ISM MP, Mathématiques
Année 2024/2025

Feuille d'exercices n°66

Exercice 1 (***)

On effectue une succession infinie de lancers indépendants d'une pièce donnant pile avec probabilité $p \in]0;1[$ et face avec probabilité $q = 1 - p$. On dit que la première série est de longueur $n \geqslant 1$ si les n premiers lancers ont donné le même côté et le $n+1$ l'autre côté. La deuxième série commence au lancer suivant la fin de la première série et se termine (si elle se termine) au lancer précédant un changement de côté. Pour $k \geqslant 1$, on note P_k l'événement pile au k-ième lancer F_k face au k-ième lancer et L_1 et L_2 les longueurs respectives des deux séries.

1. Déterminer la loi de L_1.
2. Justifier que L_1 est d'espérance finie puis la calculer.
3. Déterminer la loi de L_2.
4. Justifier que L_2 est d'espérance finie puis la calculer.
5. Les variables aléatoires L_1 et L_2 sont-elles indépendantes ?

Corrigé : 1. On a $\mathrm{L}_1(\Omega) \subset \mathbb{N}^*$ puis

$$\forall k \geqslant 1 \qquad \{\mathrm{L}_1 = k\} = \mathrm{P}_1 \ldots \mathrm{P}_k \mathrm{F}_{k+1} \sqcup \mathrm{F}_1 \ldots \mathrm{F}_k \mathrm{P}_{k+1}$$

D'où $$\boxed{\forall k \geqslant 1 \qquad \mathbb{P}(\mathrm{L}_1 = k) = p^k q + q^k p}$$

2. Les séries entières $\sum x^n$ et $\sum n x^n$ ont même rayon de convergence égal à 1 ce qui prouve la convergence (absolue) de $\sum k \mathbb{P}(\mathrm{L}_1 = k)$. Par dérivation d'une série, on trouve

$$\forall x \in]-1;1[\qquad \mathrm{S}(x) = \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x} \quad \text{et} \quad x\mathrm{S}'(x) = \sum_{n=0}^{+\infty} n x^n = \frac{x}{(1-x)^2}$$

Ainsi $$\boxed{\mathbb{E}(\mathrm{L}_1) = \frac{p}{q} + \frac{q}{p}}$$

3. On a $\mathrm{L}_2(\Omega) \subset \mathbb{N}^*$ puis

$$\begin{aligned}\forall n \geqslant 1 \qquad \{\mathrm{L}_2 = n\} &= \bigsqcup_{k \in \mathbb{N}^*} \{\mathrm{L}_2 = n, \mathrm{L}_1 = k\} \\ &= \bigsqcup_{k \in \mathbb{N}^*} (\mathrm{P}_1 \ldots \mathrm{P}_k \mathrm{F}_{k+1} \ldots \mathrm{F}_{k+n} \mathrm{P}_{k+n+1} \sqcup \mathrm{F}_1 \ldots \mathrm{F}_k \mathrm{P}_{k+1} \ldots \mathrm{P}_{k+n} \mathrm{F}_{k+n+1})\end{aligned}$$

Ainsi $$\forall n \geqslant 1 \qquad \mathbb{P}(\mathrm{L}_2 = n) = \sum_{k=1}^{+\infty} \left[p^k q^n p + q^k p^n q\right]$$

D'où $$\boxed{\forall n \geqslant 1 \qquad \mathbb{P}(\mathrm{L}_2 = n) = p^2 q^{n-1} + q^2 p^{n-1}}$$

4. Par les mêmes arguments que pour L_1, la variable aléatoire L_2 est d'espérance finie et on trouve

$$\boxed{\mathbb{E}(\mathrm{L}_2) = \frac{p^2}{p^2} + \frac{q^2}{q^2} = 2}$$

5. On a

$$\mathbb{P}(\mathrm{L}_1=1,\mathrm{L}_2=1)=\mathbb{P}(\mathrm{L}_1=1)\mathbb{P}(\mathrm{L}_2=1) \iff pqp+qpq=2pq(p^2+q^2)$$
$$\iff p^2+q^2=\frac{1}{2} \iff p(1-p)=\frac{1}{4} \iff p=\frac{1}{2}$$

Donc, pour $p\neq\dfrac{1}{2}$, les variables aléatoires L_1 et L_2 sont dépendantes. Supposons $p=\dfrac{1}{2}$. Soit $(k,n)\in(\mathbb{N}^*)^2$. On a

$$\mathbb{P}(\mathrm{L}_1=k,\mathrm{L}_2=n)=\frac{1}{2^{k+n+1}}+\frac{1}{2^{k+n+1}}=\frac{1}{2^{k+n}}$$

et $\qquad \mathbb{P}(\mathrm{L}_1=k)=\dfrac{1}{2^{k+1}}+\dfrac{1}{2^{k+1}}=\dfrac{1}{2^k} \qquad \mathbb{P}(\mathrm{L}_2=n)=\dfrac{1}{2^{2+n-1}}+\dfrac{1}{2^{2+n-1}}=\dfrac{1}{2^n}$

Ainsi $\qquad \forall(k,n)\in(\mathbb{N}^*)^2 \qquad \mathbb{P}(\mathrm{L}_1=k,\mathrm{L}_2=n)=\mathbb{P}(\mathrm{L}_1=k)\mathbb{P}(\mathrm{L}_2=n)$

On conclut $\qquad \boxed{\text{Les variables } \mathrm{L}_{1,2} \text{ sont indépendantes si et seulement si } p=\frac{1}{2}.}$

Exercice 2 (****)

