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 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 $P : \mathcal{P}(S) \to [0,1]$ telle que :

(1) $P(S) = 1$;

(2) Si les événements $E_1, E_2, E_3, \ldots$ 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) $P\left(\coprod_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$, 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 une paire de dés équilibrés. Le résultat d'un lancer est la somme des chiffres apparents. Quelle est la probabilité qu'on obtienne 5 avant 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 $\coprod E_n$. D'autre part, la probabilité d'obtenir un 5 est $\frac{4}{36} = \frac{1}{9}$

En effet, pour obtenir un total de 5, il y a 4 façons :

$1+4$, $2+3$, $3+2$; $4+1$

sur les 36 issues équiprobables.

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

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

Ainsi, la probabilité d'obtenir un 5 ou un 7 est $\frac{10}{36}$. On a alors que $P(E_n)$ vaut

$$\left(1-\frac{10}{36}\right)^{n-1} \cdot \frac{1}{9} = \left(\frac{13}{18}\right)^{n-1} \cdot \frac{1}{9} = P(E_n)$$

car les lancers sont indépendants. Par conséquent, on a

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

$$= \frac{1}{9} \cdot \frac{1}{1-\frac{13}{18}} = \frac{1}{9} \cdot \frac{1}{\frac{5}{18}}$$

$$= \frac{1}{9} \cdot \frac{18}{5} = \frac{2}{5}$$

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 suites géométriques : a_0, $a_1 = a_0 \cdot r$, $a_2 = a_1 \cdot r$ .....

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

$$S_n = \sum_{k=0}^{n} a_k = a_0 \sum_{k=0}^{n} r^n \overset{(*)}{=} \frac{1-r^{n+1}}{1-r} \cdot a_0$$

$$\text{si } |r| < 1, \quad \lim_{n\to\infty} S_n = S_\infty = \frac{a_0}{1-r}$$

(*) démo :

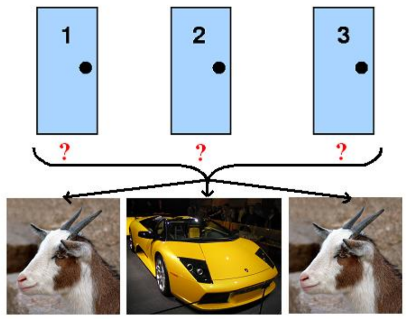
$$\begin{array}{rcl} S_n = & a_0 + a_0 r + \dots + a_0 r^n & \\ - \; r S_n = & a_0 r + \dots + a_0 r^n + a_0 r^{n+1} & \\ \hline \end{array}$$

$$S_n(1-r) = S_n - r S_n = a_0 - a_0 \cdot r^{n+1} = a_0(1-r^{n+1}) \Longrightarrow S_n = \frac{a_0(1-r^{n+1})}{1-r}$$

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.



Quelle est la probabilité de choisir la bonne porte ? $\frac{1}{3}$

Supposons maintenant que l'on sache, *avant de choisir*, qu'une chèvre se cache derrière la troisième porte. Quelle est la probabilité que la voiture soit derrière la première porte ?

Il reste 2 portes et la voiture est derrière l'une d'elle $\Rightarrow \frac{1}{2}$

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 blanche ⇒ on a tiré une des 15 billes restantes dont $\frac{1}{3}$ sont jaune et $\frac{2}{3}$ rouge.

$$\Rightarrow P(J \mid B^c) = \frac{1}{3}$$

Mathématiquement :

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

Déf. 2.2

$$P(J \mid B^c) \overset{\downarrow}{=} \frac{P(J \cap B^c)}{P(B^c)} = \frac{P(J)}{P(B^c)} = \frac{\frac{5}{25}}{\frac{15}{25}} = \frac{5}{15} = \frac{1}{3}$$

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, $P(A_i) = \frac{1}{3}$

- $P(A_1 \cap X_2) = P(A_1 \cap X_3) = \frac{1}{6}$

  car si la voiture est derrière la porte 1 et que le candidat a choisi cette porte, le présentateur peut ouvrir chacune des portes restantes.

- $P(A_2 \cap X_2) = P(A_3 \cap X_3) = 0$

  car le présentateur ne peut pas ouvrir la porte gagnante.

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

car si la voiture se trouve derrière la porte 2, le présentateur DOIT ouvrir la 3 et vice-versa.

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

$$X_3 = (X_3 \cap A_1) \amalg (X_3 \cap A_2) \amalg (X_3 \cap A_3)$$
$$\Rightarrow P(X_3) = \frac{1}{6} + \frac{1}{3} + 0 = \frac{1}{2}$$

Ainsi, $P(A_1 | X_3) \overset{\text{déf. 2.2}}{=} \frac{P(A_1 \cap X_3)}{P(X_3)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$ qui est la probabilité de gagner si on garde la porte choisie.

et $P(A_2 | X_3) = \frac{P(A_2 \cap X_3)}{P(X_3)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$.

