Université Lyon 1 Mathématiques
L3 Probabilités et Statistiques

Examen du 26 mars 2019
Durée : 2 heures.
Correction

Exercice 1 (8+1 points)
Soit $n \in \mathbf{N}^*$. On dit qu'une v.a. X suit la loi Gamma(n) si elle a pour densité

$$x \mapsto \frac{x^{n-1}}{(n-1)!}\exp(-x)\mathbf{1}_{[0,+\infty[}(x).$$

1. *Soit X une v.a. de loi* Gamma(1)*. Quelle autre nom porte la loi de X ? Calculer $\mathbf{E}[X]$ et $\mathbf{Var}(X)$.*

   (1 point) X est de loi Exponentielle de paramètre 1. $\mathbf{E}[X] = \int_0^\infty xe^{-x}dx = [x(-e^{-x})]_0^\infty + \int_0^\infty e^{-x}dx = 1$. $\mathbf{E}[X^2] = \int_0^\infty x^2e^{-x}dx = [x^2(-e^{-x})]_0^\infty + \int_0^\infty 2xe^{-x}dx = 2$, donc $\mathbf{Var}(X) = 2-1^2 = 1$.

2. *Soit (X_n) une suite de v.a. i.i.d. de loi* Gamma(1)*. Pour $n \geqslant 1$, on pose $S_n = X_1 + \cdots + X_n$.*

   (a) *Calculer l'espérance et la variance de S_n.*

   (1 point) $\mathbf{E}[S_n] = \sum_{i=1}^n \mathbf{E}[X_i] = n\mathbf{E}[X_1] = n$. Comme les variables sont i.i.d., $\mathbf{Var}(S_n) = n\mathbf{Var}(X_1) = n$.

   (b) *Est-ce que la suite (S_n/n) converge en probabilité ?*

   (1 point) Oui : X_1 est intégrable donc d'après la loi faible des grands nombres, $\frac{S_n}{n} \xrightarrow{\mathbb{P}}_n \mathbf{E}[X_1] = 1$.

   (c) *Montrer que S_n suit la loi* Gamma(n).

   (1,5 points) On montre le résultat par récurrence sur n. Pour $n = 1$ c'est vrai par définition. Fixons maintenant $n \in \mathbf{N}^*$ et supposons que S_n suit la loi Gamma(n). On va montrer que S_{n+1} est de loi Gamma($n+1$) en utilisant la méthode de la fonction muette. S_n et X_{n+1} sont indépendantes donc on peut écrire, pour toute fonction $f : \mathbb{R} \to \mathbb{R}_+$ mesurable,

$$\mathbf{E}[f(S_{n+1})] = \mathbf{E}[f(S_n + X_{n+1})] = \frac{1}{(n-1)!}\int_{\mathbb{R}_+^2} f(s+x)s^{n-1}e^{-s}e^{-x}dsdx$$
$$= \frac{1}{(n-1)!}\int_0^\infty s^{n-1}\left(\int_0^\infty f(s+x)e^{-(s+x)}dx\right)ds = \frac{1}{(n-1)!}\int_0^\infty s^{n-1}\left(\int_s^\infty f(y)e^{-y}dy\right)ds$$
$$= \frac{1}{(n-1)!}\int_0^\infty\int_0^\infty s^{n-1}f(y)e^{-y}\mathbf{1}_{s\leqslant y}dyds = \frac{1}{(n-1)!}\int_0^\infty f(y)e^{-y}\left(\int_0^y s^{n-1}ds\right)dy$$
$$= \frac{1}{n!}\int_0^\infty f(y)y^ne^{-y}dy.$$

   S_{n+1} est donc bien de loi Gamma($n+1$), ce qui conclut la preuve.

3. *Soit T une v.a. de loi* Gamma(n) *pour $n \in \mathbf{N}^*$.*

   (a) *Calculer la densité (notée h) de la variable aléatoire $Z = \log(T/n)$.*

   (1,5 points) On utilise encore la méthode de la fonction muette. Soit $f : \mathbb{R} \to \mathbb{R}_+$ mesurable.

$$\mathbf{E}[f(Z)] = \mathbf{E}[f(\log(T/n))] = \frac{1}{(n-1)!}\int_0^\infty f(\log(x/n))x^{n-1}e^{-x}dx$$
$$\overset{y=\log(x/n)}{=} \frac{1}{(n-1)!}\int_{-\infty}^\infty f(y)(ne^y)^{n-1}e^{-ne^y}ne^ydy$$

   On a donc pour $x \in \mathbb{R}$, $h(x) = \frac{n^n}{(n-1)!}e^{-n(e^x-x)}$.

