Corrigé 4

Exercice 1. (i). Soient M l'événement que la personne choisie est atteinte de la maladie, et T_i l'événement que le test $i \in \{1,2\}$ sort positif. Par les données et l'indépendance conditionnelle

$$\Pr(M) = 0.001, \quad \Pr(T_1 \cap T_2 | M) = \Pr(T_1|M)\Pr(T_2|M) = \frac{99}{100}\frac{99}{100}, \quad \Pr(T_1 \cap T_2 | M^c) = \frac{2}{100}\frac{2}{100}.$$

Par le théorème de Bayes

$$\Pr(M|T_1 \cap T_2) = \frac{\Pr(T_1 \cap T_2|M)\Pr(M)}{\Pr(T_1 \cap T_2|M)\Pr(M) + \Pr(T_1 \cap T_2|M^c)\Pr(M^c)} = \frac{\frac{1}{1000}\frac{99}{100}\frac{99}{100}}{\frac{1}{1000}\frac{99}{100}\frac{99}{100} + \frac{999}{1000}\frac{2}{100}\frac{2}{100}}$$

$$= \frac{9801}{9801 + 3996} \approx 0.71 >> 0.047.$$

(ii). La probabilité de $T_1 \cap T_2$ est le dénominateur ci-dessus, soit $(9801 + 3996)/10^7 = 0.0013797$. La probabilité de T_1 est $(99 + 1998)/10^5 = 0.02097$ comme vu en cours. La probabilité de T_2 est la même. Comme $\Pr(T_1 \cap T_2) > \Pr(T_1)\Pr(T_2)$ ils sont dépendants ; si on teste positif, la probabilité qu'on est malade est bien plus grande que la probabilité a-priori de 0.001, et donc la probabilité que le deuxième test sort positif augmente, c'est-à-dire $\Pr(T_2|T_1) > \Pr(T_2)$.

Exercice 2. (a)

$$X : \Omega \to \mathbb{R}; \qquad X(i) = i, \qquad i = 1, \ldots, 6.$$
$$Y : \Omega \to \mathbb{R}; \qquad Y(i) = \begin{cases} 0, & i = 1, \ldots, 5, \\ 1 & i = 6. \end{cases}$$

(b) Pour X on a $H = \{1, \ldots, 6\}$, pour Y on a $H = \{0, 1\}$.

(c) Pour X on a

x_i	1	2	3	4	5	6
$f(x_i) = \Pr(X = x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Pour Y on a

x_i	0	1
$f(x_i) = \Pr(Y = x_i)$	5/6	1/6

(d) Pour X on a

x_i	1	2	3	4	5	6
$F(x_i) = \Pr(X \leq x_i)$	1/6	2/6	3/6	4/6	5/6	1

Pour Y on a

x_i	0	1
$F(x_i) = \Pr(Y \leq x_i)$	5/6	1

Dans les deux cas, on a, pour $x \in [x_i, x_i + 1)$, $i = 1, \ldots, 6$, $F(x) = F(x_i)$.

Exercice 3. (a) Oui.

(b) Non, car la condition $\sum_i f(x_i) = 1$ n'est pas satisfaite.

(c) Oui.

(d) Non, car $f(x_i) = \Pr(X = x_i)$ implique que $0 \leq f(x_i) \leq 1$, $i = 1, \ldots, 11$, ce qui n'est pas le cas ici. Par ailleurs, la condition $\sum_i f(x_i) = 1$ n'est pas satisfaite.

Exercice 4. (a) On sait que $\sum_{x_i} f(x_i) = \sum_{i=0}^{\infty} P(X = i) = 1$ et on a

$$\sum_{i=0}^{\infty} P(X = i) = c \sum_{i=0}^{\infty} (1-p)^i = \frac{c}{1-(1-p)} = \frac{c}{p}.$$

Il faut donc que $c = p$.

(b) $\Pr(X = 0) = p(1-p)^0 = p$.

(c) $\Pr(X > 2) = 1 - \Pr(X = 0) - \Pr(X = 1) - \Pr(X = 2) = 1 - p - p(1-p) - p(1-p)^2 =$
$= (1-p)(1-p-p(1-p)) = (1-p)^2(1-p) = (1-p)^3$.

(d) Par définition de la probabilité conditionnelle on a

$$\Pr(X \geq n+m \,|\, X \geq n) = \frac{\Pr(\{X \geq n+m\} \cap \{X \geq n\})}{\Pr(X \geq n)}.$$

