Examen du cours de Probabilités

Mardi 7 Janvier 2020 (8h30-11h30)

Polycopié et notes de cours autorisés. Tout objet électronique est interdit.

Barême indicatif. L'ensemble de l'examen sera noté sur un barême entre 25 et 30 points, et les points seront approximativement répartis de la façon suivante : 1/3 pour l'exercice 1, 1/6 pour l'exercice 2 et 1/2 pour le problème. Il n'est donc pas nécessaire de tout faire pour avoir 20.

Exercice 1. Soit $(X_i)_{i\geq 1}$ une suite i.i.d. de variables aléatoires suivant la loi uniforme sur $[-1,1]$.

1. Montrer que e^{X_1} est intégrable et de carré intégrable. Calculer explicitement $m = \mathbb{E}[e^{X_1}]$ et $\sigma^2 = \mathbf{Var}[e^{X_1}]$ en fonction de la constante e.
2. Calculer $\tilde{\sigma}^2 = \mathbf{Var}[e^{X_1} - X_1]$ et montrer que $\tilde{\sigma} < \sigma$.

On cherche à calculer m par la méthode de Monte-Carlo. On pose $M_n = \frac{1}{n}\sum_{i=1}^n e^{X_i}$ et $\tilde{M}_n = \frac{1}{n}\sum_{i=1}^n e^{X_i} - X_i$.

3. Montrer que M_n converge vers m en un sens à préciser, puis donner le comportement asymptotique de $\sqrt{n}(M_n - m)$.
4. Construire, lorsque n est grand, un intervalle de confiance à 95% pour m, exprimé à l'aide de M_n et σ.
5. Donner le comportement asymptotique de $\tilde{M}_n$ puis de $\sqrt{n}(\tilde{M}_n - m)$. Construire un intervalle de confiance à 95% pour m, exprimé à l'aide de $\tilde{M}_n$ et $\tilde{\sigma}$. Pour une valeur de n donnée, cet intervalle est-il plus précis que le précédent ?
6. On pose $\tilde{V}_n = \frac{1}{n}\sum_{i=1}^n (e^{X_i} - X_i)^2 - \tilde{M}_n^2$. Donner le comportement asymptotique de $\tilde{V}_n$, puis de $\sqrt{\frac{n}{\tilde{V}_n}}(\tilde{M}_n - m)$. Construire un intervalle de confiance à 95% pour m, exprimé à l'aide de $\tilde{M}_n$ et $\tilde{V}_n$.
7. Comment s'appelle la méthode de réduction de variance utilisée dans cette exercice ? Cette méthode pourrait elle donner un estimateur de m meilleur que M_n ou $\tilde{M}_n$?

⋆⋆⋆

Exercice 2.

Soit X une variable aléatoire positive à densité $p : \mathbb{R} \to \mathbb{R}$ et $V \sim \beta(2,2)$ une variable aléatoire de densité $\mathbf{1}_{[0,1]}(x)6x(1-x)$, indépendante de X. On pose $Y = VX$ et $Z = (1-V)X$.

1. Calculer la loi de $1 - V$. Montrer alors que Y et Z ont même loi.
2. Calculer la loi de Y (montrer qu'il s'agit d'une loi à densité, et exprimer la densité de Y en fonction de p).
3. Calculer la loi de (Y, Z), puis retrouver la loi obtenue à la question précédente. Préciser la loi de (Y, Z) lorsque $X \sim \Gamma(4, \theta)$.

⋆⋆⋆

Problème. (Processus des records) On considère $(X_i)_{i\geq 0}$ une suite i.i.d. de variables aléatoires réelles, et on note F la fonction de répartition de X_1. On s'intéresse au processus des records $(R_n, n \geq 0)$ et des instants de records $(L_n, n \geq 0)$ définis par :

$$L_0 = 0,\ R_0 = X_0,\ L_n = \inf\{j > L_{n-1} : X_j > R_{n-1}\},\ R_n = X_{L_n}.$$

Ainsi, R_n est la valeur du $(n+1)$-ème record observé et $L_n + 1$ est le nombre de réalisations qu'il a fallu pour atteindre ce record. L'objectif de cette partie est d'étudier la loi de R_n et son comportement lorsque $n \to +\infty$.

