

Corrigé détaillé – TD3

Probabilités 1

Probabilités conditionnelles

Année universitaire 2026–2027

D'après un document de Chloé Friguet, Cédric Frambourg et Benjamin Dufée.

Proposition avec rappels progressifs et renvois entre les TD.

Rappel de cours – Notions nouvelles de ce TD

Conditionner. Pour $\mathbb{P}(B) > 0$,

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

L'évènement après « sachant » devient le nouvel univers de référence. Ce produit formalise la règle des chemins utilisée au TD2, exercice 2.

Probabilités totales. Si $B_1, \dots, B_r$ sont disjoints, couvrent Ω et ont des probabilités positives,

$$\mathbb{P}(A) = \sum_{i=1}^r \mathbb{P}(B_i)\mathbb{P}(A | B_i).$$

On additionne les contributions des différentes origines possibles de A.

Inverser le conditionnement : Bayes. Si $\mathbb{P}(A) > 0$,

$$\mathbb{P}(B_j | A) = \frac{\mathbb{P}(B_j)\mathbb{P}(A | B_j)}{\sum_i \mathbb{P}(B_i)\mathbb{P}(A | B_i)}.$$

Le dénominateur est la probabilité totale de l'information observée.

Indépendance. A et B sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$; pour $\mathbb{P}(B) > 0$, cela équivaut à $\mathbb{P}(A | B) = \mathbb{P}(A)$. Cette hypothèse doit être donnée ou explicitement posée.

Prérequis déjà acquis. Complémentaire et somme de chemins : voir TD2, rappel d'ouverture et exercice 2. Le TD3 ne les réintroduit pas.

Ce corrigé est conçu pour permettre un travail autonome. On précise les événements utilisés, on écrit les formules de probabilités avant d'effectuer les calculs et on distingue systématiquement une probabilité conditionnelle de la probabilité obtenue en inversant le conditionnement.

Exercice 1

Exercice 1 :

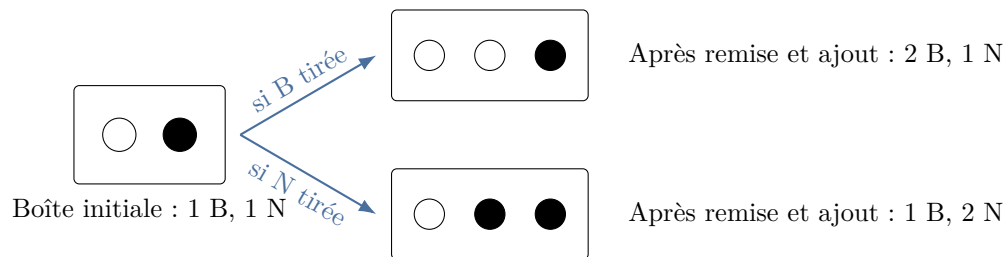
Une boîte contient une balle noire et une balle blanche. Une balle est tirée au hasard dans la boîte : on y remet celle-ci ainsi qu'une nouvelle balle de la même couleur. On tire alors une des trois balles au hasard dans la boîte.

- 1) Décrire l'expérience à l'aide d'un arbre des choix.
- 2) Quelle est la probabilité que la seconde balle tirée soit blanche ?
- 3) Quelle est la probabilité que l'une au moins des deux balles tirées soit blanche ?
- 4) Quelle est la probabilité que la première balle tirée soit blanche, sachant que l'une au moins des deux balles tirées est blanche ?

Analyse : comment choisir la méthode ?

L'expérience comporte deux tirages successifs et le premier modifie la composition de la boîte : un arbre pondéré rend cette dépendance visible. On ne peut donc pas multiplier deux probabilités égales à $1/2$ comme pour des tirages indépendants. Pour « seconde balle blanche », on réunit les deux chemins qui finissent par une blanche. Pour « au moins une blanche », le complément ne comporte qu'un chemin, contre trois pour le calcul direct. Enfin, « sachant » impose de restreindre l'univers aux issues compatibles avec l'information.

Repères de cours. Rappel d'ouverture du TD3 : conditionnement et probabilités totales ; arbre avec composition variable : voir TD2, exercice 5.



B : blanche ; N : noire. Le second tirage se fait dans la boîte modifiée.

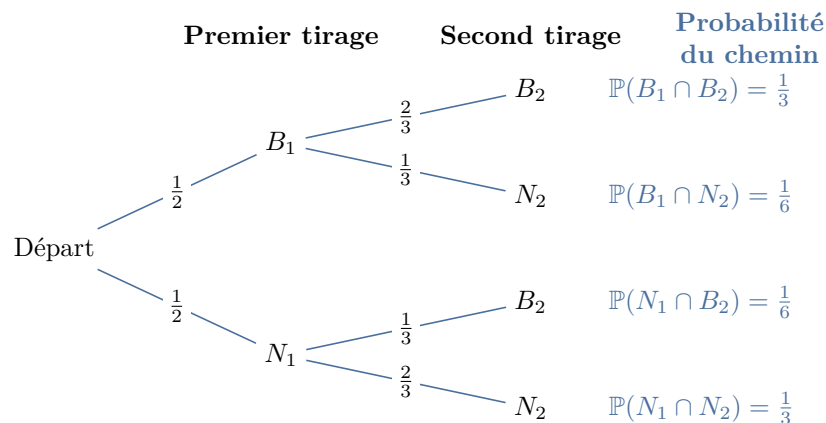
Résolution.

Notons

$$B_i = \{\text{la balle obtenue au tirage } i \text{ est blanche}\}$$

et

$$N_i = \{\text{la balle obtenue au tirage } i \text{ est noire}\}.$$



Au premier tirage,

$$\mathbb{P}(B_1) = \mathbb{P}(N_1) = \frac{1}{2}.$$

Si la première balle est blanche, on remet cette balle puis on ajoute une seconde balle blanche. La boîte contient alors deux balles blanches et une balle noire. Ainsi

$$\mathbb{P}(B_2 | B_1) = \frac{2}{3}, \quad \mathbb{P}(N_2 | B_1) = \frac{1}{3}.$$

Si la première balle est noire, la boîte contient alors une balle blanche et deux balles noires. Donc

$$\mathbb{P}(B_2 | N_1) = \frac{1}{3}, \quad \mathbb{P}(N_2 | N_1) = \frac{2}{3}.$$

Méthode à retenir

Dans un arbre pondéré, on multiplie les probabilités le long d'un chemin. Lorsque plusieurs chemins incompatibles conduisent au même événement, on additionne ensuite les probabilités de ces chemins.

