d\}$

Réponse:

abc abd acb acd adb adc bac bad bca bcd bda bdc cab cad cba
cbd cda cdb dab dac dba dbc dca dcb

Il y a 24 permutations,
Attention l'ordre intervient !

Permutations

La permutation fait référence aux différentes façons d'organiser un ensemble d'objets dans un ordre séquentiel.

Combinaisons et permutation II

Donnez toutes les combinaisons de 3 éléments de $\{a, b, c, d\}$

Réponse:

abc abd acd bcd

Il y a 4 combinaisons, $\binom{4}{3} = \frac{4!}{3! \cdot 1!} = \frac{4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \cdot (1)}$

Combinaison

La combinaison fait référence à plusieurs manières de choisir des éléments dans un grand ensemble d'objets, de sorte que leur ordre n'a pas d'importance.

Combinaisons et permutation III

Comptez le nombre de façons d'obtenir exactement 3 têtes en 10 tirages d'une pièce de monnaie

Réponse:

$$\binom{10}{3} = \frac{10!}{3! \cdot 7!} = \frac{10 \times 9 \times 8 \cdots 2 \times 1}{(3 \times 2 \times 1) \cdot (7 \times 6 \times 5 \cdots 2 \times 1)} = 120$$

Pour une pièce de monnaie équitable, quelle est la probabilité d'obtenir exactement 3 face en 10 tirages ?

Il y a 2^{10} résultats possibles à partir de 10 lancers

$$\begin{matrix} \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} & \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \end{matrix}$$

Pour une pièce de monnaie équitable, chaque résultat a la même probabilité, de sorte que la probabilité d'obtenir exactement 3 face est la suivante

$$\frac{1}{2^{10}} \binom{10}{3} = \frac{120}{1024} \approx 11.7\%$$

Combinaisons et permutation IV

Combien y-a-t-il de possibilités de permuter les lettres de MATH ?

Il y a $4! = 24$ permutations

MATH MAHT MTAH MTHA MHAT MHTA AMTH AMHT
AHMT AHTM ATHM ATMH TAMH TAHM TMAH TMHA
THAM THMA HAMT HATM HMAT HMTA HTAM HTMA

Combien y-a-t-il de possibilités de permuter les lettres de STAT ?

Il y a 12 permutations

STAT SATT STTA TSAT ASTT TSTA TAST ATST TTSA
TATS ATTS TTAS

Formellement, on obtient le nombre de permutations possibles quand il y a des répétitions $\frac{4!}{2!1!1!}$, car (S,2), (A,1) et (T,1)

Combinaisons et permutation V

Combien y-a-t-il de possibilités de permuter STATISTIQUES ?

Il y a 12 lettres, (S,3) (T,3) (A,1) (I,2) (Q,1) (U,1) (E,1)

$$\frac{12!}{3!3!2!1!1!1!1!} = 6,652,800$$

Combinaisons et permutation VI

Combien peut-on écrire mots à partir de PARALLELE ? Combien de mots n'ont pas deux A consécutifs ?

Il y a 9 lettres, (A,2), (E,2), (L,3), (P,1), (R,1)

$$\frac{9!}{2!3!2!} = 15,120 \text{ mots}$$

Avec AA, il y a 8! arrangements possibles, avec les (E,2), (L,3)

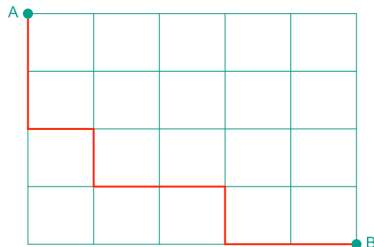
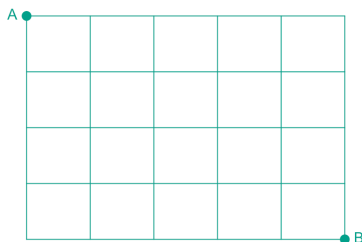
$$\frac{8!}{3!2!} = 3,360 \text{ mots}$$

Le nombre de mots sans AA est

$$\frac{9!}{2!3!2!} - \frac{8!}{3!2!} = 11,760 \text{ mots}$$

Combinaisons et permutation VII

Considérons deux points A et B sur la grille ci-dessous. **Combien y-a-t-il de chemins de A à B ?**



Un chemin est une succession de 9 pas, 4 vers le bas (B), 5 vers la droite (D). Le chemin de droite est BBDBDDDBDD

Il y a $\frac{9!}{5!4!} = 126$ chemins (ou “plus courts chemins”)

Combinaisons et permutation VIII

De combien de façons cinq garçons et cinq filles peuvent-ils s'aligner, si aucun deux garçons ou filles ne peuvent se suivre ?