Soit $(\Omega,\mathscr{A},\mathbb{P})$ un espace probabilisé, $(\mathrm{X}_n)_n$ et X des variables aléatoires réelles discrètes telles que

$$\forall\varepsilon>0 \qquad \mathbb{P}(|\mathrm{X}_n-\mathrm{X}|\geqslant\varepsilon)\xrightarrow[n\to\infty]{}0$$

Pour x point de continuité de $\mathrm{F}_\mathrm{X}: t\mapsto\mathbb{P}(\mathrm{X}\leqslant t)$, montrer que

$$\mathrm{F}_{\mathrm{X}_n}(x)\xrightarrow[n\to\infty]{}\mathrm{F}_\mathrm{X}(x)$$

Corrigé : Soit $\varepsilon>0$. On a

$$\mathbb{P}(\mathrm{X}_n\leqslant x)=\mathbb{P}(\mathrm{X}_n\leqslant x,|\mathrm{X}_n-\mathrm{X}|<\varepsilon)+\underbrace{\mathbb{P}(\mathrm{X}_n\leqslant x,|\mathrm{X}_n-\mathrm{X}|\geqslant\varepsilon)}_{=\mathrm{o}(1)}$$

Notons $\mathrm{A}_n=\{|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}$. On a donc

$$\mathrm{F}_{\mathrm{X}_n}(x)=\mathbb{P}(\{\mathrm{X}_n\leqslant x\}\cap\mathrm{A}_n)+\mathrm{o}(1)$$

Puis $\qquad \{\mathrm{X}_n\leqslant x,|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}\subset\{\mathrm{X}_n\leqslant x,\mathrm{X}<\mathrm{X}_n+\varepsilon\}\subset\{\mathrm{X}\leqslant x+\varepsilon\}$

d'où $\qquad \mathrm{F}_{\mathrm{X}_n}(x)\leqslant\mathrm{F}_\mathrm{X}(x+\varepsilon)+\mathrm{o}(1)$

Comme $\{|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}\subset\{\mathrm{X}_n<\mathrm{X}+\varepsilon\}$, on a également

$$\begin{aligned}\{\mathrm{X}\leqslant x-\varepsilon,|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}&\subset\{\mathrm{X}\leqslant x-\varepsilon,\mathrm{X}_n<\mathrm{X}+\varepsilon,|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}\\&\subset\{\mathrm{X}_n\leqslant x,|\mathrm{X}_n-\mathrm{X}|<\varepsilon\}\end{aligned}$$

d'où $\qquad \mathbb{P}(\{\mathrm{X}\leqslant x-\varepsilon\}\cap\mathrm{A}_n)\leqslant\mathbb{P}(\{\mathrm{X}_n\leqslant x\}\cap\mathrm{A}_n)=\mathrm{F}_{\mathrm{X}_n}(x)+\mathrm{o}(1)$

Enfin $\qquad \mathrm{F}_\mathrm{X}(x-\varepsilon)=\mathbb{P}(\{\mathrm{X}\leqslant x-\varepsilon\}\cap\mathrm{A}_n)+\underbrace{\mathbb{P}(\{\mathrm{X}\leqslant x-\varepsilon\}\cap\overline{\mathrm{A}_n})}_{=\mathrm{o}(1)}$

d'où $\qquad \mathrm{F}_\mathrm{X}(x-\varepsilon)+\mathrm{o}(1)\leqslant\mathrm{F}_{\mathrm{X}_n}(x)\leqslant\mathrm{F}_\mathrm{X}(x+\varepsilon)+\mathrm{o}(1)$

Soit $\eta>0$. Par continuité de F_X en x, on peut choisir $\varepsilon>0$ tel que

$$|\mathrm{F}_\mathrm{X}(x)-\mathrm{F}_\mathrm{X}(x-\varepsilon)|\leqslant\eta \quad \text{et} \quad |\mathrm{F}_\mathrm{X}(x)-\mathrm{F}_\mathrm{X}(x+\varepsilon)|\leqslant\eta$$

Enfin, on choisit N suffisamment grand pour que les termes en o(1) soient bornés par η pour $n \geqslant \mathrm{N}$. Ainsi, on a trouvé N entier tel que

$$\forall n \geqslant \mathrm{N} \qquad |\mathrm{F}_{\mathrm{X}_n}(x) - \mathrm{F}_{\mathrm{X}}(x)| \leqslant 2\eta$$

Ainsi $\boxed{\text{Pour } x \text{ point de continuité de } \mathrm{F}_{\mathrm{X}}, \text{ on a } \mathrm{F}_{\mathrm{X}_n}(x) \xrightarrow[n\to\infty]{} \mathrm{F}_{\mathrm{X}}(x).}$

Exercice 3 (****)

Soient $(\mathrm{X}_n)_n$ et X des variables aléatoires à valeurs dans $\mathbb{N}$. On note

$$\forall (n,k) \in \mathbb{N}^2 \qquad p_{k,n} = \mathbb{P}(\mathrm{X}_n = k) \quad \text{et} \quad p_k = \mathbb{P}(\mathrm{X} = k)$$

1. Définir la fonction génératrice G_{X} et justifier qu'elle est de classe $\mathscr{C}^\infty$ sur $[\,0\,;1\,[$.
2. On suppose que $p_{k,n} \xrightarrow[n\to\infty]{} p_k$ pour tout k entier. Montrer que la suite de fonctions $(\mathrm{G}_{\mathrm{X}_n})_n$ converge simplement vers G_{X} sur $[\,0\,;1\,[$.
3. Étudier la réciproque.

