Exercice| [5064] | 1| Couple de variables aléatoires | HEC|ESSEC 2011 Filière ECS

Toutes les variables aléatoires qui apparaissent dans ce problème sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$.
On note F_Z la fonction de répartition d'une variable aléatoire Z et, si cette variable aléatoire admet une densité, on note f_Z une densité de Z.
Sous réserve d'existence, on note $E(Z)$ et $v(Z)$ respectivement, l'espérance et la variance d'une variable aléatoire réelle Z, $\operatorname{Cov}(Z_1, Z_2)$ la covariance de deux variables aléatoires Z_1 et Z_2.
La fonction exponentielle est notée $\exp$ et la partie entière d'un réel x est notée $\lfloor x \rfloor$?
On admet les résultats suivants :

- la définition et les propriétés de la covariance et du coefficient de corrélation linéaire de deux vzariables aléatoires discrètes, s'appliquent au cas de variables aléatoires à densité ;
- si Z_1 et Z_2 sont deux variables aléatoires indépendantes admettant une espérance, alors la variable aléatoire Z_1Z_2 admet une espérance et $E(Z_1Z_2) = E(Z_1)E(Z_2)$.

Dans tout le problème, on considère une variable aléatoire X de fonction de répartition F_X et admettant une densité f_X.
Les solutions éventuelles de l'équation $F_X(x) = \dfrac{1}{2}$ s'appellent les *médianes théoriques* de X.
Pour n entier de $\mathbb{N}^*$, on considère un $n-$échantilon $(X_1, \ldots, X_n)$ i.i.d. (indépendant, identiquement distribué) de la loi de X, et on définit la variable aléatoire $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$, qui est la *moyenne empirique de l'échantillon* $(X_1, \ldots, X_n)$.
On admet l'existence de variables aléatoires à densité $Y_1, \ldots, Y_n$ telles que, pour tout $\omega \in \Omega$, les réels $Y_1(\omega), \ldots, Y_n(\omega)$ constituent un réarrangement par ordre croissant des réels $X_1(\omega), \ldots, X_n(\omega)$, de telle sorte que, pour tout $\omega \in \Omega$: $Y_1(\omega) \leq Y_2(\omega) \leq \ldots \leq Y_n(\omega)$.
En particulier, $Y_1 = \inf(X_1, \ldots, X_n)$, et $Y_n = \sup(X_1, \ldots, X_n)$. Plus généralement, pour tout $k \in [\![1;n]\!]$, il existe une fonction ψ_k définie et continue sur $\mathbb{R}^n$ à valeurs réelles, telle que $Y_n = \psi_k(X_1, \ldots, X_n)$.
Si n est un entier impair ($n = 2\ell + 1$ avec $\ell \in \mathbb{N}$), alors la variable aléatoire $Y_{\ell+1}$ est appelée la *médiane empirique de l'échantillon* $(X_1, \ldots, X_n)$.
La partie II du problème est indépendante de la partie I.

Partie I - Quelques propriétés des statistiques d'ordre

Pour tout réel x et tout entier $k \in [\![1;n]\!]$, on note $J_k(x)$ la variable aléatoire de Bernoulli définie par :

$$J_k(x) = \begin{cases} 1 \text{ si } [X_k \leq x] \text{ est réalisé} \\ 0 \text{ si } [X_k > x] \text{ est réalisé} \end{cases} \quad \text{et on pose : } S_n(x) = \sum_{k=1}^{n} J_k(x)$$

(1).(a). Montrer que les fonctions f_{Y_1} et f_{Y_n} définies pour tout réel x par : $f_{Y_1}(x) = n(1 - F_X(x))^{n-1} f_X(x)$ et $f_{Y_n}(x) = n(F_X(x))^{n-1} f_X(x)$ sont des densités de Y_1 et Y_n respectivement.

(b). Quelle est la loi de probabilité de la variable aléatoire $S_n(x)$?

(c). Justifier l'égalité entre événements suivante : $[Y_k \leq x] = [S_n(x) \geq k]$.

(d). Établir la relation : pour tout x réel, $F_{Y_k}(x) = \displaystyle\sum_{j=k}^{n} \binom{n}{j} (F_X(x))^j (1 - F_X(x))^{n-j}$.