Remarquons que $X \geq n+m$ implique que $X \geq n$. Donc $\{X \geq n+m\} \subseteq \{X \geq n\}$ et $\Pr(\{X \geq n+m\} \cap \{X \geq n\}) = \Pr(X \geq n+m)$. Par conséquent,

$$\Pr(X \geq n+m \,|\, X \geq n) = \frac{\Pr(X \geq n+m)}{\Pr(X \geq n)}.$$

Il nous reste donc à calculer $\Pr(X \geq k)$ pour $k = 0, 1, 2, \ldots$. D'après la loi de la variable X, on a

$$\Pr(X \geq k) = \sum_{i=k}^{\infty} \Pr(X = i) = 1 - \sum_{i=0}^{k-1} \Pr(X = i) = 1 - \sum_{i=0}^{k-1} p(1-p)^i = 1-(1-(1-p)^k) = (1-p)^k,$$

où l'avant dernière égalité se trouve à l'aide de l'indication $\sum_{i=0}^{k-1} x^i = \frac{1-x^k}{1-x}$.
Donc,

$$\Pr(X \geq n+m \,|\, X \geq n) = \frac{\Pr(X \geq n+m)}{\Pr(X \geq n)} = \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m = \Pr(X \geq m).$$

Remarque : Cette loi s'appelle la loi géométrique. On l'utilise par exemple pour modéliser la situation suivante : on réalise des essais indépendants, chacun avec la même probabilité de succès p. La variable aléatoire X qui compte le nombre d'essais jusqu'au premier succès suit une loi géométrique de paramètre p.

La propriété qu'on a démontré dans la partie (d) est caractéristique de la loi géométrique : si on a déjà observé n échecs, la probabilité d'en observer encore m ou plus est la même que la probabilité d'observer au moins m échecs dès le début. On dit que la loi géométrique est "sans mémoire".

Exercice 5. (a) Non, car $F(x) = \Pr(X \leq x)$ implique que $\lim_{x \to \infty} F(x) = 1$.

(b) Oui.

(c) Non, car la fonction doit être non décroissante.

Exercice 6. La loi de Bernoulli, $Y \sim \mathcal{B}(1/6)$.

Exercice 7. (a) La loi binomiale, $X \sim \mathcal{B}(n, p)$.

(b) Pour un système à 5 composants, la variable aléatoire X suit la loi $\mathcal{B}(5, p)$, et le système fonctionne si au moins 3 composants fonctionnent. La probabilité que le système fonctionne est donc

$$\Pr(X \geq 3) = \Pr(X = 3)+\Pr(X = 4)+\Pr(X = 5) = \binom{5}{3}p^3(1-p)^2+\binom{5}{4}p^4(1-p)+\binom{5}{5}p^5.$$

Pour un système à 3 composants, la variable aléatoire X suit la loi $\mathcal{B}(3, p)$, et le système fonctionne si au moins 2 composants fonctionnent. La probabilité que le système fonctionne est donc

$$\Pr(X \geq 2) = \Pr(X = 2) + \Pr(X = 3) = \binom{3}{2}p^2(1 - p) + \binom{3}{3}p^3.$$

On cherche p tel que

$$\binom{5}{3}p^3(1 - p)^2 + \binom{5}{4}p^4(1 - p) + \binom{5}{5}p^5 > \binom{3}{2}p^2(1 - p) + \binom{3}{3}p^3,$$

i.e. tel que

$$10p^3(1 - p)^2 + 5p^4(1 - p) + p^5 - 3p^2(1 - p) - p^3 > 0.$$

On peut résoudre cette inégalité par exemple comme suit :

$$\begin{aligned}
& 10p^3(1 - p)^2 + 5p^4(1 - p) + p^5 - 3p^2(1 - p) - p^3 > 0 \\
\Leftrightarrow\ & (1 - p)[10p(1 - p) + 5p^2 - 3] + p^3 - p > 0 \quad \text{(on a divisé par } p^2 > 0) \\
\Leftrightarrow\ & (1 - p)[10p(1 - p) + 5p^2 - 3] - p(1 - p)(1 + p) > 0 \\
\Leftrightarrow\ & 10p(1 - p) + 5p^2 - 3 - p(1 + p) > 0 \quad \text{(on a divisé par } (1 - p) > 0) \\
\Leftrightarrow\ & -6p^2 + 9p - 3 > 0 \\
\Leftrightarrow\ & (2p - 1)(-3p + 3) > 0 \\
\Leftrightarrow\ & (2p - 1) > 0 \ \text{, car le fait que } 0 < p < 1 \text{ implique que } (-3p + 3) > 0 \\
\Leftrightarrow\ & p > 1/2.
\end{aligned}$$

Exercice 8. Considérons les événements

$$\begin{aligned}
C &= \text{la clé est dans le manteau court,} \\
L &= \text{la clé est dans le manteau long,} \\
V &= \text{la clé est dans la veste trench,} \\
T_C &= \text{la clé est trouvée dans le manteau court,} \\
T_L &= \text{la clé est trouvée dans le manteau long,} \\
T_V &= \text{la clé est trouvée dans la veste trench,}
\end{aligned}$$

On sait que

$$\begin{aligned} \Pr(C) = p_c, \quad & \Pr(L) = p_l, \quad \Pr(V) = p_v, \\ \Pr(T_C|C) = \alpha_c \quad & \Pr(T_L|L) = \alpha_l, \quad \Pr(T_V|V) = \alpha_v. \end{aligned}$$