Il faut alterner, on a soit BGBGBGBGBG, soit GBGBGBGBGB.

Dans le premier cas, les garçons peuvent être disposés de 5! façons, tout comme les filles, soit un total de 5! façons.

Soit un total de $5! \times 5!$ arrangements.

Pareil pour le deuxième (symétrique)

pour un total supplémentaire de $5! \times 5!$ supplémentaires.

Il peuvent s'aligner dans $2 \times (5!)^2 = 28,800$ façons

On notera que les probabilités faibles surviennent,

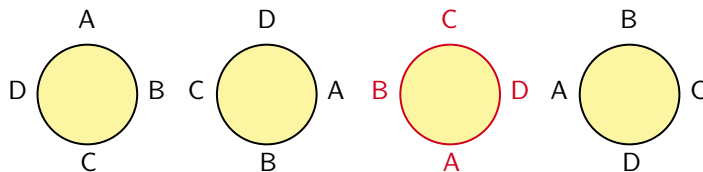
$$\frac{1}{28800} = 0.000035$$

(souvent, ce sont les probabilités relatives qui nous intéressent)

Combinaisons et permutation IX

De combien de façons peut-on asseoir 9 personnes autour d'une table circulaire ?

Attention, avec 4 personnes, les arrangements ci-dessous sont identiques



On choisit arbitrairement une personne, puis on place les 8 personnes qui restent à sa droite, de $8! = 40,320$ façons.

Jeux de cartes I

Jeux de 52 cartes,

- ▶ 13 rangs : $\{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\}$
- ▶ 4 couleurs (enseignes) : $\{\spadesuit, \clubsuit, \heartsuit, \diamondsuit\}$

Une main (au poker) est un ensemble de 5 cartes (différentes)

Example $\{4\spadesuit, 10\spadesuit, 7\clubsuit, D\heartsuit, V\diamondsuit\}$

Une main ayant “une paire” contient 2 cartes de même rang, les 3 autres étant de rang distinct

Example $\{4\spadesuit, 10\spadesuit, D\clubsuit, D\heartsuit, V\diamondsuit\}$

Quelle est la probabilité d’avoir “une paire” ?

Jeux de cartes II

► méthode 1 (sans tenir compte de l'ordre)

1. choisir le rang de la paire : 13 rangs différents
2. choisir les couleurs de la paire : 2 couleurs parmi 4, $\binom{4}{2}$
3. choisir 3 rangs parmi les $13 - 1 = 12$ restants (car il ne faut pas d'autre paire) $\binom{13}{3}$
4. choisir les 3 couleurs associées aux 3 rangs, soit 4^3

Le nombre de mains avec “une paire” est

$$13 \times \binom{4}{2} \times \binom{13}{3} \times 4^3 = 1098240$$

Le nombre de mains possibles est $\binom{52}{5} = 2598960$

Jeux de cartes III

donc une probabilité $\frac{1098240}{2598960} \approx 42.257\%$.

► méthode 2 (en tenant compte de l'ordre)

1. choisir les positions dans la main pour la paire : $\binom{5}{2}$
2. Placer une carte dans la première position de la paire : 52 cartes, donc 52 façons de faire.
3. Placer une carte dans la deuxième position de la paire : comme elle doit correspondre au rang première carte, il n'y a que 3 façons de faire.
4. Placer une carte dans le premier emplacement libre : elle ne peut pas correspondre à la paire, il y a donc $52 - 4 = 48$ façons de faire.

Jeux de cartes IV

5. Placer une carte dans l'emplacement libre suivant : elle ne peut pas correspondre à la paire ou à la carte précédente, il y a donc $48 - 4 = 44$ façons de faire.
6. Placez une carte dans le dernier emplacement libre : il y a $44 - 4 = 40$ façons de faire.

Le nombre de mains avec “une paire” est (en tenant compte de l'ordre)

$$\binom{5}{2} \times 52 \times 3 \times 48 \times 44 \times 40 = 131788800$$

Le nombre de mains est $52 \times 51 \times 50 \times 49 \times 48 = 311875200$

donc une probabilité $\frac{131788800}{311875200} \approx 42.257\%$.

