Corrigé série 5

Exercice 1 (10 points)

a) On a

$$\mathrm{E}(X) = \sum_{i=1}^{6} i\frac{1}{6} = \frac{7}{2},$$

et

$$\mathrm{E}(X^2) = \sum_{i=1}^{6} i^2\frac{1}{6} = \frac{91}{6}.$$

Ainsi

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{35}{12}.$$

b) Ici, $\mathrm{E}(X) = \frac{1}{2}$ et, comme $\mathrm{E}(X) = \mathrm{E}(X^2)$, on a

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

c)

$$\begin{aligned}
\mathrm{E}(X) &= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1,\\
\mathrm{E}(X^2) &= 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} = \frac{3}{2},\\
\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{1}{2}.
\end{aligned}$$

d)

$$\begin{aligned}
\mathrm{E}(X) &= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2},\\
\mathrm{E}(X^2) &= 0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = 3,\\
\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{3}{4}.
\end{aligned}$$

Exercice 2 (5 points)

$$
\begin{aligned}
\mathrm{E}(X) &= 7\cdot\frac{1}{60}+8\cdot\frac{3}{60}+9\cdot\frac{6}{60}+10\cdot\frac{9}{60}+11\cdot\frac{14}{60}+12\cdot\frac{11}{60}\\
&\quad+13\cdot\frac{7}{60}+14\cdot\frac{6}{60}+15\cdot\frac{2}{60}+16\cdot\frac{1}{60}=\frac{341}{30},\\
\mathrm{E}(X^2) &= 7^2\cdot\frac{1}{60}+8^2\cdot\frac{3}{60}+9^2\cdot\frac{6}{60}+10^2\cdot\frac{9}{60}+11^2\cdot\frac{14}{60}+12^2\cdot\frac{11}{60}\\
&\quad+13^2\cdot\frac{7}{60}+14^2\cdot\frac{6}{60}+15^2\cdot\frac{2}{60}+16^2\cdot\frac{1}{60}=\frac{797}{6},\\
\mathrm{Var}(X) &= \mathrm{E}(X^2)-\mathrm{E}(X)^2=\frac{3269}{900}.
\end{aligned}
$$

Exercice 3 (5 points)

Soit X la variable aléatoire qui donne le plus grand numéro tiré et Y la variable aléatoire qui donne le plus petit numéro tiré. On calcule :

$$
\begin{aligned}
\mathrm{E}(X) &= 2\cdot\frac{2}{20}+3\cdot\frac{4}{20}+4\cdot\frac{6}{20}+5\cdot\frac{8}{20}=4\\
\mathrm{E}(Y) &= 1\cdot\frac{8}{20}+2\cdot\frac{6}{20}+3\cdot\frac{4}{20}+4\cdot\frac{2}{20}=2
\end{aligned}
$$

Pour calculer $\mathrm{E}(XY)$, il faut calculer la probabilité de tous les couples $(X=i, Y=j)$. Remarquons que si $j \geq i$, la probabilité est nulle. On a donc

$$\mathrm{E}(XY)=2\cdot1\cdot\frac{2}{20}+3\cdot1\cdot\frac{2}{20}+3\cdot2\cdot\frac{2}{20}+4\cdot1\cdot\frac{2}{20}+4\cdot2\cdot\frac{2}{20}+4\cdot3\cdot\frac{2}{20}+5\cdot1\cdot\frac{2}{20}+5\cdot2\cdot\frac{2}{20}+5\cdot3\cdot\frac{2}{20}+5\cdot4\cdot\frac{2}{20}=8,5$$

Ainsi, la covariance est égale à

$$\mathrm{Cov}(X,Y)=8,5-4\cdot2=0,5.$$

Exercice 4 (5 points)

On commence par calculer les espérances $E(X)$, $E(Y)$ et $E(XY)$ pour ensuite calculer la covariance.

