MAP361 26 juin 2023

Examen Final (2h)

Documents autorisés : polycopié de MAP361, notes personnelles de PC et de cours, transparents imprimés, dictionnaire. Tous les appareils électroniques doivent être éteints.
Remarque : il y a 21 points dans le barème.

Exercice 1 [Poisson et convergence. (5pts)]

Soit $(\lambda_\ell)_{\ell\geq 1}$ une suite de réels strictement positifs. Soit $(R_\ell)_{\ell\geq 1}$ une suite de variables aléatoires indépendantes positives telles que, pour tout ℓ, R_ℓ suit la loi de Poisson de paramètre λ_ℓ. Pour $n \geq 1$ on pose $S_n = \sum_{\ell=1}^n R_\ell$.

L'objectif de l'exercice est d'étudier la convergence de la suite $(S_n)_{n\geq 1}$.

1. On rappelle que la fonction caractéristique de R_ℓ a l'expression suivante : $\mathbb{E}[\exp(iuR_\ell)] = \exp(\lambda_\ell(e^{iu}-1))$ pour tout réel u. Calculer pour tout $n \geq 1$ la fonction caractéristique de S_n, et en déduire la loi de S_n.
2. Justifier rapidement que S_n converge presque sûrement vers une variable aléatoire, notée S_∞, à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

À partir de maintenant on suppose que $\sum_{\ell=1}^{+\infty} \lambda_\ell < +\infty$ et on pose $\Lambda = \sum_{\ell=1}^{+\infty} \lambda_\ell$.

3. Soit $k \geq 0$ un entier, démontrer que
$$\mathbb{P}(S_\infty \leq k) = \lim_{n\to+\infty} \mathbb{P}(S_n \leq k).$$
4. Donner la loi de S_∞, et vérifier que S_∞ est finie presque-sûrement.
5. Est-ce que $(S_n)_{n\geq 1}$ converge dans L^1 ?

Solutions :

1.
$$\begin{aligned}\mathbb{E}[\exp(iuS_n)] &= \mathbb{E}[\exp(iuR_1)\dots\exp(iuR_n)] \\ &= \exp(\lambda_1(e^{iu}-1)) \times \dots \times \exp(\lambda_n(e^{iu}-1)) \\ &= \exp\left(\Sigma_{\ell=1}^n \lambda_\ell (e^{iu}-1)\right),\end{aligned}$$
qui la fonction caractéristique de la Poisson$(\sum_{\ell=1}^n \lambda_\ell)$. 1pt
2. Pour tout ω la série $\sum_{\ell=1}^n R_\ell(\omega)$ est à termes positifs donc elle converge vers $S_\infty = \sum_{\ell=1}^{+\infty} R_\ell(\omega)$, à valeurs dans $\mathbb{N} \cup \{+\infty\}$. 1pt
3. Pour k fixé, la suite d'événements $(\{S_n \leq k\})_{n\geq 1}$ est décroissante en n. Donc (Prop. 1.9 (iii) du Poly)
$$\mathbb{P}(\cap_{n\geq 1}\{S_n \leq k\}) = \lim_{n\to+\infty} \searrow \mathbb{P}(S_n \leq k).$$
Par ailleurs la suite d'entiers $(S_n) \nearrow S_\infty$ donc $\{S_\infty \leq k\} = \cap_{n\geq 1}\{S_n \leq k\}$. 1pt
4. D'après la Question 1 on a 0.5pt (calcul)
$$\begin{aligned}\mathbb{P}(S_n \leq k) &= \mathbb{P}\left(\text{Poisson}(\Sigma_{\ell=1}^n \lambda_\ell) \leq k\right) \\ &= \sum_{i=1}^k \frac{\left(\Sigma_{\ell=1}^n \lambda_\ell\right)^i \exp(-\Sigma_{\ell=1}^n \lambda_\ell)}{i!} \\ &\underset{n\to+\infty}{\to} \sum_{i=1}^k \frac{\Lambda^i \exp(-\Lambda)}{i!}\end{aligned}$$
Et ainsi S_∞ a la même fonction de répartition qu'une Poisson(Λ), et est donc finie p.s. 0.5p (finie ps