Formalisme probabiliste I

On note Ω tous les résultats possibles

$\Omega = \{ \text{Janvier, Février, Mars, Avril, Mai, Juin, Juillet, Août,} \\ \text{Septembre, Octobre, Novembre, Décembre} \}$

Combien y-a-t-il d'éléments dans Ω ?

Réponse : $n = 12$

On note \mathcal{F} l'ensemble des parties de Ω

Mois qui ont 31 jours :

$T = \{ \text{Janvier, Mars, Mai, Juin, Juillet, Août, Octobre, Décembre} \}$

$T \in \mathcal{F}$

Combien y-a-t-il d'éléments dans \mathcal{F} ?

Réponse : 2^n , soit ici 4096

$\emptyset \in \mathcal{F}$ et $\Omega \in \mathcal{F}$

Formalisme probabiliste II

Une probabilité est une fonction $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$.

Que vérifie la fonction \mathbb{P} ?

- ▶ $\mathbb{P}(\emptyset) = 0$ et $\mathbb{P}(\Omega) = 1$
- ▶ si $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$
- ▶ $\mathbb{P}(A) + \mathbb{P}(\bar{A}) = 1$
- ▶ si $A \cap B = \emptyset$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
- ▶ $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$

Probabilité (version simplifiée)

Une probabilité est une fonction $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ telle que

- ▶ $\mathbb{P}(\emptyset) = 0$ et $\mathbb{P}(\Omega) = 1$
- ▶ si $A \cap B = \emptyset$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Formalisme probabiliste III

On lance 3 pièces de monnaie

$$\Omega = \{PPP, PPF, PFP, FPP, FFP, FPF, PFF, FFF\}$$

$$\begin{cases} A_0 = \text{obtenir aucune fois 'face' sur 3 lancers} \\ A_1 = \text{obtenir exactement 1 fois 'face' sur 3 lancers} \\ A_2 = \text{obtenir exactement 2 fois 'face' sur 3 lancers} \\ A_3 = \text{obtenir exactement 3 fois 'face' sur 3 lancers} \end{cases}$$

Si la pièce n'est pas truquée

$$\begin{cases} A_0 = \{PPP\} \in \mathcal{F}, \text{ et } \mathbb{P}(A_0) = 1/8 \\ A_1 = \{FPP, PFP, PPF\} \in \mathcal{F}, \text{ et } \mathbb{P}(A_1) = 3/8 \\ A_2 = \{PFF, FPF, FFP\} \in \mathcal{F}, \text{ et } \mathbb{P}(A_2) = 3/8 \\ A_3 = \{FFF\} \in \mathcal{F}, \text{ et } \mathbb{P}(A_3) = 1/8 \end{cases}$$

$$A_0 \cup \dots \cup A_3 = \Omega \text{ et } \forall i, j, A_i \cap A_j = \emptyset$$

Formalisme probabiliste IV

Partition

A_1, \dots, A_k forme une partition de Ω si

- ▶ $A_1 \cup \dots \cup A_k = \Omega$
- ▶ $\forall i, j, A_i \cap A_j = \emptyset$

Partition

Si A_1, \dots, A_k forme une partition de Ω

$$\mathbb{P}(A_1) + \dots + \mathbb{P}(A_k) = \sum_{i=1}^k \mathbb{P}(A_i) = 1$$

Espace probabilisé

Un espace probabilisé est $(\Omega, \mathcal{F}, \mathbb{P})$ avec

- ▶ un ensemble d'évènements Ω
- ▶ l'ensemble \mathcal{F} des parties de Ω
- ▶ une probabilité $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$

Formalisme probabiliste VI

Dans une classe de 50 étudiants,

il y a 20 hommes, 25 personnes brunes

On tire une personne au hasard, A est l'évènement "la personne est un homme", B est l'évènement "la personne est brune".

Que vaut $\mathbb{P}(A \cup B)$?

$$\mathbb{P}(A \cup B) = \underbrace{\mathbb{P}(A)}_{20/50} + \underbrace{\mathbb{P}(B)}_{25/50} - \underbrace{\mathbb{P}(A \cap B)}_{?}$$

- ▶ $\mathbb{P}(A \cap B)$ est minimale si $A \cup B = \emptyset$ et $\mathbb{P}(A \cap B) = 0$
 - ▶ $\mathbb{P}(A \cap B)$ est maximale si $A \subset B$ et $\mathbb{P}(A \cap B) = \mathbb{P}(A) = 40\%$
- donc $\mathbb{P}(A \cup B) \in [50\%; 90\%]$.

