Oral ESCP - Europe n° 1

★

A. Le sujet

A.1. Exercice avec préparation

Dans tout le problème, N désigne un entier naturel non nul.

Soit $(U_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et suivant toutes la loi uniforme sur $[\![1, N]\!]$.

On note T_n et Z_n les deux variables aléatoires définies, pour tout entier naturel n supérieur ou égal à 1, par :

$$T_n = \sup(U_1, U_2, \dots, U_n) \quad \text{et} \quad Z_n = \inf(U_1, U_2, \dots, U_n).$$

On pose $S_n = T_n + Z_n - 1$. On pose enfin, pour tout entier naturel n non nul :

$$a_n(N) = \begin{cases} \displaystyle\sum_{k=1}^{N-1} \left(\frac{k}{N}\right)^n & \text{si } N \geqslant 2 \\ 0 & \text{si } N = 1 \end{cases}.$$

1. Soit Y une variable aléatoire à valeurs dans $[\![1, N]\!]$. Établir la relation suivante :
$$\mathbb{E}(Y) = \sum_{k=0}^{N-1} \mathbb{P}(Y > k).$$

Dans la suite de l'exercice, on cherche à estimer le paramètre N.

2. (a) Calculer, pour tout entier naturel k appartenant à $[\![1, N]\!]$, $\mathbb{P}(T_n \leqslant k)$.
   (b) En déduire la loi de T_n.
   (c) Calculer $\mathbb{E}(T_n)$ en fonction de N et de $a_n(N)$.
3. (a) Calculer, pour tout entier naturel k appartenant à $[\![0, N-1]\!]$, la probabilité $\mathbb{P}(Z_n > k)$.
   (b) En déduire $\mathbb{E}(Z_n)$ en fonction de $a_n(N)$.
4. (a) $T_n + a_n(N)$ est-il un estimateur sans biais de N ?
   (b) Montrer que T_n est un estimateur asymptotiquement sans biais de N.
   (c) Montrer que S_n est un estimateur sans biais de N.

A.2. Exercice sans préparation

Soit E un espace euclidien de dimension n. Soient f et g deux endomorphismes de E tels que $f \circ g = g \circ f$. On note S (resp. T) la matrice de f (resp. g) dans une base orthonormale $\mathcal{B}$ de E. On suppose que S est symétrique et T antisymétrique.

Montrer que :

$$\forall x \in E,\ \|(f-g)(x)\| = \|(f+g)(x)\|\,.$$

B. Indications et méthodes de résolution

B.1. Exercice avec préparation

1. Pour tout $k \in [\![1, N]\!]$, exprimer $\mathbb{P}(Y = k)$ en fonction de $\mathbb{P}(Y > k-1)$ et $\mathbb{P}(Y > k)$.
2. (a) Remarquer que $[T_n \leqslant k]$ se réalise si et seulement si, pour tout $i \in [\![1, n]\!]$, $[U_i \leqslant k]$ se réalise puis utiliser l'indépendance des variables aléatoires $U_1, \ldots, U_n$.
3. (a) Remarquer que $[Z_n > k]$ se réalise si et seulement si, pour tout $i \in [\![1, n]\!]$, $[U_i > k]$ se réalise puis utiliser l'indépendance des variables aléatoires $U_1, \ldots, U_n$.

B.2. Exercice sans préparation

Exprimer $\|(f-g)(x)\|^2 - \|(f+g)(x)\|^2$ en fonction de $\langle f(x), g(x)\rangle$ puis de S et T et utiliser le fait que S soit une matrice symétrique et T une matrice antisymétrique pour établir que ce produit scalaire est nul.

C. Analyse et correction du sujet

C.1. Exercice avec préparation

1. Comme Y prend un nombre fini de valeurs, elle admet une espérance et on a :

$$\mathbb{E}(Y) = \sum_{k=1}^{N} k\, \mathbb{P}(Y = k)$$

soit encore, comme Y prend ses valeurs dans $\mathbb{N}$:

