Corrections TD 3 : Variables discrètes

Exercice 6 : 1. On introduit la suite de variables aléatoires (X_i), $i = 1, \dots, 20$ où $X_i = 1$ si le candidat donne la bonne réponse à la $i$ème question et zéro sinon. Les questions étant indépendantes et le même nombre de réponses étant proposé à chaque fois, on en déduit que les X_i sont indépendantes et de même loi de Bernoulli de paramètre $1/k$. Donc,

$$X = \sum_{i=1}^{20} X_i$$

et par conséquent X est de loi Binomiale de paramètres $(20, 1/k)$.
2. On introduit à présent la suite de variables aléatoires (Y_i), $i = 1, \dots, 20$, où, sachant que $\{X_i = 0\}$, $Y_i = 1$ si le candidat donne la bonne réponse lors du second choix pour la $i$ème question et zéro sinon. Sachant que $\{X_i = 0\}$, les Y_i sont indépendantes et de loi de Bernoulli de paramètre $1/(k-1)$ (il ne reste en effet que $(k-1)$ réponses possibles et une seule est bonne). On a de plus

$$\begin{aligned}
\mathbb{P}(Y_i = 1) &= \mathbb{P}(Y_i = 1, X_i = 0) + \mathbb{P}(Y_i = 1, X_i = 1) \\
&= \mathbb{P}(X_i = 0)\mathbb{P}(Y_i = 1 | X_i = 0) + \mathbb{P}(X_i = 1)\mathbb{P}(Y_i = 1 | X_i = 1) \\
&= \left(1 - \frac{1}{k}\right) \times \frac{1}{k-1} + \frac{1}{k} \times 0 = \frac{1}{k} \\
\mathbb{P}(Y_i = 0) &= \mathbb{P}(Y_i = 0, X_i = 0) + \mathbb{P}(Y_i = 0, X_i = 1) \\
&= \mathbb{P}(X_i = 0)\mathbb{P}(Y_i = 0 | X_i = 0) + \mathbb{P}(X_i = 1)\mathbb{P}(Y_i = 0 | X_i = 1) \\
&= \left(1 - \frac{1}{k}\right) \times \left(1 - \frac{1}{k-1}\right) + \frac{1}{k} \times 1 \\
&= \frac{k-2}{k} + \frac{1}{k} = \frac{k-1}{k} = 1 - \frac{1}{k}.
\end{aligned}$$

Donc les variables Y_i sont indépendantes et de loi de Bernoulli de paramètre $1/k$. On en déduit donc que

$$Y = \sum_{i=1}^{20} Y_i$$

et par conséquent Y est de loi Binomiale de paramètres $(20, 1/k)$.
3. On a

$$\begin{aligned}
\mathbb{P}(X_i + Y_i = 1) &= \mathbb{P}(\{X_i = 1\} \text{ ou } \{X_i = 0, Y_i = 1\}) \\
&= \mathbb{P}(X_i = 1) + \mathbb{P}(X_i = 0, Y_i = 1) \text{ car les 2 événements sont disjoints entre eux} \\
&= \mathbb{P}(X_i = 1) + \mathbb{P}(X_i = 0)\mathbb{P}(Y_i = 1 | X_i = 0) \\
&= \frac{1}{k} + \left(1 - \frac{1}{k}\right) \times \frac{1}{k-1} = \frac{2}{k} \\
\mathbb{P}(X_i + Y_i = 0) &= \mathbb{P}(X_i = 0 \text{ et } Y_i = 0) \\
&= \mathbb{P}(X_i = 0)\mathbb{P}(Y_i = 0 | X_i = 0) \\
&= \left(1 - \frac{1}{k}\right) \times \left(1 - \frac{1}{k-1}\right) = \frac{k-2}{k} = 1 - \frac{2}{k}
\end{aligned}$$

donc X_i+Y_i est de loi Bernoulli$(2/k)$. Les variables X_i+Y_i sont indépendantes, donc $X+Y=\sum_{i=1}^{20}(X_i+Y_i)$ est de loi Binomiale$(20,2/k)$.
X et Y ne sont pas indépendantes car si c'était le cas la somme $X+Y$ serait de loi Binomiale$(40,1/k)$ d'après l'exercice 9 ci-dessous.
On peut également dire que $\mathbb{P}(X=20,Y=20)=0$ ce qui est clairement différent de $\mathbb{P}(X=20)*\mathbb{P}(Y=20)$.
4. On a $S=X+Y/2$. Donc, $\mathbb{E}(S)=\mathbb{E}(X)+\mathbb{E}(Y)/2=20/k+10/k=30/k$ car l'espérance d'une variable de loi Binomiale(n,p) est np. Donc si $k=6$, le candidat obtient en moyenne $5/20$.