(e). En déduire que pour tout $k \in [\![1;n]\!]$, la fonction f_{Y_k} définie pour x réel par :

$$f_{Y_k}(x) = k\binom{n}{k} (F_X(x))^{k-1} (1 - F_X(x))^{n-k} f_X(x)$$

est une densité de Y_k.

(f). Montrer que si X admet un moment d'ordre r ($r \in \mathbb{N}^*$), alors pour tout $k \in [\![1;n]\!]$, Y_k admet un moment d'ordre r.

Exemple : dans les questions **(2) à (4)**, on suppose que la fonction de répartition F_X est donnée par :

$$F_X(x) = \begin{cases} 1 - \dfrac{1}{\sqrt{x}} \text{ si } x \geq 1 \\ 0 \text{ si } x < 1 \end{cases}$$

(2).(a). Tracer la courbe représentative de F_X dans le plan rapporté à un repère orthonormé.
Préciser la demi-tangente à droite au point d'abscisse $x = 1$.
Justifier que X est une variable aléatoire à densité et préciser une densité f_X de X.

(b). Montrer que X n'admet aucun moment.

(c). Établir l'unicité de la médiane théorique M de X. Calculer M.

(d). Expliciter, pour tout $k \in [\![1;n]\!]$ et tout x réel, l'expression $F_{Y_k}(x)$ d'une densité de Y_k.
En déduire un équivalent de $f_{Y_k}(x)$ lorsque x tend vers $+\infty$.

(3). On suppose dans cette question que $n \geq 3$.

(a). Montrer que pour tout $k \in [\![1;n-2]\!]$, Y_k admet une espérance.

(b). En justifiant l'emploi du changement de variable $t = \dfrac{1}{\sqrt{x}}$, établir pour tout $k \in [\![1;n-2]\!]$, la formule :

$$E(Y_k) = k\binom{n}{k} \int_0^1 t^{n-k-2} (1-t)^{k-1} \, \mathrm{d}t$$

(c). Pour tout couple $(r,s) \in (\mathbb{N}^*)^2$, on pose $I_{r,s} = \displaystyle\int_0^1 t^{r-1}(1-t)^{s-1} \, \mathrm{d}t$.

Montrer que, pour tout couple $(r,s) \in (\mathbb{N}^*)^2$, on a $I_{r,s} = \dfrac{(r-1)!(s-1)!}{(r+s-1)!}$.

(d). En déduire l'expression de $E(Y_k)$ pour tout $k \in [\![1;n-2]\!]$.

(e). On suppose que n est impair et supérieur ou égal à 5, et on pose $n = 2\ell+1$. Justifier la définition de la médiane empirique $Y_{\ell+1}$ d'un échantillon, et établir l'égalité : $E(Y_{\ell+1}) = 4 + \dfrac{6}{\ell-1}$. Commenter.

(4). On pose, pour tout $n \in \mathbb{N}^*$, $Z_n = \dfrac{1}{n^2}\sup(X_1, \ldots, X_n) = \dfrac{Y_n}{n^2}$.

(a). Calculer, pour tout x réel, $F_{Z_n}(x)$.

(b). On définit la fonction φ_Z par : $\varphi_Z(x) = \begin{cases} \exp\left(-\dfrac{1}{\sqrt{x}}\right) \text{ si } x > 0 \\ 0 \text{ si } x \leq 0 \end{cases}$.

Montrer que φ_Z est la fonction de répartition d'une variable aléatoire Z à densité.

(c). Montrer que la suite de variables aléatoires $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers Z.

Partie II - Existence et unicité d'un estimateur optimal

Dans cette partie, X suit la loi normale d'espérance θ et de variance égale à 1. On suppose que le paramètre réel θ est inconnu.
On note Φ la fonction de répartition de la loi normale centrée réduite.
On rappelle que pour n entier de $\mathbb{N}^*$, $(X_1, \ldots, X_n)$ est un $n-$échantillon i.i.d. de la loi de X.

(1). Quelle est la loi de $\overline{X_n}$? Montrer que $\overline{X_n}$ est un estimateur sans biais et convergent du paramètre θ.

