Université Paris I, Panthéon - Sorbonne

Première Année Licence M.A.S.H.S. 2018 – 2019

Probabilités

Examen final, Mai 2019

Examen de 2h00. Tout document ou calculatrice est interdit.

Exercice 1 (Sur 9 points)

On note X et X' deux variables aléatoires définies sur le même espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$, suivant respectivement les lois $\mathcal{B}(p)$ et $\mathcal{B}(p')$, où $0 \leq p, p' \leq 1$.

1. Déterminer $\mathbb{E}(X)$ et $\mathbb{E}(X')$ **(0.5pts)**.
2. Quelle sont les valeurs prises par le produit $X\,X'$ **(0.5pts)**? En déduire le type de loi suivie par $X\,X'$ **(0.5pts)**.
3. Montrer que si $\mathrm{cov}(X, X') = 0$ alors on peut explicitement donner la loi de $X\,X'$ **(1pt)**. En déduire que $\mathbb{P}(X = 1 \cap X' = 1) = \mathbb{P}(X = 1)\,\mathbb{P}(X' = 1)$ **(1pt)** et $\mathbb{P}(X = 0 \cap X' = 1) = \mathbb{P}(X = 0)\,\mathbb{P}(X' = 1)$ **(1.5pts)**, puis que X et X' sont indépendantes **(1pt)**.
4. Déterminer la loi de $Y = 2X - 1$ **(0.5pts)**. Les variables Y et X sont-elles indépendantes **(1pt)**? Et Y^2 et X **(1.5pts)**?

Proof.

1. On sait que $\mathbb{E}(X) = p$ et $\mathbb{E}(X') = p'$.
2. On a $X\,X'$ produit de 0 et de 1, donc $X\,X'$ prend ses valeurs dans $\{0, 1\}$.
   On en déduit que $X\,X'$ suit une loi de Bernoulli.
3. Si $\mathrm{cov}(X, X') = 0$ alors $\mathbb{E}(X\,X') = \mathbb{E}(X)\,\mathbb{E}(X')$, donc $\mathbb{E}(X\,X') = pp'$ et ainsi $X\,X'$ suit une loi de Bernoulli $\mathcal{B}(pp')$.
   On a $\mathbb{P}(X = 1 \cap X' = 1) = \mathbb{P}(X\,X' = 1) = pp' = \mathbb{P}(X = 1)\,\mathbb{P}(X' = 1)$.
   De même $\mathbb{P}(X = 0 \cap X' = 1) + \mathbb{P}(X = 1 \cap X' = 1) = \mathbb{P}(X' = 1)$ d'après la formule des probabilités totales, donc $\mathbb{P}(X = 0 \cap X' = 1) = \mathbb{P}(X' = 1) - \mathbb{P}(X = 1 \cap X' = 1) = p' - pp' = (1-p)p' = \mathbb{P}(X = 0)\,\mathbb{P}(X' = 1)$.
   Par symétrie, $\mathbb{P}(X = 1 \cap X' = 0) = \mathbb{P}(X = 1)\,\mathbb{P}(X' = 0)$ et par la formule des probabilités totales $\mathbb{P}(X = 0 \cap X' = 0) = 1 - pp' - (1-p)p' - p(1-p') = (1-p)(1-p') = \mathbb{P}(X = 0)\,\mathbb{P}(X' = 0)$: donc dans tous les cas $\mathbb{P}(X \in A \cap X' \in A') = \mathbb{P}(X \in A)\,\mathbb{P}(X' \in A')$: les deux variables sont indépendantes.
4. On a $Y = 2X - 1$ donc Y prend ses valeurs dans $\{-1, 1\}$ et $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X' = -1) = 1 - p)$.
   Les deux variables ne sont pas indépendantes car $\mathbb{P}(X = 1 \cap Y = -1) = 0 \neq \mathbb{P}(X = 1)\,\mathbb{P}(Y = -1) = p(1-p)$.
   On a $Y^2 = 1$. Donc pour toutes fonctions g et h, $\mathbb{E}(g(X)h(Y^2)) = \mathbb{E}(g(X)h(1)) = h(1)\,\mathbb{E}(g(X)) = \mathbb{E}(g(X))\,\mathbb{E}(h(Y^2))$: les deux variables sont bien indépendantes.