Conditionnement I

Probabilité conditionnelle

Soient deux évènements A et B (dans \mathcal{F}) tels que $\mathbb{P}(B) \neq 0$, on note

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

$$\text{Aussi, } \mathbb{P}(A \cap B) = \begin{cases} \mathbb{P}(A|B) \times \mathbb{P}(B) \\ \mathbb{P}(B|A) \times \mathbb{P}(A) \end{cases}$$

Conditionnement II

On lance une pièce 4 fois, $\begin{cases} A = \text{au moins 3 fois 'pile'} \\ B = \text{'face' au premier tirage} \end{cases}$

$$\begin{cases} \Omega = \{FFFF, FFFP, \dots, FPPP, PPPP\} \\ A = \{FPPP, PFPP, PPFP, PPPF, PPPP\} \\ B = \{FPPP, FPPF, FPFP, FFPP, FPFF, FFPF, FPFF, FFFF\} \\ A \cap B = \{FPPP\} \end{cases}$$

Conditionnement III

On lance une pièce 4 fois, $\begin{cases} A = \text{au moins 3 fois 'pile'} \\ B = \text{'face' au premier tirage} \end{cases}$

Que vaut $\mathbb{P}(A|B)$?

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{8}$$

Que vaut $\mathbb{P}(B|A)$?

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{1}{5}$$

Formule des probabilités totales et formule de Bayes I

Si B_1, \dots, B_k forme une partition de Ω , i.e.

► $B_1 \cup \dots \cup B_k = \Omega$

► $\forall i, j, B_i \cap B_j = \emptyset$

$$\mathbb{P}(A) = \sum_{i=1}^k \mathbb{P}(A \cap B_i) = \sum_{i=1}^k \mathbb{P}(A|B_i) \times \mathbb{P}(B_i)$$

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

Indépendance I

Indépendance

Deux évènements A et B (dans \mathcal{F}) sont indépendants si

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

On note $A \perp\!\!\!\perp B$

Il s'agit de la définition de l'indépendance (on verra d'autres caractérisations plus simples par la suite)

Notons que si $A \subset B$ avec $A \neq B$, $A \cap B = A \neq B$, et $B \neq \Omega$,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) > \mathbb{P}(A) \times \mathbb{P}(B)$$

donc $A \not\perp\!\!\!\perp B$

Indépendance II

On lance deux fois une même pièce, possiblement truquée

$$\mathbb{P}(\{PP, FF\}) \stackrel{?}{\gtrless} \mathbb{P}(\{FP, PF\})$$

Soit $p = \mathbb{P}(\{P\})$ (et $1 - p = \mathbb{P}(\{F\})$). Si les tirages sont indépendants,

$$\begin{cases} \mathbb{P}(\{PP\}) = \mathbb{P}(\{P\}) \cdot \mathbb{P}(\{P\}) = p^2 \\ \mathbb{P}(\{FF\}) = \mathbb{P}(\{F\}) \cdot \mathbb{P}(\{F\}) = (1 - p)^2 \\ \mathbb{P}(\{FP\}) = \mathbb{P}(\{F\}) \cdot \mathbb{P}(\{P\}) = (1 - p)p \\ \mathbb{P}(\{PF\}) = \mathbb{P}(\{P\}) \cdot \mathbb{P}(\{F\}) = p(1 - p) \end{cases}$$

Notons qu'on ne sait pas trompé, ces 4 valeurs somment à 1

$$\mathbb{P}(\{PP, FF, FP, PF\}) = p^2 + (1 - p)^2 + 2p(1 - p) = 1$$

Indépendance III

Maintenant, rappelons que pour tout a et b , $a^2 + b^2 \geq 2ab$, car $a^2 + b^2 - 2ab = (a - b)^2 \geq 0$.

Si $a = p$ et $b = 1 - p$, on a alors $p^2 + (1 - p)^2 \geq 2p(1 - p)$

$$\mathbb{P}(\{PP\}) + \mathbb{P}(\{FF\}) \geq \mathbb{P}(\{PF\}) + \mathbb{P}(\{FP\})$$

donc $\mathbb{P}(\{PP, FF\}) \geq \mathbb{P}(\{FP, PF\})$.

