Corrigé de la séance 5 - MP+ - 13/02/26

Exercice 1 (***)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $(\mathrm{A}_n)_n$ une suite d'événements. On note

$$\mathrm{A} = \text{« une infinité d'événements } \mathrm{A}_n \text{ est réalisée »}$$

1. Montrer que A est un événement.
2. Si la série $\sum \mathbb{P}(\mathrm{A}_k)$ converge, montrer que $\mathbb{P}(\mathrm{A}) = 0$.
3. Soit $(\mathrm{X}_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles discrètes telles que $\sum \mathbb{P}(|\mathrm{X}_n| \geqslant \varepsilon)$ converge pour tout $\varepsilon > 0$.
   (a) Justifier que $\left\{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0\right\}$ est un événement.
   (b) Montrer $$\mathrm{X}_n \xrightarrow[n\to\infty]{} 0 \quad \text{p.s.}$$

Corrigé : 1. On a
$$\mathrm{A} = \bigcap_{\mathrm{N}\in\mathbb{N}} \bigcup_{n\geqslant \mathrm{N}} \mathrm{A}_n$$

Par stabilité par intersection et union dénombrables, il s'ensuit

$$\boxed{\text{L'ensemble A est un événement.}}$$

2. Par continuité décroissante, on a

$$\mathbb{P}(\mathrm{A}) = \lim_{\mathrm{N}\to+\infty} \mathbb{P}\left(\bigcup_{n\geqslant \mathrm{N}} \mathrm{A}_n\right)$$

et d'après l'inégalité de Boole
$$\mathbb{P}\left(\bigcup_{n\geqslant \mathrm{N}} \mathrm{A}_n\right) \leqslant \sum_{n=\mathrm{N}}^{+\infty} \mathbb{P}(\mathrm{A}_n)$$

le majorant étant le reste d'une série convergente donc de limite nulle. Par comparaison, il vient

$$\boxed{\mathbb{P}(\mathrm{A}) = 0}$$

Remarque : Ce résultat est un deuxième lemme dit de *Borel-Cantelli.*

3.(a) Soit $\omega \in \left\{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0\right\}$. On a

$$\forall \varepsilon > 0 \qquad \exists \mathrm{N} \in \mathbb{N} \quad | \quad \forall n \geqslant \mathrm{N} \qquad |\mathrm{X}_n(\omega)| < \varepsilon$$

ce qui équivaut à écrire
$$\omega \in \bigcap_{\varepsilon>0} \bigcup_{\mathrm{N}\in\mathbb{N}} \bigcap_{n\geqslant \mathrm{N}} \{|\mathrm{X}_n| < \varepsilon\}$$

Le problème dans cette écriture est que l'intersection sur $\varepsilon > 0$ porte sur un ensemble non dénombrable. On contourne cette difficulté en remarquant qu'il est équivalent d'écrire

$$\forall k \geqslant 1 \qquad \exists \mathrm{N} \in \mathbb{N} \quad | \quad \forall n \geqslant \mathrm{N} \qquad |\mathrm{X}_n(\omega)| < \frac{1}{k}$$

Ainsi
$$\left\{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0\right\} = \bigcap_{k\geqslant 1} \bigcup_{\mathrm{N}\in\mathbb{N}} \bigcap_{n\geqslant \mathrm{N}} \left\{|\mathrm{X}_n| < \frac{1}{k}\right\}$$

Par stabilité par union et intersection dénombrable, on conclut

$$\boxed{\left\{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0\right\} \in \mathscr{A}}$$

3.(b) Soit $\varepsilon > 0$. On pose
$$\mathrm{A}_\varepsilon = \bigcap_{\mathrm{N}\in\mathbb{N}} \bigcup_{n\geqslant \mathrm{N}} \{|\mathrm{X}_n| \geqslant \varepsilon\}$$

D'après le résultat de la deuxième question, on a $\mathbb{P}(\mathrm{A}_\varepsilon) = 0$ d'où $\overline{\mathrm{A}_\varepsilon}$ est presque sûr. Il s'ensuit que l'intersection dénombrable $\bigcap\limits_{k\geqslant 1} \overline{\mathrm{A}_{1/k}}$ est presque sûre et avec l'égalité

$$\left\{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0\right\} = \bigcap_{k\geqslant 1} \overline{\mathrm{A}_{1/k}}$$

On conclut
$$\boxed{\mathrm{X}_n \xrightarrow[n\to\infty]{} 0 \quad \text{p.s.}}$$

Exercice 2 (***)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $(\mathrm{X}_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles discrètes indépendantes de même loi dans L^2. On note $m = \mathbb{E}(\mathrm{X}_1)$, $\sigma = \sqrt{\mathbb{V}(\mathrm{X}_1)}$ et on pose

$$\forall n \geqslant 1 \qquad \mathrm{Y}_n = \frac{1}{n}\sum_{k=1}^{n} \mathrm{X}_k - m$$

1. Montrer $$\mathrm{Y}_{n^2} \xrightarrow[n\to\infty]{} 0 \quad \text{p.s.}$$

2. On note $\varphi(n) = \lfloor \sqrt{n} \rfloor$ pour n entier. Montrer $$\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2} \xrightarrow[n\to\infty]{} 0 \qquad \text{p.s.}$$

3. Conclure $$\frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i \xrightarrow[n\to\infty]{} \mathbb{E}(\mathrm{X}_1) \qquad \text{p.s.}$$

Corrigé : 1. La variable aléatoire Y_n admet un moment d'ordre 2 comme combinaison linéaire de variables aléatoires dans L^2. Pour $\varepsilon > 0$, il vient d'après l'inégalité de Bienaymé-Tchebychev et par indépendance des X_k

$$\mathbb{P}(|\mathrm{Y}_n| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2}\mathbb{V}(\mathrm{Y}_n) = \frac{\sigma^2}{\varepsilon^2 n}$$

Par suite, la série $\sum \mathbb{P}(|\mathrm{Y}_{n^2}| \geqslant \varepsilon)$ converge pour tout $\varepsilon > 0$ et, par application de la troisième question du premier exercice, on conclut

$$\boxed{\mathrm{Y}_{n^2} \xrightarrow[n\to\infty]{} 0 \qquad \text{p.s.}}$$

2. Soit $n \geqslant 1$. On a $$\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2} = \frac{1}{n}\sum_{k=\varphi(n)^2+1}^{n} (\mathrm{X}_k - m)$$