Première partie. Dans toute cette partie, on suppose $X_i \sim \mathcal{U}([0,1])$. Soient $U \sim \mathcal{U}([0,1])$ une variable aléatoire indépendante de Z qui suit une loi de densité $q(z)$ telle que $q(z) > 0$ si $z \in]0,1[$ et $q(z) = 0$ si $z \notin]0,1[$. On suppose Z indépendante de $(X_i)_{i\geq 1}$.

1. Soit $x \in]0,1[$. Donner la loi de $V(x) = x + (1-x)U$.
2. Calculer la loi jointe de $(Z, V(Z))$, puis calculer exprimer la densité q' de la loi de $Z' = V(Z)$ à l'aide de la fonction q. Z et Z' sont elles indépendantes ?
3. Pour $x \in]0,1[$, on définit $T(x) = \inf\{j \in \mathbb{N}^*, X_j > x\}$. Donner sans calcul la loi de $T(x)$. Pour $f : \mathbb{R}^2 \to \mathbb{R}$ bornée mesurable, montrer que
$$f(Z, X_{T(Z)}) = \mathbf{1}_{X_1 > Z} f(Z, X_1) + \sum_{n=2}^{\infty} \mathbf{1}_{X_1 \leq Z, \ldots, X_{n-1} \leq Z, X_n > Z} f(Z, X_n).$$
En déduire que $(Z, X_{T(Z)})$ suit la même loi que $(Z, V(Z))$.
4. On admet que R_n et $(X_{L_n+j}, j \geq 1)$ sont indépendantes. En déduire que (R_n, R_{n+1}) et $(R_n, V(R_n))$ ont même loi. En déduire que R_n suit la loi de densité q_n, où
$$q_n(z) = \frac{(-\log(1-z))^n}{n!}, \; z \in]0,1[, \text{ et } q_n(z) = 0 \text{ si } z \notin]0,1[. \tag{1}$$

Deuxième partie. On suppose toujours $X_i \sim \mathcal{U}([0,1])$, et on s'intéresse au comportement asymptotique de la suite $(R_n)_{n\geq 1}$. *Cette partie et la suivante sont indépendantes de la première partie, on aura uniquement à utiliser la formule* (1) *que l'on pourra admettre.*

5. Montrer que pour $x \in]0,1[$, $\mathbb{P}(R_n \leq x) \underset{n\to+\infty}{\to} 0$.
6. En déduire que R_n converge en probabilité vers une limite à préciser, puis montrer que R_n converge presque sûrement vers cette limite.
7. Calculer la loi de $-\log(1-R_n)$, et reconnaître cette loi. En déduire que $-\log(1-R_n)$ a même loi que $E_1 + \cdots + E_{n+1}$, où $(E_i)_{i\geq 1}$ est une suite i.i.d. de variables aléatoires exponentielles de paramètre 1.
8. A l'aide de la question précédente, montrer que $\frac{-\log(1-R_n)}{n}$ converge en probabilité vers 1, et donner le comportement asymptotique de $\sqrt{n}(\frac{-\log(1-R_n)}{n} - 1)$.

Troisième partie. On suppose désormais que X_i suit une loi de densité p continue sur $[a,b]$ (avec $a < b$) et telle que $p(z) > 0$ pour $z \in [a,b]$ et $p(z) = 0$ pour $z \notin [a,b]$. On note F la fonction de répartition de X_i.