Corrigé : 1. On définit la fonction génératrice de X notée G_{X} par

$$\boxed{\forall t \in [\,0\,;1\,] \qquad \mathrm{G}_{\mathrm{X}}(t) = \mathbb{E}(t^{\mathrm{X}}) = \sum_{n=0}^{+\infty} t^n \mathbb{P}(\mathrm{X} = n)}$$

La série entière définissant G_{X} a un rayon de convergence supérieur ou égal à 1 puisque $0 \leqslant \mathbb{P}(\mathrm{X} = k) \leqslant 1$ pour tout k entier. D'après le théorème de dérivation des séries entières, on obtient

$$\boxed{\mathrm{G}_{\mathrm{X}} \in \mathscr{C}^\infty([\,0\,;1\,[\,,\mathbb{R})}$$

2. Soit $t \in [\,0\,;1\,[$. On note $u_k : n \mapsto \mathbb{P}(\mathrm{X}_n = k)t^k$. On a $u_k(n) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathrm{X} = k)t^k$ et $|u_k(n)| \leqslant t^k$ avec $\sum t^k$ convergente. Ainsi, la série $\sum u_k$ converge normalement donc uniformément et d'après le théorème de double limite, on obtient

$$\sum_{k=0}^{+\infty} u_k(n) \xrightarrow[n\to\infty]{} \sum_{k=0}^{+\infty} \lim_{n\to+\infty} u_k(n)$$

Ainsi $\boxed{\mathrm{G}_{\mathrm{X}_n} \xrightarrow[n\to\infty]{\mathrm{CS}} \mathrm{G}_{\mathrm{X}} \quad \text{sur} \quad [\,0\,;1\,[}$

3. Supposons $\mathrm{G}_{\mathrm{X}_n} \xrightarrow[n\to\infty]{\mathrm{CS}} \mathrm{G}_{\mathrm{X}}$ sur $[\,0\,;1\,[$. On va utiliser un procédé diagonal sur la suite de suites $\big((p_{k,n})_n\big)_k$. Soit k entier. La suite $(p_{k,n})_n$ est à valeurs dans $[\,0\,;1\,]$. Ainsi, on dispose de φ_1 extractrice telle que

$$p_{1,\varphi_1(n)} \xrightarrow[n\to\infty]{} q_1 \in [\,0\,;1\,]$$

puis de φ_2 extractrice telle que

$$p_{2,\varphi_1\circ\varphi_2(n)} \xrightarrow[n\to\infty]{} q_2 \in [\,0\,;1\,]$$

et en itérant ce procédé, on dispose de φ_k tel que

$$p_{k,\varphi_1\circ\ldots\circ\varphi_k(n)} \xrightarrow[n\to\infty]{} q_k \in [\,0\,;1\,]$$

On pose $\qquad \forall n \in \mathbb{N} \qquad \varphi(n) = \varphi_1 \circ \ldots \circ \varphi_n(n)$

L'application φ est clairement une extractrice (injection de $\mathbb{N} \to \mathbb{N}$ strictement croissante). Pour $n \geqslant k$, on a

$$\varphi(n) = \varphi_1 \circ \ldots \varphi_k(\psi_k(n)) \quad \text{avec} \quad \psi_k(n) = \varphi_{k+1} \circ \ldots \varphi_n(n) \xrightarrow[n\to\infty]{} +\infty$$

Ainsi $$\forall k \in \mathbb{N} \qquad p_{k,\varphi(n)} \xrightarrow[n\to\infty]{} q_k$$

On pose $$\forall t \in [\,0\,;1\,[\qquad \mathrm{G}(t) = \sum_{n=0}^{+\infty} q_k t^k$$

La fonction G est bien définie sur $[\,0\,;1\,[$ puisqu'on $0 \leqslant q_k \leqslant 1$ pour tout k entier. D'après le résultat de la deuxième question, on obtient

$$\mathrm{G}_{\mathrm{X}_{\varphi(n)}} \xrightarrow[n\to\infty]{\mathrm{CS}} \mathrm{G} \quad \text{sur} \quad [\,0\,;1\,[$$

Or, la suite de fonctions $\left(\mathrm{G}_{\mathrm{X}_{\varphi(n)}}\right)_n$ est extraite de la suite $(\mathrm{G}_{\mathrm{X}_n})_n$ simplement convergente. Par unicité de la limite pour la convergence simple, il s'ensuit $\mathrm{G} = \mathrm{G}_\mathrm{X}$, autrement dit

$$\forall t \in [\,0\,;1\,[\qquad \sum_{n=0}^{+\infty} q_k t^k = \sum_{n=0}^{+\infty} p_k t^k$$

Par unicité du développement en série entière, il vient

$$\forall k \in \mathbb{N} \qquad q_k = p_k$$

Par récurrence forte sur k, on obtient que la suite bornée $(p_{k,n})_n$ admet p_k pour unique valeur d'adhérence. L'initialisation pour $k = 1$ est vraie puisque $p_1 = q_1$ avec q_1 une valeur d'adhérence de $(p_{1,n})_n$. On suppose le résultat vrai jusqu'au rang $k-1$. Les suite $(p_{1,n})_n$, …, $(p_{k-1,n})_n$ sont bornées avec une unique valeur d'adhérence donc convergentes et on choisit comme extractrice $\varphi_1 = \ldots = \varphi_{k-1} = \mathrm{id}$ et φ_k une extractrice quelconque telle que $(p_{k,\varphi_1\circ\ldots\circ\varphi_k(n)})_n = (p_{k,\varphi_k(n)})_n$ converge. L'hérédité suit et on conclut

$$\boxed{\forall k \in \mathbb{N} \qquad p_{k,n} \xrightarrow[n\to\infty]{} p_k}$$