(2). Soit α un réel tel que $0 < \alpha < 1$. On appelle *marge d'erreur* associée à un intervalle de confiance de θ au risque α, le réel positif noté $\mu(\alpha)$, égalà la demi-longueur de cet intervalle.

(a). Justifier l'existence de la fonction réciproque Φ^{-1} de la fonction Φ.

(b). Déterminer un intervalle de confiance du paramètre θ au risque α dont le milieu est $\overline{X_n}$.

Vérifier que $\mu(\alpha) = -\dfrac{\Phi^{-1}\left(\dfrac{\alpha}{2}\right)}{\sqrt{n}}$.

(c). On considère un risque β $(\beta \neq \alpha)$ tel que $\mu(\beta) = b\mu(\alpha)$ avec $0 < b < 1$. Exprimer β en fonction de α. Comparer α et β. Commenter.

(3). On note $\mathcal{E}_\theta$ l'ensemble des statistiques $U_n = g_n(X_1, \ldots, X_n)$, où g_n désigne une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, qui sont des estimateurs sans biais de θ et qui admettent une variance.
Sous réserve d'existence, on dit qu'un élément $Z_n$de $\mathcal{E}_\theta$ est un estimateur optimal dans $\mathcal{E}_\theta$, si pour tout U_n de $\mathcal{E}_\theta$, on a $v(Z_n) \leq v(U_n)$.
On admet que pour tout $U_n \in \mathcal{E}_\theta$, on a $\text{Cov}\left(\overline{X_n}, U_n - \overline{X_n}\right) = 0$.

(a). Montrer que $\mathcal{E}_\theta$ n'est pas vide.

(b). Montrer que $\overline{X_n}$ est optimal dans $\mathcal{E}_\theta$.

(c). Soit Z_n un estimateur optimal dans $\mathcal{E}_\theta$. On pose pour tout $U_n \in \mathcal{E}_\theta$, et pour tout réel λ : $A_n(\lambda) = Z_n + \lambda(U_n - Z_n)$.
Montrer que $A_n(\lambda)$ est un élément de $\mathcal{E}_\theta$. Calculer $v(A_n(\lambda))$. En déduire que $\text{Cov}(Z_n, U_n - Z_n) = 0$.

(d). On suppose l'existence de deux estimateurs optimaux $\overline{X_n}$ et Z_n dans $\mathcal{E}_\theta$.
Montrer que $Z_n = \overline{X_n}$ presque sûrement. Conclure.

(4).(a). Justifier l'existence et l'unicité de la médiane théorique M de X, et exprimer M en fonction de θ.

(b). Calculer $f_X(M)$. Montrer que pour tout réel x, on a $F_X(2M - x) = 1 - F_X(x)$.
En déduire une relation entre $f_X(2M - x)$ et $f_X(x)$.

(c). Établir pour tout $k \in [\![1; n]\!]$, la relation $E(Y_k - M) = E(M - Y_{n-k+1})$.

(d). En supposant que $n = 2\ell + 1$ avec $\ell \in \mathbb{N}$, calculer $E(Y_{\ell+1})$, puis justifier que $v(Y_{\ell+1}) \geq \dfrac{1}{n}$. Commenter.

Partie III - Résultats asymptotiques

*Le contexte de cette partie est identique à celui de la **partie II**.*
Dans cette partie, on note T une variable aléatoire qui suit la loi normale centrée réduite.
Si U est une variable aléatoire et s un réel tels que la variable aléatoire $\exp(sU)$ admette une espérance, on note $L_u(s) = E(\exp(sU))$.

(1). Soit J une variable aléatoire qui suit la loi de Bernoulli de paramètre p $(0 < p < 1)$.

(a). Calculer pour tout s réel, $L_J(s)$.

(b). Établir pour tout s réel, l'existence de $L_T(s)$.

(c). Calculer pour tout s réel, $L_T(s)$. En déduire que pour tout couple $(\theta, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$ et pour tout s réel, on a $L_{\sigma T+\theta}(s) = \exp\left(\sigma^2 \dfrac{s^2}{2} + \theta s\right)$.

