Plaquette 3 – Conditionnement : probabilités composées, totales et formule de Bayes

Probabilités finies — Prépa MPSI / MP2I (1re année)

Corrections détaillées

Boîte à outils

Conditionnement.

— $\mathbb{P}_A = \mathbb{P}(\cdot \mid A)$ est une probabilité sur Ω (pour $\mathbb{P}(A) > 0$).

— **Probabilités composées** :

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)\cdots\mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

— **Probabilités totales** : (C_i) système complet d'événements de probabilités non nulles, $\mathbb{P}(B) = \sum_i \mathbb{P}(C_i)\,\mathbb{P}(B \mid C_i)$.

— **Bayes** :

$$\mathbb{P}(C_j \mid B) = \frac{\mathbb{P}(C_j)\,\mathbb{P}(B \mid C_j)}{\sum_i \mathbb{P}(C_i)\,\mathbb{P}(B \mid C_i)}$$

— **Suites de conditionnements** : si $p_{n+1} = ap_n + b$ avec $|a| < 1$, alors $p_n = \ell + a^n(p_0 - \ell)$ avec $\ell = \frac{b}{1-a}$.

Correction de l'exercice 1 – L'urne de Pólya

a) Après un premier tirage blanc (B_1), l'urne contient 2 blanches et 1 noire ; après un noir, 1 blanche et 2 noires. Probabilités totales :

$$\mathbb{P}(B_2) = \frac{1}{2} \times \frac{2}{3} + \frac{1}{2} \times \frac{1}{3} = \frac{1}{2}$$

b) Bayes : $\mathbb{P}(B_1 \mid B_2) = \dfrac{\mathbb{P}(B_1)\mathbb{P}(B_2 \mid B_1)}{\mathbb{P}(B_2)} = \dfrac{\frac{1}{2} \times \frac{2}{3}}{\frac{1}{2}} = \frac{2}{3}$.

c) $\mathbb{P}(B_1 \cap B_2) + \mathbb{P}(\overline{B_1} \cap \overline{B_2}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$. Le mécanisme renforce la couleur tirée : les tirages ne sont pas indépendants, bien que chaque tirage, pris isolément, soit blanc avec probabilité $\frac{1}{2}$.

Réponse. $\mathbb{P}(B_2) = \frac{1}{2}$; $\mathbb{P}(B_1 \mid B_2) = \frac{2}{3}$; même couleur : $\frac{2}{3}$.

Correction de l'exercice 2 – Probabilités composées : tirages sans remise

a) On conditionne étape par étape :

$$\mathbb{P}(R_1 \cap R_2 \cap N_3) = \mathbb{P}(R_1)\,\mathbb{P}(R_2 \mid R_1)\,\mathbb{P}(N_3 \mid R_1 \cap R_2) = \frac{4}{10} \times \frac{3}{9} \times \frac{6}{8} = \frac{1}{10}$$

b) Les trois ordres RRN, RNR, NRR ont la même probabilité (mêmes facteurs dans un autre ordre) : $3 \times \frac{1}{10} = \frac{3}{10}$. Directement : $\dfrac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = \frac{36}{120} = \frac{3}{10}$.

c) Par symétrie, toutes les boules ont la même chance d'être tirée en troisième position : $\mathbb{P}(R_3) = \frac{4}{10}$. Par le calcul, en sommant sur les 4 chemins :

— $\dfrac{4 \times 3 \times 2 + 4 \times 6 \times 3 + 6 \times 4 \times 3 + 6 \times 5 \times 4}{10 \times 9 \times 8} = \dfrac{24 + 72 + 72 + 120}{720} = \dfrac{288}{720} = \dfrac{2}{5}$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. $\frac{1}{10}$; exactement deux rouges : $\frac{3}{10}$; $\mathbb{P}(R_3) = \frac{2}{5}$.