D'après l'inégalité de Bienaymé-Tchebychev et l'indépendance des X_k, on obtient pour $\varepsilon > 0$

$$\mathbb{P}\left(\left|\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2}\right| \geqslant \varepsilon\right) \leqslant \frac{1}{n^2\varepsilon^2}\mathbb{V}\left(\sum_{k=\varphi(n)^2+1}^{n} \mathrm{X}_k\right) = \frac{(n-\varphi(n)^2)\sigma^2}{n^2\varepsilon^2}$$

Par ailleurs, comme on a $\varphi(n) \leqslant \sqrt{n} < \varphi(n) + 1$, on en déduit $\sqrt{n} - 1 < \varphi(n)$ d'où

$$n - \varphi(n)^2 < 2\sqrt{n} - 1$$

Il s'ensuit $$\mathbb{P}\left(\left|\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2}\right| \geqslant \varepsilon\right) = \mathrm{O}\left(\frac{1}{n^{\frac{3}{2}}}\right)$$

Par critère de Riemann, la série $\sum \mathbb{P}\left(\left|\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2}\right| \geqslant \varepsilon\right)$ converge pour tout $\varepsilon > 0$ et d'après le résultat du premier exercice, on conclut

$$\boxed{\mathrm{Y}_n - \frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2} \xrightarrow[n\to\infty]{} 0 \qquad \text{p.s.}}$$

3. On a $\varphi(n) \leqslant \sqrt{n} < \varphi(n) + 1$ pour n entier d'où $\varphi(n) \underset{n\to+\infty}{\sim} \sqrt{n}$ et par suite $\varphi(n)^2 \underset{n\to+\infty}{\sim} n$. Combiné avec le résultat de la première question, on obtient

$$\frac{\varphi(n)^2}{n}\mathrm{Y}_{\varphi(n)^2} = \mathrm{O}(1)\mathrm{o}(1) \xrightarrow[n\to\infty]{} 0 \qquad \text{p.s.}$$

En effet, soit $\omega \in \left\{ \mathrm{Y}_{n^2} \xrightarrow[n\to\infty]{} 0 \right\}$. Puis, soit $\varepsilon > 0$. On dispose de N entier tel que pour $n \geqslant \mathrm{N}$, on a $|\mathrm{Y}_{n^2}(\omega)| \leqslant \varepsilon$. Comme $\varphi(n)^2 \xrightarrow[n\to\infty]{} +\infty$, on dispose de p entier tel que pour $n \geqslant p$, on a $\varphi(n)^2 \geqslant \mathrm{N}$ ce qui implique $\left|\mathrm{Y}_{\varphi(n)^2}\right| \leqslant \varepsilon$. Ceci prouve l'inclusion

$$\left\{ \mathrm{Y}_{n^2} \xrightarrow[n\to\infty]{} 0 \right\} \subset \left\{ \mathrm{Y}_{\varphi(n)^2} \xrightarrow[n\to\infty]{} 0 \right\}$$

et comme l'événement à gauche est presque sûr, celui qui le contient l'est aussi. Avec le résultat de la question précédente, il vient

$$\mathrm{Y}_n \xrightarrow[n\to\infty]{} 0 \qquad \text{p.s.}$$

Autrement dit $$\boxed{\frac{1}{n}\sum_{k=1}^{n} \mathrm{X}_k \xrightarrow[n\to\infty]{} m \qquad \text{p.s.}}$$

Remarque : Il s'agit de *la loi forte des grands nombres*. On peut affaiblir l'hypothèse L^2 en supposant que les variables sont d'espérance finie mais la démonstration est notablement plus difficile.

Exercice 3 (****)

On pose $$\forall x > 0 \qquad f(x) = \sum_{k=1}^{+\infty} \frac{x^k}{k!\sqrt{k}}$$

1. Montrer $$f(x) \underset{x\to+\infty}{=} \mathrm{o}(\mathrm{e}^x)$$

2. Pour $x > 0$ et Y_x variable aléatoire de loi $\mathscr{P}(x)$, montrer
$$\forall \varepsilon > 0 \qquad \mathbb{P}\left(|\mathrm{Y}_x - x| \geqslant \varepsilon x\right) \underset{x\to+\infty}{=} \mathrm{O}\left(\frac{1}{x}\right)$$

3. En déduire $$f(x) \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^x}{\sqrt{x}}$$

Corrigé : 1. Pour $x > 0$, on a $$\frac{x^k}{k!\sqrt{k}} \underset{k\to+\infty}{=} \mathrm{o}\left(\frac{x^k}{k!}\right)$$
et par comparaison à un terme général de série exponentielle convergente, on en déduit que la fonction f est bien définie. Montrons $\mathrm{e}^{-x} f(x) \xrightarrow[x\to+\infty]{} 0$. Soit n entier, $x > 0$ et $\mathrm{R}_n(x)$ le reste d'ordre n de la série définissant $\mathrm{e}^{-x}f(x)$. On a
$$0 \leqslant \mathrm{R}_n(x) = \mathrm{e}^{-x} \sum_{k=n+1}^{+\infty} \frac{x^k}{k!\sqrt{k}} \leqslant \frac{\mathrm{e}^{-x}}{\sqrt{n+1}} \underbrace{\sum_{k=n+1}^{+\infty} \frac{x^k}{k!}}_{\leqslant \mathrm{e}^x} \leqslant \frac{1}{\sqrt{n+1}}$$
On dispose donc d'un contrôle uniforme du reste d'où $\|\mathrm{R}_n\|_\infty \xrightarrow[n\to\infty]{} 0$ et d'après le théorème de double limite, il vient
$$\lim_{x\to+\infty} \mathrm{e}^{-x} f(x) = \sum_{k=1}^{+\infty} \lim_{x\to+\infty} \mathrm{e}^{-x} \frac{x^k}{k!\sqrt{k}} = 0$$
Ainsi $$\boxed{f(x) \underset{x\to+\infty}{=} \mathrm{o}(\mathrm{e}^x)}$$

2. D'après l'inégalité de Bienaymé-Tchebychev
$$\mathbb{P}\left(|\mathrm{Y}_x - \mathbb{E}(\mathrm{Y}_x)| \geqslant \varepsilon x\right) \leqslant \frac{1}{(\varepsilon x)^2} \mathbb{V}(\mathrm{Y}_x) = \frac{1}{\varepsilon^2 x}$$
Ainsi $$\boxed{\forall \varepsilon > 0 \qquad \mathbb{P}\left(|\mathrm{Y}_x - x| \geqslant \varepsilon x\right) \underset{x\to+\infty}{=} \mathrm{O}\left(\frac{1}{x}\right)}$$