2) Probabilité que la seconde balle soit blanche

Les événements B_1 et N_1 forment une partition de l'univers. Par la formule des probabilités totales,

$$\begin{aligned}\mathbb{P}(B_2) &= \mathbb{P}(B_1)\mathbb{P}(B_2 \mid B_1) + \mathbb{P}(N_1)\mathbb{P}(B_2 \mid N_1) \\ &= \frac{1}{2} \times \frac{2}{3} + \frac{1}{2} \times \frac{1}{3} \\ &= \frac{1}{3} + \frac{1}{6} \\ &= \boxed{\frac{1}{2}}.\end{aligned}$$

3) Au moins une des deux balles est blanche

Il est plus simple de raisonner avec l'événement contraire.

L'événement contraire de "au moins une balle est blanche" est "les deux balles sont noires". Ainsi

$$\mathbb{P}(N_1 \cap N_2) = \mathbb{P}(N_1)\mathbb{P}(N_2 \mid N_1) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}.$$

Par conséquent,

$$\mathbb{P}(\text{au moins une blanche}) = 1 - \frac{1}{3} = \boxed{\frac{2}{3}}.$$

4) Première balle blanche sachant qu'au moins une est blanche

Notons

$$A = \{\text{au moins une des deux balles est blanche}\}.$$

On cherche

$$\mathbb{P}(B_1 \mid A).$$

Par définition,

$$\mathbb{P}(B_1 \mid A) = \frac{\mathbb{P}(B_1 \cap A)}{\mathbb{P}(A)}.$$

Or, si la première balle est blanche, alors A est automatiquement réalisé. Ainsi

$$B_1 \cap A = B_1.$$

Donc

$$\mathbb{P}(B_1 \mid A) = \frac{\mathbb{P}(B_1)}{\mathbb{P}(A)} = \frac{\frac{1}{2}}{\frac{2}{3}} = \boxed{\frac{3}{4}}.$$

Point de vigilance

Après le premier tirage, la composition de la boîte a changé. Les probabilités du second tirage ne sont donc plus toutes égales à $1/2$.

Exercice 2

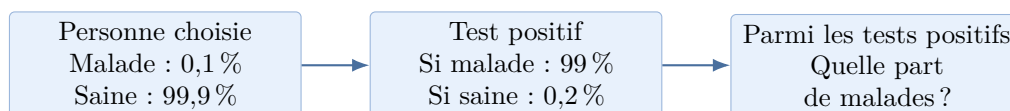
Exercice 2 :

Une maladie est présente dans la population, dans la proportion d'une personne malade sur 1 000. Un responsable d'un grand laboratoire pharmaceutique vante son nouveau test de dépistage : si une personne est malade, le test est positif à 99%. Néanmoins, sur une personne non malade, le test est positif à 0,2%. Ces chiffres ont l'air excellents, on ne peut qu'en convenir. Toutefois, ce qui est réellement intéressant, plus que les résultats présentés par le laboratoire, c'est la probabilité qu'une personne soit effectivement malade lorsque son test est positif. Calculer cette probabilité.

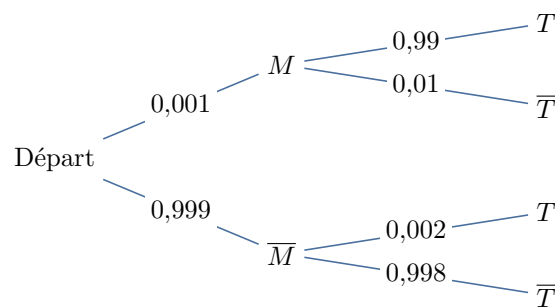
Analyse : comment choisir la méthode ?

Les données vont de l'état de santé vers le résultat du test, tandis que la question remonte du résultat vers l'état de santé : c'est une inversion du conditionnement, traitée par Bayes. Il faut d'abord réunir vrais et faux positifs. Un tableau sur une population fictive rend le rôle de la rareté de la maladie visible ; reprendre directement le taux de 99 % répondrait à une autre question.

Repères de cours. Voir TD3, exercice 1 pour calculer une probabilité totale ; formule de Bayes : rappel d'ouverture du TD3.



Résolution.



M : malade ; T : test positif.

Notons

$$M = \{\text{la personne est malade}\}, \quad T = \{\text{le test est positif}\}.$$

Les données s'écrivent

$$\begin{aligned} \mathbb{P}(M) &= 0,001, & \mathbb{P}(\overline{M}) &= 0,999, \\ \mathbb{P}(T \mid M) &= 0,99, & \mathbb{P}(T \mid \overline{M}) &= 0,002. \end{aligned}$$

Méthode à retenir

On connaît ici la probabilité d'un test positif sachant que la personne est malade, mais on cherche la probabilité inverse : être malade sachant que le test est positif. Il faut donc calculer d'abord $\mathbb{P}(T)$, puis appliquer la formule de Bayes.

Lecture sur une population fictive de 1 000 000 de personnes. Ces effectifs illustrent exactement les proportions du modèle ; ce ne sont pas des observations.

	T	$\bar{T}$	Total
M	990	10	1 000
$\bar{M}$	1 998	997 002	999 000
Total	2 988	997 012	1 000 000

Parmi les 2 988 positifs, 990 sont malades : le quotient $990/2\,988$ retrouve Bayes.

Par la formule des probabilités totales,

$$\begin{aligned}
 \mathbb{P}(T) &= \mathbb{P}(M)\mathbb{P}(T \mid M) + \mathbb{P}(\bar{M})\mathbb{P}(T \mid \bar{M}) \\
 &= 0,001 \times 0,99 + 0,999 \times 0,002 \\
 &= 0,00099 + 0,001998 \\
 &= 0,002988.
 \end{aligned}$$

D'autre part,

$$\mathbb{P}(M \cap T) = \mathbb{P}(M)\mathbb{P}(T \mid M) = 0,001 \times 0,99 = 0,00099.$$

Ainsi

$$\begin{aligned}
 \mathbb{P}(M \mid T) &= \frac{\mathbb{P}(M \cap T)}{\mathbb{P}(T)} \\
 &= \frac{0,00099}{0,002988} \\
 &= \frac{55}{166} \\
 &\approx 0,3313.
 \end{aligned}$$

Donc

$$\boxed{\mathbb{P}(M \mid T) \approx 33,1\%}.$$

Remarque

Le test est pourtant très performant. Le résultat s'explique par le fait que la maladie est extrêmement rare : dans une population très majoritairement saine, les faux positifs deviennent nombreux relativement aux vrais positifs.

Point de vigilance

Ne pas confondre

$$\mathbb{P}(T \mid M) = 99\%$$

avec

$$\mathbb{P}(M \mid T).$$

Inverser les deux événements change en général complètement la probabilité.

Exercice 3

Exercice 3 :

Vous souhaitez aller à La Rochelle pour vos vacances. La SNCF vous propose le trajet suivant : Vannes → Paris en TGV, puis Paris → La Rochelle en TGV.