(b) *Montrer que pour tout $x \geqslant 0$, on a $h(x) \leqslant h(-x)$.*
Indication. *On pourra utiliser l'inégalité $x \leqslant \frac{e^x - e^{-x}}{2}$ pour $x \geqslant 0$ (bonus si vous la démontrez)*

(1 point) Soit $x \geqslant 0$. On a $h(x) > 0$ donc on peut calculer $h(-x)/h(x) = e^{n(e^x - e^{-x} - 2x)}$. D'après l'indication, le terme dans l'exponentielle est positif, donc $h(-x)/h(x) \geqslant 1$.
Montrons l'indication **(1 point)**. Notons $f(x) = e^x - e^{-x} - 2x$. $f'(x) = e^x + e^{-x} - 2$, $f''(x) = e^x - e^{-x}$. $f'' \geqslant 0$ pour $x \geqslant 0$, donc f' est croissante sur $\mathbf{R}_+$, donc pour $x \geqslant 0$ $f'(x) \geqslant f'(0) = 0$. Donc f est croissante sur $\mathbf{R}_+$ et pour $x \geqslant 0$ $f(x) \geqslant f(0) = 0$.

(c) ($\star$) *En déduire que $\mathbf{P}(T \geqslant n) \leqslant 1/2$.*

(1 point) $\mathbf{P}(T \geqslant n) = \mathbf{P}(Z \geqslant 0) = \int_0^\infty h(x)dx \leqslant \int_0^\infty h(-x)dx$ d'après la question précédente. Or $\int_0^\infty h(-x)dx = \int_{-\infty}^0 h(x)dx = \mathbf{P}(Z < 0) = 1 - \mathbf{P}(Z \geqslant 0)$. Donc $\mathbf{P}(T \geqslant n) \leqslant 1 - \mathbf{P}(T \geqslant n)$ et par conséquent $\mathbf{P}(T \geqslant n) \leqslant 1/2$.

Exercice 2 (7 points)

Soit X une v.a. à valeurs dans $\mathbb{N}$. On rappelle que sa fonction génératrice est $G(s) = \sum_{k=0}^\infty \mathbf{P}(X = k)s^k$ pour $s \in [0,1]$.

1. *On suppose que X est bornée. Exprimer $\mathbf{E}[X]$ en fonction de G et montrer que*

$$\mathbf{Var}(X) = G''(1) + G'(1)(1 - G'(1)).$$

(1,5 points) Comme X est bornée, il existe $N \in \mathbb{N}^*$ tel que $\mathbf{P}(X = k) = 0$ pour tout $k \geqslant N$, donc G est un polynôme. $G'(1) = \sum_{k=0}^N k\mathbf{P}(X = k) = \mathbf{E}[X]$. $G''(1) = \sum_{k=1}^N k(k-1)\mathbf{P}(X = k) = \mathbf{E}[X^2] - \mathbf{E}[X]$, donc $\mathbf{Var}(X) = G''(1) + \mathbf{E}[X] - \mathbf{E}[X]^2 = G''(1) + G'(1)(1 - G'(1))$.

2. *Soit $p \in [0, \frac{1}{2}]$ et X une v.a. dans $\{0,1,2\}$ telle que $\mathbf{P}(X = 1) = \mathbf{P}(X = 2) = p$. On note Z_n le nombre d'individus à la génération n d'un processus de branchement où on suppose $Z_0 = 1$ et $Z_1 \sim_{\text{loi}} X$.*

(a) *Calculer $\mathbf{E}[X]$ et $\mathbf{Var}(X)$.*

(1 point) $\mathbf{E}[X] = \mathbf{P}(X = 1) + 2\mathbf{P}(X = 2) = 3p$. $\mathbf{Var}(X) = 1^2\mathbf{P}(X = 1) + 2^2\mathbf{P}(X = 2) - (3p)^2 = 5p - 9p^2$.