9. Montrer que $F : [a,b] \to [0,1]$ est strictement croissante et bijective. (On pourra utiliser ensuite $F^{-1} : [0,1] \to [a,b]$ la fonction réciproque). Exprimer $F(R_n)$ et L_n à l'aide de $\tilde{R}_n$ et $\tilde{L}_n$, où $\tilde{L}_0 = 0$, $\tilde{R}_0 = F(X_0)$, $\tilde{L}_n = \inf\{j > \tilde{L}_{n-1} : F(X_j) > R_{n-1}\}$, $\tilde{R}_n = X_{\tilde{L}_n}$.
10. Calculer la loi de $F(X_i)$. En déduire la loi de $F(R_n)$ puis calculer la densité de R_n.
11. Justifier que pour n grand, $\mathbb{P}(\sqrt{n}(\frac{-\log(1-F(R_n))}{n} - 1) \geq -1,96) \approx 97,5\%$.
Application : le record du monde de saut en longueur féminin (actuellement de 7,52m et détenu par Galina Chistyakova depuis le 11 juin 1988) a été amélioré 36 fois depuis que l'IAAF enregistre les records d'athlétisme (source : Wikipedia). On donne $\exp(-36 + 1,96 \times 6) \approx 3 \times 10^{-11}$. Donner, avec une confiance de $97,5\%$ un majorant de la probabilité que le prochain saut lors d'une compétition officielle soit un record du monde ? Commenter.

Corrigé

EXERCICE 1.

1. La variable aléatoire e^{X_1} est bornée ($|e^{X_i}| \leq e$), elle est donc intégrable et de carré intégrable. On a $m = \mathbb{E}[e^{X_1}] = \frac{1}{2}\int_{-1}^{1} e^x dx = \frac{1}{2}(e - e^{-1})$ et $\mathbb{E}[e^{2X_1}] = \frac{1}{2}\int_{-1}^{1} e^{2x} dx = \frac{1}{4}(e^2 - e^{-2})$, ce qui donne $\sigma^2 = \frac{1}{4}(e^2 - e^{-2} - e^2 - e^{-2} + 2) = \frac{1}{2}(1 - e^{-2})$.
2. On a $\mathbb{E}[e^{X_1} - X_1] = m - \mathbb{E}[X_1] = m$ et
$$\begin{aligned}\mathbb{E}[(e^{X_1} - X_1)^2] &= \mathbb{E}[e^{2X_1}] - 2\mathbb{E}[X_1 e^{X_1}] + \mathbb{E}[X_1^2] \\ &= \frac{1}{4}(e^2 - e^{-2}) - \int_{-1}^{1} x e^x dx + \frac{1}{2}\int_{-1}^{1} x^2 dx \\ &= \frac{1}{4}(e^2 - e^{-2}) - 2e^{-1} + \frac{1}{3},\end{aligned}$$
ce qui donne $\tilde{\sigma}^2 = \frac{1}{2}(1 - e^{-2}) - 2e^{-1} + \frac{1}{3}$. On a clairement $\frac{1}{3} - 2e^{-1} < 0$ et donc $\tilde{\sigma} < \sigma$.
3. Comme e^{X_1} est intégrable, on peut appliquer la loi forte des grands nombres et ainsi M_n converge presque sûrement vers m. Comme e^{X_1} est de carré intégrable, le théorème de la limite centrale assure que $\sqrt{n}(M_n - m)$ converge en loi vers $\mathcal{N}(0, \sigma^2)$.
4. On a $\mathbb{P}(|\sqrt{n}(M_n - m)| \leq 1,96\sigma) \approx 95/100$ et ainsi $[M_n - \frac{1{,}96}{\sqrt{n}}\sigma, M_n + \frac{1{,}96}{\sqrt{n}}\sigma]$ est un intervalle de confiance à 95% pour m.