5. Si (S_n) converge dans L^1, ça ne peut être que vers S_∞ (qui est dans L^1 d'après la question précédente : $\mathbb{E}[|S_\infty|] = \mathbb{E}[S_\infty] = \Lambda$). Par monotonie la convergence a bien lieu dans L^1 : 1pt

$$\mathbb{E}\left[|S_n - S_\infty|\right] = \mathbb{E}\left[S_\infty - S_n\right] = \Lambda - \sum_{\ell=1}^{n} \lambda_\ell \to 0.$$

(Les élèves peuvent aussi justifier par convergence dominée en dominant $|S_n - S_\infty|$ par S_∞.)

Exercice 2 [Extrêmes de gaussiennes. (4pts)]

Soit $M \geq 1$ un entier et $X_1, \ldots, X_M$ des variables aléatoires suivant toutes la loi $\mathcal{N}(0,1)$. (On ne suppose pas que les X_k sont indépendantes entre elles.) L'objectif de l'exercice est de majorer $\mathbb{E}\left[\max_{1 \leq k \leq M} X_k\right]$.

1. Soit β un réel strictement positif. Justifier que $\exp(\beta X_1)$ est intégrable et démontrer que

$$\mathbb{E}[\exp(\beta X_1)] = \exp\left(\beta^2/2\right).$$

(On pourra utiliser l'identité $\beta x - \frac{x^2}{2} = \frac{\beta^2}{2} - \frac{1}{2}(x-\beta)^2$.*)*

2. On admet que $\log\left(\sum_{k=1}^M \exp(\beta X_k)\right)$ est une variable aléatoire intégrable. Justifier que

$$\mathbb{E}\left[\log\left(\sum_{k=1}^M \exp(\beta X_k)\right)\right] \leq \log\left(\mathbb{E}\left[\sum_{k=1}^M \exp(\beta X_k)\right]\right).$$

3. Vérifier que pour tous réels $x_1, \ldots, x_M$ et tout $\beta > 0$ on a

$$\beta \max_{1 \leq k \leq M} x_k \leq \log\left(\sum_{k=1}^M \exp(\beta x_k)\right).$$

4. Déduire des questions précédentes que

$$\mathbb{E}\left[\max_{1 \leq k \leq M} X_k\right] \leq \sqrt{2\log(M)}.$$

Solutions :

1. La v.a. $e^{\beta X}$ est positive donc on peut se lancer dans le calcul, on verra après pour l'intégrabilité. (0.5pt : intégrabilité) (0.5pt : expression)

$$\begin{aligned}
\mathbb{E}[\exp(\beta X_1)] &= \frac{1}{\sqrt{2\pi}} \int \exp\left(\beta x - \tfrac{x^2}{2}\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int \exp\left(\frac{\beta^2}{2} - \frac{1}{2}(x-\beta)^2\right) dx \\
&= \exp\left(\frac{\beta^2}{2}\right) \frac{1}{\sqrt{2\pi}} \int \exp\left(-\frac{1}{2}(x-\beta)^2\right) dx \\
&= \exp\left(\beta^2/2\right) \times \text{Intégrale de la densité d'une } \mathcal{N}(\beta,1) \\
&= \exp\left(\beta^2/2\right).
\end{aligned}$$

Ceci démontre en particulier l'intégrabilité.

2. On utilise l'inégalité de Jensen avec $x \mapsto \log(x)$ qui est concave. (1pt)
Remarque. Ce n'est pas hyper intéressant mais si on veut justifier l'intégrabilité de $\log\left(\sum_{k=1}^M \exp(\beta X_k)\right)$ alors il suffit d'utiliser que $|\log(x)| \leq x + 1/x$.