La probabilité recherchée est donc

$$\Pr(C|T_C^c) = \frac{\Pr(C \cap T_C^c)}{\Pr(T_C^c)} = \frac{\Pr(T_C^c|C)\Pr(C)}{1 - \Pr(T_C)} =^* \frac{(1 - \Pr(T_C|C))\Pr(C)}{1 - \Pr(T_C \cap C)} = \frac{(1 - \Pr(T_C|C))\Pr(C)}{1 - \Pr(T_C|C)\Pr(C)} = \frac{p_c - \alpha_c p_c}{1 - \alpha_c p_c}.$$

* ici on a utilisé la formule des probabilités totales pour écrire $\Pr(T_C) = \Pr(T_C \cap C) + \Pr(T_C \cap C^c)$. Mais $\Pr(T_C \cap C^c) = 0$, parce que Julie ne peut trouver la clé que dans la poche où cette dernière se trouve. Donc $\Pr(T_C) = \Pr(T_C \cap C)$.

Exercice 9. Considérons les événements

$$\begin{aligned} O &= \text{le mathématicien a oublié son parapluie dans un magasin,} \\ O_i &= \text{le mathématicien a oublié son parapluie dans le } i^{\text{ième}}\text{magasin}, \quad i = 1, 2, 3, 4. \end{aligned}$$

Nous savons que

$$\begin{aligned} \Pr(O_1) &= \frac{1}{4}, \\ \Pr(O_2|O_1^c) &=^* \frac{1}{4}, \\ \Pr(O_3|O_1^c \cap O_2^c) &=^* \frac{1}{4}, \\ \Pr(O_4|O_1^c \cap O_2^c \cap O_3^c) &=^* \frac{1}{4}. \end{aligned}$$

* On ne peut oublier son parapluie dans un magasin que si on est entré le magasin avec le parapluie.

Selon la formule de Bayes la probabilité cherchée est

$$\Pr(O_4|O) = \frac{\Pr(O|O_4)\Pr(O_4)}{\sum_{i=1}^4 \Pr(O|O_i)\Pr(O_i)} =^{**} \frac{\Pr(O_4)}{\sum_{i=1}^4 \Pr(O_i)}.$$

** L'égalité découle du fait que $\Pr(O|O_i) = 1$ pour $i = 1, 2, 3, 4$.

Pour calculer $\Pr(O_2)$ on utilise la formule des probabilités totales :

$$\Pr(O_2) = \Pr(O_2 \cap O_1^c) + \Pr(O_2 \cap O_1) =^{***} \Pr(O_2 \cap O_1^c) = \Pr(O_2|O_1^c)\Pr(O_1^c) = \Pr(O_2|O_1^c)(1 - \Pr(O_1)) = \frac{1}{4} \times \frac{3}{4} =$$

*** $\Pr(O_2 \cap O_1) = 0$, parce qu'on ne peut pas oublier le même parapluie dans deux magasins différents.

On calcule $\Pr(O_3)$ et $\Pr(O_4)$ de la même manière :

$$\Pr(O_3) = \Pr(O_3 \cap (O_1^c \cap O_2^c)) + \Pr(O_3 \cap (O_1 \cup O_2)) = \Pr(O_3 \cap (O_1^c \cap O_2^c)) = \Pr(O_3|O_1^c \cap O_2^c)\Pr(O_1^c \cap O_2^c) =$$
$$= \Pr(O_3|O_1^c \cap O_2^c)\Pr(O_2^c|O_1^c)\Pr(O_1^c) = \Pr(O_3|O_1^c \cap O_2^c)(1 - \Pr(O_2|O_1^c))(1 - \Pr(O_1)) = \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{9}{64},$$

$$\begin{aligned}\Pr(O_4) &= \Pr(O_4 \cap (O_1^c \cap O_2^c \cap O_3^c)) + \Pr(O_4 \cap (O_1 \cup O_2 \cup O_3)) = \Pr(O_4 \cap (O_1^c \cap O_2^c \cap O_3^c)) = \\ &= \Pr(O_4|O_1^c \cap O_2^c \cap O_3^c)\Pr(O_1^c \cap O_2^c \cap O_3^c) = \Pr(O_4|O_1^c \cap O_2^c \cap O_3^c)\Pr(O_3^c|O_1^c \cap O_2^c)\Pr(O_1^c \cap O_2^c) = \\ &= \Pr(O_4|O_1^c \cap O_2^c \cap O_3^c)\Pr(O_3^c|O_1^c \cap O_2^c)\Pr(O_2^c|O_1^c)\Pr(O_1^c) = \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{256}.\end{aligned}$$

Donc

$$\Pr(O_4|O) = \frac{\Pr(O_4)}{\sum_{i=1}^4 \Pr(O_i)} = \frac{\frac{27}{256}}{\frac{1}{4} + \frac{3}{16} + \frac{9}{64} + \frac{27}{256}} = \frac{27}{175}.$$