Variante : Si on ne pense pas à la diagonale de Cantor, on peut encore aboutir au résultat mais c'est technique. Supposons $\mathrm{G}_{\mathrm{X}_n} \xrightarrow[n\to\infty]{\mathrm{CS}} \mathrm{G}_\mathrm{X}$ sur $[\,0\,;1\,[$. On va montrer qu'il y a en réalité convergence normale sur tout segment de $[\,0\,;1\,[$ pour $(\mathrm{G}_{\mathrm{X}_n})_n$ et ses dérivées. Soit $a \in \,]\,0\,;1\,[$. Les fonctions $\mathrm{G}_{\mathrm{X}_n}$ et G_X sont continues. On peut alors considérer

$$\forall n \in \mathbb{N} \qquad \delta_n = \|\mathrm{G}_{\mathrm{X}_n} - \mathrm{G}_\mathrm{X}\|_{\infty,[\,0\,;a\,]}$$

Par continuité sur un segment, on a

$$\forall n \in \mathbb{N} \qquad \exists t_n \in [\,0\,;a\,] \quad | \quad \delta_n = |\mathrm{G}_{\mathrm{X}_n}(t_n) - \mathrm{G}_\mathrm{X}(t_n)|$$

Par compacité de $[\,0\,;a\,]$, il existe φ extractrice telle que $t_{\varphi(n)} \xrightarrow[n\to\infty]{} t^\star \in [\,0\,;a\,]$. Puis, par inégalité triangulaire

$$\delta_n \leqslant \underbrace{|\mathrm{G}_{\mathrm{X}_n}(t_n) - \mathrm{G}_{\mathrm{X}_n}(t^\star)|}_{=\alpha_n} + \underbrace{|\mathrm{G}_{\mathrm{X}_n}(t^\star) - \mathrm{G}_\mathrm{X}(t^\star)|}_{=\beta_n} + \underbrace{|\mathrm{G}_\mathrm{X}(t^\star) - \mathrm{G}_\mathrm{X}(t_n)|}_{=\gamma_n}$$

Par convergence simple, on a $\beta_n = \mathrm{o}(1)$ et par continuité de G_X, on a $\gamma_{\varphi(n)} = \mathrm{o}(1)$. Puis, par linéarité du symbole Σ, convergence absolue par convergence de $\sum a^k$ et inégalité triangulaire, il vient

$$\alpha_n = \left|\sum_{k=0}^{+\infty} p_{k,n}(t_n^k - t^{\star k})\right| \leqslant \sum_{k=0}^{+\infty} \left|t_n^k - t^{\star k}\right|$$

En commençant la somme à $k = 1$ et avec l'identité de Bernoulli

$$t_n^k - t^{\star k} = (t_n - t^\star)\sum_{j=0}^{k-1} t_n^j t^{\star k-1-j} \qquad \Longrightarrow \qquad \left|t_n^k - t^{\star k}\right| \leqslant |t_n - t^\star|\, k a^{k-1}$$

La série $\sum\limits_{k\geqslant 1} ka^{k-1}$ converge (série entière dérivée) et par suite

$$\alpha_n \leqslant |t_n - t^\star| \sum_{k=1}^{+\infty} ka^{k-1}$$

d'où $\alpha_{\varphi(n)} = \mathrm{o}(1)$ et par conséquent $\delta_{\varphi(n)} = \mathrm{o}(1)$. Puis, on a

$$\forall t \in [\,0\,;1\,[\qquad 0 \leqslant \mathrm{G}_{\mathrm{X}_n}(t) \leqslant \sum_{k=0}^{+\infty} p_{k,n} = 1 \quad \text{et} \quad 0 \leqslant \mathrm{G}_{\mathrm{X}}(t) \leqslant \sum_{k=0}^{+\infty} p_k = 1$$

d'où $$\forall n \in \mathbb{N} \qquad 0 \leqslant \delta_n \leqslant 2$$

La suite $(\delta_n)_n$ est donc à valeurs dans le compact $[\,0\,;2\,]$. Soit ψ une extractrice telle que $\left(\delta_{\psi(n)}\right)_n$ converge. Le résultat précédemment établi sur $(\delta_n)_n$ vaut pour $\left(\delta_{\psi(n)}\right)_n$. Ainsi, il existe φ extractrice telle que $\delta_{\psi\circ\varphi(n)} = \mathrm{o}(1)$. Il s'ensuit que $\delta_{\psi(n)} = \mathrm{o}(1)$ et zéro est donc l'unique valeur d'adhérence de la suite $(\delta_n)_n$. Ainsi, on a $\delta_n = \mathrm{o}(1)$. On procède ensuite par récurrence au rang ℓ pour établir $\delta_n^{(\ell)} = \mathrm{o}(1)$ avec $\delta_n^{(\ell)} = \|\mathrm{G}_{\mathrm{X}_n}^{(\ell)} - \mathrm{G}_{\mathrm{X}}^{(\ell)}\|_{\infty,[\,0\,;a\,]}$. On suit un schéma de preuve analogue et le seul terme délicat à contrôler est

$$\alpha_n^{(\ell)} = \left|\sum_{k=\ell}^{+\infty} p_{k,n} \frac{\ell!}{(k-\ell)!}\left(t_n^{k-\ell} - t^{\star k-\ell}\right)\right| = \left|\sum_{k=0}^{+\infty} \frac{(\ell+k)!}{k!}\left(t_n^{k} - t^{\star k}\right)\right|$$