*Dans les questions **(2) et (3)**, x est un réel fixé.*

(2). On pose pour tout $n \in \mathbb{N}^*$, $y_n = M + \dfrac{x}{\sqrt{n}}$, $q_n = F_X(y_n)$ et $k(n) = \left\lfloor \dfrac{n}{2} \right\rfloor + 1$.

(a). Montrer que l'on a $k(n) = \dfrac{n}{2} + o\left(\sqrt{n}\right)$ lorsque n tend vers $+\infty$.

(b). En appliquant la formule de Taylor-Young à la fonction F_X au voisinage de M, justifier la relation :

$$q_n = \frac{1}{2} + \frac{x}{\sqrt{2\pi n}} + o\left(\frac{1}{\sqrt{n}}\right)$$

(c). Soit $(u_n)_{n\geq 1}$ la suite définie par, pour tout $n \in \mathbb{N}^*$, $u_n = \dfrac{1}{\sqrt{n}}(k(n) - nq_n)$.
Montrer que la suite $(u_n)_{n\geq 1}$ est convergente et déterminer sa limite u.

(3). On pose pour tout $n \in \mathbb{N}^*$, $W_n = \dfrac{1}{\sqrt{n}}(S_n(y_n) - nq_n)$, où $S_n(y_n)$ a été définie dans le préambule de la **partie I**.

(a). Établir pour tout réel s la relation :

$$L_{W_n}(s) = \exp\left(-s\sqrt{n}q_n\right)\left(1 + q_n\exp\left(\frac{s}{\sqrt{n}}\right) - q_n\right)^n$$

(b). En utilisant un développement limité à l'ordre 2, montrer que l'on a : $\ln(L_{W_n}(s)) = \dfrac{s^2}{8} + o(1)$ lorsque n tend vers $+\infty$. En déduire l'égalité $\lim\limits_{n\to+\infty} L_{W_n}(x) = L_{\frac{T}{2}}(s)$.

On admet alors que la suite de variables aléatoires $(W_n)_{n\geq 1}$ converge en loi vers la variable aléatoire $\dfrac{T}{2}$.

(4). On suppose que $x = 0$. Quels sont les arguments qui permettent d'obtenir directement le résultat final de la question **(3)(b).** ?

(5).(a). Établir pour tout n de $\mathbb{N}^*$ et pour tout réel x,les égalités d'événements suivantes :

$$\left[\sqrt{n}\left(Y_{k(n)} - M\right) \leq x\right] = [S_n(y_n) \geq k(n)] = [W_n \geq u_n]$$

(b). Montrer l'égalité : $\lim\limits_{n\to+\infty} \mathbb{P}(W_n \geq u_n) = \mathbb{P}\left(\sqrt{T\dfrac{\pi}{2}} \leq x\right)$.

(c). En déduire que la suite de variables aléatoires $\left(\sqrt{\frac{2n}{\pi}}\left(Y_{k(n)}-M\right)\right)_{n\geq 1}$ converge en loi et préciser sa limite.

(6). On suppose dans cette question que n est impair et on pose $n=2\ell+1$ ($\ell\in\mathbb{N}$). On note ρ_n le coefficient de correlation linéaire de $Y_{k(n)}$ et $\overline{X_n}$.

(a). Que vaut $k(n)$?

(b). Préciser la valeur de $\lim\limits_{n\to+\infty} E\left(Y_{k(n)}\right)$.

(c). On admet sans démonstration que la suite réelle de terme général $E\left(\left(\sqrt{\frac{2n}{\pi}}\left(Y_{k(n)}-M\right)\right)^2\right)$ converge vers $E(T^2)$. En déduire un équivalent de $v\left(Y_{k(n)}\right)$ lorsque n tend vers $+\infty$.

(d). Déterminer la limite de ρ_n lorsque n tend vers $+\infty$ à l'aide des questions **II(1)**, **II(3)**, **II(4)** et **III(6)(c)**.