5. Par la loi forte des grands nombres, M_n converge presque sûrement vers $\mathbb{E}[e^{X_1} - X_1] = m$ et le théorème de la limite centrale donne la convergence en loi de $\sqrt{n}(\tilde{M}_n - m)$ vers $\mathcal{N}(0, \tilde{\sigma}^2)$. Ainsi,$[M_n - \frac{1{,}96}{\sqrt{n}}\tilde{\sigma}, M_n + \frac{1{,}96}{\sqrt{n}}\tilde{\sigma}]$ est un intervalle de confiance à 95% pour m. Comme $\tilde{\sigma} < \sigma$, cet intervalle a une largeur plus petite et est donc plus précis que l'intervalle de confiance obtenu à la question précédente.
6. La variable aléatoire $(e^{X_1} - X_1)^2$ et donc $\frac{1}{n}\sum_{i=1}^{n}(e^{X_i} - X_i)^2 \underset{n\to+\infty}{\to} \mathbb{E}[(e^{X_1} - X_1)^2]$ p.s., par la loi forte des grands nombres. Comme $\tilde{M}_n \underset{n\to+\infty}{\to} m$ p.s., il vient que $\tilde{V}_n \underset{n\to+\infty}{\to} \tilde{\sigma}^2$ p.s. Ensuite, le théorème de Slutsky assure la convergence en loi de $(\tilde{V}_n, \sqrt{n}(\tilde{M}_n - m))$ vers $(\tilde{\sigma}^2, \tilde{G})$, où $\tilde{G} \sim \mathcal{N}(0, \tilde{\sigma}^2)$. Enfin, la continuité de $(x, y) \in \mathbb{R}_+^* \times \mathbb{R} \mapsto y/\sqrt{x}$ donne la convergence en loi de $\sqrt{\frac{n}{\tilde{V}_n}}(\tilde{M}_n - m)$ vers $\mathcal{N}(0, 1)$. Ainsi, $[M_n - \frac{1{,}96}{\sqrt{n}}\sqrt{\tilde{V}_n}, M_n + \frac{1{,}96}{\sqrt{n}}\sqrt{\tilde{V}_n}]$ est un intervalle de confiance à 95% pour m.
7. La méthode de réduction de variance utilisée est la technique de variable de contrôle. On pourrait obtenir un meilleur estimateur de m que M_n ou $\tilde{M}_n$ en prenant $M_n^\lambda = \frac{1}{n}\sum_{i=1}^{n} e^{X_i} + \lambda X_i$, et en optimisant $\lambda \in \mathbb{R}$. M_n correspond à $\lambda = 0$ et $\tilde{M}_n$ à $\lambda = -1$. Précisément, la loi forte des grands nombre donne $M_n^\lambda \underset{n\to+\infty}{\to} m$ p.s., et le théorème de la limite centrale donne la convergence en loi de $\sqrt{n}(M_n^\lambda - m)$ vers $\mathcal{N}(0, \sigma_\lambda^2)$, où $\sigma_\lambda^2 = \mathbf{Var}(e^{X_1} + \lambda X_1) = \sigma^2 + 2\lambda \mathbf{Cov}(e^{X_1}, X_1) + \lambda^2 \mathbf{Var}(X_1) = \sigma^2 + 2\lambda e^{-1} + \frac{1}{3}\lambda^2$. Le choix de λ qui minimise la variance est $-\frac{\mathbf{Cov}(e^{X_1}, X_1)}{\mathbf{Var}(X_1)} = -3e^{-1}$. Dans les cas où l'on ne sait pas calculer explicitement le λ optimal, il faut également l'estimer. Ici $\hat{\lambda}_n = -\frac{\frac{1}{n}\sum_{i=1}^{n} e^{X_1} X_1}{\frac{1}{n}\sum_{i=1}^{n} X_1^2}$ converge p.s. vers $-\frac{\mathbf{Cov}(e^{X_1}, X_1)}{\mathbf{Var}(X_1)}$