□

Exercice 2 (Sur 10 points)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités.

1. Pour (E_1, E_2, E_3) trois événements de $(\Omega, \mathcal{A}, \mathbb{P})$, montrer que
$$\mathbb{P}(E_3 \cap E_2 \cap E_1) = \mathbb{P}(E_3 \mid E_2 \cap E_1)\mathbb{P}(E_2 \mid E_1)\mathbb{P}(E_1) \quad \textbf{(1pt)}.$$
2. On considère un jeu vidéo, où l'on peut chaque jour grimper d'un niveau ou rester au même niveau. On note A_j l'événement "Grimper d'un niveau le jour j", pour $j = 1, 2$ ou 3. Le jour 1 on a une probabilité $q > 1/2$ de grimper d'un niveau. Pour $j = 1$ ou 2, quand le jour j on a grimpé d'un niveau, cela fatigue et on a alors une probabilité $(1-q)$ de grimper d'un niveau le jour $j + 1$. Mais si on est resté au même niveau le jour j, on a la probabilité q de passer au niveau supérieur le lendemain. Déterminer la probabilité d'avoir gagné 3 niveaux en 3 jours **(1pt)**, et celle d'être resté au même niveau en 3 jours **(1.5pts)**.

3. Déterminer la probabilité d'avoir gagné 2 niveaux en 3 jours **(2pts)** et en déduire celle d'avoir gagné un seul niveau **(0.5pts)**.

4. On suppose que l'on a gagné 2 niveaux en 3 jours. Quelle est la probabilité que l'on ait gagné un niveau le premier jour **(2pts)**?

5. A quelle condition sur q le fait de gagner un niveau est-il indépendant du résultat de la veille **(1pt)**? Dans ce cas, déterminer sans calcul la formule de la probabilité de gagner 2 niveaux **(1pt)**.

Proof. 1. On a $\mathbb{P}(E_3 \cap E_2 \cap E_1) = \mathbb{P}(E_3 \mid E_2 \cap E_1)\,\mathbb{P}(E_2 \cap E_1)$ par la formule classique des probabilités conditionnelles. Et comme $\mathbb{P}(E_2 \cap E_1) = \mathbb{P}(E_2 \mid E_1)\mathbb{P}(E_1)$, on obtient le résultat.

2. On note X le nombre de niveaux gagnés après 3 jours. Alors $\mathbb{P}(X = 3) = \mathbb{P}(A_3 \cap A_2 \cap A_1) = \mathbb{P}(A_3 \mid A_2 \cap A_1)\,\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_1) = (1-q) \times (1-q) \times q = (1-q)^2 q$.
De la même manière on a $\mathbb{P}(X = 0) = \mathbb{P}(\overline{A_3} \cap \overline{A_2} \cap \overline{A_1}) = \mathbb{P}(\overline{A_3} \mid \overline{A_2} \cap \overline{A_1})\,\mathbb{P}(\overline{A_2} \mid \overline{A_1})\mathbb{P}(\overline{A_1}) = (1-q) \times (1-q) \times (1-q) = (1-q)^3$.

3. On cherche $\mathbb{P}(X = 2)$. Il n'y a que 3 possibilités pour gagner 2 niveaux, qui sont de ne pas avoir grimper un des 3 jours et d'avoir grimper les 2 autres. D'où

$$\begin{aligned}
\mathbb{P}(X = 2) &= \mathbb{P}(\overline{A_3} \cap A_2 \cap A_1) + \mathbb{P}(A_3 \cap \overline{A_2} \cap A_1) + \mathbb{P}(A_3 \cap A_2 \cap \overline{A_1}) \\
&= \mathbb{P}(\overline{A_3} \mid A_2 \cap A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(A_3 \mid \overline{A_2} \cap A_1)\mathbb{P}(\overline{A_2} \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(A_3 \mid A_2 \cap \overline{A_1})\mathbb{P}(A_2 \mid \overline{A_1})\mathbb{P}(\overline{A_1}) \\
&= q \times (1-q) \times q + q \times q \times q + (1-q) \times q \times (1-q) \\
&= q\,(q + (1-q)^2).
\end{aligned}$$