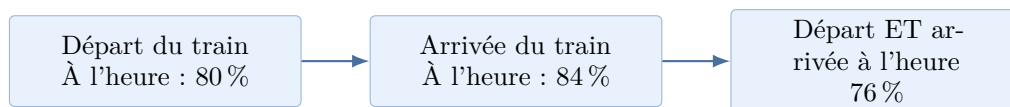
Des statistiques empiriques (inventées dans le cadre de cet exercice) indiquent, pour les TGV que 80% des départs se font à l'heure, 84% des arrivées se font à l'heure, et 76% des trajets respectent à la fois les horaires de départ et d'arrivée.

- 1) Dans chacun des cas suivants, écrire la probabilité cherchée en fonction d'événements que l'on définira, puis la calculer :
 - a) probabilité qu'un train arrive à l'heure sachant qu'il est parti à l'heure.
 - b) probabilité qu'un train soit parti à l'heure sachant qu'il est arrivé à l'heure.
 - c) probabilité qu'un train n'arrive pas à l'heure sachant qu'il n'est pas parti à l'heure.
 - d) probabilité qu'un train n'arrive pas à l'heure sachant qu'il est parti à l'heure.
- 2) La SNCF déclare autour de 1% de retard critique ($>1h00$) sur les TGV. Un tel retard sur le train Vannes → Paris fait manquer la correspondance à Paris pour aller à La Rochelle ensuite. En considérant l'ensemble du trajet, calculer la probabilité d'arriver à l'heure à La Rochelle.

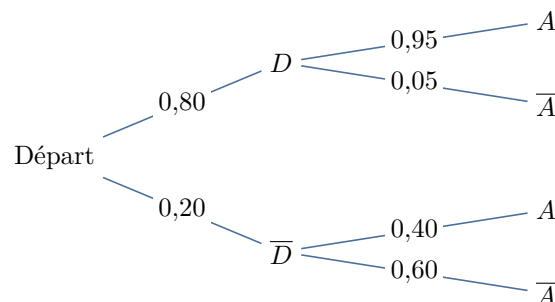
Analyse : comment choisir la méthode ?

On connaît deux probabilités marginales et une intersection, mais aucune probabilité conditionnelle : un tableau croisé se complète directement par soustractions. Un arbre demanderait de calculer au préalable ses branches. Chaque « sachant » désignera ensuite la ligne ou la colonne de référence. Pour la correspondance, on distingue un retard quelconque d'un retard critique ; le produit final exige une hypothèse supplémentaire sur le second trajet.

Repères de cours. Voir TD3, exercice 1 pour le dénominateur d'une probabilité conditionnelle ; indépendance : rappel d'ouverture du TD3.



Résolution.



D : départ à l'heure ; A : arrivée à l'heure. Les branches du second niveau sont calculées à partir des intersections données.

Notons

$$D = \{\text{le train part à l'heure}\}, \quad A = \{\text{le train arrive à l'heure}\}.$$

Les données sont

$$\mathbb{P}(D) = 0,80, \quad \mathbb{P}(A) = 0,84, \quad \mathbb{P}(D \cap A) = 0,76.$$

Méthode à retenir

Avec deux événements et plusieurs questions conditionnelles, un tableau croisé est souvent le moyen le plus sûr d'organiser les données.

On complète d'abord les probabilités manquantes :

$$\mathbb{P}(D \cap \bar{A}) = 0,80 - 0,76 = 0,04,$$

$$\mathbb{P}(\bar{D} \cap A) = 0,84 - 0,76 = 0,08,$$

puis

$$\mathbb{P}(\bar{D} \cap \bar{A}) = 1 - (0,76 + 0,04 + 0,08) = 0,12.$$

On obtient ainsi

	A	$\bar{A}$	Total
D	0,76	0,04	0,80
$\bar{D}$	0,08	0,12	0,20
Total	0,84	0,16	1

1a) Arriver à l'heure sachant que le train est parti à l'heure

$$\begin{aligned}\mathbb{P}(A \mid D) &= \frac{\mathbb{P}(A \cap D)}{\mathbb{P}(D)} \\ &= \frac{0,76}{0,80} \\ &= \boxed{0,95}.\end{aligned}$$

Donc

$$\boxed{\mathbb{P}(A \mid D) = 95\%}.$$

1b) Être parti à l'heure sachant que le train est arrivé à l'heure

$$\begin{aligned}\mathbb{P}(D \mid A) &= \frac{\mathbb{P}(D \cap A)}{\mathbb{P}(A)} \\ &= \frac{0,76}{0,84} \\ &= \frac{19}{21} \\ &\approx 0,9048.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(D \mid A) \approx 90,5\%}.$$

1c) Ne pas arriver à l'heure sachant que le train n'est pas parti à l'heure

$$\begin{aligned}\mathbb{P}(\bar{A} \mid \bar{D}) &= \frac{\mathbb{P}(\bar{A} \cap \bar{D})}{\mathbb{P}(\bar{D})} \\ &= \frac{0,12}{0,20} \\ &= \boxed{0,60}.\end{aligned}$$

Donc

$$\boxed{\mathbb{P}(\bar{A} \mid \bar{D}) = 60\%}.$$

1d) Ne pas arriver à l'heure sachant que le train est parti à l'heure

$$\begin{aligned}\mathbb{P}(\bar{A} \mid D) &= \frac{\mathbb{P}(\bar{A} \cap D)}{\mathbb{P}(D)} \\ &= \frac{0,04}{0,80} \\ &= \boxed{0,05}.\end{aligned}$$

Donc

$$\boxed{\mathbb{P}(\bar{A} \mid D) = 5\%}.$$

2) Trajet complet Vannes–Paris–La Rochelle

Notons

$$C = \{\text{le premier train subit un retard critique}\}.$$

L'énoncé donne

$$\mathbb{P}(C) = 0,01, \quad \mathbb{P}(\bar{C}) = 0,99.$$

Pour obtenir une valeur numérique à partir des seules données de l'énoncé, on modélise la ponctualité du second train comme indépendante de l'existence d'un retard critique sur le premier train.

Pour arriver à l'heure à La Rochelle, il faut alors :

- ne pas subir de retard critique sur Vannes–Paris ;
- arriver à l'heure avec le TGV Paris–La Rochelle.

Donc

$$\mathbb{P}(\text{arriver à l'heure}) = 0,99 \times 0,84 = 0,8316.$$

Ainsi

$$\boxed{\mathbb{P}(\text{arriver à l'heure à La Rochelle}) = 83,16\%}.$$

Remarque

L'hypothèse d'indépendance utilisée dans la dernière question est une hypothèse de modélisation. Elle n'est pas une conséquence des données 80%, 84% et 76%.

Point de vigilance

Dans une probabilité conditionnelle, le dénominateur est toujours la probabilité de l'événement placé après le mot “sachant”.