3. On pose $$\forall u \geqslant 0 \qquad \varphi(u) = \begin{cases} \dfrac{1}{\sqrt{u}} & \text{si } u > 0 \\ 0 & \text{si } u = 0 \end{cases}$$
Soit $x > 0$. Par transfert, la convergence étant assurée par l'existence de f, on a
$$\mathbb{E}\left(\varphi\left(\frac{\mathrm{Y}_x}{x}\right)\right) = \sum_{k=0}^{+\infty} \varphi\left(\frac{k}{x}\right) \mathbb{P}(\mathrm{Y}_x = k) = \sum_{k=1}^{+\infty} \frac{\sqrt{x}}{\sqrt{k}} \mathrm{e}^{-x} \frac{x^k}{k!} = \mathrm{e}^{-x}\sqrt{x} f(x)$$
Puis $$\mathrm{e}^{-x}\sqrt{x} f(x) - 1 = \mathbb{E}\left(\varphi\left(\frac{\mathrm{Y}_x}{x}\right) - \varphi(1)\right)$$
Par inégalité triangulaire dans L^1, il vient

$$\left|\mathbb{E}\left(\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right)\right| \leqslant \mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\right)$$

On localise avec des fonctions indicatrices et on obtient

$$\mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\right) = \mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\mathbb{1}_{|\mathrm{Y}_x-x|<\varepsilon x}\right) + \mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\mathbb{1}_{|\mathrm{Y}_x-x|\geqslant\varepsilon x}\right)$$

Comme la variable Y_x est à valeurs dans $\mathbb{N}$, on observe que

$$\varphi\left(\frac{\mathrm{Y}_x}{x}\right) = \begin{cases} 0 & \text{si } \mathrm{Y}_x = 0 \\ \sqrt{\dfrac{x}{\mathrm{Y}_x}} \leqslant \sqrt{x} & \text{si } \mathrm{Y}_x \geqslant 1 \end{cases}$$

Par conséquent
$$\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-1\right| \leqslant 1+\sqrt{x}$$

et par suite

$$\mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\mathbb{1}_{|\mathrm{Y}_x-x|\geqslant\varepsilon x}\right) \leqslant (\sqrt{x}+1)\mathbb{P}\left(|\mathrm{Y}_x-x|\geqslant\varepsilon x\right) \underset{x\to+\infty}{=} \mathrm{o}(1)$$

Par continuité de φ en 1, pour $\delta > 0$, on peut choisir $\varepsilon > 0$ tel que

$$\forall u > 0 \qquad |u-1| < \varepsilon \quad \Longrightarrow \quad |\varphi(u)-\varphi(1)| < \delta$$

Ainsi
$$\mathbb{E}\left(\left|\varphi\left(\frac{\mathrm{Y}_x}{x}\right)-\varphi(1)\right|\mathbb{1}_{|\mathrm{Y}_x-x|<\varepsilon x}\right) \leqslant \delta$$

On peut donc rendre la quantité $\mathrm{e}^{-x}\sqrt{x}f(x)-1$ arbitrairement petite pour $x \to +\infty$ et on conclut

$$\boxed{f(x) \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^x}{\sqrt{x}}}$$

Exercice 4 (***)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé, Y une variable aléatoire réelle discrète centrée telle que $\mathrm{Y}(\Omega) \subset [\,\alpha\,;\beta\,]$, $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des variables aléatoires réelles discrètes indépendantes à valeurs dans $[\,a\,;b\,]$. On note $\mathrm{S}_n = \sum\limits_{i=1}^{n} \mathrm{X}_i$.

1. Soit s réel. Montrer

$$\forall y \in [\,\alpha\,;\beta\,] \qquad \mathrm{e}^{sy} \leqslant \frac{\beta - y}{\beta - \alpha}\mathrm{e}^{s\alpha} + \frac{y - \alpha}{\beta - \alpha}\mathrm{e}^{s\beta}$$

En déduire
$$\mathbb{E}(\mathrm{e}^{s\mathrm{Y}}) \leqslant q\mathrm{e}^{s\alpha} + p\mathrm{e}^{s\beta}$$

avec
$$p = \frac{-\alpha}{\beta - \alpha} \quad \text{et} \quad q = \frac{\beta}{\beta - \alpha}$$

2. On pose $\qquad \forall s \in \mathbb{R} \qquad \psi(s) = s\alpha + \ln\left(q + p\mathrm{e}^{s(\beta-\alpha)}\right)$

Justifier que la fonction ψ est deux fois dérivable puis établir

$$\forall s \in \mathbb{R} \qquad \psi''(s) = (\beta - \alpha)^2 \frac{qp\mathrm{e}^{s(\beta-\alpha)}}{\left(q + p\mathrm{e}^{s(\beta-\alpha)}\right)^2}$$

3. En déduire pour s réel $\qquad \psi''(s) \leqslant \dfrac{(\beta - \alpha)^2}{4}$

puis
$$\mathbb{E}(\mathrm{e}^{s\mathrm{Y}}) \leqslant \exp\left(\frac{(\beta - \alpha)^2 s^2}{8}\right)$$

4. Soient $\varepsilon, s > 0$. Montrer

$$\mathbb{P}\left(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\right) \leqslant \exp\left(-s\varepsilon + n\frac{(b - a)^2 s^2}{8}\right)$$

En déduire
$$\mathbb{P}\left(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\right) \leqslant \exp\left(-\frac{2\varepsilon^2}{n(b - a)^2}\right)$$

Corrigé : 1. Soit $y \in [\,\alpha\,;\beta\,]$ et s réel. On a

$$sy = \frac{\beta - y}{\beta - \alpha}s\alpha + \frac{y - \alpha}{\beta - \alpha}s\beta \quad \text{avec} \quad \frac{\beta - y}{\beta - \alpha} + \frac{y - \alpha}{\beta - \alpha} = 1 \quad \text{et} \quad \frac{\beta - y}{\beta - \alpha}\frac{y - \alpha}{\beta - \alpha} \geqslant 0$$

Par convexité de l'exponentielle, il vient

$$\boxed{\forall y \in [\,\alpha\,;\beta\,] \qquad \mathrm{e}^{sy} \leqslant \frac{\beta - y}{\beta - \alpha}\mathrm{e}^{s\alpha} + \frac{y - \alpha}{\beta - \alpha}\mathrm{e}^{s\beta}}$$