3. Puisque tout est positif, on a (1pt)

$$\sum_{k=1}^{M} \exp(\beta x_k) \geq \max_{1\leq k\leq M} \exp(\beta x_k),$$

$$\begin{aligned}\log\left(\sum_{k=1}^{M} \exp(\beta x_k)\right) &\geq \log\left(\max_{1\leq k\leq M} \exp(\beta x_k)\right), \\ &= \max_{1\leq k\leq M} \log\left(\exp(\beta x_k)\right) \qquad \text{(car log est croissante)} \\ &= \max_{1\leq k\leq M} \beta x_k.\end{aligned}$$

4. En combinant les questions précédentes on a que pour tout $\beta > 0$

$$\begin{aligned}\mathbb{E}\left[\max_{1\leq k\leq M} X_k\right] &\leq \frac{1}{\beta}\log\left(\mathbb{E}\left[\sum_{k=1}^{M}\exp(\beta X_k)\right]\right) \\ &\leq \frac{1}{\beta}\log\left(M\mathbb{E}\left[\exp(\beta X_1)\right]\right) \\ &\leq \frac{1}{\beta}\left(\log(M) + \frac{\beta^2}{2}\right).\end{aligned}$$

En prenant $\beta = \sqrt{2\log(M)}$ on a le résultat (c'est d'ailleurs le β qui minimise le terme de droite). (1pt)
Remarque. Cette borne est optimale car si les X_k sont indépendantes on a, lorsque $M \to +\infty$,

$$\mathbb{E}\left[\max_{1\leq k\leq M} X_k\right] \geq \sqrt{2\log(M)}(1 - o(1)).$$

Exercice 3 [Calculs de loi. (4pts)]

Soit (X, Y) un couple de variables aléatoires de densité jointe

$$f(x, y) = \frac{1}{\pi\sqrt{2}} \exp\left(-x^2 - y^2 + xy\sqrt{2}\right).$$

1. Soit $Z = Y\sqrt{2} - X$. Démontrer que X et Z sont indépendantes et suivent toutes les deux la loi $\mathcal{N}(0, 1)$.
2. En utilisant la Question 1, démontrer que

$$\mathbb{P}(X > 0 \ \cap \ Y > 0) = \frac{3}{8}.$$

3. Sans aucun calcul supplémentaire, donner la loi de Y.

Solutions :

1. **Méthode 1 : Fonction muette.** Soit $\phi : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue bornée, en posant (1.5pt)

$$\begin{cases} u = x, \\ v = y\sqrt{2} - x \end{cases}$$

qui est clairement bijectif sur $\mathbb{R}^2$ (c'est une application linéaire de déterminant $\neq 0$) et de |jacobien| égal à $\sqrt{2}$. On a donc

$$\begin{aligned}\mathbb{E}[\phi(X, Y\sqrt{2} - X)] &= \int_{(x,y)\in\mathbb{R}^2} \phi(x, y\sqrt{2} - x)\frac{1}{\pi\sqrt{2}} \exp\left(-x^2 - y^2 + xy\sqrt{2}\right) dxdy \\ &= \int_{(u,v)\in\mathbb{R}^2} \phi(u, v)\frac{1}{\pi\sqrt{2}} \exp\left(-(u^2 + v^2)/2\right) \frac{dudv}{\sqrt{2}} \\ &= \int_{(u,v)\in\mathbb{R}^2} \phi(u, v)\frac{e^{-u^2/2}}{\sqrt{2\pi}}\frac{e^{-v^2/2}}{\sqrt{2\pi}} dudv \\ &= \mathbb{E}[\phi(U, V)],\end{aligned}$$

où U, V sont deux $\mathcal{N}(0,1)$ indépendantes. On conclut avec la Proposition 6.3 du Poly.