Exercice 7 : Dans cet exercice, l'espace des états est $\Omega=\{1,\ldots,N\}^n$ et donc $\mathrm{card}(\Omega)=N^n$.
1. L'évènement $\{X\geq 1\}$ est l'évènement certain et donc $\mathbb{P}(X\geq 1)=1$. Calculons $\mathbb{P}(X\geq 2)$. C'est en fait la probabilité de ne jamais tirer la boule numérotée "1". Il y a $(N-1)^n$ façons pour cela donc,

$$\mathbb{P}(X\geq 2)=\left(\frac{N-1}{N}\right)^n.$$

Le même raisonnement conduit à

$$\mathbb{P}(X\geq x)=\left(\frac{N-x+1}{N}\right)^n.$$

Il y a en effet $N-x+1$ nombres entiers compris entre x et N.
Donc, pour $x=1,\ldots,N-1$,

$$\mathbb{P}(X=x)=\mathbb{P}(X\geq x)-\mathbb{P}(X\geq x+1)=\left(\frac{N-x+1}{N}\right)^n-\left(\frac{N-x}{N}\right)^n.$$

De plus, pour $\mathbb{P}(X=N)=1/N^n$ et donc la formule ci-dessous marche encore pour $x=N$.
2. Calculons la loi de Y. Il est clair que $\mathbb{P}(Y\leq 1)=\mathbb{P}(Y=1)=1/N^n$. Comme précédemment, on calcule d'abord $\mathbb{P}(Y\leq y)$ pour $y\in\{1,\ldots,N\}$. Il est facile de voir que l'évènement $\{Y\leq y\}$ correspond à l'évènement "tirer à chacun des n tirages une boule parmi celles numérotées de 1 à y". Il y a y^n possibilités pour cela et donc

$$\mathbb{P}(Y\leq y)=\left(\frac{y}{N}\right)^n.$$

Ainsi, pour $y\in\{2,\ldots,N\}$,

$$\mathbb{P}(Y=y)=\mathbb{P}(Y\leq y)-\mathbb{P}(Y\leq y-1)=\left(\frac{y}{N}\right)^n-\left(\frac{y-1}{N}\right)^n.$$

La formule ci-dessus marche encore pour $y=1$.
3. Remarquons pour commencer que si $y\leq x$ alors $\mathbb{P}(X>x,Y\leq y)=0$. Si $y\geq x+1$ alors l'évènement $\{X>x\cap Y\leq y\}$ correspond à l'évènement "tirer à chaque tirage uniquement des boules numérotées entre $x+1$ et y". Donc,

$$\mathbb{P}(X>x,Y\leq y)=\begin{cases}0 & \text{si } y\leq x,\\ (y-x)^n/N^n & \text{sinon.}\end{cases}$$

Exercice 8 : On sait que $\mathbb{P}(X=k)=Ck$ pour $k\in K:=\{1,2,3,6\}$. Pour trouver la constante C, on résoud simplement l'équation :

$$\sum_{k\in K}\mathbb{P}(X=k)=C+2C+3C+6C=1\Leftrightarrow 12C=1\Leftrightarrow C=1/12.$$

On connait donc à présent la loi de X. Ainsi,

$$\mathbb{E}(X) = \sum_{k \in K} k\, \mathbb{P}(X = k) = \frac{1}{12} + 2 \times \frac{2}{12} + 3 \times \frac{3}{12} + 6 \times \frac{6}{12} = \frac{50}{12} = \frac{25}{6}.$$