On applique cette inégalité avec Y qui est à valeurs dans $[\,\alpha\,;\beta\,]$. Passant à l'espérance, on trouve par linéarité

$$\boxed{\mathbb{E}(\mathrm{e}^{s\mathrm{Y}}) \leqslant q\mathrm{e}^{s\alpha} + p\mathrm{e}^{s\beta}}$$

2. Par croissance de l'espérance, on a

$$\alpha \leqslant \mathbb{E}(\mathrm{Y}) = 0 \leqslant \beta$$

et comme $\beta > \alpha$, on en déduit que p et q sont positifs dont l'un strictement. Par conséquent, on a $q + p\mathrm{e}^{s(\beta-\alpha)} > 0$ pour s réel et il s'ensuit que l'application ψ est deux fois dérivable. Par dérivation, il vient

$$\forall s \in \mathbb{R} \qquad \psi'(s) = \alpha + (\beta - \alpha)\frac{p\mathrm{e}^{s(\beta-\alpha)}}{q + p\mathrm{e}^{s(\beta-\alpha)}} = \alpha + (\beta - \alpha)\left[1 - \frac{q}{q + p\mathrm{e}^{s(\beta-\alpha)}}\right]$$

Puis

$$\boxed{\forall s \in \mathbb{R} \qquad \psi''(s) = (\beta - \alpha)^2 \frac{qp\mathrm{e}^{s(\beta-\alpha)}}{(q + p\mathrm{e}^{s(\beta-\alpha)})^2}}$$

3. Pour u et v réels, on a $(u+v)^2 \geqslant 4uv$ puisque $(u+v)^2 - 4uv = (u-v)^2 \geqslant 0$. Avec $u = q$ et $v = p\mathrm{e}^{s(\beta-\alpha)}$, on conclut

$$\boxed{\forall s \in \mathbb{R} \qquad \psi''(s) \leqslant \frac{(\beta-\alpha)^2}{4}}$$

D'après la formule de Taylor reste intégral, on a pour s réel

$$\psi(s) = \psi(0) + \psi'(0)s + \int_0^s \psi''(t)(s-t)\,\mathrm{d}t \leqslant \frac{(\beta-\alpha)^2}{4}\int_0^s (s-t)^2\,\mathrm{d}t$$

c'est-à-dire

$$\forall s \in \mathbb{R} \qquad \psi(s) \leqslant \frac{(\beta-\alpha)^2 s^2}{8}$$

Passant à l'exponentielle, on trouve pour s réel

$$\exp(\psi(s)) = \mathrm{e}^{s\alpha}\left(q + p\mathrm{e}^{s(\beta-\alpha)}\right) = q\mathrm{e}^{s\alpha} + p\mathrm{e}^{s\beta} \leqslant \exp\left(\frac{(\beta-\alpha)^2 s^2}{8}\right)$$

et avec le résultat de la première question, on conclut

$$\boxed{\forall s \in \mathbb{R} \qquad \mathbb{E}(\mathrm{e}^{s\mathrm{Y}}) \leqslant \exp\left(\frac{(\beta-\alpha)^2 s^2}{8}\right)}$$

4. Soient $\varepsilon, s > 0$. On a $\qquad \{\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\} = \left\{\mathrm{e}^{s(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n)} \geqslant \mathrm{e}^{s\varepsilon}\right\}$

D'après l'inégalité de Markov avec la variable $\mathrm{e}^{s(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n))}$ positive et par indépendance des X_i, il vient

$$\mathbb{P}\left(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\right) \leqslant \mathrm{e}^{-s\varepsilon}\mathbb{E}\left(\prod_{i=1}^n \mathrm{e}^{s(\mathrm{X}_i - \mathbb{E}(\mathrm{X}_i))}\right) = \mathrm{e}^{-s\varepsilon}\prod_{i=1}^n \mathbb{E}\left(\mathrm{e}^{s(\mathrm{X}_i - \mathbb{E}(\mathrm{X}_i))}\right)$$

Les variables $\mathrm{X}_i - \mathbb{E}(\mathrm{X}_i)$ sont dans l'intervalle $[\, a - \mathbb{E}(\mathrm{X}_i)\,; b - \mathbb{E}(\mathrm{X}_i)\,]$. Ainsi, par application de l'inégalité obtenue à la question précédente, on obtient

$$\boxed{\mathbb{P}\left(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\right) \leqslant \exp\left(-s\varepsilon + n\frac{(b-a)^2 s^2}{8}\right)}$$

Enfin, on choisit la valeur de s qui minimise le trinôme dans l'exponentielle et on conclut

$$\boxed{\mathbb{P}\left(\mathrm{S}_n - \mathbb{E}(\mathrm{S}_n) \geqslant \varepsilon\right) \leqslant \exp\left(-\frac{2\varepsilon^2}{n(b-a)^2}\right)}$$

Remarque : Il s'agit de *l'inégalité de Hoeffding.*

Exercice 5 (***)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des variables aléatoires réelles discrètes indépendantes centrées dans L^2. On note $\mathrm{S}_k = \sum\limits_{i=1}^{k} \mathrm{X}_i$ pour $k \in [\![1\,;\, n]\!]$. Soit $\varepsilon > 0$. On pose

$$\forall k \in [\![1\,;\, n]\!] \qquad \mathrm{A}_k = \bigcap_{j=1}^{k-1} \{|\mathrm{S}_j| < \varepsilon\} \cap \{|\mathrm{S}_k| \geqslant \varepsilon\}$$

1. Justifier l'indépendance des variables aléatoires $\mathrm{S}_k \mathbb{1}_{\mathrm{A}_k}$ et $\mathrm{S}_n - \mathrm{S}_k$ pour $k \in [\![1\,;\, n]\!]$.