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{i,j=1}^{6}(i+j)\cdot\frac{1}{36}=2\cdot\frac{1}{36}+3\cdot\frac{2}{36}+4\cdot\frac{3}{36}+\ldots+10\cdot\frac{3}{36}+11\cdot\frac{2}{36}+12\cdot\frac{1}{36}=7,\\
\mathrm{E}(Y) &= \sum_{i,j=1}^{6}(i-j)\cdot\frac{1}{36}=0\cdot\frac{6}{36}+1\cdot\frac{5}{36}+\ldots+5\cdot\frac{1}{36}+(-1)\cdot\frac{5}{36}+\ldots+(-5)\cdot\frac{1}{36}=0,
\end{aligned}
$$

Pour calculer $\mathrm{E}(XY)$, il faut calculer la probabilité de tous les couples $(X = i, Y = j)$. Par exemple, $P(X = 2, Y = 0) = \frac{1}{36}$ et $P(X = 2, Y = j) = 0$ si $j \neq 0$. On obtient alors

$$\mathrm{E}(XY) = \sum_{i,j=1}^{6} (i + j)(i - j) \cdot \frac{1}{36} = 0.$$

Ainsi, la covariance est égale à

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = 0.$$

Exercice 5 (10 points)

a) Faux. Les événements sont indépendants.

b) Vrai. Si X est l'événement *7, 23 et 45 sortent cette semaine* et Y est l'événement *7, 23 et 45 sortent la semaine prochaine*, alors

$$P(X \cap Y) = P(X)P(Y) \leq P(X).$$

c) Faux. Sinon la variance serait une notion triviale.

d) Vrai. Conséquence du théorème 3.4 des notes de cours (chapitre IV).

e) Faux. Soient a et b sont deux variables réelles. On peut facilement vérifier avec le théorème 3.4 des notes de cours que

$$\mathrm{E}(aX + b) = a\mathrm{E}(X) + b,$$

Ainsi,

$$\begin{aligned}
\mathrm{Var}(aX + b) &= \mathrm{E}((aX + b)^2) - \mathrm{E}(aX + b)^2 \\
&= \mathrm{E}(a^2X^2 + b^2 + 2abX) - (a\mathrm{E}(X) + b)^2 \\
&= a^2\mathrm{E}(X^2) + b^2 + 2ab\mathrm{E}(X) - (a^2\mathrm{E}(X)^2 + b^2 + 2ab\mathrm{E}(X)) \\
&= a^2(\mathrm{E}(X^2) - \mathrm{E}(X)^2) = a^2\mathrm{Var}(X).
\end{aligned}$$

f) Faux. Construisons un contre-exemple. Soit X le résultat du lancer d'un dé. On a déjà vu dans l'exercice que $\mathrm{Var}(X) = \frac{35}{12}$. Or, comme $\mathrm{E}(X^2) = \frac{91}{6}$ et $\mathrm{E}(X^4) = \frac{2275}{6}$, on a

$$\mathrm{Var}(X^2) = \frac{2275}{6} - \left(\frac{91}{6}\right)^2 = \frac{5369}{36}$$

qui est différent de $\left(\frac{35}{12}\right)^2 = \frac{1225}{144}$.

Exercice 6 (5 points)

a) On a

$$P(X=2\cap Y=1)=0, \text{ et}$$
$$P(X=2)P(Y=1)=\frac{1}{3}\cdot\frac{1}{3}\neq 0.$$

b)

$$\mathrm{E}(X)=-2\cdot\frac{1}{3}+0\cdot\frac{1}{3}+2\cdot\frac{1}{3}=0,$$
$$\mathrm{E}(Y)=0\cdot\frac{2}{3}+1\cdot\frac{1}{3}=\frac{1}{3},$$
$$\mathrm{E}(XY)=0.$$

c)

$$\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)=0-0=0.$$

Exercice 7 (5 points)

Soit X le nombre de carpes pêchées dans l'étang. On a, pour $i\in\{0,\dots,20\}$

$$P(X=i)=\frac{\binom{30}{i}\binom{70}{20-i}}{\binom{100}{20}}.$$

Avec cela, on peut calculer

$$\mathrm{E}(X)=\sum_{i=0}^{20} iP(X=i)=6,$$
$$\mathrm{E}(X^2)=\sum_{i=0}^{20} i^2P(X=i)=\frac{1300}{33}, \text{ et}$$
$$\mathrm{Var}(X)=\mathrm{E}(X^2)-\mathrm{E}(X)^2=\frac{112}{33}.$$

Exercice 8 (5 points)

Soit X_i la variable aléatoire représentant le nombre de boules blanches tirées dans l'urne i. Soit Y la variable aléatoire représentant le nombre de boules blanches tirées au total.
On a $Y=\sum_i X_i$, donc

$$\mathrm{E}(Y)=\sum_i \mathrm{E}(X_i)=\frac{1}{6}+\frac{1}{2}+\frac{3}{5}+\frac{1}{4}+\frac{3}{10}=\frac{109}{60}=1.81\overline{6}.$$

