

### III. Probabilités conditionnelles

Le concept de probabilité conditionnelle est l'un des plus importants de cette théorie puisque l'on cherche souvent à savoir quelle est la probabilité d'un événement alors qu'on dispose d'une information partielle. La probabilité qu'il neige un matin n'est pas la même si l'on sait qu'il a fait 30 degrés la veille ou -5 degrés.

## 1 Rappels sur les probabilités

L'ensemble fondamental d'une expérience est l'ensemble de toutes les issues possibles d'une expérience.

**Définition 1.1.** Soit  $S$  l'ensemble fondamental d'une expérience.

Une *probabilité* est une application  $P : \mathcal{P}(S) \rightarrow [0, 1]$  telle que :

(1)  $\underline{P(S) = 1};$

(2) Si les événements  $E_1, E_2, E_3, \dots$  sont disjoints, alors  $P\left(\coprod_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$

Le nombre réel  $P(E)$  est appelé *probabilité* de l'événement  $E$ .

Nous avons vu la semaine passée que

- a)  $P(\emptyset) = 0$ , c'est-à-dire que la probabilité que rien ne se passe est nulle, notre expérience a toujours une issue ;
- b)  $P(\coprod_{i=1}^n E_i) = \sum_{i=1}^n P(E_i)$ , donc l'axiome (2) est vrai aussi si on traite le cas d'une union disjointe *finie* d'événements ;
- c)  $P(E^c) = 1 - P(E)$  ;
- d) si  $F \subset E$  des événement, alors  $P(F) \leq P(E)$ .

**Exemple 1.2.** On lance plusieurs fois de suite une paire de dés équilibrés et on s'intéresse à la somme des chiffres obtenus.

Quelle est la probabilité d'obtenir une somme égale à 5 avant une somme égale à 7 ?

Soit  $E_n$  l'événement "Pendant les  $n - 1$  premières épreuves ni 5 ni 7 ne sont obtenus, puis à la  $n$ -ème on obtient 5". La probabilité cherchée est la probabilité de la réunion disjointe  $\bigcup E_n$ .

La probabilité d'obtenir un total de 5 vaut  $\frac{4}{36} = \frac{1}{9}$ .

En effet, pour obtenir un total de 5,

il faut lancer soit

1+4, 2+3; 3+2; 4+1 soit

4 façons sur les 36 issues équiprobables.

De même, celle d'obtenir un total de 7 vaut  $\frac{6}{36} = \frac{1}{6}$ .

Ainsi, la probabilité d'obtenir un total de 5 ou de 7 est

Ainsi, puisque les lancers sont indépendants,

$$P(E_n) = \left(\frac{13}{18}\right)^{n-1} \cdot \frac{1}{9}$$

	dé 1					
+	1	2	3	4	5	6
1	2	3	4	5	5	6
2	3	4	5	5	7	
3	4	5	5	7		
4	5		7			
5	6	7				
6	7					

$\frac{1}{9} + \frac{1}{6} = \frac{10}{36}$  et la probabilité d'un total autre que 5 ou 7 vaut  $1 - \frac{10}{36} = \frac{26}{36} = \frac{13}{18}$

Par conséquent, on a

$$\text{La probabilité d'obtenir un total de 5 avant un total de 7 vaut } P\left(\bigcup_{n=1}^{\infty} E_n\right) \stackrel{\text{A2}}{=} \sum_{n=1}^{\infty} P(E_n) = \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} \cdot \frac{1}{9}$$

$$= \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} = \frac{1}{9} \cdot 1 \frac{1}{1 - \frac{13}{18}} = \frac{1}{9} \cdot \frac{1}{\frac{5}{18}} = \frac{2}{5}$$

Somme d'une P.G.