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{k=1}^{N} k\,[\mathbb{P}(Y > k-1) - \mathbb{P}(Y > k)] \\
&= \sum_{k=1}^{N} k\, \mathbb{P}(Y > k-1) - \sum_{k=1}^{N} k\, \mathbb{P}(Y > k) \\
&= \sum_{k=0}^{N-1} (k+1)\, \mathbb{P}(Y > k) - \sum_{k=1}^{N} k\, \mathbb{P}(Y > k) \\
&= \sum_{k=0}^{N-1} (k+1-k)\, \mathbb{P}(Y > k) - N\, \mathbb{P}(Y > N) + 0 \times \mathbb{P}(Y > 0)
\end{aligned}$$

et donc, comme $\mathbb{P}(Y > N) = 0$ (car Y prend ses valeurs dans $[\![1, N]\!]$) :

$$\boxed{\mathbb{E}(Y) = \sum_{k=0}^{N-1} \mathbb{P}(Y > k)}$$

Alerte, résultat classique !
Bien retenir ce résultat, ainsi que sa preuve. On pourra également retenir que l'on peut établir, si X est une variable aléatoire à valeurs dans $\mathbb{N}$, que X admet une espérance si et seulement si la série $\sum \mathbb{P}(X > n)$ converge et que, dans ce cas :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n).$$

Rappel de cours
Si X est une variable aléatoire prenant ses valeurs dans $\mathbb{Z}$, alors :

$$\begin{aligned}\forall k \in \mathbb{Z},\ \mathbb{P}(X = k) &= \mathbb{P}(X \leqslant k) - \mathbb{P}(X \leqslant k-1)\\ &= \mathbb{P}(X > k-1) - \mathbb{P}(X > k).\end{aligned}$$

2. (a) On a :

$$\forall k \in [\![1, N]\!]\,,\ [T_n \leqslant k] = \bigcap_{i=1}^{n} [U_i \leqslant k]$$

donc, comme $U_1, \dots, U_n$ sont indépendantes :

$$\forall k \in [\![1, N]\!]\,,\ \mathbb{P}(T_n \leqslant k) = \prod_{i=1}^{n} \mathbb{P}(U_i \leqslant k)$$

et comme $U_1, \dots, U_n$ suivent toutes la loi uniforme sur $[\![1, N]\!]$:

$$\boxed{\forall k \in [\![1, N]\!]\,,\ \mathbb{P}(T_n \leqslant k) = \left(\frac{k}{N}\right)^n}$$

Rappel de cours
Si $(X_i)_{1 \leqslant i \leqslant n}$ est une famille de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors :

- $S_n = \sup(X_1, \dots, X_n)$ *est la variable aléatoire définie par :*

$$\forall \omega \in \Omega,\ S_n(\omega) = \max(X_1(\omega), \dots, X_n(\omega)),$$

- $I_n = \inf(X_1, \dots, X_n)$ *est la variable aléatoire définie par :*

$$\forall \omega \in \Omega,\ I_n(\omega) = \min(X_1(\omega), \dots, X_n(\omega)).$$

Méthode
Plus généralement, retenir que, si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes et de même loi, on a toujours, en notant $S_n = \sup(X_1, \ldots, X_n)$:

$$\begin{aligned}\forall x \in \mathbb{R},\ \mathbb{P}(S_n \leqslant x) &= \mathbb{P}\left(\bigcap_{k=1}^{n} [X_k \leqslant x]\right) \\ &= [\mathbb{P}(X_1 \leqslant x)]^n .\end{aligned}$$

(b) Comme $U_1, \ldots, U_n$ prennent leurs valeurs dans $[\![1, N]\!]$, T_n prend ses valeurs dans $[\![1, N]\!]$ (donc en particulier dans $\mathbb{N}$) et on a :

$$\forall k \in [\![1, N]\!]\,,\ \mathbb{P}(T_n = k) = \mathbb{P}(T_n \leqslant k) - \mathbb{P}(T_n \leqslant k-1)$$

et donc, d'après le résultat de la question précédente et en remarquant que $\mathbb{P}(T_n \leqslant 0) = 0 = \left(\dfrac{0}{N}\right)^n$:

$$\boxed{\begin{array}{c}T_n \text{ prend ses valeurs dans } [\![1, N]\!] \text{ et sa loi est caractérisée par :} \\ \forall k \in [\![1, N]\!]\,,\ \mathbb{P}(T_n = k) = \left(\dfrac{k}{N}\right)^n - \left(\dfrac{k-1}{N}\right)^n .\end{array}}$$