⋆ ⋆ ⋆

Exercice 2.

1. La variable aléatoire $1-V$ suit également une loi $\beta(2,2)$ et est indépendante de X. Ainsi, $(1-V,X)$ a même loi que (V,X) et donc $(1-V)X$ suit la même loi que VX.
2. Par indépendance, le vecteur aléatoire (V,X) a pour densité $\mathbf{1}_{[0,1]}(v)6v(1-v)p(x)$. Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction bornée mesurable. On a, grâce au théorème de Fubini,

$$\begin{aligned}\mathbb{E}[f(VX)] &= \int_0^\infty \left(\int_0^1 f(vx)6v(1-v)du\right)p(x)dx = \int_0^\infty \int_0^x f(y)\frac{6y(x-y)}{x^3}dyp(x)dx \\ &= \int_0^\infty \int_0^\infty \mathbf{1}_{y<x}\frac{6y(x-y)}{x^3}f(y)\frac{p(x)}{x}dxdy = \int_0^\infty f(y)\left(\int_y^\infty \frac{6y(x-y)p(x)}{x^3}dx\right)dy.\end{aligned}$$

   Donc Y suit la loi de densité $\mathbf{1}_{y>0}\int_y^\infty \frac{6y(x-y)p(x)}{x^3}dx$.
3. On pose $\phi(v,x)=(vx,(1-v)x)$. C'est un C^1-difféomorphisme de $]0,1[\times\mathbb{R}_+^*$ dans $\mathbb{R}_+^*\times\mathbb{R}_+^*$, de réciproque $\phi^{-1}(y,z)=(\frac{y}{y+z},y+z)$. On vérifie en effet facilement que $\phi(]0,1[\times\mathbb{R}_+^*)\subset\mathbb{R}_+^*\times\mathbb{R}_+^*$ et $\phi^{-1}(\mathbb{R}_+^*\times\mathbb{R}_+^*)\subset]0,1[\times\mathbb{R}_+^*$. On a

$$\mathrm{Jac}\phi^{-1}(y,z) = \det\begin{pmatrix}\frac{z}{(y+z)^2} & -\frac{y}{(y+z)^2} \\ 1 & 1\end{pmatrix} = \frac{1}{y+z}.$$

   Soit $f:\mathbb{R}^2\to\mathbb{R}$ bornée mesurable. Grâce au théorème de changement de variable, on obtient

$$\begin{aligned}\mathbb{E}[f(Y,Z)] &= \int_0^\infty\int_0^1 f(\phi(v,x))p(x)6v(1-v)dvdx \\ &= \int_0^\infty\int_0^\infty f(y,z)p(y+z)\frac{6yz}{(y+z)^3}dydz.\end{aligned}$$

   Ainsi, (Y,Z) suit la loi de densité $\frac{6yzp(y+z)}{(y+z)^3}\mathbf{1}_{y>0}\mathbf{1}_{z>0}$. En appliquant la formule des densités marginales, on retrouve que Y suit la loi de densité $\mathbf{1}_{y>0}\int_0^\infty \frac{6yzp(y+z)}{(y+z)^3}dz = \mathbf{1}_{y>0}\int_y^\infty \frac{6y(x-y)p(x)}{x^3}dx$.
   Pour $X\sim\Gamma(4,\theta)$, on a $p(x)=\frac{\theta^4}{6}x^3e^{-\theta x}$ et donc (Y,Z) suit la loi de densité $\theta^4\left(ye^{-\theta y}\mathbf{1}_{y>0}\right)\times\left(ze^{-\theta z}\mathbf{1}_{z>0}\right)$. Ainsi, Y et Z sont indépendantes et suivent la loi $\Gamma(2,\theta)$. On retrouve ainsi un résultat du cours.

⋆⋆⋆

Problème.

1. $V(x)$ suit la loi $\mathcal{U}([x,1])$.
2. On applique la méthode de la fonction muette. Soit $f:\mathbb{R}^2\to\mathbb{R}$ bornée mesurable. On a

$$\begin{aligned}\mathbb{E}[f(Z,V(Z))] &= \int_0^1\left(\int_0^1 f(z,z+(1-z)u)du\right)q(z)dz \\ &= \int_0^1\left(\int_0^1 f(z,z+(1-z)u)du\right)q(z)dz \\ &= \int_0^1\left(\int_z^1 f(z,z')\frac{1}{1-z}dz'\right)q(z)dz = \int_0^1\int_0^1 f(z,z')\mathbf{1}_{z<z'}\frac{q(z)}{1-z}dzdz'\end{aligned}$$

Ainsi (Z,Z') suit la loi de densité $\mathbf{1}_{0<z<z'<1}\frac{q(z)}{1-z}$. Elle ne s'écrit pas sous forme produit donc Z et Z' ne sont pas indépendantes. La formule de la densité marginale assure

$$q'(z') = \int_0^1 \mathbf{1}_{z<z'}\frac{q(z)}{1-z}dz = \int_0^{z'} \frac{q(z)}{1-z}dz,\ z' \in [0,1].$$