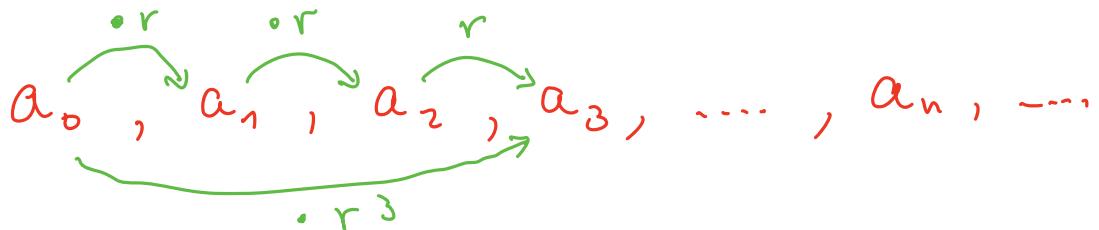
Nous avons terminé le cours de la semaine passée sur le résultat suivant :

**Théorème 1.3.** Soient  $E$  et  $F$  deux événements. Alors

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

# Rappel sur les progressions géométriques

$$a_0, \quad a_1 = a_0 \cdot r, \quad a_2 = a_1 \cdot r$$



$$a_n = a_0 \cdot r^n$$

$$\begin{aligned} S_n &= \sum_{k=0}^n a_k = a_0 + a_1 + \dots + a_n \\ &= a_0 + a_0 \cdot r + a_0 \cdot r^2 + \dots + a_0 \cdot r^n \\ &= a_0 (1 + r + r^2 + \dots + r^n) \end{aligned}$$

$$r \cdot S_n = a_0 (r + r^2 + \dots + r^n + r^{n+1})$$

$$\Rightarrow S_n - r S_n = a_0 (1 - r^{n+1})$$

$$\Leftrightarrow S_n (1 - r) = a_0 (1 - r^{n+1})$$

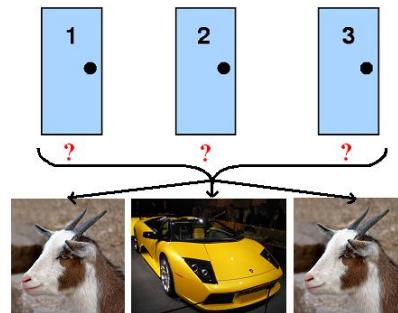
$$\Leftrightarrow \begin{matrix} r \neq 1 \\ S_n = a_0 \frac{1 - r^{n+1}}{1 - r} \end{matrix}$$

$$|r| < 1 \Rightarrow \lim_{n \rightarrow \infty} S_n = a_0 \frac{1}{1 - r} \quad \text{car } \lim_{n \rightarrow \infty} r^{n+1} = 0$$

## 2 Probabilité conditionnelle

Commençons par un exemple élémentaire, que nous reprendrons par la suite pour expliquer le "paradoxe" du problème de Monty Hall.

**Exemple 2.1.** Supposez que vous êtes sur le plateau d'un jeu télévisé face à trois portes et que vous devez choisir d'en ouvrir une seule, en sachant que derrière l'une d'elles se trouve une voiture et derrière les deux autres des chèvres.



La probabilité de choisir la bonne porte vaut donc  $\frac{1}{3}$ .

Supposons maintenant que l'on sache, *avant de choisir*, qu'une chèvre se cache derrière la troisième porte. Dans ce cas, la probabilité que la voiture soit derrière la première porte vaut  $\frac{1}{2}$ :

*Il reste 2 portes et la voiture est derrière l'une d'elles*

On peut aussi raisonner comme suit : la probabilité que la voiture soit derrière la première porte et que simultanément une chèvre se trouve derrière la troisième porte vaut  $\frac{1}{3}$  car si la voiture est derrière la première porte alors forcément, il y a une chèvre derrière la troisième porte.

Par ailleurs, la probabilité qu'une chèvre se trouve derrière la troisième porte vaut  $\frac{2}{3}$  si bien que si l'on sait qu'une chèvre se trouve derrière la troisième porte, l'ensemble fondamental est réduit au  $\frac{2}{3}$  de son état initial.

Ainsi, la probabilité que la voiture soit derrière la première porte sachant qu'une chèvre est derrière la troisième porte vaut

$$\frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

**Définition 2.2.** Soient  $E$  et  $F$  deux événements. On suppose que  $P(F) > 0$ . Alors, la *probabilité conditionnelle*  $P(E|F)$  de  $E$  sachant que  $F$  est réalisé est

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Voici un exemple simple où l'on pourrait aussi raisonner directement.

**Exemple 2.3.** Une urne se trouve dans une pièce obscure et contient 10 billes rouges, 5 billes jaunes et 10 billes blanches lumineuses. On tire une boule et constate qu'elle n'est pas lumineuse. Quelle est la probabilité qu'elle soit jaune ?