Exercice 4

Exercice 4 :

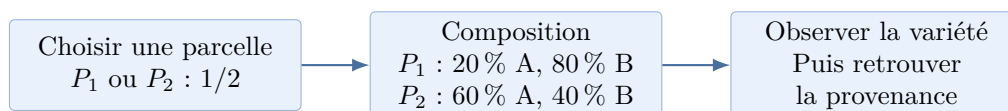
Un agriculteur plante des pommes de terre de deux variétés A et B sur deux parcelles, il plante 20% de la variété A et 80% de la variété B . Sur l'autre parcelle il plante 60% de la variété A et 40% de la variété B . Quelques mois plus tard, l'agriculteur a oublié quelle parcelle correspond à la première ou à la seconde. Il choisit au hasard (équiprobabilité) une parcelle, et prélève une pomme de terre.

- 1) Définir clairement les événements (préciser les notations utilisées), puis reporter les informations de l'énoncé à l'aide des notations choisies.
- 2) Calculer la probabilité que la pomme de terre soit de la variété B .
- 3) La pomme de terre prélevée est de la variété B . Quelle est la probabilité qu'elle soit issue de la première parcelle (parcelle contenant 20% de A et 80% de B) ?

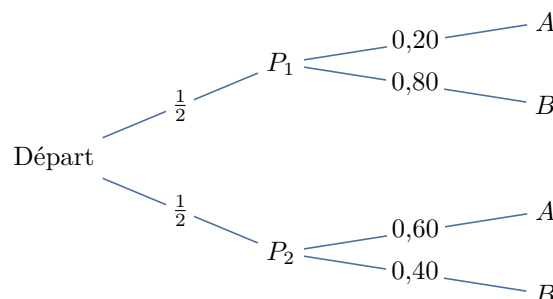
Analyse : comment choisir la méthode ?

Le choix se fait en deux étapes : parcelle, puis variété. Un arbre suit cet ordre et les probabilités totales réunissent les deux façons d'obtenir B . La moyenne simple des proportions fonctionne ici parce que les parcelles sont équiprobables ; elle ne serait pas valable avec des poids différents. La dernière question remonte de B vers la parcelle : on utilise Bayes.

Repères de cours. Voir TD3, exercice 2 pour remonter d'une observation à son origine par Bayes.



Résolution.



Premier niveau : parcelle choisie ; second niveau : variété obtenue.

Notons

$$\begin{aligned}
 P_1 &= \{\text{la première parcelle est choisie}\}, \\
 P_2 &= \{\text{la seconde parcelle est choisie}\}, \\
 A &= \{\text{la pomme de terre est de variété } A\}, \\
 B &= \{\text{la pomme de terre est de variété } B\}.
 \end{aligned}$$

La parcelle est choisie au hasard, donc

$$\mathbb{P}(P_1) = \mathbb{P}(P_2) = \frac{1}{2}.$$

Les données de composition donnent

$$\mathbb{P}(B \mid P_1) = 0,80, \quad \mathbb{P}(A \mid P_1) = 0,20,$$

et

$$\mathbb{P}(B \mid P_2) = 0,40, \quad \mathbb{P}(A \mid P_2) = 0,60.$$

Tableau des probabilités conjointes. Chaque case est le produit du poids de la parcelle par la proportion de la variété.

	<i>A</i>	<i>B</i>	Total
P_1	0,10	0,40	0,50
P_2	0,30	0,20	0,50
Total	0,40	0,60	1

Pour la dernière question, on se restreint à la colonne B : 0,40/0,60.

2) Probabilité d'obtenir la variété *B*

Par la formule des probabilités totales,

$$\begin{aligned}
 \mathbb{P}(B) &= \mathbb{P}(P_1)\mathbb{P}(B \mid P_1) + \mathbb{P}(P_2)\mathbb{P}(B \mid P_2) \\
 &= \frac{1}{2} \times 0,80 + \frac{1}{2} \times 0,40 \\
 &= 0,40 + 0,20 \\
 &= \boxed{0,60}.
 \end{aligned}$$

Ainsi

$$\mathbb{P}(B) = 60\%.$$

3) Provenance sachant que la pomme de terre est de variété *B*

On cherche

$$\mathbb{P}(P_1 \mid B).$$

Or

$$\mathbb{P}(P_1 \cap B) = \mathbb{P}(P_1)\mathbb{P}(B \mid P_1) = \frac{1}{2} \times 0,80 = 0,40.$$

Donc

$$\begin{aligned}
 \mathbb{P}(P_1 \mid B) &= \frac{\mathbb{P}(P_1 \cap B)}{\mathbb{P}(B)} \\
 &= \frac{0,40}{0,60} \\
 &= \boxed{\frac{2}{3}}.
 \end{aligned}$$

Méthode à retenir

Pour remonter d'une observation finale vers sa cause possible, on calcule d'abord la probabilité totale de l'observation, puis on utilise la formule de Bayes.

Point de vigilance

Avant le prélèvement, les deux parcelles sont équiprobables. Après avoir observé une pomme de terre de variété *B*, elles ne le sont plus.

Exercice 5

Exercice 5 :

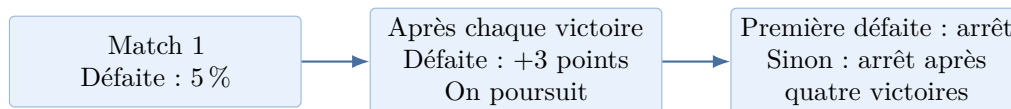
Pour progresser physiquement, un joueur de tennis entame un marathon dans lequel il va jouer 4 matchs dans sa journée jusqu'à ce qu'il perde. Ses adversaires étant très largement à sa portée, il est estimé que la probabilité qu'il perde un match est de 5%. Cependant, la fatigue avançant, à chaque fois qu'il remporte un match, cette probabilité de défaite augmente de 3 points (par exemple de 5 à 8% entre le match 1 et le match 2, etc.).

Calculer la probabilité qu'il sorte vaincu de cette journée.

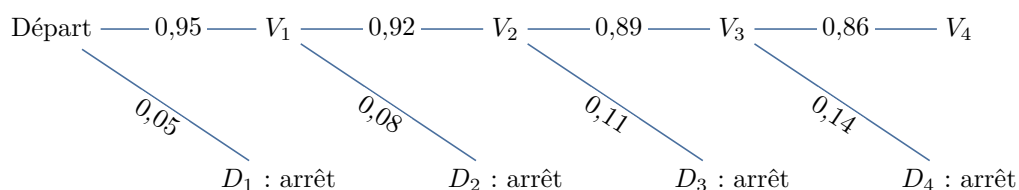
Analyse : comment choisir la méthode ?

Le joueur s'arrête à la première défaite : on représente seulement les matchs effectivement atteints. Les taux annoncés sont conditionnels aux victoires précédentes. Le complément de « sortir vaincu » est l'unique parcours de quatre victoires ; il est plus court à calculer que les quatre parcours de défaite. Une puissance ne convient pas puisque les probabilités changent à chaque match.