Exercice 9 (5 points)
Donnons un nom à chacune des personnes : la première s'appelle 1, la deuxième s'appelle 2, etc. Soient $i, j, k \in \{1, \ldots, 100\}$ trois personnes distinctes. On note $X_{\{i,j,k\}}$ la variable aléatoire valant 1 si i, j, k ont leur anniversaire le même jour (et personne d'autre qu'eux n'a son anniversaire ce jour-là), et 0 sinon.
L'espérance de $X_{\{i,j,k\}}$ est

$$P(X_{\{i,j,k\}} = 1) = \left(\frac{1}{365}\right)^3 \left(\frac{364}{365}\right)^{97}.$$

Soit Y la variable aléatoire représentant le nombre de jours où exactement trois personnes ont leur anniversaire.
Comme

$$Y = \sum_{\{i,j,k\} \subseteq \{1,\ldots,100\}} X_{\{i,j,k\}},$$

on a

$$E(Y) = \sum_{\{i,j,k\} \subseteq \{1,\ldots,100\}} E(X_{\{i,j,k\}}) = \binom{100}{3} \left(\frac{1}{365}\right)^3 \left(\frac{364}{365}\right)^{97} \approx 0.93.$$

Exercice 10 (10 points)
On sait que l'équation de la droite passant par (a, b) et (c, d) est

$$y = \frac{d-b}{c-a}(x-a) + b.$$

Il serait "moral" que le calcul de la droite de régression linéaire passant par (a, b) et (c, d) nous conduise à cette droite. Procédons :
Posons

$$X = (a, c) \quad \text{et} \quad Y = (b, d).$$

La droite de régression linéaire est alors, par un théorème du cours,

$$y = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)}(x - \mathrm{E}(X)) + \mathrm{E}(Y).$$

On a

$$\begin{aligned} \mathrm{E}(X) &= \frac{a+c}{2}, & \mathrm{E}(X^2) &= \frac{a^2+c^2}{2}, \\ \mathrm{E}(Y) &= \frac{b+d}{2}, & \mathrm{E}(Y^2) &= \frac{b^2+d^2}{2}, \\ \mathrm{E}(XY) &= \frac{ab+cd}{2}. \end{aligned}$$

Ainsi,

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{a^2+c^2}{2} - \left(\frac{a+c}{2}\right)^2 = \frac{(a-c)^2}{4},\\ \mathrm{Cov}(X,Y) &= \mathrm{E}(XY) - \mathrm{E}(X)\cdot\mathrm{E}(Y) = \frac{ab+cd}{2} - \frac{a+c}{2}\cdot\frac{b+d}{2}\\ &= \frac{(a-c)(b-d)}{4},\end{aligned}$$

et donc

$$\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} = \frac{\frac{(a-c)(b-d)}{4}}{\frac{(a-c)^2}{4}} = \frac{b-d}{a-c}.$$

L'équation de la droite de régression est alors

$$y = \frac{b-d}{a-c}\left(x - \frac{a+c}{2}\right) + \frac{b+d}{2} = \frac{b-d}{a-c}x - \frac{b-d}{a-c}\frac{a+c}{2} + \frac{b+d}{2}.$$

On vérifie facilement que

$$-\frac{b-d}{a-c}\frac{a+c}{2} + \frac{b+d}{2} = \frac{d-b}{c-a}(-a) + b,$$

ainsi, l'équation de la droite de régression est

$$y = \frac{d-b}{c-a}(x-a) + b,$$

en accord avec notre intuition de début.

Exercice 11 (5 points)

a) Si l'on fait les calculs avec $X = (1,2,3,4,5,6,7)$ et $Y = (210,200,2000,180,155,125,80)$, on obtient

$\mathrm{E}(X) = 4,$ $\mathrm{E}(Y) \cong 164,29$

$\mathrm{E}(X^2) = 20,$ $\mathrm{E}(Y^2) \cong 28935,71$

$\mathrm{Var}(X) = E(X^2) - E(X)^2 = 4,$ $\mathrm{Var}(Y) \cong 1945,92$

$\mathrm{E}(XY) = 573,57,$ $\mathrm{Cov}(X,Y) = E(XY) - E(X)\cdot E(Y) = -83.57$

$\mathrm{Corr}(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}} = -0.947$

b) Comme $|\mathrm{Corr}(X,Y)|$ est très proche de 1, une approximation linéaire peut sans doute être utilisée pour représenter Y en fonction de X.