Intuitivement : *la bille n'est pas lumineuse  $\Rightarrow$  elle est rouge ou jaune.*

*il y a 5 billes jaunes sur les 15 jaunes ou rouges*

$$\Rightarrow P(Y | B^c) = \frac{5}{15} = \frac{1}{3}$$

Mathématiquement, avec la définition 2.2 :

Soit  $J$  l'événement produit par le tirage d'une bille jaune et  $B$  celui produit par le tirage d'une bille *blanche brillante*.

$$P(Y | B^c) \stackrel{\text{def. 2.2}}{=} \frac{P(Y \cap B^c)}{P(B^c)} = \frac{P(Y)}{P(B^c)} = \frac{\frac{5}{25}}{\frac{15}{25}} = \frac{5}{15} = \frac{1}{3} .$$

Enfin, revenons au problème de Monty Hall.

Le candidat est placé devant trois portes fermées. Derrière l'une d'elles se trouve une voiture et derrière chacune des deux autres se trouve une chèvre. Il doit tout d'abord désigner une porte. Puis le présentateur, qui sait quelle est la bonne porte dès le début, doit ouvrir une porte qui n'est ni celle choisie par le candidat, ni celle cachant la voiture. Le candidat a alors le droit soit d'ouvrir la porte qu'il a choisie initialement, soit d'ouvrir la troisième porte.

Quelle stratégie doit-il adopter pour maximiser la probabilité de gagner la voiture ?

**Exemple 2.4.** Supposons que l'on choisisse la première porte.

Appelons  $A_i$  l'événement "la voiture se trouve derrière la  $i$ -ème porte", pour  $1 \leq i \leq 3$ .

Appelons  $X_j$  l'événement "le présentateur ouvre la  $j$ -ème porte" pour  $2 \leq j \leq 3$ .

Calculons toutes les probabilités  $P(A_i \cap X_j)$ .

D'abord, remarquons que pour tout  $i$ ,  $P(A_i) = \frac{1}{3}$

$$\bullet P(A_1 \cap X_2) = P(A_1 \cap X_3) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

*car si la voiture est derrière la porte 1, qui celle choisie par le candidat, le présentateur peut ouvrir chacune des deux portes restantes.*

$$\bullet P(A_2 \cap X_3) = P(A_3 \cap X_3) = 0$$

*le présentateur ne peut pas ouvrir la porte gagnante.*

$$\bullet P(A_2 \cap X_3) = P(A_3 \cap X_2) = \frac{1}{3} \cdot 1$$

Car si la voiture se trouve derrière la porte 2, l'animateur DOIT ouvrir la porte 3 et vice versa.

Ceci nous permet aussi de calculer la probabilité que le présentateur ouvre la troisième porte.

$$X_3 = (X_3 \cap A_1) \sqcup (X_3 \cap A_2) \sqcup (X_3 \cap A_3)$$

$$\Rightarrow P(X_3) = \frac{1}{6} + \frac{1}{3} + 0 = \frac{1}{2}$$

$$\text{Ainsi, } P(A_1 | X_3) \stackrel{\text{def 2.7.}}{=} \frac{P(A_1 \cap X_3)}{P(X_3)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

$\frac{1}{3}$  est la probabilité de gagner si on garde la porte initialement  
 $\frac{2}{3}$  est la probabilité de gagner si on change de porte. choisie

Il faut changer de porte pour avoir le plus de chance de gagner la voiture!

### 3 La formule des probabilités totales

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

Dans la section précédente, nous avons vu comment calculer une probabilité conditionnelle.

Parfois il peut être utile de retourner cette formule et de l'appliquer pour calculer la probabilité d'un événement. Nous commençons par la formule de multiplication.

#### Proposition 3.1. Formule de multiplication

Soient  $A$  et  $B$  deux événements. Alors  $P(A \cap B) = P(A) \cdot P(B|A)$ .