Repères de cours. Voir TD2, exercice 6 pour le complémentaire ; la règle du produit ne suppose pas des probabilités constantes.



Résolution.



V_i : victoire au match i ; D_i : défaite au match i . On ne poursuit qu'après une victoire et on s'arrête après le quatrième match.

Les probabilités successives de défaite, conditionnellement au fait d'avoir gagné tous les matchs précédents, sont :

0,05, 0,08, 0,11, 0,14.

Les probabilités successives de victoire sont donc

0,95, 0,92, 0,89, 0,86.

Méthode à retenir

L'événement contraire de « sortir vaincu » est particulièrement simple : le joueur doit gagner les quatre matchs. Calculer le contraire évite d'additionner quatre chemins différents de défaite.

Les cinq parcours terminaux. V signifie victoire et D défaite. Aucun match n'est joué après D.

Parcours	Produit des probabilités conditionnelles	Probabilité
<i>D</i>	0,05	0,05
<i>VD</i>	$0,95 \times 0,08$	0,076
<i>VVD</i>	$0,95 \times 0,92 \times 0,11$	0,09614
<i>VVVD</i>	$0,95 \times 0,92 \times 0,89 \times 0,14$	0,1089004
<i>VVVV</i>	$0,95 \times 0,92 \times 0,89 \times 0,86$	0,6689596

Les quatre premières lignes décrivent « sortir vaincu ». Leur somme vaut 0,3310404 ; la somme des cinq lignes vaut 1.

La probabilité de gagner les quatre matchs est

$$0,95 \times 0,92 \times 0,89 \times 0,86.$$

On obtient

$$\mathbb{P}(\text{quatre victoires}) = 0,6689596.$$

Par conséquent,

$$\begin{aligned}\mathbb{P}(\text{sortir vaincu}) &= 1 - 0,6689596 \\ &= 0,3310404.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(\text{sortir vaincu}) \approx 33,10\%}.$$

On peut vérifier ce résultat en additionnant les quatre chemins possibles :

$$\begin{aligned}\mathbb{P}(\text{défaite}) &= 0,05 \\ &\quad + 0,95 \times 0,08 \\ &\quad + 0,95 \times 0,92 \times 0,11 \\ &\quad + 0,95 \times 0,92 \times 0,89 \times 0,14.\end{aligned}$$

Point de vigilance

Il ne faut pas calculer

$$1 - 0,95^4.$$

La probabilité de gagner un match n'est pas constante : elle diminue au fur et à mesure que la fatigue augmente.

Exercice 6

Exercice 6 :

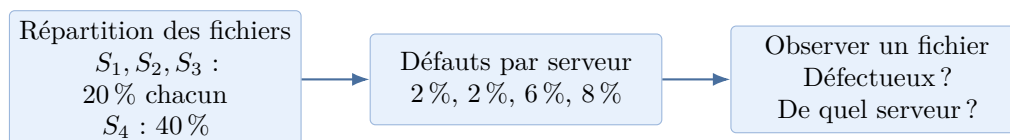
Un logiciel Peer-to-Peer utilise 4 serveurs S_1 , S_2 , S_3 et S_4 de listes de fichiers partagés. S_4 est le plus gros des serveurs et recense 40% des données disponibles. Les données restantes sont distribuées équitablement entre les 3 autres serveurs. Sur la masse des fichiers disponibles, un certain nombre d'entre eux sont défectueux : soit parce que leur contenu n'est pas conforme à la description qui en est donnée, soit parce qu'ils contiennent des virus. Les pourcentages de fichiers défectueux sont : 8% pour S_4 , 6% pour S_3 , 2% pour S_2 et 2% pour S_1 .

- 1) Définir clairement les événements. Préciser les notations utilisées, représenter l'arbre des choix ; et reporter les informations de l'énoncé à l'aide de vos notations.
- 2) Pour les questions suivantes, un fichier est téléchargé au hasard. Pour chaque calcul demandé, écrire la probabilité cherchée et les principales étapes du calcul.
 - a) Quelle est la probabilité que ce fichier soit défectueux ?
 - b) Quelle est la probabilité qu'il provienne du serveur S_4 et qu'il soit défectueux ?
 - c) Le fichier provient du serveur S_4 . Quelle est la probabilité qu'il soit défectueux ?
 - d) Sachant que le fichier est défectueux, quelle est la probabilité qu'il provienne du serveur S_4 ?

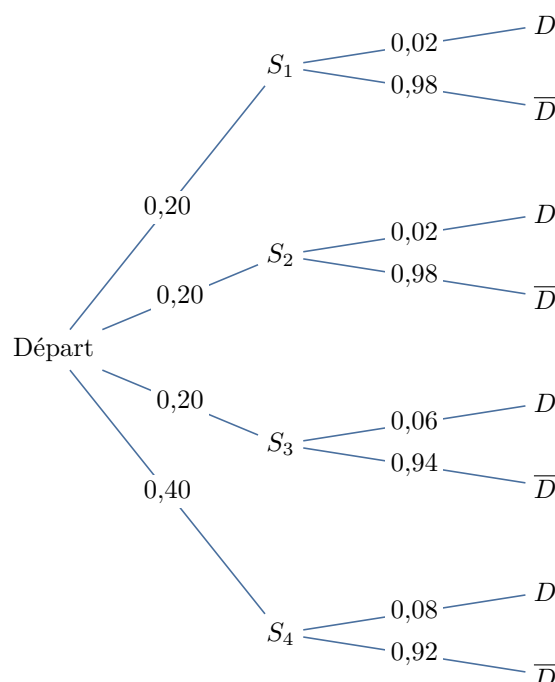
Analyse : comment choisir la méthode ?

L'origine du fichier partage la population en quatre groupes de poids différents. Un tableau pondéré permet de calculer la contribution de chaque serveur aux défauts, puis de les additionner. La moyenne simple des quatre taux ignorerait que le serveur 4 contient deux fois plus de données. Observer un défaut conduit ensuite à inverser le conditionnement par Bayes.

Repères de cours. Voir TD3, exercices 2 et 4 pour les probabilités totales puis Bayes ; ici, la partition possède quatre groupes.



Résolution.



S_i : serveur d'origine ; D : fichier défectueux.