2. Montrer
$$\sum_{k=1}^{n} \mathbb{E}\left(\mathrm{S}_k^2 \mathbb{1}_{\mathrm{A}_k}\right) \leqslant \mathbb{E}(\mathrm{S}_n^2)$$

3. En déduire
$$\mathbb{P}\left(\operatorname*{Max}_{k \in [\![1\,;\, n]\!]} |\mathrm{S}_k| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \mathbb{V}(\mathrm{X}_i)$$

Corrigé : 1. Soit $k \in [\![1\,;\, n]\!]$. On a $\mathrm{S}_n - \mathrm{S}_k = \sum\limits_{i=k+1}^{n} \mathrm{X}_i$. La variable S_k et l'évenement A_k sont définis à partir des variables aléatoire $\mathrm{X}_1, \ldots, \mathrm{X}_k$. Il n'y a donc pas de chevauchement d'indice et par indépendance des X_i, on conclut

$$\boxed{\text{Pour tout } k \in [\![1\,;\, n]\!], \text{ les variables aléatoires } \mathrm{S}_k \mathbb{1}_{\mathrm{A}_k} \text{ et } \mathrm{S}_n - \mathrm{S}_k \text{ sont indépendantes.}}$$

2. Les événements A_k sont incompatibles d'où

$$\sum_{k=1}^{n} \mathbb{1}_{\mathrm{A}_k} = \mathbb{1}_{\mathrm{B}_k} \leqslant \mathbb{1}_{\Omega} = 1 \quad \text{avec} \quad \mathrm{B}_k = \bigsqcup_{k=1}^{n} \mathrm{A}_k$$

En multipliant par S_n^2 et par linéarité de l'espérance, les variables concernées étant d'espérance finie, il vient

$$\mathbb{E}\left(\sum_{k=1}^{n} \mathrm{S}_n^2 \mathbb{1}_{\mathrm{A}_k}\right) = \sum_{k=1}^{n} \mathbb{E}\left(\mathrm{S}_n^2 \mathbb{1}_{\mathrm{A}_k}\right) \leqslant \mathbb{E}(\mathrm{S}_n^2)$$

En décomposant $\mathrm{S}_n = \mathrm{S}_k + \mathrm{S}_n - \mathrm{S}_k$, on obtient par linéarité de l'espérance

$$\mathbb{E}\left(\mathrm{S}_n^2 \mathbb{1}_{\mathrm{A}_k}\right) = \mathbb{E}\left(\mathrm{S}_k^2 \mathbb{1}_{\mathrm{A}k}\right) + 2\mathbb{E}\left((\mathrm{S}_n - \mathrm{S}_k)\mathrm{S}_k \mathbb{1}_{\mathrm{A}_k}\right) + \mathbb{E}\left((\mathrm{S}_n - \mathrm{S}_k)^2 \mathbb{1}_{\mathrm{A}_k}\right)$$

D'après l'indépendance établie à la première question puis le caractère centré des X_i, il vient

$$\mathbb{E}\left((\mathrm{S}_n - \mathrm{S}_k)\mathrm{S}_k \mathbb{1}_{\mathrm{A}_k}\right) = \mathbb{E}\left(\mathrm{S}_n - \mathrm{S}_k\right) \mathbb{E}\left(\mathrm{S}_k \mathbb{1}_{\mathrm{A}_k}\right) = 0$$

Enfin, comme $\mathbb{E}\left((\mathrm{S}_n - \mathrm{S}_k)^2 \mathbb{1}_{\mathrm{A}_k}\right) \geqslant 0$, on conclut

$$\boxed{\sum_{k=1}^{n} \mathbb{E}\left(\mathrm{S}_k^2 \mathbb{1}_{\mathrm{A}_k}\right) \leqslant \mathbb{E}(\mathrm{S}_n^2)}$$

3. Soit $\omega \in \left\{\operatorname*{Max}_{k \in [\![1\,;\, n]\!]} |\mathrm{S}_k| \geqslant \varepsilon\right\}$. En considérant le plus petit indice $k \in [\![1\,;\, n]\!]$ tel que $|\mathrm{S}_k(\omega)| \geqslant \varepsilon$, on obtient

$$\left\{\operatorname*{Max}_{k \in [\![1\,;\, n]\!]} |\mathrm{S}_k| \geqslant \varepsilon\right\} \subset \bigsqcup_{k=1}^{n} \mathrm{A}_k$$

et l'inclusion réciproque est immédiate d'où

$$\left\{\operatorname*{Max}_{k \in [\![1\,;\, n]\!]} |\mathrm{S}_k| \geqslant \varepsilon\right\} = \bigsqcup_{k=1}^{n} \mathrm{A}_k$$

et par sous-additivité
$$\mathbb{P}\left(\underset{k\in[\![1\,;\,n]\!]}{\mathrm{Max}} |S_k| \geqslant \varepsilon\right) \leqslant \sum_{k=1}^{n} \mathbb{P}(A_k)$$

Or, par définition des A_k, on a
$$\forall k \in [\![1\,;\,n]\!] \qquad \mathbb{E}(S_k^2 \mathbb{1}_{A_k}) \geqslant \varepsilon^2 \mathbb{E}(\mathbb{1}_{A_k}) = \varepsilon^2 \mathbb{P}(A_k)$$

Comme les X_i sont centrées, on a $\mathbb{E}(S_n^2) = \mathbb{V}(S_n)$ et par indépendances des X_i et l'inégalité établie à la question précédente, on conclut
$$\boxed{\mathbb{P}\left(\underset{k\in[\![1\,;\,n]\!]}{\mathrm{Max}} |S_k| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \mathbb{V}(X_i)}$$

Remarque : Il s'agit de *l'inégalité de Kolmogorov*. Cette inégalité est un élément clé de la preuve de Kolmogorov pour établir la loi forte des grands nombres.

Exercice 6 (****)

1. Montrer que $x \mapsto \mathrm{e}^{-2x}$ est limite uniforme sur $\mathbb{R}_+$ d'une suite de fonctions de la forme $x \mapsto \mathrm{P}(x)\mathrm{e}^{-x}$ avec $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$.
2. On considère E l'espace des fonctions de $\mathbb{R}_+$ dans $\mathbb{R}$, continues et de limite nulle en $+\infty$. On munit E de la norme $\|\cdot\|_\infty$. Soit $f : \mathbb{R}_+ \to \mathbb{R}$, continue à support compact. Soit $\lambda \geqslant 0$ et $(\mathrm{X}_k^{(\lambda)})_{k\geqslant 1}$ une suite de variables aléatoires indépendantes suivant toutes une loi de Poisson de paramètre λ (on généralise pour $\lambda = 0$). Montrer que la suite de fonctions $(g_n)_{n\geqslant 1}$ avec
$$\forall \lambda \geqslant 0 \qquad g_n(\lambda) = \mathbb{E}\left(f\left(\frac{1}{n}\sum_{k=1}^{n} \mathrm{X}_k^{(\lambda)}\right)\right)$$
converge uniformément sur $\mathbb{R}_+$.
3. Montrer que $\{x \in \mathbb{R}_+ \mapsto \mathrm{P}(x)\mathrm{e}^{-x}, \mathrm{P} \in \mathbb{R}[\mathrm{X}]\}$ est dense dans E.