(c) Comme T_n prend ses valeurs dans $[\![1, N]\!]$, le résultat de la question 1 s'applique et :

$$\begin{aligned}\mathbb{E}(T_n) &= \sum_{k=0}^{N-1} \mathbb{P}(T_n > k) \\ &= \sum_{k=0}^{N-1} [1 - \mathbb{P}(T_n \leqslant k)]\end{aligned}$$

et donc :

$$\boxed{\mathbb{E}(T_n) = N - a_n(N)}$$

3. (a) On a :

$$\forall k \in [\![0, N-1]\!]\,,\ [Z_n > k] = \bigcap_{i=1}^{n} [U_i > k]$$

donc, comme $U_1, \ldots, U_n$ sont indépendantes :

$$\forall k \in [\![0, N-1]\!]\,,\ \mathbb{P}(Z_n > k) = \prod_{i=1}^{n} \mathbb{P}(U_i > k)$$

et comme $U_1, \ldots, U_n$ suivent toutes la loi uniforme sur $[\![1, N]\!]$:

$$\boxed{\forall k \in [\![0, N-1]\!]\,,\ \mathbb{P}(Z_n > k) = \left(1 - \frac{k}{N}\right)^n}$$

Méthode
Plus généralement, retenir que, si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes et de même loi, on a toujours, en notant $I_n = \inf(X_1, \ldots, X_n)$:

$$\begin{aligned}\forall x \in \mathbb{R},\ \mathbb{P}(I_n > x) &= \mathbb{P}\left(\bigcap_{k=1}^{n} [X_k > x]\right)\\ &= [\mathbb{P}(X_1 > x)]^n\,.\end{aligned}$$

(b) Comme $U_1, \ldots, U_n$ prennent leurs valeurs dans $[\![1, N]\!]$, Z_n prend ses valeurs dans $[\![1, N]\!]$, le résultat de la question 1 s'applique et :

$$\begin{aligned}\mathbb{E}(Z_n) &= \sum_{k=0}^{N-1} \mathbb{P}(Z_n > k)\\ &= \sum_{k=0}^{N-1} \left(\frac{N-k}{N}\right)^n\\ &= \sum_{i=1}^{N} \left(\frac{i}{N}\right)^n\end{aligned}$$

et donc :

$$\boxed{\mathbb{E}(Z_n) = 1 + a_n(N)}$$

4. (a) Comme $T_n + a_n(N)$ dépend de N, ce n'est pas un estimateur de N, donc :

$$\boxed{\text{Non}}$$

Alerte, erreur fréquente !
Quand l'énoncé demande de justifier qu'une variable aléatoire T_n est un estimateur sans biais d'un paramètre réel θ, trop de candidat se contentent de justifier que T_n admet une espérance et que celle-ci est égale à θ : il ne faut pas oublier de justifier également que T_n est un estimateur de θ.

Rappel de cours
Soient θ un paramètre réel inconnu appartenant à un ensemble Θ et g une application de Θ dans $\mathbb{R}$.
Un estimateur de $g(\theta)$ est une variable aléatoire qui peut s'écrire sous la forme $T_n = \varphi(X_1, \ldots, X_n)$ où $X_1, \ldots, X_n$ sont des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes, de même loi qu'une variable aléatoire X dont la loi dépend de θ et φ est une application de $(X(\Omega))^n$ dans $\mathbb{R}$ ***indépendante de θ****.*
Si $(T_n)_{n\in\mathbb{N}^*}$ est une suite d'estimateurs de $g(\theta)$, on dit que :

- *T_n est un estimateur sans biais de $g(\theta)$ si T_n admet une espérance, égale) $g(\theta)$,*
- *la suite $(T_n)_{n\in\mathbb{N}^*}$ est une suite d'estimateurs asymptotiquement sans biais de $g(\theta)$ si T_n admet une espérance pour tout $n \in \mathbb{N}^*$ et si : $\lim\limits_{n\to+\infty} \mathbb{E}(T_n) = g(\theta)$,*
- *la suite $(T_n)_{n\in\mathbb{N}^*}$ est une suite convergentes d'estimateurs de $g(\theta)$ si : $T_n \xrightarrow{\mathbb{P}} g(\theta)$.*

(b) On peut remarquer que $T_n = \varphi(U_1, \ldots, U_n)$ où

$$\varphi : (x_1, \ldots, x_n) \mapsto \sup(x_1, \ldots, x_n)$$

est une fonction définie sur $\mathbb{R}^n$ et indépendante de N donc, comme $(U_1, \ldots, U_n)$ est un n-échantillon i.i.d. dont la loi dépend de N, T_n est un estimateur de N.
De plus, on a vu que T_n admet une espérance et que :

$$\mathbb{E}(T_n) = N - \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n$$

et donc, comme la somme est finie et comme chacun de ses termes tend vers 0 lorsque n tend vers $+\infty$ (ce sont des termes généraux de suites géométriques dont la raison $\frac{k}{N}$ appartient à $[0, 1[$) :

$$\lim_{n\to+\infty} \mathbb{E}(T_n) = N,$$

ce qui nous permet de conclure :

$(T_n)_{n\in\mathbb{N}^*}$ est une suite d'estimateurs asymptotiquement sans biais de N.