On en déduit par la formule des probabilités totales que

$$\mathbb{P}(X = 1) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 2) - \mathbb{P}(X = 3) = 1 - (1-q)^3 - q\,(1-q)^2 - q\,(q + (1-q)^2) = q(1-q^2).$$

4. On cherche $\mathbb{P}(A_1 \mid X = 2)$. On a

$$\begin{aligned}
\mathbb{P}(A_1 \mid X = 2) &= \frac{\mathbb{P}(A_1 \cap X = 2)}{\mathbb{P}(X = 2)} \\
&= \frac{\mathbb{P}(A_1)\,\mathbb{P}(X = 2 \mid A_1)}{\mathbb{P}(X = 2)} \\
&= \frac{\mathbb{P}(A_1)}{\mathbb{P}(X = 2)} \big(\mathbb{P}(A_3 \cap \overline{A_2} \mid A_1) + \mathbb{P}(\overline{A_3} \cap A_2 \mid A_1)\big) \\
&= \frac{\mathbb{P}(A_1)}{\mathbb{P}(X = 2)} \big(\mathbb{P}(A_3 \mid \overline{A_2})\mathbb{P}(\overline{A_2} \mid A_1) + \mathbb{P}(\overline{A_3} \mid A_2)\mathbb{P}(A_2 \mid A_1)\big) \\
&= \frac{1}{q + (1-q)^2} \big(q \times q + q \times (1-q)\big) = \frac{q}{q + (1-q)^2}.
\end{aligned}$$

5. S'il y a indépendance $\mathbb{P}(A_{j+1} \mid A_j) = 1 - q = \mathbb{P}(A_{j+1} \mid \overline{A_j}) = q$, d'où $q = 1/2$.
Dans le cas $q = 1/2$, alors on en déduit que X suit une loi binomiale $\mathcal{B}(3, 1/2)$.

□

Exercice 3 (Sur 10 points)

On s'intéresse au temps (en secondes) mis pour se connecter à un réseau, sachant qu'à partir d'un certain temps inconnu il y a déconnexion automatique. On modélise le temps de connexion T comme une variable aléatoire continue de densité

$$f_T(t) = \begin{cases} C\,\sqrt{\tau - t} & \text{si } 0 \le t \le \tau \\ 0 & \text{sinon.} \end{cases}$$

où $\tau > 0$ et C est un réel,

1. Montrer que $C = \frac{3}{2}\,\tau^{-3/2}$ **(1pt)**.

2. Montrer que $\mathbb{E}(T) = \frac{2}{5}\,\tau$ (on pourra utiliser le changement de variable $t' = \tau - t$) **(2pts)**.

3. Montrer que $\mathrm{var}(T) \le \tau^2$ **(1pt)**. On peut montrer (ne pas le faire !) que $\mathrm{var}(T) = \frac{12}{175}\,\tau^2$.

4. On observe maintenant 100 variables aléatoires indépendantes et identiquement distribuées suivant la loi de densité f_T. En notant $q \simeq 1.96$ et $\overline{T}_{100} = \frac{1}{100}\sum_{k=1}^{100} T_k$, montrer que

$$\mathbb{P}\Big(-q\sqrt{\frac{12}{175}}\,\tau \leq 10(\overline{T}_{100} - \frac{2}{5}\,\tau) \leq q\sqrt{\frac{12}{175}}\,\tau\Big) \simeq 0.95 \quad \textbf{(1.5pts)}.$$

En déduire un intervalle de confiance à 95% sur τ (on utilisera le fait que $q\sqrt{\frac{12}{175}} \simeq 1/2$) **(2.5pts)**.

5. On note par classes les différentes valeurs observées:

Temps en secondes	Effectifs
[0 , 4]	34
[4 , 8]	18
[8 , 12]	23
[12 , 16]	14
[16 , 20]	11

Tracer l'histogramme relatif à ces données **(0.5pts)**. Déterminer une approximation de la moyenne empirique **(1pt)**. En déduire l'intervalle de confiance à 95% sur τ **(0.5pts)**.