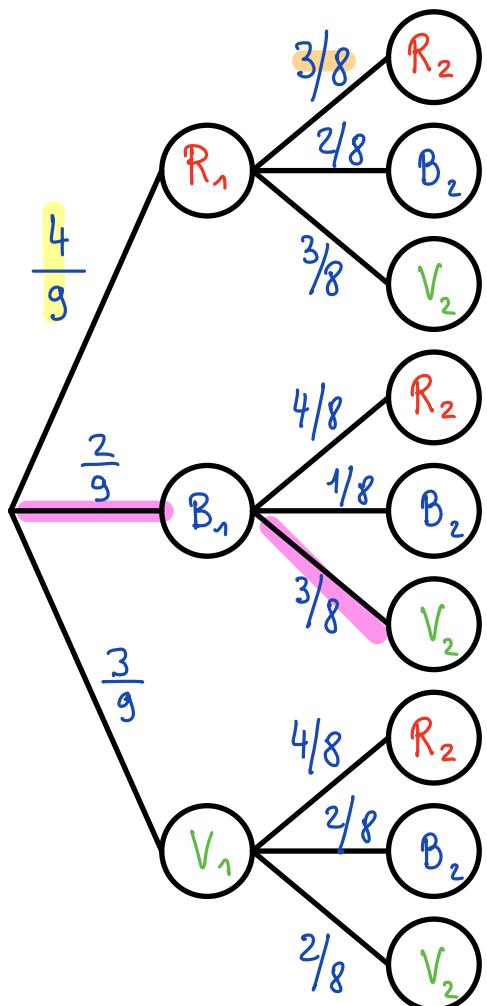
**Exemple 3.2.** Un sac contient quatre billes rouges, deux billes bleues et trois billes vertes.

On cherche la probabilité des événements suivants :  $A$  : "les deux billes tirées sont rouges",  $B$  : "la première est bleue et la seconde est verte" et  $C$  : "l'une des billes est rouge et l'autre bleue".

On désigne par  $R_1$  l'événement "la première bille tirée est rouge", par  $R_2$  "la seconde est rouge", par  $B_1$  "la première est bleue", etc.

On dessine un diagramme en arbre où la *racine* indique le début de l'expérience, le premier niveau de branches indique le premier tirage, le second niveau le deuxième tirage.

On écrit la probabilité de chaque tirage sur la branche et la couleur tirée aux noeuds.



$$\begin{aligned}
 P(A) &= P(R_1 \cap R_2) \\
 &= P(R_1) \cdot P(R_2 | R_1) \\
 &= \frac{4}{9} \cdot \frac{3}{8} = \frac{1}{6}
 \end{aligned}$$

$$\begin{aligned}
 P(B) &= P(B_1 \cap V_2) \\
 &= \frac{2}{9} \cdot \frac{3}{8} = \frac{1}{12}
 \end{aligned}$$

$$\begin{aligned}
 P(C) &= P(R_1 \cap B_2) + P(B_1 \cap R_2) \\
 &= \frac{4}{9} \cdot \frac{2}{8} + \frac{2}{9} \cdot \frac{4}{8} \\
 &= \frac{2}{9}
 \end{aligned}$$

Le théorème suivant est une simple adaptation de cette formule.

### Théorème 3.3. des probabilités totales

Soit  $E_1, \dots, E_n$  des événements tels que  $S = E_1 \sqcup E_2 \sqcup \dots \sqcup E_n$ . Alors pour tout événement  $A$ ,

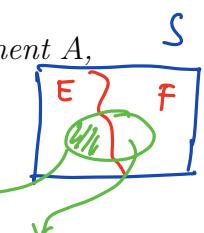
$$P(A) = \sum_{i=1}^n P(A|E_i)P(E_i).$$

Démonstration. Montrons simplement le cas de deux événements :  $S = E \sqcup F$ .

$$\begin{aligned}
 A &= A \cap S = A \cap (E \sqcup F) = (A \cap E) \perp\!\!\!\perp (A \cap F) \\
 \Rightarrow P(A) &\stackrel{(A2)}{=} P(A \cap E) + P(A \cap F)
 \end{aligned}$$

$$\stackrel{\text{Prop 3.1}}{=} P(A|E) \cdot P(E) + P(A|F) \cdot P(F)$$

□



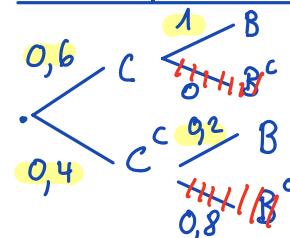
**Exemple 3.4.** Un inspecteur de police est convaincu à 60% que le suspect principal de son enquête sur le vol d'un tableau de Picasso est coupable. A ce stade de l'enquête on trouve un cheveu blond et court sur la scène du crime. Il se trouve que le suspect est blond ! Quelle est la probabilité qu'il ait volé le tableau sachant que 20% de la population a des cheveux blonds ?