Plus précisément, si $p = 1/2$ (pièce équilibrée)

$$\mathbb{P}(\{PP, FF\}) = \mathbb{P}(\{FP, PF\})$$

alors que si $p \neq 1/2$ (pièce déséquilibrée)

$$\mathbb{P}(\{PP, FF\}) > \mathbb{P}(\{FP, PF\})$$

Indépendance IV

Probabilité conditionnelle et indépendance

Deux évènements A et B (dans \mathcal{F}) sont indépendants

$$\begin{cases} \text{si } \mathbb{P}(A) \neq 0, \mathbb{P}(B|A) = \mathbb{P}(B) \\ \text{si } \mathbb{P}(B) \neq 0, \mathbb{P}(A|B) = \mathbb{P}(A) \end{cases}$$

On lance deux pièces. $A = \{ \text{face pour la première pièce} \}$ et $B = \{ \text{face pour la seconde pièce} \}$, $\mathbb{P}[B|A] = \mathbb{P}[B]$

Indépendance V

On lance deux dés (à 6 faces), $\begin{cases} A = \text{le premier dé est un 3} \\ B = \text{la somme des dés vaut 6} \\ C = \text{la somme des dés vaut 7} \end{cases}$

$$\begin{cases} A = \{31, 32, 33, 34, 35, 36\} \\ B = \{15, 24, 33, 42, 51\} \\ C = \{16, 25, 34, 43, 52, 61\} \\ A \cap B = \{33\} \text{ et } A \cap C = \{34\} \end{cases}$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{5} \neq \mathbb{P}(A), \quad A \not\perp B$$

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{1}{6} = \mathbb{P}(A), \quad A \perp C$$

Indépendance VI

Parmi les évènements suivants, quand a-t-on $A \perp\!\!\!\perp B$?

On lance deux dés, $A = \{ \text{premier dé est 6} \}$ et $B = \{ \text{somme} > 6 \}$

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde } \clubsuit \}$

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde est 10} \}$

Indépendance VII

Parmi les évènements suivants, quand a-t-on $A \perp\!\!\!\perp B$?

On lance deux dés, $A = \{ \text{premier dé est 6} \}$ et $B = \{ \text{somme} > 6 \}$

Non indépendants ($A \subset B$)

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde } \clubsuit \}$

Non indépendants $\mathbb{P}[A \cap B] = \frac{1}{4} \cdot \frac{13}{51} > \frac{1}{4} \cdot \frac{13}{52} = \mathbb{P}[A] \cdot \mathbb{P}[B]$

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde est 10} \}$

Indépendants $\mathbb{P}[A \cap B] = \frac{1}{52} = \mathbb{P}[A] \cdot \mathbb{P}[B]$

Indépendance VIII

On lance deux dés, $A = \{ \text{premier dé est 6} \}$ et $B = \{ \text{somme} > 6 \}$

Notons que $A = \{61, 62, 63, 64, 65, 66\}$ alors que $B =$

$\{16, 25, 26, 34, 35, 36, 43, 44, 45, 46, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66\}$

et le cardinal de Ω est $6 \times 6 = 36$. Aussi

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\text{Card}(\{61, 62, 63, 64, 65, 66\})}{\text{Card}(B)} = \frac{6}{21}$$

qui n'est pas égal à $\mathbb{P}[A]$, qui vaut $1/6$. De même

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]} = \frac{\text{Card}(\{61, 62, 63, 64, 65, 66\})}{\text{Card}(A)} = \frac{6}{6} = 1$$

qui n'est pas égal à $\mathbb{P}[B]$, qui vaut $21/36$. De même

$$\mathbb{P}[A] \cdot \mathbb{P}[B] = \frac{1}{6} \cdot \frac{21}{36} \neq \mathbb{P}[A \cap B] = \frac{6}{36}$$

Indépendance IX

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde } \clubsuit \}$

$$\mathbb{P}[B|A] = \frac{13}{51} \neq \mathbb{P}[B] = \frac{1}{4}$$

$$\mathbb{P}[A|B] = \frac{13}{51} \neq \mathbb{P}[A] = \frac{1}{4}$$

$$\mathbb{P}[B \cap A] = \mathbb{P}[B|A] \cdot \mathbb{P}[A] = \frac{12}{51} \cdot \frac{1}{4} \neq \mathbb{P}[A] \cdot \mathbb{P}[B]$$

autrement dit A et B ne sont pas indépendants.

Intuitivement, savoir qu'on a eu un \heartsuit (et pas un \clubsuit) au premier tirage donne une information importante pour prédire *la couleur* de la carte qu'on aura au second tirage.