Toujours avec l'identité de Bernoulli, on obtient

$$\alpha_n^{(\ell)} \leqslant |t_n - t^\star| \sum_{k=0}^{+\infty} \frac{(\ell+k)!}{k!} ka^{k-1} = |t_n - t^\star| \sum_{k=0}^{+\infty} \frac{(\ell+1+k)!}{k!} a^k$$

grâce à la converge de la série $\sum \dfrac{(\ell+1+k)!}{k!} a^k$, série entière dérivée à l'ordre $\ell+1$. L'hérédité s'ensuit. On en déduit

$$\forall k \in \mathbb{N} \qquad \frac{\mathrm{G}_{\mathrm{X}_n}^{(k)}(0)}{k!} \xrightarrow[n\to\infty]{} \frac{\mathrm{G}_{\mathrm{X}}^{(k)}(0)}{k!}$$

Et finalement $$\boxed{\forall k \in \mathbb{N} \qquad p_{k,n} \xrightarrow[n\to\infty]{} p_k}$$

Exercice 4 (***)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et Y une variable aléatoire réelle discrète centrée telle que $\mathrm{Y}(\omega) \subset [\,a\,;b\,]$. Montrer

$$\forall \lambda \in \mathbb{R} \qquad \ln \mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}}) \leqslant \frac{\lambda^2 (b-a)^2}{8}$$

Corrigé : Soit $\mathrm{Z} = \mathrm{Y} - \dfrac{a+b}{2}$. On trouve

$$\mathbb{V}(\mathrm{Y}) = \mathbb{V}(\mathrm{Z}) \leqslant \mathbb{E}(\mathrm{Z}^2) \leqslant \frac{(b-a)^2}{4}$$

Soit λ réel. On pose $$\forall \mathrm{A} \in \mathscr{A} \qquad \mathbb{Q}(\mathrm{A}) = \mathbb{E}\left(\mathbb{1}_{\mathrm{A}} \frac{\mathrm{e}^{\lambda \mathrm{Y}}}{\mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})}\right)$$

On vérifie sans difficulté que $\mathbb{Q}$ définit une probabilité sur $(\Omega, \mathscr{A})$. Soit $f : \mathbb{R} \to \mathbb{R}$ telle que $f(\mathrm{Y})$ est d'espérance finie. Par transfert, il vient

$$\mathbb{E}_{\mathbb{Q}}(f(\mathrm{Y})) = \sum_{y \in \mathrm{Y}(\Omega)} f(y) \mathbb{Q}(\mathrm{Y} = y) = \sum_{y \in \mathrm{Y}(\Omega)} f(y) \mathrm{e}^{\lambda y} \frac{\mathbb{P}(\mathrm{Y} = y)}{\mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})} = \frac{\mathbb{E}(f(\mathrm{Y}) \mathrm{e}^{\lambda \mathrm{Y}})}{\mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})}$$

Puis, d'après la relation de König-Huygens

$$\mathbb{V}_{\mathbb{Q}}(\mathrm{Y}) = \mathbb{E}_{\mathbb{Q}}(\mathrm{Y}^2) - \mathbb{E}_{\mathbb{Q}}(\mathrm{Y})^2 = \frac{\mathbb{E}(\mathrm{Y}^2 \mathrm{e}^{\lambda \mathrm{Y}})}{\mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})} - \left(\frac{\mathbb{E}(\mathrm{Y}\mathrm{e}^{\lambda \mathrm{Y}})}{\mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})}\right)^2$$

On pose $\qquad \forall (n, \lambda) \in \mathbb{N} \times \mathbb{R} \qquad u_n(\lambda) = \mathrm{e}^{\lambda y_n} \mathbb{P}(\mathrm{Y} = y_n)$

Les u_n sont de classe $\mathscr{C}^k$ avec k entier et

$$\forall n \in \mathbb{N} \qquad u_n^{(k)}(\lambda) = y_n^k \mathrm{e}^{\lambda y_n}$$

La fonction $(\lambda, y) \mapsto \mathrm{e}^{\lambda y}$ est bornée sur le compact $[\,\alpha\,;\beta\,] \times [\,a\,;b\,]$ d'où

$$\forall n \in \mathbb{N} \qquad \|u_n^{(k)}\|_{\infty,[\,\alpha\,;\beta\,]} = \mathrm{O}(\mathbb{P}(\mathrm{Y} = y_n))$$

ce qui prouve la convergence normale et donc uniforme de $\sum u_n^{(k)}$. Par dérivation d'une série de fonction, il vient

$$\frac{\mathrm{d}^k}{\mathrm{d}\lambda^k} \mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}}) = \frac{\mathrm{d}^k}{\mathrm{d}\lambda^k} \sum_{n=0}^{+\infty} \mathrm{e}^{\lambda y_n} \mathbb{P}(\mathrm{Y} = y_n) = \sum_{n=0}^{+\infty} y_n^k \mathrm{e}^{\lambda y_n} \mathbb{P}(\mathrm{Y} = y_n) = \mathbb{E}(\mathrm{Y}^k \mathrm{e}^{\lambda \mathrm{Y}})$$

Puis on trouve $\qquad \mathbb{V}_{\mathbb{Q}}(\mathrm{Y}) = \varphi''(\lambda) \quad \text{avec} \quad \varphi(\lambda) = \ln \mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}})$