Correction de l'exercice 3 – Un filtre anti-spam bayésien

a) Probabilités totales :

$$\mathbb{P}(G) = 0{,}4 \times 0{,}3 + 0{,}6 \times 0{,}02 = 0{,}12 + 0{,}012 = 0{,}132$$

b) Bayes : $\mathbb{P}(S \mid G) = \frac{0{,}12}{0{,}132} = \frac{10}{11} \approx 0{,}909$.

c) $\mathbb{P}(\overline{G}) = 0{,}868$ et $\mathbb{P}(S \cap \overline{G}) = 0{,}4 \times 0{,}7 = 0{,}28$, donc $\mathbb{P}(S \mid \overline{G}) = \frac{0{,}28}{0{,}868} \approx 0{,}323$.

d) Le mot fait passer la probabilité de spam de 40 % à 91 % : c'est un bon indice. Mais son absence ne l'abaisse qu'à 32 % : la plupart des spams ne contiennent pas ce mot. Un vrai filtre combine des centaines de mots (en les supposant indépendants sachant la classe : « bayésien naïf »).

Réponse. $\mathbb{P}(G) = 0{,}132$; $\mathbb{P}(S \mid G) = \frac{10}{11}$; $\mathbb{P}(S \mid \overline{G}) \approx 0{,}323$.

Correction de l'exercice 4 – Trois machines, une pièce défectueuse

a) Système complet (A, B, C), probabilités totales :

$$\mathbb{P}(D) = 0{,}5 \times 0{,}01 + 0{,}3 \times 0{,}02 + 0{,}2 \times 0{,}05 = 0{,}005 + 0{,}006 + 0{,}010 = 0{,}021$$

b) Bayes, en divisant chaque terme par 0,021 :

— $\mathbb{P}(A \mid D) = \frac{5}{21} \approx 0{,}24$, $\mathbb{P}(B \mid D) = \frac{6}{21} \approx 0{,}29$, $\mathbb{P}(C \mid D) = \frac{10}{21} \approx 0{,}48$.

c) La machine C, qui ne produit que 20 % des pièces, est responsable de près de la moitié des défauts. Avec 1 % : $\mathbb{P}(D) = 0{,}005 + 0{,}006 + 0{,}002 = 0{,}013$, soit une baisse de 38 % environ.

Réponse. $\mathbb{P}(D) = 2{,}1\,\%$; $\frac{5}{21}$, $\frac{6}{21}$, $\frac{10}{21}$; améliorer C : 1,3 %.

Correction de l'exercice 5 – Transmission d'un bit par relais

a) Le bit est correct après $n+1$ relais si et seulement s'il était correct après n et que le relais $n+1$ ne l'inverse pas, ou s'il était faux et qu'il est inversé (probabilités totales) :

$$q_{n+1} = (1-p)q_n + p(1-q_n) = (1-2p)q_n + p$$

b) Point fixe : $\ell = (1-2p)\ell + p$, soit $\ell = \frac{1}{2}$. Alors $q_{n+1} - \frac{1}{2} = (1-2p)\left(q_n - \frac{1}{2}\right)$, suite géométrique : $q_n = \frac{1}{2} + (1-2p)^n\left(1 - \frac{1}{2}\right)$.

c) $p = 0{,}1$: $q_2 = \frac{1}{2} + \frac{1}{2} \times 0{,}64 = 0{,}82$ (c'est $0{,}9^2 + 0{,}1^2$) ; $q_{10} = \frac{1}{2} + \frac{1}{2} \times 0{,}8^{10} \approx 0{,}554$. Comme $|1-2p| < 1$, $q_n \to \frac{1}{2}$: après beaucoup de relais, le bit reçu n'apporte plus aucune information (pile ou face).

Réponse. $q_{n+1} = (1-2p)q_n + p$; $q_n = \frac{1}{2} + \frac{1}{2}(1-2p)^n$; $q_2 = 0{,}82$, $q_{10} \approx 0{,}554$, limite $\frac{1}{2}$.