(c) On a : $S_n = \psi(U_1, \ldots, U_n)$ où

$$\psi : (x_1, \ldots, x_n) \mapsto \sup(x_1, \ldots, x_n) + \inf(x_1, \ldots, x_n) - 1$$

est une fonction définie sur $\mathbb{R}^n$ et indépendante de N donc, comme $(U_1, \ldots, U_n)$ est un n-échantillon i.i.d. dont la loi dépend de N, S_n est un estimateur de N.

De plus, on a vu que T_n et Z_n admettent une espérance, donc S_n admet une espérance et, par linéarité de l'espérance, on a :

$$\mathbb{E}(S_n) = \mathbb{E}(T_n) + \mathbb{E}(Z_n) - 1 = N,$$

ce qui nous permet de conclure :

$$\boxed{S_n \text{ est un estimateur sans biais de } N}$$

C.2. Exercice sans préparation

◇ Soit $x \in E$. On peut déjà remarquer que :

$$\begin{aligned} ||(f-g)(x)||^2 &= ||f(x) - g(x)||^2 \\ &= ||f(x)||^2 - 2\langle f(x), g(x)\rangle + ||g(x)||^2 \\ &= ||f(x) + g(x)||^2 - 4\langle f(x), g(x)\rangle \\ &= ||(f+g)(x)||^2 - 4\langle f(x), g(x)\rangle . \end{aligned}$$

et donc :

$$||(f-g)(x)||^2 - ||(f+g)(x)||^2 = -4\langle f(x), g(x)\rangle .$$

◇ On note alors X la colonne des coordonnées de x dans la base $\mathcal{B}$. Alors SX et TX sont les colonnes des coordonnées respectives de $f(x)$ et $g(x)$ dans la base $\mathcal{B}$ et comme celle-ci est orthonormale :

$$\begin{aligned} \langle f(x), g(x)\rangle &= {}^t(SX)\,(TX) \\ &= {}^tX\,{}^tS\,TX \end{aligned}$$

et donc, comme ${}^tS = S$:

$$\begin{aligned} \langle f(x), g(x)\rangle &= {}^t(SX)\,(TX) \\ &= {}^tX\,ST\,X. \end{aligned}$$

De plus, comme f et g commutent et comme S et T sont les matrices associées à f et g dans une même base, S et T commutent, donc :

$$\begin{aligned} \langle f(x), g(x)\rangle &= {}^tX\,TS\,X \\ &= {}^t({}^tTX)\,(SX) \end{aligned}$$

et comme ${}^tT = -T$:

$$\begin{aligned} \langle f(x), g(x)\rangle &= -{}^t(TX)\,(SX) \\ &= -\langle g(x), f(x)\rangle , \end{aligned}$$

et donc, par symétrie du produit scalaire :

$$\langle f(x), g(x)\rangle = 0,$$

ce qui nous permet de conclure :

$$\boxed{\forall x \in E,\ ||(f-g)(x)||^2 = ||(f+g)(x)||^2}$$

ESCP - Europe

Rappel de cours

*Soient E un espace euclidien et $\mathcal{B}$ une base **orthonormale** de E et f un endomorphisme de E, x et y deux vecteurs de E. Si on note A la matrice représentative de f dans la base $\mathcal{B}$, X et Y les colonnes des coordonnées respectives de x et y dans $\mathcal{B}$, alors on a :*

$$\langle x, y\rangle = {}^tXY = {}^tYX \quad \text{et} \quad \langle f(x), y\rangle = {}^t(AX)Y = {}^tYAX.$$

Alerte, résultat classique !

On aurait aussi pu utiliser le fait que, si E est un espace euclidien, alors :

$$\forall (x,y) \in E^2,\ \langle x, y\rangle = \frac{||x+y||^2 - ||x-y||^2}{4} = \frac{||x+y||^2 - ||x||^2 - ||y||^2}{2}.$$