En observant $\varphi(0) = \varphi'(0) = 0$ et avec $\mathbb{V}_{\mathbb{Q}}(\mathrm{Y}) \leqslant \dfrac{(b-a)^2}{4}$, il vient par intégration

$$\varphi'(\lambda) = \int_0^\lambda \varphi''(s)\,\mathrm{d}s \leqslant \int_0^\lambda \frac{(b-a)^2}{4}\,\mathrm{d}s = \frac{(b-a)^2 \lambda}{4}$$

et $\qquad \varphi(\lambda) = \int_0^\lambda \varphi'(s)\,\mathrm{d}s \leqslant \int_0^\lambda \frac{(b-a)^2 \lambda}{4}\,\mathrm{d}s = \frac{(b-a)^2 \lambda^2}{8}$

On conclut $\qquad \boxed{\forall \lambda \in \mathbb{R} \qquad \ln \mathbb{E}(\mathrm{e}^{\lambda \mathrm{Y}}) \leqslant \dfrac{\lambda^2 (b-a)^2}{8}}$

Remarque : Il s'agit du résultat intermédiaire délicat pour établir l'inégalité de concentration dite *inégalité de Hoeffding*

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

avec $\mathrm{X}_1, \ldots, \mathrm{X}_n$ variables aléatoires discrètes à valeurs dans $[\,a\,;b\,]$, $\mathrm{S}_n = \sum\limits_{i=1}^{n} \mathrm{X}_i$ et $\varepsilon > 0$.

Exercice 5 (****)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Par commodité, on confond vecteur et matrice colonne. Soit $\mathrm{U} = (\mathrm{U}_1, \ldots, \mathrm{U}_n)$ un vecteur aléatoire discret. Le vecteur U est *d'espérance finie* si les U_i le sont et on pose $\mathbb{E}(\mathrm{U}) = (\mathbb{E}(\mathrm{U}_1), \ldots, \mathbb{E}(\mathrm{U}_n))$. Le vecteur U est dans L^2 si les $\mathrm{U}_i \mathrm{U}_j$ sont d'espérance finie et on définit $\mathbb{V}(\mathrm{U})$ sa *matrice de covariance* par

$$\mathbb{V}(\mathrm{U}) = \mathbb{E}\left[(\mathrm{U} - \mathbb{E}(\mathrm{U}))(\mathrm{U} - \mathbb{E}(\mathrm{U}))^{\mathrm{T}}\right]$$

On considère le *modèle de régression* décrit par

$$\mathrm{Y} = \mathrm{A}\theta + \varepsilon$$

avec $\mathrm{A} \in \mathscr{M}_{n,p}(\mathbb{R})$ telle que $\operatorname{rg} \mathrm{A} = p$, $\theta \in \mathscr{M}_{p,1}(\mathbb{R})$ un paramètre inconnu et ε un vecteur aléatoire discret dans L^2 avec $\mathbb{E}(\varepsilon) = 0$ et $\mathbb{V}(\varepsilon) = \sigma^2 \mathrm{I}_n$. La grandeur d'intérêt dans ce modèle est le paramètre θ. On observe le vecteur Y et on cherche à estimer finement θ à partir de cette observation Y. On appelle *estimateur sans biais* de θ une vecteur aléatoire $\mathrm{T(Y)}$ avec $\mathrm{T} : \mathscr{M}_{n,1}(\mathbb{R}) \to \mathscr{M}_{p,1}(\mathbb{R})$ tel que $\mathrm{T(Y)}$ est d'espérance finie avec $\mathbb{E}(\mathrm{T(Y)}) = \theta$.

1. Déterminer le vecteur aléatoire $\widehat{\theta}$ qui minimise $\|\mathrm{Y} - \mathrm{AX}\|^2$ pour $\mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})$.
2. Montrer que $\widehat{\theta}$ est un estimateur sans biais de θ.
3. Soit $\widetilde{\theta} = \mathrm{LY}$ avec $\mathrm{L} \in \mathscr{M}_{p,n}(\mathbb{R})$. Déterminer une condition nécessaire et suffisante sur la matrice L pour que $\widetilde{\theta}$ soit un estimateur sans biais de θ et ce pour tout $\theta \in \mathscr{M}_{p,1}(\mathbb{R})$.
4. On suppose que la matrice L vérifie la condition précédemment établie. Établir
$$\mathbb{V}(\widetilde{\theta}) - \mathbb{V}(\widehat{\theta}) \in \mathscr{S}_n^+(\mathbb{R})$$
5. Conclure que $$\mathbb{E}(\|\widetilde{\theta} - \theta\|^2) \geqslant \mathbb{E}(\|\widehat{\theta} - \theta\|^2)$$