Correction de l'exercice 6 – Les coffrets de Bertrand

a) On note C_1, C_2, C_3 les coffrets et O : « la pièce tirée est en or ». Probabilités totales :

$$\mathbb{P}(O) = \frac{1}{3} \times 1 + \frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times 0 = \frac{1}{2}$$

— L'autre pièce est en or si et seulement si le coffret est C_1 : $\mathbb{P}(C_1 \mid O) = \frac{\frac{1}{3} \times 1}{\frac{1}{2}} = \frac{2}{3}$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

b) Sachant O, les coffrets C_1 et C_2 ne sont pas équiprobables : le coffret C_1 a deux fois plus de chances de produire une pièce d'or. Autre vue : il y a 3 pièces d'or équiprobables ; deux d'entre elles sont dans C_1.

Réponse. $\mathbb{P}(C_1 \mid O) = \frac{2}{3}$.

Correction de l'exercice 7 – Garçons et filles

Univers : $\{GG, GF, FG, FF\}$ (aîné en premier), équiprobable.

a) $\mathbb{P}(GG \mid \text{aîné } G) = \dfrac{\mathbb{P}(GG)}{\mathbb{P}(\{GG, GF\})} = \dfrac{1/4}{1/2} = \dfrac{1}{2}$.

b) « Au moins un garçon » $= \{GG, GF, FG\}$:

$$\mathbb{P}(GG \mid \text{au moins un } G) = \frac{1/4}{3/4} = \frac{1}{3}$$

c) L'information b) est plus faible : elle n'élimine que FF et laisse trois cas équiprobables, dont un seul est GG. En a), on sait **lequel** des enfants est un garçon, ce qui élimine deux cas.

Réponse. a) $\frac{1}{2}$; b) $\frac{1}{3}$.

Correction de l'exercice 8 – Premier as à la troisième carte

a) Probabilités composées : non-as, non-as, as :

$$\frac{48}{52} \times \frac{47}{51} \times \frac{4}{50} = \frac{9\,024}{132\,600} \approx 0{,}068$$

b) $\mathbb{P}(T = k) = \frac{48}{52} \times \frac{47}{51} \times \cdots \times \frac{48-k+2}{52-k+2} \times \frac{4}{52-k+1}$ pour $1 \le k \le 49$.

c) $\frac{4}{52} \times \frac{3}{51} = \frac{12}{2\,652} = \frac{1}{221} \approx 0{,}0045$.

Réponse. $\approx 0{,}068$; formule produit ; $\frac{1}{221}$.

Correction de l'exercice 9 – Savoir ou deviner ?

a) $\mathbb{P}(J) = 0{,}6 \times 1 + 0{,}4 \times \frac{1}{4} = 0{,}7$.

b) Bayes :

$$\mathbb{P}(K \mid J) = \frac{0{,}6}{0{,}7} = \frac{6}{7} \approx 0{,}857$$

c) Avec 2 réponses : $\mathbb{P}(J) = 0{,}6 + 0{,}4 \times \frac{1}{2} = 0{,}8$, et $\mathbb{P}(K \mid J) = \frac{0{,}6}{0{,}8} = 0{,}75$. Moins il y a de choix, plus une bonne réponse peut être due au hasard, et moins elle prouve la connaissance.

Réponse. $\mathbb{P}(J) = 0{,}7$; $\mathbb{P}(K \mid J) = \frac{6}{7}$; avec 2 choix : 0,75.

Correction de l'exercice 10 – Une chaîne météorologique

a) Probabilités totales avec le système complet (beau, pluie) au jour n :

$$p_{n+1} = 0{,}8 p_n + 0{,}4(1 - p_n) = 0{,}4 p_n + 0{,}4$$

b) Point fixe $\ell = 0{,}4\ell + 0{,}4$, soit $\ell = \frac{2}{3}$. Suite arithmético-géométrique :
$p_n = \frac{2}{3} + 0{,}4^n\left(p_0 - \frac{2}{3}\right) \to \frac{2}{3}$, quelle que soit la météo initiale.

c) $p_0 = 1$: $p_3 = \frac{2}{3} + 0{,}064 \times \frac{1}{3} = \frac{2}{3} + 0{,}0213 \approx 0{,}688$. Contrôle par récurrence : $p_1 = 0{,}8$, $p_2 = 0{,}72$, $p_3 = 0{,}688$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. $p_n = \frac{2}{3} + 0{,}4^n(p_0 - \frac{2}{3}) \to \frac{2}{3}$; $p_3 = 0{,}688$.