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

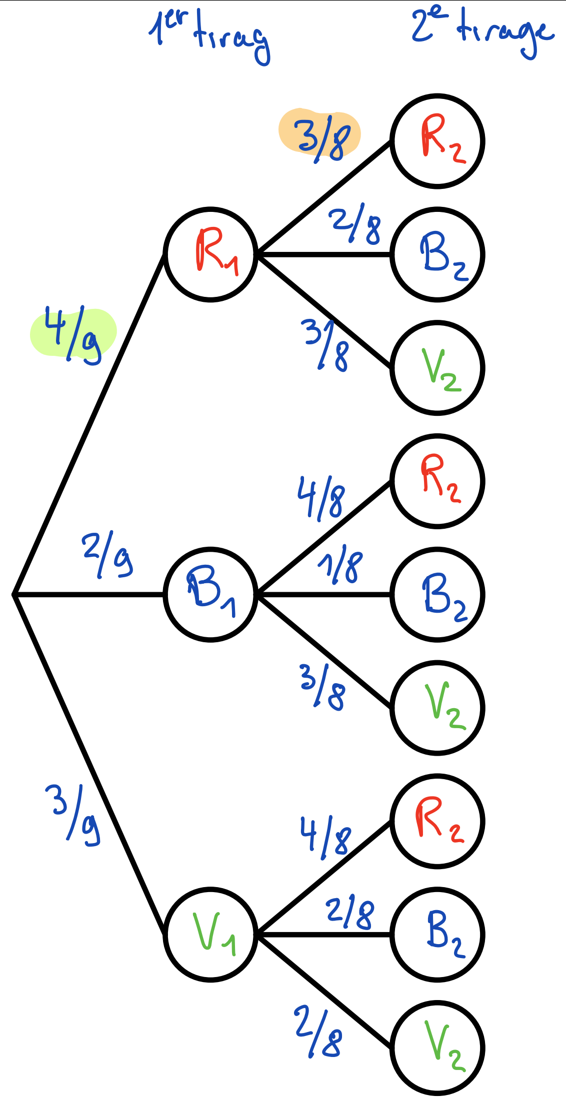
3 La formule des probabilités totales

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 $P(A \cap B) = P(A) \cdot P(B|A)$. $\overset{\text{def 2.2 avec } P(A) \neq 0}{\Longleftrightarrow}$ $P(B|A) = \frac{P(A \cap B)}{P(A)}$

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.



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

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

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

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 $S = E_1 \coprod E_2 \coprod \cdots \coprod E_n$. 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 : $S = E \coprod F$.

$$A = A \cap S = A \cap (E \amalg F) = (A \cap E) \amalg (A \cap F)$$

$$\Longrightarrow P(A) \overset{A.2}{=} P(A \cap E) + P(A \cap F)$$

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

□

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

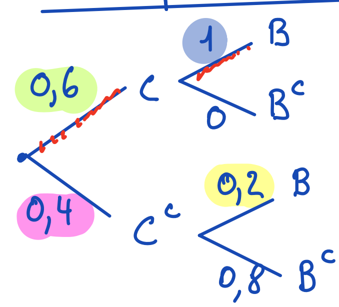
$$P(B) = P(B|C)P(C) + P(B|C^c)P(C^c) = 1 \cdot 0{,}6 + 0{,}2 \cdot 0{,}4 = 0{,}68$$
$$= 1 \cdot \frac{3}{5} + \frac{1}{5} \cdot \frac{2}{5} = \frac{17}{25}$$

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

$$P(C|B) = \frac{P(C \cap B)}{P(B)} = \frac{0{,}6 \cdot 1}{\frac{17}{25}} = \frac{\frac{3}{5}}{\frac{17}{25}} = \frac{15}{17} = 88{,}2\%$$

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

Arbre pour aider :



Théorème 3.5. Théorème de Bayes
Soit $E_1, \ldots, E_n$ des événements tels que $S = E_1 \coprod E_2 \coprod \cdots \coprod E_n$. Alors pour tout événement A de probabilité non nulle et tout k, on a

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

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". $R \amalg R^c = S$
Par la formule des probabilités totales, on a

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

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. S'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 :

$$P(E_n|R^c) \overset{(2)}{=} P(E_{n-1}) + \frac{1}{n-1} P(E_{n-2})$$

$$\overset{(1)}{\Longrightarrow} P(E_n) = \frac{n-1}{n} P(E_n|R^c) \overset{(2)}{=} \frac{n-1}{n} P(E_{n-1}) + \frac{1}{n} P(E_{n-2})$$

Nous avons obtenu une formule de récurrence !
On peut commencer à calculer puisque $P(E_1) = 0$ et $P(E_2) = \dfrac{1}{2}$.
Il se trouve que la formule donne

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

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