Corrigé : 1. On pose $\mathrm{P}_n = \sum_{k=0}^{n} (-1)^k \dfrac{\mathrm{X}^k}{k!}$ pour n entier et on a
$$\forall x \geqslant 0 \qquad \mathrm{P}_n(x)\mathrm{e}^{-x} \xrightarrow[n\to\infty]{} \mathrm{e}^{-2x}$$
D'après l'inégalité de Taylor-Lagrange, on trouve
$$\forall x \geqslant 0 \qquad |\mathrm{P}_n(x) - \mathrm{e}^{-x}| \leqslant \frac{x^{n+1}}{(n+1)!}$$
d'où $\quad \forall n \geqslant x \qquad |\mathrm{P}_n(x)\mathrm{e}^{-x} - \mathrm{e}^{-2x}| \leqslant h_n(x) \quad$ avec $\quad \forall x \geqslant 0 \qquad h_n(x) = \mathrm{e}^{-x}\dfrac{x^{n+1}}{(n+1)!}$

On trouve
$$\|h_n\|_\infty = h_n(n+1) = \mathrm{e}^{-(n+1)}\frac{(n+1)^{n+1}}{(n+1)!}$$
Avec l'équivalent de Stirling, on trouve
$$\|h_n\|_\infty \underset{n\to+\infty}{\sim} \frac{1}{\sqrt{2\pi n}}$$
Ainsi
$$\boxed{\exists (\mathrm{P}_n)_n \in \mathbb{R}[\mathrm{X}]^{\mathbb{N}} \quad | \quad \operatorname*{Sup}_{x\geqslant 0} |\mathrm{P}_n(x)\mathrm{e}^{-x} - \mathrm{e}^{-2x}| \xrightarrow[n\to\infty]{} 0}$$

2. On pose
$$\forall \lambda \geqslant 0 \qquad \forall n \geqslant 1 \qquad \bar{\mathrm{X}}_n^{(\lambda)} = \frac{1}{n}\sum_{k=1}^{n} \mathrm{X}_k^{(\lambda)}$$
On note M la borne supérieure du support de f (si f est nulle, le résultat est trivial).
$$\forall \lambda \geqslant 0 \qquad \forall \eta > 0 \qquad \mathrm{A}_{\lambda,\eta} = \left\{\left|\bar{\mathrm{X}}_n^{(\lambda)} - \lambda\right| \geqslant \eta\right\}$$
• Supposons $\lambda \leqslant 2\mathrm{M}$. Pour $\eta > 0$, il vient
$$\begin{aligned}\left|f(\lambda) - \mathbb{E}(f(\bar{\mathrm{X}}_n^{(\lambda)}))\right| &\leqslant \mathbb{E}\left(\left|f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)})\right|\right)\\ &\leqslant \mathbb{E}\left(\left|f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)})\right|\mathbb{1}_{\mathrm{A}_{\lambda,\eta}}\right) + \mathbb{E}\left(\left|f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)})\right|\mathbb{1}_{\Omega\smallsetminus\mathrm{A}_{\lambda,\eta}}\right)\end{aligned}$$
Puis, avec l'inégalité de Bienaymé-Tchebycheff, les X_k^λ ayant un moment d'ordre 2, il vient
$$\begin{aligned}\mathbb{E}\left(\left|f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)})\right|\mathbb{1}_{\mathrm{A}_{\lambda,\eta}}\right) &\leqslant 2\|f\|_\infty\mathbb{P}\left(\left|\bar{\mathrm{X}}_n^{(\lambda)} - \lambda\right| \geqslant \eta\right)\\ &\leqslant \frac{2\|f\|_\infty}{\eta^2}\mathbb{V}(\bar{\mathrm{X}}_n^\lambda) = \frac{2\|f\|_\infty\lambda}{n\eta^2} \leqslant \frac{4\|f\|_\infty\mathrm{M}}{n\eta^2}\end{aligned}$$

Soit $\varepsilon > 0$. D'après le théorème de Heine, la fonction f est uniformément continue sur $[\,0\,;2\mathrm{M}\,]$. On choisit $\eta \in \,]\,0\,;\mathrm{M}\,[$ tel que

$$\forall (x,y) \in [\,0\,;2\mathrm{M}\,]^2 \qquad |x-y| \leqslant \eta \quad \Longrightarrow \quad |f(x)-f(y)| \leqslant \varepsilon$$

Ainsi
$$\Omega \smallsetminus \mathrm{A}_{\lambda,\eta} \subset \left\{ \left| f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)}) \right| \leqslant \varepsilon \right\}$$

En effet, si $\bar{\mathrm{X}}_n^{(\lambda)} \in [\,0\,;2\mathrm{M}\,]$, l'inclusion résulte de l'uniforme continuité. Sinon, on a

$$\begin{cases} \bar{\mathrm{X}}_n^{(\lambda)} > 2\mathrm{M} \\ \left| \bar{\mathrm{X}}_n^{(\lambda)} - \lambda \right| \leqslant \eta < \mathrm{M} \end{cases} \Longrightarrow -\mathrm{M} + 2\mathrm{M} < -\mathrm{M} + \bar{\mathrm{X}}_n^{(\lambda)} < \lambda$$

d'où
$$f(\lambda) = f(\bar{\mathrm{X}}_n^{(\lambda)}) = 0$$

Par ailleurs, on dispose d'un seuil N tel que pour $n \geqslant \mathrm{N}$, on a $\dfrac{4\|f\|_\infty \mathrm{M}}{n\eta^2} \leqslant \varepsilon$ et on a donc prouvé

$$\forall \varepsilon > 0 \qquad \exists \mathrm{N} \in \mathbb{N} \quad | \quad \forall n \geqslant \mathrm{N} \qquad \operatorname*{Sup}_{\lambda \in [\,0\,;2\mathrm{M}\,]} \left| f(\lambda) - \mathbb{E}(f(\bar{\mathrm{X}}_n^{(\lambda)})) \right| \leqslant 2\varepsilon$$

• Supposons $\lambda > 2\mathrm{M}$. Il vient

$$\begin{aligned} \left| f(\lambda) - \mathbb{E}(f(\bar{\mathrm{X}}_n^{(\lambda)})) \right| &\leqslant \mathbb{E}\left(\left| f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)}) \right| \right) \\ &\leqslant \mathbb{E}\left(\left| f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)}) \right| \mathbb{1}_{\mathrm{A}_{\lambda,\lambda/2}} \right) + \mathbb{E}\left(\left| f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)}) \right| \mathbb{1}_{\Omega \smallsetminus \mathrm{A}_{\lambda,\lambda/2}} \right) \end{aligned}$$