Méthode 2 : En reconnaissant un vecteur gaussien. En appliquant la formule de la densité pour les vecteurs gaussiens (Prop. 8.17 du Poly) on voit que

$$(X,Y) \sim \mathcal{N}(\mathbf{Id}_2, \mathbf{C}), \qquad \text{avec} \qquad \mathbf{C} = \begin{pmatrix} 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1 \end{pmatrix}.$$

On écrit alors

$$\begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & \sqrt{2} \end{pmatrix} \times \begin{pmatrix} X \\ Y \end{pmatrix} =: A \times \begin{pmatrix} X \\ Y \end{pmatrix}.$$

On en déduit que (X,Z) est un vecteur gaussien de moyenne nulle et de matrice de covariance

$$A^T \mathbf{C} A = \mathbf{Id}_2.$$

2. On commence par écrire

$$\begin{aligned}
\mathbb{P}(X>0\ ;\ Y>0) &= \mathbb{P}(X>0\ ;\ \frac{1}{\sqrt{2}}(Z+X)>0) \\
&= \mathbb{P}(X>0\ ;\ Z+X>0) \\
&= \int_{x=0}^{+\infty} \left(\int_{z=-x}^{+\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \right) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx.
\end{aligned}$$

Calcul de l'intégrale : méthode directe. (1.5pt)

On pose $\psi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$ la densité de la $\mathcal{N}(0,1)$ et $\phi(x)$ sa fonction de répartition. On a alors

$$\begin{aligned}
\int_{x=0}^{+\infty} \left(\int_{z=-x}^{+\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \right) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx &= \int_{x=0}^{+\infty} \phi(x)\psi(x) dx \\
&= \int_{x=0}^{+\infty} \phi(x)\phi'(x) dx \\
&= \left[\tfrac{1}{2}\phi(x)^2\right]_{x=0}^{+\infty} \\
&= \tfrac{1}{2}\phi(+\infty)^2 - \tfrac{1}{2}\phi(0)^2 \\
&= \frac{1}{2} - \frac{1}{8} = \frac{3}{8}.
\end{aligned}$$

Calcul de l'intégrale : astuce diabolique. © M.Doumic

$$\int\int_{x>0, z>-x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dxdz = \int\int_{K \cup H} f(x,z) dxdz,$$

où

— K est le quart de plan en haut à droite,

— H est le huitième de plan $-x < z < 0$.

Puisque f est une fonction de densité invariante par rotation, on obtient que

$$\int_K f(x,z) dxdz = \frac{1}{4}, \qquad \int_H f(x,z) dxdz = \frac{1}{8},$$

d'où le résultat.

Calcul de l'intégrale : passage en polaires.

$$\begin{aligned}
\int_{x=0}^{+\infty} \left(\int_{z=-x}^{+\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \right) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx &= \int_{\theta=-\pi/4}^{\pi/2} \int_{r=0}^{+\infty} r \frac{e^{-r^2/2}}{2\pi} dr d\theta \\
&= \frac{3\pi}{4} \int_{r=0}^{+\infty} r \frac{e^{-r^2/2}}{2\pi} dr \\
&= \frac{3\pi}{4} \left[-\frac{e^{-r^2/2}}{2\pi} \right]_{r=0}^{+\infty} \\
&= \frac{3\pi}{4} \frac{1}{2\pi} = \frac{3}{8}.
\end{aligned}$$

3. On peut utiliser la symétrie de la densité f pour dire que $(X,Y) \overset{\mathcal{L}}{=} (Y,X)$, et donc en particulier $Y \overset{\mathcal{L}}{=} X$. On peut aussi dire que puisque X, Z sont indépendantes (1pt)

$$Y = \frac{1}{\sqrt{2}}(Z+X) \overset{\mathcal{L}}{=} \mathcal{N}(0, \frac{1}{\sqrt{2}^2}(1+1)) = \mathcal{N}(0,1).$$

Problème. [Mélange de lois et estimation (8pts)]

L'objectif de ce problème est d'étudier les *mélanges* de lois de probabilités. Ils sont souvent utilisés en modélisation probabiliste et en statistique.