Notons

$$S_i = \{\text{le fichier provient du serveur } S_i\},$$

et

$$D = \{\text{le fichier est défectueux}\}.$$

On a

$$\mathbb{P}(S_4) = 0,40.$$

Les 60% restants sont répartis en trois parts égales, donc

$$\mathbb{P}(S_1) = \mathbb{P}(S_2) = \mathbb{P}(S_3) = 0,20.$$

Les taux de défauts donnent

$$\mathbb{P}(D \mid S_1) = 0,02,$$

$$\mathbb{P}(D \mid S_2) = 0,02,$$

$$\mathbb{P}(D \mid S_3) = 0,06,$$

et

$$\mathbb{P}(D \mid S_4) = 0,08.$$

On peut résumer les données par

Serveur	$\mathbb{P}(S_i)$	$\mathbb{P}(D \mid S_i)$	$\mathbb{P}(S_i \cap D)$
S_1	0,20	0,02	0,004
S_2	0,20	0,02	0,004
S_3	0,20	0,06	0,012
S_4	0,40	0,08	0,032

2a) Probabilité qu'un fichier soit défectueux

Les quatre événements S_1, S_2, S_3, S_4 forment une partition. Par la formule des probabilités totales,

$$\begin{aligned} \mathbb{P}(D) &= \mathbb{P}(S_1)\mathbb{P}(D \mid S_1) + \mathbb{P}(S_2)\mathbb{P}(D \mid S_2) \\ &\quad + \mathbb{P}(S_3)\mathbb{P}(D \mid S_3) + \mathbb{P}(S_4)\mathbb{P}(D \mid S_4). \end{aligned}$$

Donc

$$\begin{aligned} \mathbb{P}(D) &= 0,20 \times 0,02 + 0,20 \times 0,02 \\ &\quad + 0,20 \times 0,06 + 0,40 \times 0,08 \\ &= 0,004 + 0,004 + 0,012 + 0,032 \\ &= \boxed{0,052}. \end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(D) = 5,2\%}.$$

2b) Le fichier provient de S_4 et est défectueux

$$\begin{aligned} \mathbb{P}(S_4 \cap D) &= \mathbb{P}(S_4)\mathbb{P}(D \mid S_4) \\ &= 0,40 \times 0,08 \\ &= \boxed{0,032}. \end{aligned}$$

Donc

$$\boxed{\mathbb{P}(S_4 \cap D) = 3,2\%}.$$

2c) Défaut sachant que le fichier provient de S_4

Cette probabilité est directement donnée par l'énoncé :

$$\mathbb{P}(D \mid S_4) = 0,08 = 8\%.$$

2d) Provenance de S_4 sachant que le fichier est défectueux

On cherche maintenant

$$\mathbb{P}(S_4 \mid D).$$

Par définition,

$$\begin{aligned}\mathbb{P}(S_4 \mid D) &= \frac{\mathbb{P}(S_4 \cap D)}{\mathbb{P}(D)} \\ &= \frac{0,032}{0,052} \\ &= \frac{8}{13} \\ &\approx 0,6154.\end{aligned}$$

Donc

$$\mathbb{P}(S_4 \mid D) \approx 61,5\%.$$

Méthode à retenir

Une question du type “sachant que le fichier est défectueux, d'où vient-il?” inverse le sens du conditionnement. Il faut alors calculer le dénominateur par la formule des probabilités totales.

Point de vigilance

Les deux quantités

$$\mathbb{P}(D \mid S_4) = 8\%$$

et

$$\mathbb{P}(S_4 \mid D) \approx 61,5\%$$

n'ont pas la même signification.

Exercice 7**Exercice 7 :**

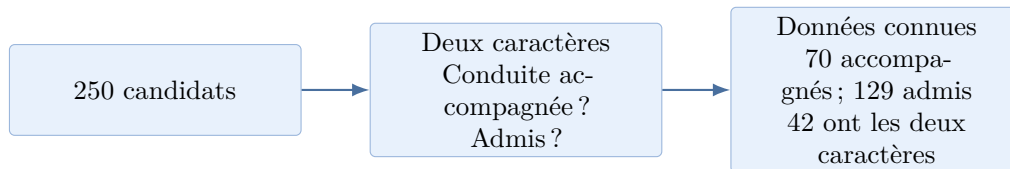
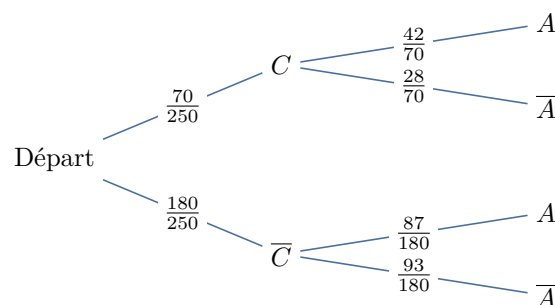
Pour étudier la réussite à l'épreuve pratique du permis de conduire, on observe 250 candidats d'une auto-école se présentant pour la 1^{ère} fois. Au total, 129 candidats sont admis. 70 candidats ont pratiqué la conduite accompagnée avant de se présenter à l'examen du permis de conduire. Il y a 42 candidats ayant pratiqué la conduite accompagnée et qui sont admis. On suppose qu'on peut étendre ces résultats, obtenus à partir d'un échantillon représentatif, à l'ensemble des candidats au permis de conduire de l'auto-école.

- 1) Définir clairement les événements (Préciser les notations utilisées). Représenter l'arbre de probabilités. Reporter les informations de l'énoncé à l'aide de vos notations.
- 2) Pour les questions suivantes, un candidat est choisi au hasard.
 - a) Quelle est la probabilité que ce candidat ne soit pas admis à l'épreuve pratique du permis de conduire ?
 - b) Sachant que le candidat est admis à l'épreuve pratique du permis de conduire, quelle est la probabilité qu'il ait fait la conduite accompagnée ?
 - c) Le candidat a fait la conduite accompagnée. Quelle est la probabilité qu'il ne soit pas admis à l'épreuve pratique du permis de conduire ?
 - d) Le candidat n'a pas fait la conduite accompagnée. Quelle est la probabilité qu'il soit

Analyse : comment choisir la méthode ?

Les données sont des effectifs croisant deux caractères : formation suivie et réussite. Un tableau d'effectifs permet de retrouver les quatre groupes sans convertir prématurément en probabilités. Un arbre serait possible, mais exigerait d'abord ces mêmes effectifs pour pondérer ses branches. Dans chaque question conditionnelle, on divise par l'effectif du groupe indiqué après « sachant ».

Repères de cours. Voir TD3, exercice 3 pour organiser deux caractères dans un tableau et identifier le groupe de référence.

**Résolution.**

C : conduite accompagnée ; A : admission. Les effectifs utilisés sont détaillés ci-dessous.

Notons

$$C = \{\text{le candidat a pratiqué la conduite accompagnée}\}$$

et

$$A = \{\text{le candidat est admis}\}.$$

Méthode à retenir

Lorsque l'énoncé fournit des effectifs croisant deux caractères, commencer par construire un tableau d'effectifs. Les probabilités conditionnelles deviennent ensuite de simples quotients.