Correction de l'exercice 11 – Le sophisme du procureur

a) Il a calculé $\mathbb{P}(I \mid \overline{C}) = 10^{-6}$, et l'a confondu avec $\mathbb{P}(\overline{C} \mid I)$, qui est la quantité pertinente.

b) Bayes, avec $\mathbb{P}(I \mid C) = 1$:

$$\mathbb{P}(C \mid I) = \frac{\mathbb{P}(C)}{\mathbb{P}(C) + \mathbb{P}(\overline{C}) \times 10^{-6}} = \frac{\frac{1}{2 \times 10^6}}{\frac{1}{2 \times 10^6} + \left(1 - \frac{1}{2 \times 10^6}\right)10^{-6}} \approx \frac{0{,}5}{0{,}5 + 1} = \frac{1}{3}$$

— Intuitivement : dans la ville, on s'attend à environ 2 innocents présentant l'indice, plus le coupable. Le suspect est l'un de ces 3.

c) Un indice rare n'est pas une preuve lorsqu'on l'a cherché dans une grande population : il faut tenir compte de la probabilité **a priori** (d'autres éléments de l'enquête).

Réponse. Confusion $\mathbb{P}(I \mid \overline{C})$ / $\mathbb{P}(\overline{C} \mid I)$; $\mathbb{P}(C \mid I) \approx \frac{1}{3}$.

Correction de l'exercice 12 – Un dé choisit l'urne

a) Probabilités totales :

$$\mathbb{P}(B) = \sum_{k=1}^{6} \frac{1}{6} \times \frac{k}{6} = \frac{21}{36} = \frac{7}{12}$$

b) Bayes : $\mathbb{P}(D = 6 \mid B) = \frac{\frac{1}{6} \times 1}{\frac{7}{12}} = \frac{2}{7}$; $\mathbb{P}(D = 1 \mid B) = \frac{\frac{1}{6} \times \frac{1}{6}}{\frac{7}{12}} = \frac{1}{21}$.

c) $\mathbb{P}(D = k \mid B) = \frac{\frac{k}{36}}{\frac{21}{36}} = \frac{k}{21}$: c'est la loi du dé truqué de la plaquette précédente. Somme : $\frac{1 + \ldots + 6}{21} = 1$. L'observation d'une boule blanche « tire » la loi du dé vers les grandes valeurs.

Réponse. $\mathbb{P}(B) = \frac{7}{12}$; $\mathbb{P}(6 \mid B) = \frac{2}{7}$, $\mathbb{P}(1 \mid B) = \frac{1}{21}$; $\mathbb{P}(k \mid B) = \frac{k}{21}$.

Correction de l'exercice 13 – Dépistage répété

a) $\mathbb{P}(T_1) = 0{,}01 \times 0{,}95 + 0{,}99 \times 0{,}05 = 0{,}0095 + 0{,}0495 = 0{,}059$, et

$$\mathbb{P}(M \mid T_1) = \frac{0{,}0095}{0{,}059} \approx 0{,}161$$

b) Par indépendance conditionnelle, $\mathbb{P}(T_1 \cap T_2 \mid M) = 0{,}95^2$ et $\mathbb{P}(T_1 \cap T_2 \mid \overline{M}) = 0{,}05^2$:

— $\mathbb{P}(M \mid T_1 \cap T_2) = \frac{0{,}01 \times 0{,}9025}{0{,}01 \times 0{,}9025 + 0{,}99 \times 0{,}0025} = \frac{0{,}009025}{0{,}0115} \approx 0{,}785$.

c) Avec l'a priori $\pi = 0{,}161$: $\frac{\pi \times 0{,}95}{\pi \times 0{,}95 + (1 - \pi) \times 0{,}05} \approx \frac{0{,}153}{0{,}153 + 0{,}042} \approx 0{,}785$. Bayes s'applique **séquentiellement** : chaque résultat met à jour la probabilité, qui sert d'a priori pour le suivant.