On commence par quelques notations. Soient μ, ν deux lois de probabilités sur $\mathbb{R}$, et soit $p \in [0,1]$ un paramètre réel. On considère les variables aléatoires suivantes :

— X de loi μ ;

— Y de loi ν ;

— B de loi de Bernoulli de moyenne p.

On suppose que X, Y, B sont indépendantes et on pose $Z = BX + (1-B)Y$. Autrement dit,

$$Z = \begin{cases} X & \text{si } B = 1, \\ Y & \text{si } B = 0. \end{cases}$$

La loi de Z est appelée *mélange entre μ et ν de taux p* et est notée $p\mu \oplus (1-p)\nu$.

Partie 1 : Étude générale des mélanges

1. On suppose que X et Y (respectivement de lois μ et ν) sont toutes les deux des variables aléatoires intégrables. On note $m_X = \mathbb{E}[X]$, $m_Y = \mathbb{E}[Y]$. Soit Z une variable aléatoire de loi $p\mu \oplus (1-p)\nu$.
   Montrer que Z est intégrable et exprimer $\mathbb{E}[Z]$ en fonction de p, m_X, m_Y.
2. On suppose que les lois μ, ν admettent toutes les deux une densité. On note f_μ et f_ν ces densités, que l'on suppose continues. Démontrer que la loi $p\mu \oplus (1-p)\nu$ est également à densité et exprimer cette densité en fonction de p, f_μ, f_ν.
   (Indication : On peut utiliser la méthode de la fonction muette ou les fonctions de répartition.)

Partie 2 : Estimation pour un modèle de mélange

On considère maintenant l'expérience statistique suivante. On observe grâce à des capteurs des temps de vie d'atomes radioactifs. Il se trouve que ces capteurs sont parfois défaillants et renvoient alors une valeur aberrante dans l'intervalle $[0,1]$. Toute la difficulté du problème est que lorsque l'on observe une valeur entre 0 et 1 on ne sait pas si le capteur a fonctionné ou pas.

On modélise donc les observations par le modèle statistique suivant. Soit $((U_i, E_i, B_i))_{i \geq i}$ une suite i.i.d. de variables aléatoires où, pour tout $i \geq 1$,

— B_i suit la loi de Bernoulli de moyenne p ;

— U_i suit la loi uniforme $\mathcal{U}$ sur $[0,1]$;

— E_i suit la loi exponentielle $\mathcal{E}_\theta$ de paramètre θ, c'est-à-dire la loi de densité $x \mapsto \theta \exp(-\theta x)\mathbf{1}_{x>0}$.

— Les variables U_i, E_i, B_i sont indépendantes.

Ainsi pour tout i la variable $Z_i = B_i U_i + (1-B_i)E_i$ suit la loi du mélange $p\,\mathcal{U} \oplus (1-p)\mathcal{E}_\theta$, et les variables $(Z_i)_{i\geq 1}$ forment une suite i.i.d.

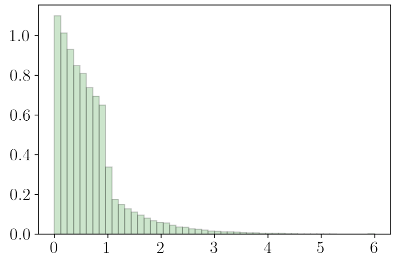

Figure : *Histogramme de* 50000 *réalisations de variables* Z_i*, pour* $p = 0.4$ *et* $\theta = 1.3$*. On voit qu'à cause des défaillances des capteurs les valeurs entre* 0 *et* 1 *sont sur-représentées.*

On suppose que les paramètres p et θ sont inconnus. L'objectif de la fin du problème est de construire, à partir d'un échantillon $(Z_1, \dots, Z_n)$, deux estimateurs pour p et θ.