Exercice 12 (5 points)

Posons X la variable aléatoire représentant la taille, et Y la variable aléatoire représentant le poids.

a) $\mathrm{E}(X) = \dfrac{566}{5} = 113.2$ et $E(Y) = \dfrac{203}{10} = 20.3$

b) $\mathrm{Var}(X) = \dfrac{869}{25} = 34.76$, donc $\sigma_X = \dfrac{\sqrt{869}}{5} \approx 5.896$
$\mathrm{Var}(Y) = \dfrac{761}{100} = 7.61$, donc $\sigma_Y = \dfrac{\sqrt{761}}{10} \approx 2.759$

c) $\mathrm{Corr}(X, Y) = \dfrac{732}{\sqrt{761}\sqrt{869}} \approx 0.900$

d) L'équation de la droite est alors

$$\begin{aligned} y &= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}(x - \mathrm{E}(X)) + \mathrm{E}(Y) = \frac{366\left(x - \frac{566}{5}\right)}{869} + \frac{203}{10} \\ &= \frac{732x - 47581}{1738} \approx 0.421x - 27.377. \end{aligned}$$

e) Il suffit de poser $x = 120$ dans la dernière équation :

$$y = \frac{732 \cdot 120 - 47581}{1738} = \frac{40259}{1738} \approx 23.16.$$

Exercice 13 (5 points)

On utilise le théorème 3.4 des notes de cours, avec $g(x) = ax + b$. On a

$$\begin{aligned} \mathrm{E}(aX + b) = \mathrm{E}(g(X)) &= \sum_x g(x)P(X = x) = \sum_x (ax + b)P(X = x) \\ &= \sum_x axP(X = x) + \sum_x bP(X = x) = a\mathrm{E}(X) + b. \end{aligned}$$

Exercice 14 (5 points)

$$\begin{aligned} \mathrm{Var}(X + Y) &= \mathrm{E}((X + Y)^2) - (\mathrm{E}(X + Y))^2 \\ &= \mathrm{E}(X^2 + 2XY + Y^2) - (\mathrm{E}(X) + \mathrm{E}(Y))^2 \\ &= \mathrm{E}(X^2) + 2\mathrm{E}(XY) + \mathrm{E}(Y^2) - \mathrm{E}(X)^2 - 2\mathrm{E}(X)\mathrm{E}(Y) - \mathrm{E}(Y)^2 \\ &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 + \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 + 2(\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y). \end{aligned}$$

Exercice 15 (10 points)

a) $\mathrm{E}(X_i) = \dfrac{1}{N}$ et $\mathrm{E}(X) = N \cdot \dfrac{1}{N} = 1$.

b) Comme $\mathrm{E}(X_i^2) = \dfrac{1}{N}$, on a

$$\mathrm{Var}(X_i) = \frac{1}{N} - \frac{1}{N^2} = \frac{N-1}{N^2}.$$

De plus,

$$\mathrm{Cov}(X_i, X_j) = \mathrm{E}(X_i X_j) - \mathrm{E}(X_i)\mathrm{E}(X_j) = \frac{1}{N}\frac{1}{N-1} - \frac{1}{N^2} = \frac{1}{N^2(N-1)}.$$

c) On utilise l'exercice pour, au moyen d'une récurrence, démontrer que

$$\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathrm{Var}(X_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \mathrm{Cov}(X_i, X_j).$$

Ainsi,

$$\mathrm{Var}(X) = N \cdot \frac{N-1}{N^2} + 2\binom{N}{2}\frac{1}{N^2(N-1)} = 1.$$

Exercice 16 (5 points)

On pose $Y = aX + b$. La formule du calcul de la corrélation donne

$$\begin{aligned}\rho(X,Y) = \mathrm{Corr}(X,Y) &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{a\mathrm{Cov}(X,X) + \mathrm{Cov}(X,b)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(aX+b)}} \\ &= \frac{a\mathrm{Var}(X) + 0}{\sqrt{\mathrm{Var}(X)}\sqrt{a^2\mathrm{Var}(X)}} = \frac{a}{\sqrt{a^2}} \\ &= \begin{cases} 1 & \text{si } a > 0 \\ -1 & \text{sinon.} \end{cases}\end{aligned}$$

Si $a = 0$, alors la variance de Y est nulle et, donc, la corrélation n'est pas définie.