3. $T(x)$ est l'instant de premier succès et suit une loi géométrique de paramètre $1-x$. Par la méthode du rejet, $X_{T(x)} \sim \mathcal{U}([x,1])$. On décompose ensuite selon les événements disjoints :

$$\{T(Z)=n\} = \{Z < X_1\} \cup (\cup_{n\geq 1}\{X_1 \leq Z\} \cap \cdots \cap \{X_n \leq Z\} \cap \{Z < X_n\}),$$

ce qui donne

$$f(Z, X_{T(Z)}) = \mathbf{1}_{X_1>Z}f(Z,X_1) + \sum_{n=2}^{\infty} \mathbf{1}_{X_1\leq Z,\ldots,X_{n-1}\leq Z, X_n>Z}f(Z,X_n).$$

Par indépendance des X_i et de Z, il vient $\mathbb{E}[\mathbf{1}_{X_1>Z}f(Z,X_1)] = \int_0^1\int_0^1 f(z,u)\mathbf{1}_{u>z}q(z)dudz$ et $\mathbb{E}[\mathbf{1}_{X_1\leq Z,\ldots,X_{n-1}\leq Z,X_n>Z}f(Z,X_n)] = \int_0^1 \ldots \int_0^1 f(z,u_n)\mathbf{1}_{u_n>z}q(z)\mathbf{1}_{u_1\leq z,\ldots,u_{n-1}\leq z,u_n>z}du_1\ldots du_n dz = \int_0^1\int_0^1 f(z,u_n)\mathbf{1}_{u_n>z}q(z)z^{n-1}du_n dz$. Ainsi, on obtient par le théorème de Fubini

$$\mathbb{E}[f(Z,X_{T(Z)})] = \int_0^1\int_0^1 f(z,u)\mathbf{1}_{u>z}q(z)\sum_{n=1}^{\infty} z^{n-1}dudz = \int_0^1\int_0^1 f(z,u)\mathbf{1}_{u>z}\frac{q(z)}{1-z}dudz = \mathbb{E}[f(Z,V(Z))].$$

Cela prouve l'égalité en loi par le théorème de la fonction muette.

4. On a $L_{n+1} = \inf\{j > L_n : X_j > R_n\} = L_n + \inf\{j \geq 1 : X_{L_n+j} > R_n\}$ et $R_{n+1} = X_{L_{n+1}}$. On est donc exactement dans la situation de la question précédente, avec $Z = R_n$, ce qui donne l'égalité en loi. On procède ensuite par récurrence. On a clairement $q_0(z) = \mathbf{1}_{[0,1]}(z)$. Ce qui prouve la formule pour $n = 0$. Ensuite, on a par la question 2 $q_{n+1}(z') = \int_0^{z'}\frac{q_n(z)}{1-z}dz = \int_0^{z'}\frac{(-\log(1-z))^n}{n!(1-z)}dz = \frac{(-\log(1-z'))^{n+1}}{(n+1)!}$.

5. Pour $x \in]0,1[$, $\mathbb{P}(R_n \leq x) = \int_0^x q_n(z)dz \leq xq_n(x)$ par croissance de la fonction q_n. Or, $q_n(x) \underset{n\to+\infty}{\to} 0$ ce qui donne le résultat. Comme $\mathbb{P}(R_n \leq 1) = 1$, on en déduit que $R_n \to 1$ en probabilité.

6. On a par définition $R_n < R_{n+1}$, et donc R_n converge p.s. car toute suite croissante de réels converge. Comme la convergence p.s. entraîne celle en probabilité, il vient que R_n converge p.s. vers 1.

7. Soit $g : \mathbb{R} \to \mathbb{R}$ bornée mesurable. On a

$$\begin{aligned}\mathbb{E}[g(-\log(1-R_n))] &= \int_0^1 g(-\log(1-r))\frac{(-\log(1-r))^n}{n!}dr \\ &= \int_0^{\infty} g(x)\frac{x^n}{n!}e^{-x}dx.\end{aligned}$$

Il s'agit de la loi $\Gamma(n+1,1)$. La loi exponentielle est une loi $\Gamma(1,1)$, et un théorème du cours donne la loi de la somme de deux lois gamma indépendantes de même paramètre θ : ce théorème assure que $E_1 + \cdots + E_{n+1} \sim \Gamma(n+1,1)$.