On remarque
$$\begin{cases} \lambda > 2\mathrm{M} \\ \left| \bar{\mathrm{X}}_n^{(\lambda)} - \lambda \right| < \lambda/2 \end{cases} \quad \Longrightarrow \quad \bar{\mathrm{X}}_n^{(\lambda)} > \frac{\lambda}{2} > \mathrm{M}$$

Ainsi
$$\mathbb{E}\left(\left| f(\lambda) - f(\bar{\mathrm{X}}_n^{(\lambda)}) \right| \mathbb{1}_{\Omega \smallsetminus \mathrm{A}_{\lambda,\lambda/2}} \right) = 0$$

De nouveau par inégalité de Bienaymé-Tchebycheff, on obtient

$$\left| f(\lambda) - \mathbb{E}(f(\bar{\mathrm{X}}_n^{(\lambda)})) \right| \leqslant \frac{2\|f\|_\infty}{(\lambda/2)^2} \mathbb{V}(\bar{\mathrm{X}}_n^{(\lambda)}) = \frac{8\|f\|_\infty}{n\lambda} \leqslant \frac{4\|f\|_\infty}{\mathrm{M}n}$$

d'où
$$\operatorname*{Sup}_{\lambda > 2\mathrm{M}} \left| f(\lambda) - \mathbb{E}(f(\bar{\mathrm{X}}_n^{(\lambda)})) \right| \xrightarrow[n\to\infty]{} 0$$

On a établit la convergence uniforme et sur $[\,0\,;2\mathrm{M}\,]$ et $]\,2\mathrm{M}\,;+\infty\,[$. On conclut

$$\boxed{\|g_n - f\|_\infty \xrightarrow[n\to\infty]{} 0}$$

3. Soit $f \in \mathrm{E}$ et $\varepsilon > 0$. Soit $\mathrm{M} \geqslant 0$ tel que $|f(x)| \leqslant \varepsilon$ pour $x \geqslant \mathrm{M}$. Soit g coincidant avec f sur $[\,0\,;\mathrm{M}\,]$, affine sur $[\,\mathrm{M}\,;\mathrm{M}+1\,]$ et nulle ensuite. La fonction g est à support compact et vérifie $\|f-g\|_\infty \leqslant 2\varepsilon$. Enfin, on peut choisir g_n pour n assez grand tel que $\|g-g_n\|_\infty \leqslant \varepsilon$ et la fonction g_n est de la forme souhaitée par transfert. Ainsi, on dispose d'un seuil N tel que, pour $n \geqslant \mathrm{N}$

$$\|f - g_n\|_\infty \leqslant \|f-g\|_\infty + \|g-g_n\|_\infty \leqslant 3\varepsilon$$

On conclut $\boxed{\text{L'ensemble } \{x \in \mathbb{R}_+ \mapsto \mathrm{P}(x)\mathrm{e}^{-x}, \mathrm{P} \in \mathbb{R}[\mathrm{X}]\} \text{ est dense dans E.}}$

Exercice 7 (****)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et $(\mathrm{X}_k)_{k\geqslant 1}$ une suite de variables indépendantes de loi uniforme sur $[\![1\,;\, r]\!]$ avec $r \geqslant 1$ et $\mathrm{Z}_1, \ldots, \mathrm{Z}_r$ des variables indépendantes avec $\mathrm{Z}_1 = 1$ et $\mathrm{Z}_i \sim \mathscr{G}\left(\dfrac{r-i+1}{r}\right)$ pour tout $i \in [\![2\,;\, r]\!]$. On pose $\mathrm{C}_r = \sum\limits_{i=1}^{r} \mathrm{Z}_i$ puis, pour $i \in [\![1\,;\, r]\!]$

$$\forall \omega \in \Omega \qquad \mathrm{T}_i(\omega) = \inf\{n \geqslant 1 \mid \operatorname{Card}\{\mathrm{X}_1(\omega), \ldots, \mathrm{X}_n(\omega)\} = i\}$$

et $\mathrm{Y}_i = \mathrm{T}_i - \mathrm{T}_{i-1}$ avec la convention $\mathrm{T}_0 = 0$.

1. Préciser espérance et variance des variables Z_i.
2. Montrer que
$$\mathbb{E}(\mathrm{C}_r) = r \times \mathrm{H}_r \quad \text{et} \quad \mathbb{V}(\mathrm{C}_r) = -r\mathrm{H}_r + r^2 \left(\sum_{k=1}^{r} \frac{1}{k^2}\right) \quad \text{avec} \quad \mathrm{H}_r = \sum_{k=1}^{r} \frac{1}{k}$$
3. Justifier $$\mathrm{H}_r \underset{r\to+\infty}{\sim} \ln r$$
4. Établir $$\forall \varepsilon > 0 \qquad \mathbb{P}\left(\left|\frac{\mathrm{C}_r}{r \ln r} - 1\right| \geqslant \varepsilon\right) \xrightarrow[r\to+\infty]{} 0$$
5. Déterminer la loi de $(\mathrm{Y}_1, \ldots, \mathrm{Y}_r)$.
6. En déduire $$\mathrm{T}_r \sim \mathrm{C}_r$$

Corrigé : 1. On a

$$\boxed{\forall i \in [\![1\,;\, r]\!] \qquad \mathbb{E}(\mathrm{Z}_i) = \frac{r}{r-i+1} \quad \text{et} \quad \mathbb{V}(\mathrm{Z}_i) = \frac{i-1}{r}\left(\frac{r}{r-i+1}\right)^2}$$

2. Par linéarité de l'espérance, il vient

$$\mathbb{E}(\mathrm{C}_r) = \mathbb{E}\left(\sum_{i=1}^{r} \mathrm{Z}_i\right) = \sum_{i=1}^{r} \frac{r}{r-i+1}$$

Les variables Z_i étant indépendantes, il vient

$$\mathbb{V}(\mathrm{C}_r) = \mathbb{V}\left(\sum_{i=1}^{r} \mathrm{Z}_i\right) = \sum_{i=1}^{r} \frac{i-1}{r} \times \frac{r^2}{(r-i+1)^2} = r\sum_{i=1}^{r}\left[-\frac{1}{r-i+1} + \frac{r}{(r-i+1)^2}\right]$$

Avec le changement d'indice $k = r - i + 1$, on obtient

$$\boxed{\mathbb{E}(\mathrm{C}_r) = r \times \mathrm{H}_r \quad \text{et} \quad \mathbb{V}(\mathrm{C}_r) = -r\mathrm{H}_r + r^2 \left(\sum_{k=1}^{r} \frac{1}{k^2}\right)}$$