3. On considère le programme `python` suivant :

```
import numpy
def SimulationMelange(p,theta):
   # Entrées : 0<p<1, theta>0
   # Sortie : simulation de la variable Z1
   Uniforme = numpy.random.rand() # simule une v.a. uniforme dans [0,1]
   if Uniforme &&&:
      return ***
   else:
      return np.random.exponential()*(1/theta) # simule une v.a. exponentielle
                                                 # de paramètre theta
```

   Par quoi faut-il remplacer `&&&` et `***` pour que cette fonction renvoie une simulation de la variable Z_1 ? (Aucune justification n'est demandée.)
4. Calculer $\mathbb{E}[Z_1]$ et $\mathbb{P}(Z_1 > 1)$.
5. Démontrer que
$$\mathbb{E}\left[(Z_1 - 1)\mathbf{1}_{Z_1>1}\right] = (1-p)\frac{\exp(-\theta)}{\theta}.$$

Pour tout $n \geq 1$ on pose
$$F_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{Z_i>1}, \qquad G_n = \frac{1}{n}\sum_{i=1}^{n}(Z_i - 1)\mathbf{1}_{Z_i>1}, \qquad \overline{Z_n} = \frac{1}{n}\sum_{i=1}^{n} Z_i.$$

6. Démontrer que F_n converge presque-sûrement vers une constante, et déterminer cette constante.
7. Démontrer que, presque-sûrement,
$$\hat{\theta}_n = \frac{F_n}{G_n + 1/n} \overset{n\to\infty}{\to} \theta.$$

Le paramètre θ est toujours inconnu mais on suppose maintenant que $0 < \theta < 2$.

8. Écrire p en fonction de θ et $\mathbb{E}[Z_1]$. En déduire un estimateur $\hat{p}_n$ qui converge presque-sûrement vers p.

Solutions :

1. On a tout d'abord (0.5pt : intégrabilité)
$$|Z| \leq |X| + |Y|,$$
qui est donc intégrable. Ensuite, (0.5pt : calcul)
$$\mathbb{E}[BX + (1-B)Y] = p\mathbb{E}[X] + (1-p)\mathbb{E}[Y] = pm_X + (1-p)m_Y.$$

2. **Méthode 1 : Fonction muette.** Soit ϕ une fonction continue bornée, on commence par remarquer la "formule magique" (1pt)

$$\phi\left(BX+(1-B)Y\right)=B\phi\left(X\right)+(1-B)\phi\left(Y\right).$$

Et ainsi (en utilisant l'indépendance)

$$\begin{aligned}\mathbb{E}\left[\phi\left(BX+(1-B)Y\right)\right]&=\mathbb{E}\left[B\phi\left(X\right)+(1-B)\phi\left(Y\right)\right]\\&=p\mathbb{E}\left[\phi\left(X\right)\right]+(1-p)\mathbb{E}\left[\phi\left(Y\right)\right]\qquad\text{(indépendance)}\\&=p\int\phi(x)f_X(x)dx+(1-p)\int\phi(y)f_Y(y)dy\\&=\int\phi(x)\left(pf_X(x)+(1-p)f_Y(x)\right)dx.\end{aligned}$$

Comme ceci est vrai pour toute fonction continue bornée ϕ, on a démontré que Z admet pour densité $x\mapsto pf_X(x)+(1-p)f_Y(x)$.

Méthode 2 : Fonction de répartition. Soit $t\in\mathbb{R}$,

$$\begin{aligned}\mathbb{P}\left(BX+(1-B)Y\leq t\right)&=\mathbb{P}\left(BX+(1-B)Y\leq t\ \cap B=1\right)+\mathbb{P}\left(BX+(1-B)Y\leq t\ \cap B=0\right),\\&=\mathbb{P}\left(X\leq t\ \cap B=1\right)+\mathbb{P}\left(Y\leq t\ \cap B=0\right),\\&=p\mathbb{P}\left(X\leq t\right)+(1-p)\mathbb{P}\left(Y\leq t\right),\qquad\text{(indépendance)}\\&=pF_X(t)+(1-p)F_Y(t).\end{aligned}$$

Comme $F_X(t)=\int_{-\infty}^t f_X(s)ds$ et $F_Y(t)=\int_{-\infty}^t f_Y(s)ds$ on trouve que le mélange est à densité $pf_X+(1-p)f_Y$.