Éléments de correction

[pb]

Partie I - Quelques propriétés des statistiques d'ordre

(1).(a). On a $Y_1=\inf(X_1,\ldots,X_n)$. Soit alors $x\in\mathbb{R}$. Par définition $F_{Y_1}(x)=\mathbb{P}(Y_1\leq x)$. Ainsi :

$$\begin{aligned}
F_{Y_1}(x) &= \mathbb{P}(\inf(X_1,\ldots,X_n)\leq x)\\
&= 1-\mathbb{P}(\inf(X_1,\ldots,X_n)>x)\\
&= 1-\mathbb{P}([X>x]\cap\ldots\cap[X_n>x])\\
&\underset{\text{indépendance}}{=} 1-\mathbb{P}(X_1>x)\times\ldots\times\mathbb{P}(X_n>x)\\
&\underset{\text{même loi que } X}{=} 1-(\mathbb{P}(X>x))^n\\
&= 1-(1-F_X(x))^n
\end{aligned}$$

et la continuité et le caractère $\mathcal{C}^1$sauf éventuellement en un nombre fini de points de F_X assure l'identique pour F_{Y_1}, qui donne ainsi que Y_1 est une variable aléatoire à densité. Par dérivation, on en déduit qu'une densité de Y_1 est donc, en tout point x où F_{Y_1} est dérivable :

$$f_{Y_1}(x)=nf_X(x)(1-F_X(x))^{n-1}$$

De même :

$$\begin{aligned}
F_{Y_n}(x) &= \mathbb{P}(\sup(X_1,\ldots,X_n)\leq x)\\
&= \mathbb{P}([X_1\leq x]\cap\ldots\cap[X_n\leq x])\\
&\underset{\text{indépendance}}{=} \mathbb{P}(X_1\leq x)\times\ldots\times\mathbb{P}(X_n\leq x)\\
&\underset{\text{même loi que } X}{=} (\mathbb{P}(X\leq x))^n\\
&= (F_X(x))^n
\end{aligned}$$

et la continuité et le caractère $\mathcal{C}^1$sauf éventuellement en un nombre fini de points de F_X assure l'identique pour F_{Y_n}, qui donne ainsi que Y_n est une variable aléatoire à densité. Par dérivation, on en déduit qu'une densité de Y_n est donc, en tout point x où F_{Y_n} est dérivable :

$$f_{Y_1}(x)=nf_X(x)(F_X(x))^{n-1}$$

(b). Soit x réel fixé. Pour tout $k\geq 1$, $J_k(x)$ suit une loi de Bernoulli de paramètre $\mathbb{P}(X_k\leq x)$. Les événements $[X_1\leq x]$, ..., $[X_n\leq x]$ étant indépendants car $(X_1,\ldots,X_n)$ sont indépendantes (car $n-$échantillon i.i.d.), les variables aléatoires $(Y_1,\ldots,Y_n)$ le sont aussi, et elles ont toutes le même paramètre égal à $p(x)=\mathbb{P}(X\leq x)=F_X(x)$ car les $(X_1,\ldots,X_n)$ est un $n-$échantillon de loi parente X.

Ainsi, $S_n(x)$ est la somme de n variables aléatoires indépendantes suivant toutes la même loi de Bernoulli de paramètre $p(x)$. Par conséquent, $S_n(x)\hookrightarrow\mathcal{B}(n;F_X(x))$.

(c). Soit $(x,k)\in\mathbb{R}\times[\![1;n]\!]$.

$[Y_k\leq x]$ est réalisé si, et seulement si, le k^{e} élément de la suite des valeurs prises par $X_1,\ldots,X_n$ rangées dans l'ordre croissant est inférieure ou égale à x, c'est à dire qu'il y a au moins k valeurs prises par les variables $X_1,\ldots,X_n$ qui sont inférieures ou égales à x.

De part sa définition, $S_n(x)$ est une variable aléatoire qui compte le nombre de variables aléatoires X_k qui prennent une valeur inférieure ou égale à x.

Ainsi, $[S_n(x)\geq k]=[Y_k\leq x]$.