3. Par comparaison série/intégrale ou en considérant la série téléscopique $\sum (u_{r+1} - u_r)$ avec $u_r = \mathrm{H}_r - \ln r$, on obtient

$$\boxed{\mathrm{H}_r \underset{r\to+\infty}{\sim} \ln r}$$

4. Soit $\varepsilon > 0$. On a l'inclusion

$$\left\{\left|\frac{\mathrm{C}_r}{r \ln r} - 1\right| \geqslant \varepsilon\right\} \subset \left\{\left|\frac{\mathrm{C}_r}{r \ln r} - \frac{\mathrm{H}_r}{\ln r}\right| \geqslant \frac{\varepsilon}{2}\right\} \cup \left\{\left|\frac{\mathrm{H}_r}{\ln r} - 1\right| \geqslant \frac{\varepsilon}{2}\right\}$$

et $\left\{\left|\dfrac{\mathrm{H}_r}{\ln r} - 1\right| \geqslant \dfrac{\varepsilon}{2}\right\} = \varnothing$ pour r assez grand. Par suite, en utilisant l'inégalité de Bienaymé-Tchebycheff, on obtient

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\mathrm{C}_r}{r\ln r}-1\right|\geqslant\varepsilon\right)&\leqslant\mathbb{P}\left(\left|\frac{\mathrm{C}_r}{r\ln r}-\frac{\mathrm{H}_r}{\ln r}\right|\geqslant\tfrac{\varepsilon}{2}\right)+\mathrm{o}(1)\\&\leqslant\frac{4}{\varepsilon^2}\mathbb{V}\left(\frac{\mathrm{C}_r}{r\ln r}\right)+\mathrm{o}(1)\end{aligned}$$

et $$\mathbb{V}\left(\frac{\mathrm{C}_r}{r\ln r}\right)=\frac{\mathrm{O}(r^2)}{r^2\ln^2 r}=\mathrm{O}\left(\frac{1}{\ln^2 r}\right)=\mathrm{o}(1)$$

On conclut $$\boxed{\forall\varepsilon>0\qquad\mathbb{P}\left(\left|\frac{\mathrm{C}_r}{r\ln r}-1\right|\geqslant\varepsilon\right)\xrightarrow[r\to+\infty]{}0}$$

5. On a clairement $\mathrm{Y}_i(\Omega)\subset\mathbb{N}^*$ pour tout $i\in[\![1\,;\,r]\!]$. Soit $y_1=1$ et $(y_2,\ldots,y_r)\in(\mathbb{N}^*)^{r-1}$. On pose $t_i=\sum\limits_{k=1}^{i}y_k$ pour $i\in[\![1\,;\,r]\!]$. Il vient

$$\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{Y}_i=y_i\}\right)=\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{T}_i=t_i\}\right)$$

puis

$$\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{Y}_i=y_i\}\right)=$$
$$\sum_{\sigma\in\mathrm{S}_r}\mathbb{P}\left(\bigcap_{k=1}^{r-1}\{\mathrm{X}_{t_k}=\sigma(k),\mathrm{X}_{t_k+1}\in\sigma([\![1\,;\,k]\!]),\ldots,\mathrm{X}_{t_{k+1}-1}\in\sigma([\![1\,;\,k]\!])\}\cap\{\mathrm{X}_{t_r}=\sigma(r)\}\right)$$

Ainsi, par indépendance des X_i, on obtient

$$\begin{aligned}\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{Y}_i=y_i\}\right)&=\sum_{\sigma\in\mathrm{S}_r}\frac{1}{r}\prod_{k=1}^{r-1}\left[\frac{1}{r}\left(\frac{k}{r}\right)^{t_{k+1}-t_k-1}\right]\\&=\frac{r!}{r}\prod_{k=2}^{r}\left[\frac{1}{r}\left(\frac{k-1}{r}\right)^{y_k-1}\right]\\&=\prod_{k=2}^{r}\left[\frac{r-(k-1)}{r}\left(\frac{k-1}{r}\right)^{y_k-1}\right]\\\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{Y}_i=y_i\}\right)&=\mathbb{P}\left(\bigcap_{i=1}^{r}\{\mathrm{Z}_i=y_i\}\right)\end{aligned}$$

Ainsi $$\boxed{(\mathrm{Y}_1,\ldots,\mathrm{Y}_r)\sim(\mathrm{Z}_1,\ldots,\mathrm{Z}_r)}$$

Remarque : On a passé sous silence le fait que les fonctions T_i et donc aussi les Y_i sont bien des variables aléatoires discrètes. Soit $i\in[\![1\,;\,r]\!]$. On a *a priori* $\mathrm{T}_i(\Omega)\subset\mathbb{N}^*\cup\{\infty\}$ puis pour $t_i\geqslant 1$

$$\{\mathrm{T}_i=t_i\}=\{\mathrm{Card}\,\{\mathrm{X}_1,\ldots,\mathrm{X}_{t_i}\}=i\}\cap\{\mathrm{Card}\,\{\mathrm{X}_1,\ldots,\mathrm{X}_{t_i-1}\}=i-1\}$$

et $$\{\mathrm{T}_i=\infty\}=\overline{\{\mathrm{T}_i<\infty\}}\quad\text{et}\quad\{\mathrm{T}_i<\infty\}=\bigcup_{t_i\geqslant 1}\{\mathrm{T}_i=t_i\}$$

Pour n entier, la quantité $\mathrm{Card}\,\{\mathrm{X}_1,\ldots,\mathrm{X}_n\}$ (nulle si $n=0$) est fonction de $(\mathrm{X}_1,\ldots,\mathrm{X}_n)$ donc est une variable aléatoire. Ainsi, l'ensemble $\{\mathrm{T}_i=t_i\}$ est un événement puis $\{\mathrm{T}_i<\infty\}$ également par union dénombrable et $\{\mathrm{T}_i=\infty\}$ par complémentation ce qui prouve que les T_i sont des variables aléatoires discrètes et les Y_i également.

6. D'après le résultat de la question précédente, on conclut

$$\boxed{\mathrm{T}_r = \sum_{i=1}^{r} \mathrm{Y}_i \sim \sum_{i=1}^{r} \mathrm{Z}_i = \mathrm{C}_r}$$

Commentaire : Le problème est connu sous l'appellation *problème du collectionneur de vignettes*. La variable X_i est la vignette obtenue lors du i-ème achat de votre boîte de céréales favorite. La variable T_i désigne donc le nombre d'achats à effectuer pour posséder i vignettes différentes et T_r le nombre d'achat pour avoir la collection complète.