On connaît

$$\text{Card}(A) = 129, \quad \text{Card}(C) = 70, \quad \text{Card}(A \cap C) = 42.$$

Par conséquent,

$$\text{Card}(C \cap \bar{A}) = 70 - 42 = 28,$$

et

$$\text{Card}(\bar{C} \cap A) = 129 - 42 = 87.$$

Comme

$$\text{Card}(\bar{C}) = 250 - 70 = 180,$$

on obtient

$$\text{Card}(\bar{C} \cap \bar{A}) = 180 - 87 = 93.$$

Le tableau complet est donc

	A	$\bar{A}$	Total
C	42	28	70
$\bar{C}$	87	93	180
Total	129	121	250

2a) Le candidat n'est pas admis

Il y a 121 candidats non admis parmi 250. Ainsi

$$\mathbb{P}(\overline{A}) = \frac{121}{250} = 0,484.$$

Donc

$$\boxed{\mathbb{P}(\overline{A}) = 48,4\%}.$$

2b) Conduite accompagnée sachant que le candidat est admis

$$\begin{aligned}\mathbb{P}(C \mid A) &= \frac{\text{Card}(C \cap A)}{\text{Card}(A)} \\ &= \frac{42}{129} \\ &= \frac{14}{43} \\ &\approx 0,3256.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(C \mid A) \approx 32,6\%}.$$

2c) Non admis sachant que le candidat a pratiqué la conduite accompagnée

$$\begin{aligned}\mathbb{P}(\overline{A} \mid C) &= \frac{\text{Card}(\overline{A} \cap C)}{\text{Card}(C)} \\ &= \frac{28}{70} \\ &= \boxed{0,40}.\end{aligned}$$

Donc

$$\boxed{\mathbb{P}(\overline{A} \mid C) = 40\%}.$$

2d) Admis sachant que le candidat n'a pas pratiqué la conduite accompagnée

$$\begin{aligned}\mathbb{P}(A \mid \overline{C}) &= \frac{\text{Card}(A \cap \overline{C})}{\text{Card}(\overline{C})} \\ &= \frac{87}{180} \\ &= \frac{29}{60} \\ &\approx 0,4833.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(A \mid \overline{C}) \approx 48,3\%}.$$

Point de vigilance

Dans

$$\mathbb{P}(C \mid A),$$

le nouvel univers de référence est l'ensemble des 129 admis. Le dénominateur est donc 129, et non 250.

Exercice 8

Exercice 8 :

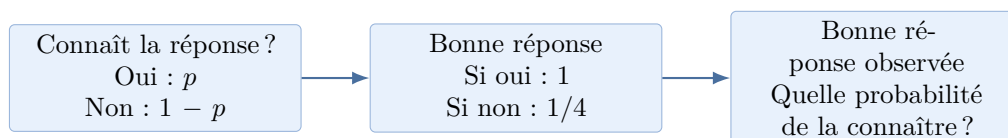
Un questionnaire à choix multiples propose 4 réponses pour chaque question. Soit p la probabilité qu'un étudiant connaisse la bonne réponse à une question donnée. S'il ignore la réponse, il choisit au hasard l'une des réponses proposées. Quelle est, pour le correcteur, la probabilité qu'un étudiant connaisse vraiment la bonne réponse lorsqu'il l'a donnée ?

Application numérique pour $p = 0,8$, $p = 0,5$ et $p = 0,2$.

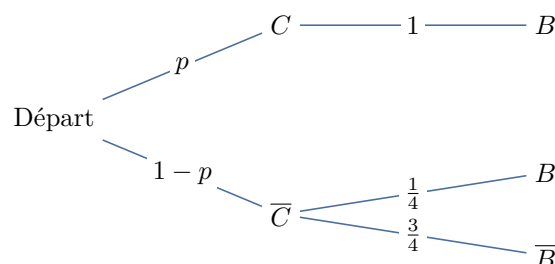
Analyse : comment choisir la méthode ?

Une bonne réponse peut venir de la connaissance ou du hasard : il faut distinguer ces deux causes avant de les réunir. Un arbre donne la probabilité totale d'une bonne réponse ; Bayes permet ensuite de retrouver la part due à la connaissance. Le taux $1/4$ ne concerne que les étudiants qui ignorent la réponse. Garder le paramètre p jusqu'au bout évite de refaire trois raisonnements séparés.

Repères de cours. Voir TD3, exercice 2 pour Bayes ; on conserve cette fois le paramètre p jusqu'au résultat final.



Résolution.



C : connaît la réponse ; B : donne la bonne réponse. Lorsque la réponse est connue, seule la branche certaine, de probabilité 1, est représentée.

Notons

$$C = \{\text{l'étudiant connaît la bonne réponse}\}$$

et

$$B = \{\text{l'étudiant donne la bonne réponse}\}.$$

On a

$$\mathbb{P}(C) = p, \quad \mathbb{P}(\overline{C}) = 1 - p.$$

Si l'étudiant connaît la réponse, il donne la bonne réponse :

$$\mathbb{P}(B \mid C) = 1.$$

S'il ne connaît pas la réponse, il choisit au hasard parmi quatre possibilités :

$$\mathbb{P}(B \mid \overline{C}) = \frac{1}{4}.$$

Méthode à retenir

La probabilité cherchée est

$$\mathbb{P}(C \mid B).$$

Il faut commencer par calculer $\mathbb{P}(B)$ à l'aide de la formule des probabilités totales.

Tableau des probabilités conjointes.

	B	$\bar{B}$	Total
C	p	0	p
$\bar{C}$	$(1-p)/4$	$3(1-p)/4$	$1-p$
Total	$(1+3p)/4$	$3(1-p)/4$	1

La colonne B est le nouvel univers de référence : la part de C dans cette colonne est $p/[(1+3p)/4]$.

On a

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(C)\mathbb{P}(B \mid C) + \mathbb{P}(\bar{C})\mathbb{P}(B \mid \bar{C}) \\ &= p + (1-p)\frac{1}{4} \\ &= \frac{4p + 1 - p}{4} \\ &= \frac{1 + 3p}{4}.\end{aligned}$$

Par ailleurs,

$$\mathbb{P}(C \cap B) = \mathbb{P}(C) = p,$$

puisque lorsqu'il connaît la réponse, il répond correctement.

Ainsi

$$\begin{aligned}\mathbb{P}(C \mid B) &= \frac{\mathbb{P}(C \cap B)}{\mathbb{P}(B)} \\ &= \frac{p}{(1+3p)/4} \\ &= \boxed{\frac{4p}{1+3p}}.\end{aligned}$$

Application numérique : $p = 0,8$

$$\mathbb{P}(C \mid B) = \frac{4 \times 0,8}{1 + 3 \times 0,8} = \frac{3,2}{3,4} = \frac{16}{17}.$$

Donc

$$\boxed{\mathbb{P}(C \mid B) \approx 94,1\%}.$$

Application numérique : $p = 0,5$

$$\mathbb{P}(C \mid B) = \frac{2}{2,5} = 0,8.$$

Donc

$$\boxed{\mathbb{P}(C \mid B) = 80\%}.$$

Application numérique : $p = 0,2$

$$\mathbb{P}(C \mid B) = \frac{0,8}{1,6} = 0,5.$$

Donc

$$\boxed{\mathbb{P}(C \mid B) = 50\%}.$$

Remarque

Lorsque $p = 0,2$, même après avoir observé une bonne réponse, le correcteur ne peut conclure qu'avec une probabilité de 50% que l'étudiant connaissait effectivement la réponse.