Pour calculer la variance, on calcule dans un premier temps :

$$\mathbb{E}(X^2) = \sum_{k \in K} k^2\, \mathbb{P}(X = k) = \frac{1}{12} + 4 \times \frac{2}{12} + 9 \times \frac{3}{12} + 36 \times \frac{6}{12} = \frac{252}{12}.$$

Donc

$$\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{252}{12} - \left(\frac{25}{6}\right)^2 = \frac{131}{36}.$$

Enfin, on rappelle que le coefficient de corrélation linéaire entre deux variables X et Y est :

$$\rho(X, Y) = \frac{\mathbb{C}\mathrm{ov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} \in [-1, 1].$$

Ce coefficient mesure la dépendance linéaire entre X et Y. Si $|\rho(X, Y)| \approx 1$ alors $Y \approx aX + b$. On a :

$$\begin{aligned}
\mathbb{C}\mathrm{ov}(X, (X-3)^2) &= \mathbb{E}(X(X-3)^2) - \mathbb{E}(X)\mathbb{E}[(X-3)^2] \\
&= \left\{4 \times \frac{1}{12} + 2 \times \frac{2}{12} + 54 \times \frac{6}{12}\right\} - \frac{25}{6}\left\{4 \times \frac{1}{12} + \frac{2}{12} + 9 \times \frac{6}{12}\right\} \\
&= \frac{332}{12} - \frac{25}{6} \times 5 = \frac{82}{12} = \frac{41}{6},
\end{aligned}$$

car

$$(X-3)^2 = \begin{cases} 4 & \text{avec probabilité } 1/12 \\ 1 & \text{avec probabilité } 2/12 \\ 0 & \text{avec probabilité } 3/12 \\ 9 & \text{avec probabilité } 6/12 \end{cases} \quad \text{et } X(X-3)^2 = \begin{cases} 4 & \text{avec probabilité } 1/12 \\ 2 & \text{avec probabilité } 2/12 \\ 0 & \text{avec probabilité } 3/12 \\ 54 & \text{avec probabilité } 6/12. \end{cases}$$

Il reste à calculer $\mathbb{V}[(X-3)^2] = \mathbb{E}[(X-3)^4] - \big(\mathbb{E}[(X-3)^2]\big)^2 = [\mathbb{E}(X-3)]^4 - 25$. Or,

$$\mathbb{E}[(X-3)^4] = 16 \times \frac{1}{12} + \frac{2}{12} + 81 \times \frac{6}{12} = 42.$$

En conclusion,

$$\rho(X, (X-3)^2) = \frac{41/6}{\sqrt{131/36 \times (42 - 25)}} \approx 0.8688.$$

Exercice 9 : On utilise pour cet exercice les fonctions génératrices qui caractérisent complètement les lois. On rappelle que si X et Y sont deux variables aléatoires indépendantes alors pour tout $s \in [0, 1]$, $G_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X)\mathbb{E}(s^Y) = G_X(s)G_Y(s)$.
Soit $X_1 \sim \mathcal{B}(n_1, p_1)$ et $X_2 \sim \mathcal{B}(n_2, p_2)$. Si X_1 et X_2 sont indépendantes alors,

$$G_{X_1+X_2}(s) = (p_1 s + 1 - p_1)^{n_1} \times (p_2 s + 1 - p_2)^{n_2}.$$

Donc, pour retrouver la fonction génératrice d'une loi binomiale, il faut que $p_1 = p_2 = p$. Dans ce cas,

$$G_{X_1+X_2}(s) = (ps + 1 - p)^{n_1 + n_2},$$

et donc $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$.

Exercice 10 : Pour commencer une petite précision sur l'énoncé: toutes les variables $X_1, \cdots, X_n$ sont de loi de Poisson, elles sont indépendantes et le paramètre de la Poisson pour la variable X_j est $\lambda_j, j = 1, \cdots, n$. Posons $S_n = X_1 + \ldots + X_n$ (la somme des n variables) et $S_{n,-j} = X_1 + \ldots X_{j-1} + X_{j+1} + \ldots X_n$ (la somme des $n-1$ variables, X_j ayant été enlevée). Calculons la loi de X_j sachant S_n. Soit $(k, s) \in \{0, 1, \ldots\}^2$:

$$\mathbb{P}(X_j = k | S_n = s) = \frac{\mathbb{P}(X_j = k \cap S_n = s)}{\mathbb{P}(S_n = s)} = \frac{\mathbb{P}(X_j = k \cap S_{n,-j} = s - k)}{\mathbb{P}(S_n = s)}$$