8. L'égalité en loi donne $\mathbb{P}(|\frac{-\log(1-R_n)}{n}-1|\le\varepsilon)=\mathbb{P}(|\frac{\sum_{i=1}^{n+1}E_i}{n}-1|\le\varepsilon)$. La LFGN assure que $\frac{\sum_{i=1}^{n+1}E_i}{n}$ converge presque sûrement vers 1 et donc en probabilité. Ainsi, $\mathbb{P}(|\frac{-\log(1-R_n)}{n}-1|\le\varepsilon)\underset{n\to+\infty}{\to}1$.
   Soit $f:\mathbb{R}\to\mathbb{R}$ continue bornée. On a

$$\mathbb{E}\left[f\left(\sqrt{n}\left(\frac{-\log(1-R_n)}{n}-1\right)\right)\right]=\mathbb{E}\left[f\left(\sqrt{n}\left(\frac{\sum_{i=1}^{n+1}E_i}{n}-1\right)\right)\right].$$

   Le TCL assure que $\sqrt{n}\left(\frac{\sum_{i=1}^{n+1}E_i}{n}-1\right)$ converge en loi vers $\mathcal{N}(0,1)$, et il en est donc de même pour $\sqrt{n}\left(\frac{-\log(1-R_n)}{n}-1\right)$.
9. On a $F(a)=0$, $F(b)=1$ et $F'(z)=p(z)>0$ pour $z\in]a,b[$: F est donc continue et strictement croissante sur $[a,b]$ ce qui donne la bijectivité. Comme F est strictement croissante, les records de la suite $(X_i)_{i\ge0}$ sont ceux de la suite $(F(X_i))_{i\ge0}$, on a donc $\tilde{L}_n=L_n$ et $\tilde{R}_n=F(X_{L_n})=F(R_n)$.
10. Soit $u\in]0,1[$. $\mathbb{P}(F(X_1)\le u)=\mathbb{P}(X_1\le F^{-1}(u))=F(F^{-1}(u))=u$. Donc $F(X_i)$ suit la loi uniforme sur $[0,1]$.
    Grâce à la question, précédente, $F(R_n)=\tilde{R}_n$ suit la loi de densité q_n donnée par (1). On utilise alors le théorème de la fonction muette. Pour $f:\mathbb{R}\to\mathbb{R}$ bornée mesurable, on a

$$\mathbb{E}[f(R_n)]=\mathbb{E}[f(F^{-1}(\tilde{R}_n))]=\int_0^1 f(F^{-1}(x))q_n(x)dx=\int_a^b f(y)q_n(F(y))p(y)dy,$$

    et donc R_n suit la loi de densité $\mathbf{1}_{[a,b]}(y)\frac{(-\log(1-F(y)))^n}{n!}p(y)$.
11. D'après la question 8, on sait que pour n grand, $\sqrt{n}(\frac{-\log(1-F(R_n))}{n}-1)\approx\mathcal{N}(0,1)$, ce qui donne $\mathbb{P}(\sqrt{n}(\frac{-\log(1-F(R_n))}{n}-1)\ge-1,96)\approx 97,5/100$. Cela donne $\mathbb{P}(F(R_n)\in[1-e^{-n+1,96\sqrt{n}},1])\approx 97,5/100$. Ainsi, avec probabilité 97,5%, la probabilité que le prochain saut soit améliore le record du monde est de 3×10^{-11}. Cette probabilité est très faible, mais notre approche ne dit rien sur ce qui se passe sur les 2,5% restants. On peut être tenté de chercher à rapprocher (voire égaliser) les deux probabilités, par exemple $\mathbb{P}(G\le-4)\approx3\times10^{-5}$ et $e^{-36+4\sqrt{36}}\approx6\times10^{-6}$, ce qui nous dit qu'avec une confiance de 99,997%, la probabilité que le prochain saut améliore le record du monde est de 6×10^{-6}. Enfin, on peut aussi discuter de l'hypothèses i.i.d. de la suite $(X_i)_{i\in\mathbb{N}}$: souvent, les records sont battus par le détenteur du record, ou peuvent être dus à des changements de technique ou de matériels.