Exercice 17 (10 points)

Partie I : Le calcul de E_1 peut se faire directement :

$$P(E_1) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = \frac{3}{10}.$$

Pour calculer E_2, notons E l'événement *la première personne est du groupe B*. On a alors, avec la formule des probabilités totales,

$$\begin{aligned} P(E_2) &= P(E_2|E)P(E) + P(E_2|E^c)P(E^c) \\ &= \frac{1}{9} \cdot \frac{2}{10} + \frac{2}{9} \cdot \frac{8}{10} = \frac{1}{5}. \end{aligned}$$

On aurait aussi pu s'aider d'un arbre pour calculer $P(E_2)$.

On voit facilement que E_3 est en fait l'événement *les trois personnes sont O ou les trois personnes sont A*, ainsi

$$P(E_3) = \frac{\binom{3}{3}}{\binom{10}{3}} + \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{1}{24}.$$

Pour calculer $P(E_4)$, on considère toutes les possibilités (sans tenir compte de l'ordre) :

$$\begin{aligned} P(E_4) &= P(O, A, B) + P(O, A, AB) + P(O, B, AB) + P(A, B, AB) \\ &= \frac{\binom{3}{1}\binom{4}{1}\binom{2}{1}}{\binom{10}{3}} + \frac{\binom{3}{1}\binom{4}{1}\binom{1}{1}}{\binom{10}{3}} + \frac{\binom{3}{1}\binom{2}{1}\binom{1}{1}}{\binom{10}{3}} + \frac{\binom{4}{1}\binom{2}{1}\binom{1}{1}}{\binom{10}{3}} = \frac{5}{12}. \end{aligned}$$

Partie II :

a) Notons G_A (resp. G_O, G_B ou G_{AB}) l'événement *la personne est du groupe A (resp. O, B ou AB)*.
   Notons aussi R_+ (resp. R_-) l'événement *la personne est Rhésus + (resp. −)*.
   Ainsi,

$$\begin{aligned} P(G_A) &= P(G_A \cap (R_+ \cup R_-)) \\ &= P((G_a \cap R_+) \cup (G_a \cap R_-)) \\ &= P(G_a \cap R_+) + P(G_a \cap R_-) \\ &= 0.4 + 0.07 = 0.47. \end{aligned}$$

   Avec la définition de la probabilité conditionnelle

$$\begin{aligned} P(G_A|R_-) &= \frac{P(G_A \cap R_-)}{P(R_-)} \\ &= \frac{P(G_A \cap R_-)}{P(R_- \cap (G_A \cup G_O \cup G_B \cup G_{AB}))} \\ &= \frac{P(G_A \cap R_-)}{P(R_- \cap G_A) + P(R_- \cap G_O) + P(R_- \cap G_B) + P(R_- \cap G_{AB})} \\ &= \frac{0.07}{0.07 + 0.06 + 0.01 + 0.01} = \frac{7}{15}. \end{aligned}$$

b) On a

$$\binom{20}{7}0.47^7(1-0.47)^{20-7} \approx 0.102.$$

La probabilité qu'il y ait cinq élèves de chaque groupe sanguin est

$$\binom{20}{5}\binom{15}{5}\binom{10}{5}\binom{5}{5} \cdot 0.47^5 \cdot 0.41^5 \cdot 0.08^5 \cdot 0.04^5 \approx 1.046 \cdot 10^{-6}.$$

c) Soit X la variable aléatoire représentant le nombre de députés du groupe A. Pour $i = 0, \ldots, 100$, on a

$$P(X = i) = \binom{100}{i}0.47^i(1-0.47)^{(100-i)}.$$

Dès lors, la probabilité cherchée est

$$\sum_{i=38}^{50}\binom{100}{i}0.47^i(1-0.47)^{(100-i)}.$$

La plupart des calculatrices de poche n'arrivent pas à évaluer correctement cette somme. Il vaut mieux effectuer le calcul avec un logiciel, par exemple *Maxima*[1].

Il suffit alors d'enter deux lignes de code :

```
P(i) := binomial(100,i) * 0.47^i * (1-0.47)^(100-i);
sum(P(i),i,38,50);
```

Maxima évalue la somme à 0.731.

1. On peut l'obtenir gratuitement sur Internet à l'adresse : http ://maxima.sourceforge.net