De part l'indépendance des X_i, on a :

$$\mathbb{P}(X_j = k | S_n = s) = \begin{cases} 0 & \text{si } s < k \\ \mathbb{P}(X_j = k)\mathbb{P}(S_{n,-j} = s - k)/\mathbb{P}(S_n = s) & \text{sinon.} \end{cases}$$

Or, comme les X_i sont indépendantes, on a :

$$G_{S_n}(s) = G_{X_1}(s) \times \ldots \times G_{X_n}(s) = \exp\left(\sum_{i=1}^{n} \lambda_i (s-1)\right).$$

Donc S_n suit une loi de Poisson de paramètre $\sum_{i=1}^n \lambda_i$. Un calcul tout à fait similaire donne que $S_{n,-j}$ suit une loi de Poisson de paramètre $\sum_{i=1, i\neq j}^n \lambda_i$. Ainsi,

$$\begin{aligned}
\frac{\mathbb{P}(X_j = k)\mathbb{P}(S_{n,-j} = s-k)}{\mathbb{P}(S_n = s)} &= \frac{\frac{e^{-\lambda_j}\lambda_j^k}{k!}\frac{e^{-\sum_{i=1,i\neq j}^n \lambda_i}(\sum_{i=1,i\neq j}^n \lambda_i)^{s-k}}{(s-k)!}}{\frac{e^{-\sum_{i=1}^n \lambda_i}(\sum_{i=1}^n \lambda_i)^s}{s!}} \\
&= C_s^k \frac{\lambda_j^k \left(\sum_{i=1,i\neq j}^n \lambda_i\right)^{s-k}}{\left(\sum_{i=1}^n \lambda_i\right)^s} \\
&= C_s^k \left(\frac{\lambda_j}{\sum_{i=1}^n \lambda_i}\right)^k \left(\frac{\sum_{i=1,i\neq j}^n \lambda_i}{\sum_{i=1}^n \lambda_i}\right)^{s-k} \\
&= C_s^k \left(\frac{\lambda_j}{\sum_{i=1}^n \lambda_i}\right)^k \left(1 - \frac{\lambda_j}{\sum_{i=1}^n \lambda_i}\right)^{s-k}.
\end{aligned}$$

En conclusion, X_j sachant $\{S_n = s\}$ suit une loi binomiale $\mathcal{B}(s, p)$ et donc X_j sachant S_n suit une loi binomiale $\mathcal{B}(S_n, p)$.

Exercice 11 : Notons P_i l'évènement "il pleut le $i$ème jour" et B_i l'évènement "il fait beau le $i$ème jour".
1. On a $\mathbb{P}(B_1) = \mathbb{P}(B_1|B_0)\mathbb{P}(B_0) + \mathbb{P}(B_1|P_0)\mathbb{P}(P_0) = (1/2) * 0 + (1/4) * 1 = 1/4$. Donc

$$\begin{aligned}
\mathbb{P}(P_2) &= \mathbb{P}(P_2 \cap P_1) + \mathbb{P}(P_2 \cap B_1) = \mathbb{P}(P_1)\mathbb{P}(P_2|P_1) + \mathbb{P}(B_1)\mathbb{P}(P_2|B_1) \\
&= \frac{3}{4} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{2} = \frac{11}{16},
\end{aligned}$$

car

$$\begin{aligned}
\mathbb{P}(B_{n+1}|B_n)+\mathbb{P}(P_{n+1}|B_n) &= \frac{\mathbb{P}(B_{n+1}\cap B_n)}{\mathbb{P}(B_n)}+\frac{\mathbb{P}(P_{n+1}\cap B_n)}{\mathbb{P}(B_n)}\\
&= \frac{\mathbb{P}(B_{n+1}\cap B_n)+\mathbb{P}(P_{n+1}\cap B_n)}{\mathbb{P}(B_n)}\\
&= \frac{\mathbb{P}(B_n)}{\mathbb{P}(B_n)}=1,
\end{aligned}$$

et donc $\mathbb{P}(P_{n+1}|B_n)=1-\mathbb{P}(B_{n+1}|B_n)=1/2$ et de même $\mathbb{P}(P_{n+1}|P_n)+\mathbb{P}(B_{n+1}|B_n)=1$ ce qui implique que $\mathbb{P}(P_{n+1}|P_n)=1-1/4=3/4$.