(d). Soit $(x,k)\in\mathbb{R}\times[\![1;n]\!]$, on a :

$$\begin{aligned}
F_{Y_k}(x) &= \mathbb{P}(Y_k\leq x)\\
&= \mathbb{P}(S_n(x)\leq k)\\
&= \sum_{j=k}^{n}\mathbb{P}(S_n(x)=j)\\
&\underset{\text{loi binomiale}}{=} \sum_{j=k}^{n}\binom{n}{j}(F_X(x))^j(1-F_X(x))^{n-j}
\end{aligned}$$

(e). Le caractère continue et de classe $\mathcal{C}^1$sauf en un nombre fini de points pour F_{Y_k} est une conséquence des opérations sur les fonctions continues et de classe $\mathcal{C}^1$sauf en un nombre fini de points apporté par F_X. Une densité f_{Y_k} de Y_k est obtenue par dérivation de la fonction de répartition $F_{Y_k}(x)=\mathbb{P}(Y_k\leq x)$ aux points où cette dernière est dérivable. Soit alors $x\in\mathbb{R}$, tel que F_{Y_k} est dérivable en x. On a alors :

$$F'_{Y_k}(x)=\sum_{j=k}^{n}\binom{n}{j}j(F_X(x))^{j-1}\times f_X(x)\times(1-F_X(x))^{n-j}+\sum_{j=k}^{n}\binom{n}{j}(F_X(x))^j\times(n-j)(1-F_X(x))^{n}$$

Il reste à simplifier l'expression de $f_{Y_k}(x)$ ainsi trouvée. Quitte à factoriser par $f_X(x)$, cela revient à simplifier

$$g_k(x)=\sum_{j=k}^{n}\binom{n}{j}j(F_X(x))^{j-1}\times(1-F_X(x))^{n-j}-\sum_{j=k}^{n}\binom{n}{j}(F_X(x))^j\times(n-j)(1-F_X(x))^{n-j-1}$$

Alors :

$$\begin{aligned}
g_k(x) &= \sum_{j=k}^{n}\binom{n}{j} j\,(F_X(x))^{j-1}\times(1-F_X(x))^{n-j} \\
&\quad - \sum_{j=k}^{n}\binom{n}{j}(F_X(x))^{j}\times(n-j)\,(1-F_X(x))^{n-j-1} \\
&\underset{\ell=j-1}{=} \sum_{\ell=k-1}^{n-1}\binom{n}{\ell+1}(\ell+1)\,(F_X(x))^{\ell}\times(1-F_X(x))^{n-\ell-1} \\
&\quad - \sum_{j=k}^{n}\binom{n}{j}(F_X(x))^{j}\times(n-j)\,(1-F_X(x))^{n-j-1} \\
&= \sum_{j=k-1}^{n-1}\binom{n}{j+1}(j+1)\,(F_X(x))^{j}\times(1-F_X(x))^{n-j-1} \\
&\quad - \sum_{j=k}^{n-1}\binom{n}{j}(F_X(x))^{j}\times(n-j)\,(1-F_X(x))^{n-j-1} \\
&= \binom{n}{k}k\,(F_X(x))^{k-1}\,(1-F_X(x))^{n-k} \\
&\quad + \sum_{j=k}^{n-1}\left(\binom{n}{j+1}(j+1)-\binom{n}{j}(n-j)\right)(F_X(x))^{j}\,(1-F_X(x))^{n-j-1}
\end{aligned}$$

Or $\binom{n}{j+1}(j+1)-\binom{n}{j}(n-j)=\frac{n!}{(j+1)(n-j-1)!}(j+1)-\frac{n!}{j!(n-j)!}(n-j)=\frac{n!}{j!(n-j-1)!}-\frac{n!}{j!(n-j-1)!}=0$, d'où le résultat attendu en multipliant par $f_X(x)$.

(f). Supposons que X admette un moment d'ordre $r\geq 1$, c'est à dire que $\int_{-\infty}^{+\infty}t^r f_X(t)\,\mathrm{d}t$ est convergente [a]. Puisqu'il y a convergence, on a nécessairement que $\int_{-\infty}^{0}t^r f_X(t)\,\mathrm{d}t$ et $\int_{0}^{+\infty}t^r f_X(t)\,\mathrm{d}t$ sont convergentes.

Il s'agit donc de montrer que $\int_{-\infty}^{+\infty}t^r f_{Y_k}(t)\,\mathrm{d}t$ est convergente, en d'autres termes, de s'assurer de la convergence de $\int_{-\infty}^{0}t^r f_{Y_k}(t)\,\mathrm{d}t$ et de $\int_{0}^{+\infty}t^r f_{Y_k}(t)\,\mathrm{d}t$.