Point de vigilance

La probabilité $1/4$ ne concerne que le cas où l'étudiant ignore la réponse. Elle ne doit pas être appliquée à tous les étudiants.

Exercice 9**Exercice 9 :**

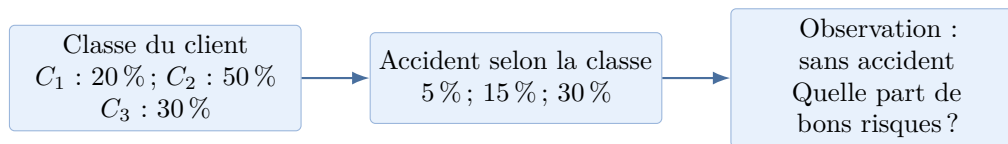
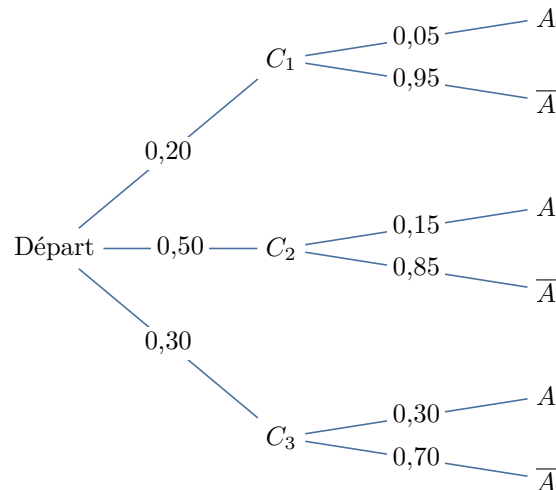
Une compagnie d'assurance répartit ses clients en trois classes C_1, C_2 et C_3 : les bons risques, les risques moyens, et les mauvais risques. Les effectifs de ces trois classes représentent 20% de la population totale pour la classe C_1 , 50% pour la classe C_2 , et 30% pour la classe C_3 . Les statistiques indiquent que les probabilités d'avoir un accident au cours de l'année pour une personne de l'une de ces trois classes sont respectivement de 0,05, 0,15 et 0,30.

- 1) Quelle est la probabilité qu'une personne choisie au hasard dans la population ait un accident dans l'année ?
- 2) Si M. Dufée n'a pas eu d'accident cette année, quelle est la probabilité qu'il soit un bon risque (en classe C_1) ?

Analyse : comment choisir la méthode ?

Les classes de risque forment trois groupes dont les proportions diffèrent : on pondère chaque taux d'accident par le poids de sa classe. Le tableau des intersections organise à la fois accidents et absences d'accident. Pour la seconde question, on sélectionne la colonne « sans accident », puis on mesure la part des bons risques ; utiliser la colonne des accidents répondrait à une autre question.

Repères de cours. Voir TD3, exercices 4 et 6 pour une partition en plusieurs origines ; complémentaire : voir TD2, exercice 5.

**Résolution.**

C_1, C_2, C_3 : classes de risque ; A : accident dans l'année.

Notons

$A = \{\text{le client a un accident dans l'année}\}.$

Les données sont

$$\mathbb{P}(C_1) = 0,20, \quad \mathbb{P}(C_2) = 0,50, \quad \mathbb{P}(C_3) = 0,30,$$

et

$$\mathbb{P}(A | C_1) = 0,05,$$

$$\mathbb{P}(A | C_2) = 0,15,$$

$$\mathbb{P}(A | C_3) = 0,30.$$

Tableau pondéré des deux issues.

	A	$\bar{A}$	Total
C_1	0,010	0,190	0,20
C_2	0,075	0,425	0,50
C_3	0,090	0,210	0,30
Total	0,175	0,825	1

Chaque case est une intersection. Pour retrouver les bons risques parmi les clients sans accident, on lit 0,190/0,825 dans la seconde colonne.

1) Probabilité d'avoir un accident

Les trois classes C_1, C_2, C_3 forment une partition de la population. Par la formule des probabilités totales,

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(C_1)\mathbb{P}(A | C_1) + \mathbb{P}(C_2)\mathbb{P}(A | C_2) \\ &\quad + \mathbb{P}(C_3)\mathbb{P}(A | C_3). \end{aligned}$$

Ainsi

$$\begin{aligned}\mathbb{P}(A) &= 0,20 \times 0,05 + 0,50 \times 0,15 + 0,30 \times 0,30 \\ &= 0,01 + 0,075 + 0,09 \\ &= \boxed{0,175}.\end{aligned}$$

Donc

$$\boxed{\mathbb{P}(A) = 17,5\%}.$$

Par conséquent,

$$\mathbb{P}(\bar{A}) = 1 - 0,175 = 0,825.$$

2) M. Dufée n'a pas eu d'accident

On cherche

$$\mathbb{P}(C_1 \mid \bar{A}).$$

Pour un bon risque,

$$\mathbb{P}(\bar{A} \mid C_1) = 1 - \mathbb{P}(A \mid C_1) = 0,95.$$

Donc

$$\mathbb{P}(C_1 \cap \bar{A}) = 0,20 \times 0,95 = 0,19.$$

Par la formule de Bayes,

$$\begin{aligned}\mathbb{P}(C_1 \mid \bar{A}) &= \frac{\mathbb{P}(C_1 \cap \bar{A})}{\mathbb{P}(\bar{A})} \\ &= \frac{0,19}{0,825} \\ &= \frac{38}{165} \\ &\approx 0,2303.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(C_1 \mid \bar{A}) \approx 23,0\%}.$$

Remarque

Avant toute information sur l'année écoulée, la probabilité d'appartenir à la classe C_1 est 20%. Après avoir appris que le client n'a pas eu d'accident, elle passe à environ 23%. L'information reçue modifie donc les probabilités des différentes classes.

Méthode à retenir

Pour une question de Bayes :

- 1) identifier clairement l'événement observé ;
- 2) calculer sa probabilité totale ;
- 3) calculer l'intersection recherchée ;
- 4) diviser l'intersection par la probabilité de l'événement observé.

Point de vigilance

Dans la seconde question, l'information donnée est "pas d'accident". Il faut donc conditionner par $\bar{A}$, et non par A .