2. Il faut calculer $\mathbb{P}(P_1|P_2)$:

$$\mathbb{P}(P_1|P_2)=\frac{\mathbb{P}(P_1)\mathbb{P}(P_2|P_1)}{\mathbb{P}(P_2)}=\frac{3/4\times 3/4}{11/16}=\frac{9}{11}.$$

3. On a par la formule des probabilités totales :

$$p_{n+1}=\mathbb{P}(P_n)\mathbb{P}(P_{n+1}|P_n)+\mathbb{P}(B_n)\mathbb{P}(P_{n+1}|B_n)=p_n\times\frac{3}{4}+(1-p_n)\times\frac{1}{2}=\frac{1}{2}+\frac{1}{4}p_n.$$

On en déduit que

$$\begin{aligned}
p_n &= \frac{1}{2}+\frac{1}{4}p_{n-1}\\
&= \frac{1}{2}+\frac{1}{4}\left(\frac{1}{2}+\frac{1}{4}p_{n-2}\right)\\
&= \frac{1}{2}+\frac{1}{2}\times\frac{1}{4}+\left(\frac{1}{4}\right)^2 p_{n-2}\\
&= \frac{1}{2}+\frac{1}{2}\times\frac{1}{4}+\left(\frac{1}{4}\right)^2\left(\frac{1}{2}+\frac{1}{4}p_{n-3}\right)\\
&= \frac{1}{2}+\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\left(\frac{1}{4}\right)^2+\left(\frac{1}{4}\right)^3 p_{n-3}\\
&= \ldots
\end{aligned}$$

On peut donc montrer facilement par récurrence que :

$$p_n=\sum_{i=0}^{n-2}\frac{1}{2}\left(\frac{1}{4}\right)^i+\left(\frac{1}{4}\right)^{n-1}p_1.$$

Comme $p_1=1-\mathbb{P}(B_1)=3/4$, on a :

$$p_n=\sum_{i=0}^{n-2}\frac{1}{2}\left(\frac{1}{4}\right)^i+\frac{3}{4^n}=\frac{2}{3}\left(1-\left(\frac{1}{4}\right)^{n-1}\right)+\frac{3}{4^n},$$

car

$$\begin{aligned}
S &= \frac{1}{2}+\frac{1}{2}\times\frac{1}{4}+\ldots+\frac{1}{2}\times\left(\frac{1}{4}\right)^{n-2}\\
\frac{S}{4} &= \qquad \frac{1}{2}\times\frac{1}{4}+\ldots+\frac{1}{2}\times\left(\frac{1}{4}\right)^{n-2}+\frac{1}{2}\times\left(\frac{1}{4}\right)^{n-1}
\end{aligned}$$

et donc en faisant la différence entre les deux lignes précédentes, on trouve $S = \frac{2}{3}[1 - \left(\frac{1}{4}\right)^{n-1}]$. Donc $p_n \to 2/3$ lorsque $n \to \infty$.

Exercice 12 :
1. Cette question a déjà été traitée avec les fonctions génératrices dans l'exercice 10.
2. On rappelle que la convergence des fonctions génératrices est équivalente à la convergence en loi. La fonction génératrice de X_n est :

$$G_{X_n}(s) = (1 + p_n(s-1))^n = \exp\left(n \ln(1 + p_n(s-1))\right).$$

Comme $p_n \to 0$, $n \ln(1 + p_n(s-1)) \sim np_n(s-1) \to \lambda(s-1)$. Donc $G_{X_n}(s) \to \exp(\lambda(s-1))$ qui est la fonction génératrice d'une loi de Poisson. Par conséquent X_n converge en loi vers une loi de Poisson de paramètre λ.