(b) *Quelle est la probabilité d'extinction du processus de branchement ?*

(1,5 points) La probabilité d'extinction est le plus petit point fixe de G dans $[0,1]$. Pour $s \in [0,1]$, $G(s) = 1 - 2p + ps + ps^2$. $G(s) = s \Leftrightarrow s^2 + (1 - 1/p)s + 1/p - 2 = 0 \Leftrightarrow (s-1)(s + 2 - 1/p) = 0$ (on peut utiliser le fait que 1 est racine de G pour factoriser). G a donc deux points fixes sur $\mathbf{R}$: 1 et $1/p - 2$. Comme $p \leqslant 1/2$, $1/p - 2 \geqslant 0$. Reste donc à savoir dans quel cas $1/p - 2 \leqslant 1$: c'est exactement quand $p \geqslant 1/3$. On a donc

$$\mathbf{P}(\text{extinction}) = \begin{cases} 1 & \text{si } p < 1/3 \\ 1/p - 2 & \text{si } p \geqslant 1/3. \end{cases}$$

(c) *Montrer que pour tout réel $a > \mathbf{E}[X]$, il existe $r \in [0,1[$ tel que*

$$\mathbf{P}(Z_n \geqslant a^n) \leqslant r^n.$$

(1,5 points) On sait (cf TD5) que $\mathbf{E}[Z_n] = (\mathbf{E}[X])^n$. L'inégalité de Markov implique donc pour tout $a > 0$ $\mathbf{P}(Z_n \geqslant a^n) \leqslant (\mathbf{E}[X]/a)^n$. On peut donc poser $r = \mathbf{E}[X]/a$, et si $a > \mathbf{E}[X]$ alors $r \in [0,1[$.

(d) *En déduire que $\left(Z_n^{1/n}\right)_{n \geqslant 1}$ est bornée p.s.*

(1,5 points) Remarquons d'abord que $Z_n \geqslant 0$. On veut montrer que $\mathbf{P}(\left(Z_n^{1/n}\right)_{n \geqslant 1} \text{ est bornée}) = 1$. Fixons $M > 0$ et pour $n \in \mathbb{N}^*$, considérons l'événement $A_n = \{Z_n^{1/n} \geqslant M\}$. Pour $M > \mathbf{E}[X]$, la question précédente implique que $\sum_n \mathbf{P}(A_n) < \infty$. D'après le premier lemme de Borel-Cantelli, on a donc $\mathbf{P}(\limsup_n A_n) = 0$, c'est-à-dire (en passant au complémentaire) $\mathbf{P}(\exists n \in \mathbb{N}^* \ \forall k \geqslant n, \ Z_k^{1/k} < M) = 1$. Posons $A = (\limsup_n A_n)^c$. Pour tout $\omega \in A$, on a donc l'existence de $n(\omega)$ tel que pour tout $k \geqslant n(\omega)$, $Z_k(\omega)^{1/k} < M$. D'après un résultat classique sur les suites réelles, ceci implique que $\left(Z_n(\omega)^{1/n}\right)_n$ est bornée (on peut prendre comme majorant $K(\omega) = \max(M, \max_{i=1,\dots,n(\omega)}(Z_i^{1/i}))$). Donc $A \subset \left\{\left(Z_n^{1/n}\right)_{n \geqslant 1} \text{ est bornée}\right\}$, et comme $\mathbf{P}(A) = 1$ on a aussi $\mathbf{P}(\left(Z_n^{1/n}\right)_{n \geqslant 1} \text{ est bornée}) = 1$.

Exercice 3 (6 points)

Soit $X_n, n \geqslant 1$ une suite de v.a. i.i.d. de loi exponentielle de paramètre 1. On rappelle que leur densité est $\mathbf{1}_{[0,+\infty[}(x)\exp(-x)$.

1. *Montrer que pour tout $\varepsilon \geqslant 0$,*

$$\mathbf{P}\left(X_n > (1+\varepsilon)\log n\right) = \frac{1}{n^{1+\varepsilon}}.$$

(1 point) On sait (ou on recalcule) que pour X une variable exponentielle de paramètre $\lambda > 0$ et $t \geqslant 0$, $\mathbf{P}(X \geqslant t) = e^{-\lambda t}$. On a donc $\mathbf{P}(X_n > (1+\varepsilon)\log n) = e^{-(1+\varepsilon)\log n} = n^{-(1+\varepsilon)}$.