Soit  $C$  l'événement "le suspect est coupable" et  $B$  "le coupable est blond".

Alors, avant d'avoir trouvé le cheveu blond, par le théorème des probabilités totales, on a

$$\begin{aligned} P(B) &= P(B|C)P(C) + P(B|C^c)P(C^c) = \\ &= 1 \cdot 0,6 + 0,2 \cdot 0,4 \\ &= 0,6 + 0,08 = 0,68 \end{aligned}$$

Arbre pour aider.



D'autre part, la probabilité conditionnelle que le suspect est coupable sachant que le coupable est blond se calcule directement avec la définition : *(situation APRÈS la découverte du cheveux blond)*

$$P(C|B) = \frac{P(C \cap B)}{P(B)} = \frac{0,6}{0,68} \approx 88,2\%$$

Dans cet exemple, nous avons utilisé le *Théorème de Bayes*, du nom de son auteur Thomas Bayes (1702–1761), mathématicien et pasteur britannique.



### Théorème 3.5. Théorème de Bayes

Soit  $E_1, \dots, E_n$  des événements tels que  $S = E_1 \sqcup E_2 \sqcup \dots \sqcup E_n$ .

Alors pour tout événement  $A$  de probabilité non nulle et tout  $k$ , on a

$$\frac{P(E_k \cap A)}{P(A)} = P(E_k|A) = \frac{\underbrace{P(A|E_k)P(E_k)}_{\text{def 2.2}}}{\underbrace{\sum_{i=1}^n P(A|E_i)P(E_i)}_{P(A) (\text{thm 3.3})}} = P(E_k \cap A) \quad (\text{Thm 3.1})$$

Pour terminer ce cours, étudions un exemple de nature théorique.

### Exemple 3.6. Problème de rencontre de Mont-mort (1708)

Lors d'une réunion de  $n$  hommes, chacun enlève son chapeau et le lance dans le vestibule. À la fin de la réunion, chacun prend un chapeau au hasard dans le tas. On dit qu'il y a *rencontre* si quelqu'un tire son propre chapeau. Quelle est la probabilité qu'il n'y ait aucune rencontre ?

Soit  $E_n$  l'événement "il n'y a aucune rencontre". L'idée est de conditionner  $P(E_n)$  selon l'événement  $R$  : "le premier homme tire son propre chapeau".

$$R \perp\!\!\!\perp R^c = S$$

Par la formule des probabilités totales, on a

$$(*) P(E_n) = \underbrace{P(E_n | R)}_{=0} + P(E_n | R^c) \cdot \underbrace{P(R^c)}_{\frac{n-1}{n}}$$

Analysons maintenant  $P(E_n | R^c)$ . C'est la probabilité qu'il n'y ait pas de rencontre lorsque  $n - 1$  hommes tirent un chapeau dans un tas de  $n - 1$  chapeaux, tas dans lequel un chapeau n'appartient à personne, celui du premier homme, et l'un manque, celui que le premier homme a tiré.

Il y a deux cas de figure sans rencontre. Soit l'homme en trop ne tire pas le chapeau en trop, soit il le tire. Si il ne le tire pas, faisons comme si ce chapeau lui appartenait, si bien que cette situation est équivalente à  $E_{n-1}$ . Mais s'il le tire, et il y a une chance sur  $n - 1$  qu'il le fasse, il reste  $n - 2$  hommes qui doivent tirer un chapeau dans un tas constitué de leurs chapeaux :

$$\begin{aligned} P(E_n | R^c) &\stackrel{(*)}{=} P(E_{n-1}) + \frac{1}{n-1} P(E_{n-2}) \\ \Rightarrow P(E_n) &\stackrel{(*)}{=} \frac{n-1}{n} P(E_n | R^c) \stackrel{(*)}{=} \frac{n-1}{n} P(E_{n-1}) + \frac{1}{n} P(E_{n-2}) \end{aligned}$$

Nous avons obtenu une formule de récurrence !

On peut commencer à calculer puisque  $P(E_1) = 0$  et  $P(E_2) = \frac{1}{2}$ .

Il se trouve que la formule donne

$$P(E_n) = \sum_{k=0}^n (-1)^k \cdot \frac{1}{k!} \xrightarrow{n \rightarrow \infty} \frac{1}{e}$$

une série qui tend vers  $e^{-1} \cong 0,368$ .