Réponse. Après un test : ≈ 16 % ; après deux : ≈ 78,5 %.

Correction de l'exercice 14 – La probabilité conditionnelle est une probabilité

a) $\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \in [0, 1]$; $\mathbb{P}_A(\Omega) = 1$; si $B \cap C = \varnothing$, $A \cap B$ et $A \cap C$ sont disjoints :

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

$$\mathbb{P}_A(B \cup C) = \frac{\mathbb{P}(A \cap B) + \mathbb{P}(A \cap C)}{\mathbb{P}(A)} = \mathbb{P}_A(B) + \mathbb{P}_A(C)$$

b) Ce sont les propriétés générales d'une probabilité, appliquées à $\mathbb{P}_A$.

c) Non : on conditionne par deux événements différents. Dé équilibré, A : « pair », B : « obtenir 6 ». Alors $\mathbb{P}(B \mid A) = \frac{1}{3}$ et $\mathbb{P}(B \mid \overline{A}) = 0 \neq 1 - \frac{1}{3}$. (C'est $\mathbb{P}(\overline{B} \mid A) = 1 - \mathbb{P}(B \mid A)$ qui est vrai.)

d) C'est la formule des probabilités totales pour la probabilité $\mathbb{P}_A$ et le système complet $(C, \overline{C})$, car $\mathbb{P}_A(B \mid C) = \mathbb{P}(B \mid A \cap C)$.

Réponse. $\mathbb{P}_A$ est une probabilité ; $\mathbb{P}(B \mid \overline{A}) \neq 1 - \mathbb{P}(B \mid A)$ en général ; probabilités totales sous $\mathbb{P}_A$.

Correction de l'exercice 15 – Synthèse : ruine du joueur en trois étapes

a) Partant de 1 : avec probabilité p, il passe à 2 et réussit ensuite avec probabilité s_2 ; avec probabilité q, il est ruiné. Probabilités totales : $s_1 = ps_2 + q \times 0$. Partant de 2 : $s_2 = p \times 1 + qs_1$.

b) En substituant : $s_1 = p(p + qs_1)$, donc $s_1(1 - pq) = p^2$:

$$s_1 = \frac{p^2}{1 - pq}, \qquad s_2 = p + qs_1 = \frac{p - p^2q + p^2q}{1 - pq} = \frac{p}{1 - pq}$$

c) $p = \frac{1}{2}$: $1 - pq = \frac{3}{4}$, $s_1 = \frac{1/4}{3/4} = \frac{1}{3}$, $s_2 = \frac{2}{3}$. $p = 0{,}6$: $1 - pq = 0{,}76$, $s_1 = \frac{0{,}36}{0{,}76} = \frac{9}{19} \approx 0{,}47$, $s_2 = \frac{0{,}6}{0{,}76} = \frac{15}{19} \approx 0{,}79$.

d) Pour un jeu équitable, la probabilité d'atteindre 3 est proportionnelle à la fortune initiale : la fortune moyenne reste constante au cours du jeu ($k = 3s_k + 0 \times (1 - s_k)$), propriété qu'on retrouve avec l'espérance.

Réponse. $s_1 = \frac{p^2}{1 - pq}$, $s_2 = \frac{p}{1 - pq}$; $p = \frac{1}{2}$: $\frac{1}{3}$, $\frac{2}{3}$; $p = 0{,}6$: $\frac{9}{19}$, $\frac{15}{19}$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).