Indépendance X

On peut regarder tous les tirages possibles (♥ = H et ♣ = C)

```
1 > N = rep(1:13, each=4)
2 > C = rep(c("D","C","H","S"), 4)
3 > X = paste0(N,C)
4 > X1 = rep(X,length(X))
5 > X2 = rep(X,each = length(X))
6 > Y = cbind(X1,X2)[which(X1 != X2),]
```

Nos deux évènements sont ici

```
1 > A = substr(Y[,1], nchar(Y[,1]), nchar(Y[,1]))=="H"
2 > B = substr(Y[,2], nchar(Y[,2]), nchar(Y[,2]))=="C"
```

et calculer les probabilité, par exemple $\mathbb{P}[A \cap B]$ et $\mathbb{P}[A] \cdot \mathbb{P}[B]$

```
1 > mean(A&B)
2 [1] 0.06372549
3 > mean(A) * mean(B)
4 [1] 0.0625
```

qui sont manifestement différentes,

Indépendance XI

ou les probabilités conditionnelles, par exemple $\mathbb{P}[B|A]$ et $\mathbb{P}[B]$

```
1 > mean(B[A])
2 [1] 0.254902
3 > mean(B)
4 [1] 0.25
```

ou $\mathbb{P}[A|B]$ et $\mathbb{P}[A]$

```
1 > mean(A[B])
2 [1] 0.254902
3 > mean(A)
4 [1] 0.25
```

qui sont manifestement différentes (comme on devrait s'y attendre).

Indépendance XII

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde } 10 \}$

Intuitivement, savoir qu'on a eu un \heartsuit au premier tirage ne donne aucune information pertinente pour prédire *la valeur* de la carte qu'on aura au second tirage.

On utilise la formule des probabilités totales en conditionnant que la première carte a été, ou pas un $10\heartsuit$. Pour simplifier, on note $\mathbb{P}[\cdot|A] = \mathbb{P}_A[\cdot]$

$$\mathbb{P}[B|A] = \mathbb{P}_A[B|\{10\heartsuit\}] \cdot \mathbb{P}_A[\{10\heartsuit\}] + \mathbb{P}_A[B|\overline{\{10\heartsuit\}}] \cdot \mathbb{P}_A[\overline{\{10\heartsuit\}}]$$

$$\mathbb{P}[B|A] = \mathbb{P}[B|\{10\heartsuit\}] \cdot \mathbb{P}[\{10\heartsuit\}|A] + \mathbb{P}_A[B|\overline{\{10\heartsuit\}}] \cdot \mathbb{P}[\overline{\{10\heartsuit\}}|A]$$

soit

$$\mathbb{P}[B|A] = \frac{3}{51} \cdot \frac{1}{13} + \frac{4}{51} \cdot \frac{12}{13} = \frac{1}{13} \text{ et } \mathbb{P}[B] = \frac{1}{13}$$

donc $\mathbb{P}[B|A] = \mathbb{P}[B]$.

Indépendance XIII

On pige deux cartes, $A = \{ \text{première } \heartsuit \}$ et $B = \{ \text{seconde } 10 \}$

De même, on peut regarder l'autre conditionnement, en notant cette fois $\{10 \heartsuit\}$ l'obtention du $10 \heartsuit$ au second tirage

$$\mathbb{P}[A|B] = \mathbb{P}_B[A|\{10 \heartsuit\}] \cdot \mathbb{P}_B[\{10 \heartsuit\}] + \mathbb{P}_B[A|\overline{\{10 \heartsuit\}}] \cdot \mathbb{P}_B[\overline{\{10 \heartsuit\}}]$$

$$\mathbb{P}[A|B] = \mathbb{P}[A|\{10 \heartsuit\}] \cdot \mathbb{P}[\{10 \heartsuit\}] + \mathbb{P}_B[A|\overline{\{10 \heartsuit\}}] \cdot \mathbb{P}[\overline{\{10 \heartsuit\}}|B]$$

soit

$$\mathbb{P}[A|B] = \frac{12}{51} \cdot \frac{1}{4} + \frac{13}{51} \cdot \frac{3}{4} = \frac{1}{4} \text{ et } \mathbb{P}[A] = \frac{1}{4}$$

donc $\mathbb{P}[A|B] = \mathbb{P}[A]$.