3. Il faut remplacer `&&&` par `< p` et `***` par `numpy.random.rand()`. (Si on remplace `***` par `Uniforme` la loi n'est pas correcte, il ne faut PAS donner les points.) (1pt)

4. En utilisant la question 1 on obtient

$$\mathbb{E}[Z_1]=p\times\frac{1}{2}+(1-p)\times\frac{1}{\theta}.$$

Par ailleurs on a $\{Z_1>1\}=\{E_1>1\}\cap\{B_1=0\}$. On a donc (1pt)

$$\mathbb{P}(Z_1>1)=(1-p)\times\mathbb{P}(E_1>1)\ =(1-p)\times\int_1^{+\infty}\theta\exp(-\theta x)dx=(1-p)\exp(-\theta).$$

5. (1pt)

$$\begin{aligned}\mathbb{E}[(Z_1-1)\mathbf{1}_{Z_1>1}]&=\mathbb{E}[(E_1-1)\mathbf{1}_{E_1>1}\mathbf{1}_{B_1=0}]\\&=(1-p)\times\int_1^{+\infty}\theta(x-1)\exp(-\theta x)dx\\&=(1-p)\times\int_0^{+\infty}\theta y\exp(-\theta(y+1))dy\\&=(1-p)\exp(-\theta)\mathbb{E}[E_1]=(1-p)\exp(-\theta)\frac{1}{\theta}.\end{aligned}$$

6. La variable $\sum_{i=1}^n\mathbf{1}_{Z_i>1}$ est une somme de n Bernoulli i.i.d. de moyenne $(1-p)\exp(-\theta)$, elles sont bien entendu intégrables puisque bornées, et la loi forte des grands nombres assure que (1pt)

$$F_n\overset{n\to+\infty}{\to}(1-p)\exp(-\theta)\qquad\text{p.s..}$$

7. (**Remarque :** Le terme $+1/n$ au dénominateur est là pour garantir que l'on ne divise pas par zéro et que l'estimateur est bien défini.)
En combinant Question 5 + LFGN + Question 6 on obtient (1pt)

$$\frac{F_n}{G_n+1/n}\overset{n\to\infty}{\to}\frac{(1-p)\exp(-\theta)}{(1-p)\exp(-\theta)/\theta+0}=\theta.$$

8. On rappelle que

$$\mathbb{E}[Z_1] = p \times \frac{1}{2} + (1-p) \times \frac{1}{\theta},$$

soit

$$p = \frac{2\theta\mathbb{E}[Z_1] - 2}{\theta - 2}.$$

Ainsi on pose (1pt)

$$\hat{p}_n = \begin{cases} \dfrac{2\overline{Z_n}\hat{\theta}_n - 2}{\hat{\theta}_n - 2} & \text{si } \hat{\theta}_n \neq 2, \\ 14 & \text{si } \hat{\theta}_n = 2 \end{cases}.$$

Si $\theta < 2$ alors $\hat{\theta}_n \to \theta \neq 2$, et donc pour n assez grand on a $\hat{\theta}_n \neq 2$.

$$\hat{p}_n \overset{n\to\infty}{\to} \Big(2\big(p/2 + (1-p)/\theta\big)\theta - 2\Big)\frac{1}{\theta - 2} = p.$$

Remarque : Il y a plus intelligent ! Puisque $\mathbb{P}(Z_1 > 1) = (1-p)\exp(-\theta)$ on a $p = 1 - e^{\theta}\mathbb{P}(Z_1 > 1)$. Donc on peut prendre comme estimateur

$$\tilde{p}_n = 1 - \exp(\hat{\theta}_n)F_n,$$

qui converge p.s. vers p quelque soit la valeur de θ !