Proof. 1. On doit avoir $\int_{-\infty}^{\infty} f_X(t)\,dt = \int_0^\tau C\sqrt{\tau - t}\,dt = 1$. D'où $C\left[-\frac{2}{3}(\tau - t)^{3/2}\right]_0^\tau = C\,\frac{2}{3}\,\tau^{3/2} = 1$, soit $C = \frac{3}{2}\,\tau^{-3/2}$.

2. On a $\mathbb{E}(T) = \frac{3}{2}\,\tau^{-3/2}\int_0^\tau t\sqrt{\tau - t}\,dt = -\frac{3}{2}\,\tau^{-3/2}\int_\tau^0 (\tau - t')\sqrt{t'}\,dt'$ avec le changement de variable $t' = \tau - t$. D'où $\mathbb{E}(T) = \frac{3}{2}\,\tau^{-3/2}\left[\frac{2}{3}\,\tau(t')^{3/2} - \frac{2}{5}\,(t')^{5/2}\right]_0^\tau = \frac{2}{5}\,\tau$.

3. On a $0 \leq T \leq \tau$. Or $\mathrm{var}(T) = \mathbb{E}(T^2) - (\mathbb{E}(T))^2 \leq \mathbb{E}(T^2) \leq \tau^2$.
On a $\mathbb{E}(T^2) = \frac{3}{2}\,\tau^{-3/2}\int_0^\tau t^2\sqrt{\tau - t}\,dt = -\frac{3}{2}\,\tau^{-3/2}\int_\tau^0 (\tau - t')^2\sqrt{t'}\,dt'$ avec le changement de variable $t' = \tau - t$. Donc $\mathbb{E}(T^2) = \frac{3}{2}\,\tau^{-3/2}\left[\frac{2}{3}\,\tau^2(t')^{3/2} - \frac{4}{5}\,\tau(t')^{5/2} + \frac{2}{7}\,(t')^{7/2}\right]_0^\tau = \frac{3}{2}\left(\frac{2}{3} - \frac{4}{5} + \frac{2}{7}\right)\tau^2 = \frac{24}{105}\,\tau^2$. On en arrive ainsi à ce que $\mathrm{var}(T) = \frac{24}{105}\,\tau^2 - \frac{4}{25}\,\tau^2 = \frac{12}{175}\,\tau^2$.

4. Le théorème de la limite centrale, pour n suffisamment grand, indique que:

$$\sqrt{n}\,\frac{\left(\overline{T}_n - \frac{2}{5}\,\tau\right)}{\sqrt{\frac{12}{175}}\,\tau} \overset{\mathcal{L}}{\sim} \mathcal{N}(0,1).$$

Comme $n = 100$ est "grand", cela se transforme en $10\,\frac{\left(\overline{T}_{100} - \frac{2}{5}\,\tau\right)}{\sqrt{\frac{12}{175}}\,\tau} \overset{\mathcal{L}}{\sim} \mathcal{N}(0,1)$. Mais pour $Z \overset{\mathcal{L}}{\sim} \mathcal{N}(0,1)$, $\mathbb{P}(-q \leq Z \leq q) \simeq 0.95$, d'où le résultat.
Si $10\big(\overline{T}_{100} - \frac{2}{5}\,\tau\big) \leq q\sqrt{\frac{12}{175}}\,\tau$ alors $\frac{20}{9}\,\overline{T}_{100} \leq \tau$. De même $10\big(\overline{T}_{100} - \frac{2}{5}\,\tau\big) \geq -q\sqrt{\frac{12}{175}}\,\tau$ alors $\frac{20}{9}\,\overline{T}_{100} \leq \tau$ soit $\geq \tau$.
On obtient ainsi comme intervalle de confiance à 95% pour τ, l'intervalle:

$$\Big[\frac{20}{9}\,\overline{T}_{100}\,,\ \frac{20}{7}\,\overline{T}_{100}\Big].$$

5. Diagramme ...
Une approximation de $\overline{T}_{100}$ est: $\frac{1}{100}\,\big(2\times 34 + 6\times 18 + 10\times 23 + 14\times 14 + 18\times 11\big) = 8$.
On en déduit que l'intervalle de confiance pour τ est: $\big[160/9\,,\ 160/7\big] \simeq [17.78\,,\ 22.86]$.

□