2. *En choisissant $\varepsilon > 0$, en déduire que p.s.,*

$$\limsup_{n\to\infty} \frac{X_n}{\log n} \leqslant 1.$$

(2 points) Pour $\varepsilon > 0$, d'après la question précédente, on a $\sum_n \mathbf{P}(X_n > (1+\varepsilon)\log n) < \infty$, donc d'après le premier lemme de Borel-Cantelli, $\mathbf{P}(\limsup_n\{X_n > (1+\varepsilon)\log n\}) = 0$. En passant au complémentaire, on obtient $\mathbf{P}(\exists n \in \mathbb{N}^* \ \forall k \geqslant n,\ \frac{X_k}{\log k} \leqslant 1+\varepsilon) = 1$. Or pour tout $\omega \in \{\exists n \in \mathbb{N}^* \ \forall k \geqslant n,\ \frac{X_k}{\log k} \leqslant 1+\varepsilon\}$, $\limsup_n \frac{X_n}{\log n} \leqslant 1+\varepsilon$. Donc $\mathbf{P}(\limsup_n \frac{X_n}{\log n} \leqslant 1+\varepsilon) = 1$ pour tout $\varepsilon > 0$.
En particulier, en choisissant $\varepsilon = 1/m$ avec $m \in \mathbb{N}^*$, on obtient $\mathbf{P}(\forall m \in \mathbb{N}^* \ \limsup_n \frac{X_n}{\log n} \leqslant 1 + 1/m) = 1$, comme probabilité d'une intersection dénombrable d'événements de proba 1. D'où $\mathbf{P}(\limsup_n \frac{X_n}{\log n} \leqslant 1) = 1$.

3. *En utilisant les questions précédentes, montrer que*

$$\limsup_{n\to\infty} \frac{X_n}{\log n} = 1.$$

(1 point) D'après la question 1, $\sum_n \mathbf{P}(X_n > \log n) = +\infty$. De plus, les événements $\{X_n > \log n\}$ sont indépendants car les X_n le sont, donc d'après le second lemme de Borel-Cantelli, $\mathbf{P}(\limsup_n\{X_n > \log n\}) = 1$. Or par définition pour $\omega \in \limsup_n\{X_n > \log n\}$, pour tout $n \in \mathbb{N}^*$ il existe $k \geqslant n$ tel que $X_k(\omega)/\log k > 1$, et donc $\limsup_n X_n/\log n \geqslant 1$. Donc $\mathbf{P}(\limsup_n X_n/\log n \geqslant 1) \geqslant \mathbf{P}(\limsup_n\{X_n > \log n\}) = 1$. Donc $\mathbf{P}(\limsup_n X_n/\log n = 1) = \mathbf{P}(\{\limsup_n X_n/\log n \geqslant 1\} \cap \{\limsup_n X_n/\log n \leqslant 1\}) = 1$ en utilisant le dernier résultat et la question 2.

4. *On pose $Z_n = \frac{\max(X_1,\cdots,X_n)}{\log n}$ pour tout $n \geqslant 2$, montrer que*

$$\liminf_{n\to\infty} Z_n \geqslant 1, \text{ p.s.}$$

(2 points) Soit $\varepsilon > 0$. Calculons $\mathbf{P}(Z_n \leqslant (1-\varepsilon)\log n) = \mathbf{P}(\forall i = 1,\ldots,n,\ X_i \leqslant (1-\varepsilon)\log n) = (1 - e^{-(1-\varepsilon)\log n})^n$ en utilisant le fait que les X_i sont i.i.d. Donc $\mathbf{P}(Z_n/\log n \leqslant 1-\varepsilon) = (1 - 1/n^{1-\varepsilon})^n \sim e^{-n^\varepsilon}$. On a donc $\sum_n \mathbf{P}(Z_n/\log n \leqslant 1-\varepsilon) < \infty$ et d'après le premier lemme de Borel-Cantelli, $\mathbf{P}(\limsup_n\{Z_n/\log n \leqslant 1-\varepsilon\}) = 0$. En passant au complémentaire, on a donc $\mathbf{P}(\exists n \in \mathbb{N}^* \ \forall k \geqslant n Z_n/\log n > 1-\varepsilon) = 1$. Or si une suite est supérieure à $1-\varepsilon$ à partir d'un certain rang, c'est le cas aussi de sa limite inférieure, donc $\mathbf{P}(\liminf_n Z_n/\log n \geqslant 1-\varepsilon) = 1$. En choisissant $\varepsilon = 1/m$ et en prenant l'intersection sur $m \in \mathbb{N}^*$, on obtient $1 = \mathbf{P}(\cap_{m\in\mathbb{N}^*}\{\liminf_n Z_n/\log n \geqslant 1 - 1/m\}) = \mathbf{P}(\liminf_n Z_n/\log n \geqslant 1)$.