Corrigé : 1. L'ensemble $\{\|\mathrm{Y} - \mathrm{AX}\|^2, \mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})\}$ est une partie non vide minorée de $\mathbb{R}$ et admet donc une borne inférieure finie. On a $\{\mathrm{AX}, \mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})\} = \mathrm{Im\ A}$ puis, par continuité et croissance de $u \mapsto u^2$ sur $\mathbb{R}_+$, il vient
$$\operatorname*{Inf}_{\mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})} \|\mathrm{Y} - \mathrm{AX}\|^2 = \left(\operatorname*{Inf}_{\mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})} \|\mathrm{Y} - \mathrm{AX}\|\right)^2 = \left(\operatorname*{Inf}_{\mathrm{Z} \in \mathrm{Im\ A}} \|\mathrm{Y} - \mathrm{Z}\|\right)^2 = \mathrm{d(Y, Im\ A)}^2$$
Le sev Im A est de dimension finie et par caractérisation métrique du projeté orthogonal, on a
$$\mathrm{d(Y, Im\ A)} = \|\mathrm{Y} - p_{\mathrm{Im\ A}}(\mathrm{Y})\|$$
Comme $p_{\mathrm{Im\ A}}(\mathrm{Y}) \in \mathrm{Im\ A}$, il existe $\widehat{\theta}$ vecteur aléatoire dans $\mathscr{M}_{p,1}(\mathbb{R})$ tel que $p_{\mathrm{Im\ A}}(\mathrm{Y}) = \mathrm{A}\widehat{\theta}$ ce qui prouve que la borne inférieure considérée est effectivement un minimum. On a l'unicité de $\widehat{\theta}$ car l'équation $\mathrm{AX} = p_{\mathrm{Im\ A}}(\mathrm{Y})$ admet une unique solution par injectivité de A puisque $\mathrm{rg\ A} = p$. D'après la caractérisation géométrique du projeté orthogonal, il vient
$$\forall \mathrm{Z} \in \mathrm{Im\ A} \qquad \langle \mathrm{Z}, \mathrm{Y} - p_{\mathrm{Im\ A}}(\mathrm{Y})\rangle = 0$$
c'est-à-dire $$\forall \mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R}) \qquad \left\langle \mathrm{AX}, \mathrm{Y} - \mathrm{A}\widehat{\theta}\right\rangle = 0$$
autrement dit, pour tout $\mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R})$, on trouve $(\mathrm{AX})^{\mathrm{T}}\left(\mathrm{Y} - \mathrm{A}\widehat{\theta}\right) = 0$ ce qui s'interprète, en munissant $\mathscr{M}_{p,1}(\mathbb{R})$ de sa structure euclidienne canonique
$$\forall \mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R}) \qquad \left\langle \mathrm{X}, \mathrm{A}^\top\mathrm{Y} - \mathrm{A}^\top\mathrm{A}\widehat{\theta}\right\rangle = 0$$
d'où $$\mathrm{A}^{\mathrm{T}}\mathrm{Y} - \mathrm{A}^{\mathrm{T}}\mathrm{A}\widehat{\theta} = 0$$
On a l'égalité $\mathrm{rg\ A}^{\mathrm{T}}\mathrm{A} = \mathrm{rg\ A} = p$ et comme $\mathrm{A}^{\mathrm{T}}\mathrm{A} \in \mathscr{M}_p(\mathbb{R})$, il en résulte que $\mathrm{A}^{\mathrm{T}}\mathrm{A} \in \mathrm{GL}_p(\mathbb{R})$ et on conclut
$$\boxed{\widehat{\theta} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{Y}}$$
2. Notons $\mathrm{B} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}\mathrm{A}^{\mathrm{T}}$. Il vient
$$\widehat{\theta} = \mathrm{BY} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{A}\theta + \mathrm{B}\varepsilon = \theta + \mathrm{B}\varepsilon$$
Le vecteur aléatoire ε est dans L^2 donc est d'espérance finie et $\mathrm{B}\varepsilon$ également par combinaison linéaire. Par linéarité de l'espérance, on obtient
$$\mathbb{E}(\mathrm{B}\varepsilon) = \mathrm{B}\mathbb{E}(\varepsilon) = 0$$
On conclut $$\boxed{\text{Le vecteur aléatoire } \widehat{\theta} \text{ est un estimateur sans biais de } \theta.}$$

3. La variable Y est d'espérance finie et $\widetilde{\theta}$ également par combinaison linéaire. On trouve par linéarité de l'espérance

$$\mathbb{E}(\widetilde{\theta}) = \mathbb{E}(\mathrm{LY}) = \mathbb{E}(\mathrm{LA}\theta + \mathrm{L}\varepsilon) = \mathrm{LA}\theta + \mathrm{L}\mathbb{E}(\varepsilon) = \mathrm{LA}\theta$$

Le vecteur aléatoire $\widetilde{\theta}$ est un estimateur sans biais de θ pour tout $\theta \in \mathscr{M}_{p,1}(\mathbb{R})$ si et seulement si

$$\forall \theta \in \mathscr{M}_{p,1}(\mathbb{R}) \qquad \mathrm{LA}\theta = \theta$$

Par conséquent $\boxed{\text{La condition nécessaire et suffisante attendue est } \mathrm{LA} = \mathrm{I}_p.}$

4. On a
$$\widetilde{\theta} = \mathrm{LY} = \mathrm{L}\,(\mathrm{A}\theta + \varepsilon) = \theta + \mathrm{L}\varepsilon$$

puis
$$\left(\widetilde{\theta} - \theta\right)\left(\widetilde{\theta} - \theta\right)^{\mathrm{T}} = (\mathrm{L}\varepsilon)(\mathrm{L}\varepsilon)^{\mathrm{T}} = \mathrm{L}\varepsilon\varepsilon^{\mathrm{T}}\mathrm{L}^{\mathrm{T}}$$