Pareil, on peut montrer que $\mathbb{P}[A \cap B] = \frac{1}{52}$

Indépendance XIV

Là encore, on peut faire du dénombrement (informatique),

```
1 > A = substr(Y[,1], nchar(Y[,1]), nchar(Y[,1]))=="H"  
2 > B = substr(Y[,2], 1, 2)=="10"
```

On peut commencer par la probabilité dite jointe, $\mathbb{P}[A \cap B]$ et $\mathbb{P}[A] \cdot \mathbb{P}[B]$

```
1 > mean(A&B)  
2 [1] 0.01923077  
3 > mean(A) * mean(B)  
4 [1] 0.01923077
```

soit 1/52.

Indépendance XV

Pour les probabilités conditionnelles, par exemple $\mathbb{P}[B|A]$ et $\mathbb{P}[B]$

```
1 > mean(B[A])
2 [1] 0.07692308
3 > mean(B)
4 [1] 0.07692308
```

on retrouve $1/13$, ou $\mathbb{P}[A|B]$ et $\mathbb{P}[A]$

```
1 > mean(A[B])
2 [1] 0.25
3 > mean(A)
4 [1] 0.25
```

on retrouve $1/4$.

Bref, quelle que soit la méthode, on retrouve le fait que A et B sont deux évènements indépendants.

Probabilité et cote I

Dans un jeu de 52 cartes, soit A l'évènement 'en tirant une carte au hasard, on a un nombre pair'. Que vaut $\mathbb{P}[A]$?

Il y a 5 cartes associés à un nombre pair (2, 4, 6, 8, 10) et 4 couleurs, donc 20 cartes en tout. Donc

$$\mathbb{P}[A] = \frac{20}{52} = \frac{5}{13}$$

Cote

La cote d'un évènement A (ou en faveur de A) est $a : \bar{a}$ où $\mathbb{P}[A] = a/(a + \bar{a})$, ou parfois $\mathbb{P}[A]/\mathbb{P}[\bar{A}]$.

- ▶ 1 : 1 correspond à un évènement de probabilité 1/2.
- ▶ 2 : 1 correspond à un évènement de probabilité 2/3.
- ▶ 1 : 2 correspond à un évènement de probabilité 1/3.

Probabilité et cote II

Dans l'exemple précédant (carte paire), **quelle est la cote de A ?**

Notons que $\mathbb{P}[A] = \frac{20}{52} = \frac{5}{13}$ et $\mathbb{P}[\bar{A}] = 1 - \frac{5}{13} = \frac{8}{13}$

La cote est alors 5 : 8.

La cote est aussi parfois simplement définie comme le rapport $\mathbb{P}[A]/\mathbb{P}[\bar{A}]$

Pour passer d'une cote $a : \bar{a}$ à $\mathbb{P}[A]$, notons que

$$\mathbb{P}[A] \propto a \text{ et } \mathbb{P}[\bar{A}] \propto \bar{a}$$

$$\text{et donc } \mathbb{P}[A] = \frac{a}{a + \bar{a}}$$

Arbres et ensembles

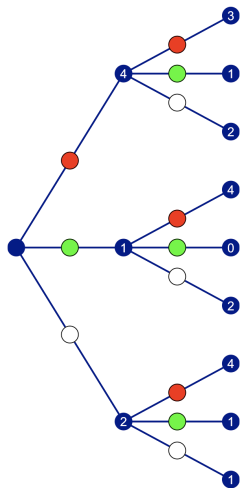
Une urne avec 10 boules, 5 rouge ● 2 vertes ● et 3 blanches ○.

La **probabilité conditionnelle** de A sachant que B est survenu est

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \text{ et } B]}{\mathbb{P}[B]}$$

ou $\mathbb{P}[A \text{ and } B] = \mathbb{P}[A|B] \cdot \mathbb{P}[B]$
(on parle de **chain rule**)

On pige 2 boules, sans remise, quelle est la probabilité d'avoir au moins une boule verte ?



Arbres et ensembles

On pige 2 boules, sans remise, quelle est la probabilité d'avoir au moins une boule verte ?