Par combinaison linéaire, il s'ensuit que $\left(\widetilde{\theta} - \theta\right)\left(\widetilde{\theta} - \theta\right)^{\mathrm{T}}$ est d'espérance finie ce qui prouve que le vecteur $\widetilde{\theta}$ est dans L^2. On remarque

$$\mathrm{BA} = \left(\mathrm{A}^{\mathrm{T}}\mathrm{A}\right)^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{A} = \mathrm{I}_p$$

donc la matrice B vérifie la même condition que la matrice L et comme pour $\widetilde{\theta}$, le vecteur $\widehat{\theta}$ est dans L^2. On trouve par linéarité de l'espérance

$$\begin{aligned}\mathbb{V}(\widetilde{\theta}) &= \mathbb{E}\left[\left(\widetilde{\theta} - \widehat{\theta} + \widehat{\theta} - \theta\right)\left(\widetilde{\theta} - \widehat{\theta} + \widehat{\theta} - \theta\right)^{\mathrm{T}}\right] \\ &= \underbrace{\mathbb{E}\left[\left(\widetilde{\theta} - \widehat{\theta}\right)\left(\widetilde{\theta} - \widehat{\theta}\right)^{\mathrm{T}}\right]}_{\mathbb{V}(\widetilde{\theta} - \widehat{\theta})} + \mathbb{E}\left[\left(\widetilde{\theta} - \widehat{\theta}\right)\left(\widehat{\theta} - \theta\right)^{\mathrm{T}}\right] + \mathbb{E}\left[\left(\widehat{\theta} - \theta\right)\left(\widetilde{\theta} - \widehat{\theta}\right)^{\mathrm{T}}\right] + \underbrace{\mathbb{E}\left[\left(\widehat{\theta} - \theta\right)\left(\widehat{\theta} - \theta\right)^{\mathrm{T}}\right]}_{\mathbb{V}(\widehat{\theta})}\end{aligned}$$

Puis
$$\begin{aligned}\mathbb{E}\left[\left(\widetilde{\theta} - \widehat{\theta}\right)\left(\widehat{\theta} - \theta\right)^{\mathrm{T}}\right] &= \mathbb{E}\left[(\theta + \mathrm{L}\varepsilon - \theta - \mathrm{B}\varepsilon)(\theta + \mathrm{B}\varepsilon - \theta)^{\mathrm{T}}\right] \\ &= (\mathrm{L} - \mathrm{B})\mathbb{E}(\varepsilon\varepsilon^{\mathrm{T}})\mathrm{B}^{\mathrm{T}} = \sigma^2(\mathrm{L} - \mathrm{B})\mathrm{B}^{\mathrm{T}} = \sigma^2(\mathrm{LB}^{\mathrm{T}} - \mathrm{BB}^{\mathrm{T}})\end{aligned}$$

On a $\quad \mathrm{LB}^{\mathrm{T}} = \mathrm{LA}(\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1} \quad$ et $\quad \mathrm{BB}^{\mathrm{T}} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{A}(\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1} = (\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}$

Par conséquent
$$\mathbb{E}\left[\left(\widetilde{\theta} - \widehat{\theta}\right)\left(\widehat{\theta} - \theta\right)^{\mathrm{T}}\right] = 0$$

et par transposition
$$\mathbb{E}\left[\left(\widehat{\theta} - \theta\right)\left(\widetilde{\theta} - \widehat{\theta}\right)^{\mathrm{T}}\right] = 0$$

Finalement
$$\mathbb{V}(\widetilde{\theta}) = \mathbb{V}(\widetilde{\theta} - \widehat{\theta}) + \mathbb{V}(\widehat{\theta})$$

Enfin, par linéarité de l'espérance, on a

$$\forall \mathrm{X} \in \mathscr{M}_{p,1}(\mathbb{R}) \qquad \mathrm{X}^{\top}\mathbb{V}(\widetilde{\theta} - \widehat{\theta})\mathrm{X} = \mathbb{E}\left(\mathrm{X}^{\mathrm{T}}(\widetilde{\theta} - \widehat{\theta})(\widetilde{\theta} - \widehat{\theta})^{\mathrm{T}}\mathrm{X}\right) = \mathbb{V}\big(\underbrace{\mathrm{X}^{\mathrm{T}}(\widetilde{\theta} - \widehat{\theta})}_{\in\mathbb{R}}\big) \geqslant 0$$

On conclut
$$\boxed{\mathbb{V}(\widetilde{\theta}) - \mathbb{V}(\widehat{\theta}) \in \mathscr{S}_n^+(\mathbb{R})}$$

5. On note $(\mathrm{E}_j)_{1 \leqslant j \leqslant p}$ la base canonique de $\mathscr{M}_{p,1}(\mathbb{R})$. On a

$$\|\widetilde{\theta} - \theta\|^2 = \sum_{j=1}^{p}\left(\widetilde{\theta}_j - \theta_j\right)^2 = \sum_{j=1}^{p}\mathrm{E}_j^{\mathrm{T}}(\widetilde{\theta} - \theta)(\widetilde{\theta} - \theta)^{\mathrm{T}}\mathrm{E}_j$$

Par linéarité de l'espérance, il vient

$$\mathbb{E}(\|\widetilde{\theta}-\theta\|^2) = \sum_{j=1}^{p} \mathrm{E}_j^{\mathrm{T}} \mathbb{V}(\widetilde{\theta}) \mathrm{E}_j$$

En procédant de même avec $\widehat{\theta}$, on obtient

$$\mathbb{E}(\|\widetilde{\theta}-\theta\|^2) - \mathbb{E}(\|\widehat{\theta}-\theta\|^2) = \sum_{j=1}^{p} \mathrm{E}_j^{\mathrm{T}} \left(\mathbb{V}(\widetilde{\theta}) - \mathbb{V}(\widehat{\theta})\right) \mathrm{E}_j \geqslant 0$$

On conclut $$\boxed{\mathbb{E}(\|\widetilde{\theta}-\theta\|^2) \geqslant \mathbb{E}(\|\widehat{\theta}-\theta\|^2)}$$

Remarque : On a démontré le *théorème de Gauss-Markov* : dans un modèle de régression linéaire avec des erreurs centrées de même variance, l'estimateur des moindres carrés $\widehat{\theta}$ est d'erreur quadratique minimale parmi les estimateurs sans biais de θ.