$$p = \mathbb{P}[X_1 = \bullet \text{ ou } X_2 = \bullet]$$

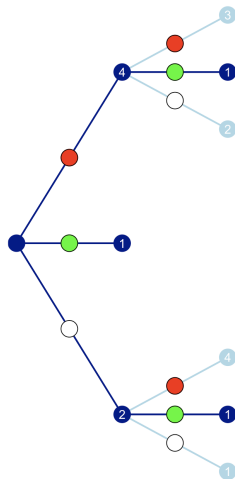
$$p = \mathbb{P}[X_1 = \bullet] + \mathbb{P}[X_2 = \bullet \text{ et } X_1 \neq \bullet]$$

$$p = \mathbb{P}[X_1 = \bullet] + \mathbb{P}[X_2 = \bullet | X_1 = \bullet] \cdot \mathbb{P}[X_1 = \bullet] \\ + \mathbb{P}[X_2 = \bullet | X_1 = \circ] \cdot \mathbb{P}[X_1 = \circ]$$

$$p = \frac{2}{10} + \frac{2}{9} \cdot \frac{5}{10} + \frac{2}{9} \cdot \frac{3}{10} = \dots = \frac{17}{45}$$

ou encore plus simple

$$p = 1 - \mathbb{P}[X_1 \in \{\bullet, \circ\} \text{ et } X_2 \in \{\bullet, \circ\}] = 1 - \frac{8}{10} \cdot \frac{7}{9} = \frac{17}{45}$$



Monty Hall

$$\begin{aligned} & \mathbb{P}(\text{trésor derrière l'autre porte}) \\ &= \mathbb{P}(\text{trésor derrière l'autre porte} | X \text{ correct}) \cdot \mathbb{P}(X \text{ correct}) \\ &+ \mathbb{P}(\text{trésor derrière l'autre porte} | X \text{ faux}) \cdot \mathbb{P}(X \text{ faux}) \\ &= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3} \end{aligned}$$

donc oui, il faut prendre l'autre porte.

Paradoxe des anniversaires

Considérons un ensemble de n personnes choisies au hasard. Si $n \geq 23$, il y a plus de 50% de chances qu'une paire d'entre elles ait le même anniversaire.

(en supposant que chaque jour de l'année a la même probabilité d'être un anniversaire)

$A = \{(\text{au moins}) \text{ une paire d'entre eux aura le même anniversaire}\}$

$\bar{A} = \{\text{aucune paire d'entre eux n'aura le même anniversaire}\}$

$$\mathbb{P}(\bar{A}) = \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \cdots \times \frac{343}{365} \simeq 49.2703\%$$

Approximation de Poisson, $\lambda = \frac{1}{365} \binom{23}{2} = \frac{253}{365} \simeq 0.6932$ i.e.

$$\mathbb{P}(X > 0) = 1 - \mathbb{P}(X = 0) \simeq 1 - e^{-0.6932} \simeq 0.500002$$

Test médicaux I

La probabilité qu'une femme ait un cancer du sein est de 1%.
Si une femme a un cancer du sein, la probabilité que le test soit positif est de 90%. Si une femme n'a pas de cancer du sein, la probabilité qu'elle soit néanmoins positive est de 9%.

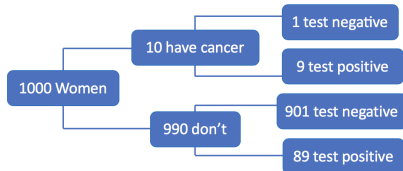
Une femme de 50 ans, sans symptômes, participe à un dépistage de routine par mammographie. Le test est positif, elle est alarmée et veut que vous lui disiez si elle est certaine d'avoir un cancer du sein ou quelles sont les chances qu'elle en ait un. En dehors des résultats du dépistage, vous ne savez rien d'autre sur cette femme. Combien de femmes dont le test est positif ont réellement un cancer du sein ? Quelle est la meilleure réponse ?

- A) 9 sur 10
- B) 8 sur 10
- C) 1 sur 10
- D) 1 sur 100



Test médicaux II

La probabilité qu'une femme ait un cancer du sein est de 1%.
Si une femme a un cancer du sein, la probabilité que le test soit positif est de 90%. Si une femme n'a pas de cancer du sein, la probabilité qu'elle soit néanmoins positive est de 9%.



$$\mathbb{P}[\text{cancer}|\text{test positif}] = \frac{9}{9 + 89} \approx \frac{1}{10}$$

See [Do doctors understand test results?](#) la moitié du groupe de 160 gynécologues a répondu que la probabilité pour la femme d'avoir un cancer était de neuf sur dix

Complément I

Ω peut être un ensemble relativement grand

- ▶ $\Omega = \{0, 1\}$ ou $\Omega = \{0, 1, \dots, n\}$: fini
- ▶ $\Omega = \mathbb{N} = \{0, 1, \dots, n, \dots\}$ (entiers naturels) : dénombrable
- ▶ $\Omega = \mathbb{R}$ (réels) : non dénombrable