— Pour $\int_{0}^{+\infty}t^r f_{Y_k}(t)\,\mathrm{d}t$, il est immédiat que, pour tout $t\in\mathbb{R}$, $0\leq (F_X(t))^{k-1}(1-F_X(t))^{n-k}\leq 1$ car $F_X(t)\in[0;1]$. Par suite, on a directement :

$$\forall t\in\mathbb{R}_+,\qquad 0\leq\left|t^r f_{Y_K}(t)\right|\leq k\binom{n}{k}\left|t^r f_X x\right| = \underbrace{k\binom{n}{k}}_{\text{constante}}\ t^r f_X(t)$$

Ainsi, la convergence de $\int_{0}^{+\infty}t^r f_X(t)\,\mathrm{d}t$, assure la convergence par le critère de comparaison des intégrales impropres de fonctions positives, de $\int_{0}^{+\infty}t^r f_{Y_k}(t)\,\mathrm{d}t$.

— Pour $\int_{-\infty}^{0}t^r f_X(t)\,\mathrm{d}t$, sur le même principe, on trouve que :

$$\forall t\in\mathbb{R}_-,\qquad 0\leq\left|t^r f_{Y_K}(t)\right|\leq k\binom{n}{k}\left|t^r f_X x\right| = k\binom{n}{k}f_X(t)\left|t^r\right| = \begin{cases}\underbrace{k\binom{n}{k}}_{\text{constante}}\ t^r f_X(t) \text{ si rest pa} \\ \underbrace{-k\binom{n}{k}}_{\text{constante}}\ t^r f_X(t) \text{ si rest im}\end{cases}$$

Ainsi, dans les deux cas, la convergence de $\int_{-\infty}^{0}t^r f_X(t)\,\mathrm{d}t$, assure la convergence par le critère de comparaison des intégrales impropres de fonctions positives, de $\int_{-\infty}^{0}t^r f_{Y_k}(t)\,\mathrm{d}t$.

(2). (a). F_X est clairement continue sur $\mathbb{R}$, dérivable sur $\mathbb{R}$ et de dérivée $F_X'(x)=\begin{cases}0 \text{ si } x\leq 1\\ \dfrac{1}{2x^{\frac{3}{2}}} \text{ si } x\geq 1\end{cases}$.

Sa dérivée à droite en 1, $(F_X')_d(1)=\frac{1}{2}$. Par ailleurs, puisque sur $[1;+\infty[$, F_X' est encore dérivable de dérivée $F_X''(x)=-\frac{3}{4x^{\frac{5}{2}}}\leq 0$, F_X est donc convexe sur cet intervalle, ce qui nous assure de son allure au final :

Par ailleurs, la fonction F_X satisfait toutes les conditions pour être la fonction de répartition d'une variable aléatoire à densité, à savoir est continue sur $\mathbb{R}$, et de classe $\mathcal{C}^1$sur $\mathbb{R}$ sauf au point $x=1$. Une densité f_X de X est alors obtenue en dérivant F_X en tout point où F_X est dérivable, soit :

$$f_X(x)=\begin{cases}\dfrac{1}{2x^{\frac{3}{2}}} \text{ si } x\geq 1\\ 0 \text{ sinon}\end{cases}$$

(b). Soit $r\in\mathbb{N}^*$. Il s'agit de montrer que $E(X^r)$ n'existe pas, dans le sens où l'intégrale $\int_{-\infty}^{+\infty}t^r f_X(t)\,\mathrm{d}t$ n'est pas convergente. Compte-tenu de l'expression de f_X, cela revient à s'intéresser à la (non)convergence de $\int_{1}^{+\infty}t^r f_X(t)\,\mathrm{d}t$, qui dans notre cas est exactement $\int_{1}^{+\infty}\frac{1}{2t^{\frac{3}{2}-r}}\,\mathrm{d}t$. Or puisque $r\geq 1$, on a clairement $\frac{3}{2}-r<1$, et de ce fait, l'intégrale $\int_{1}^{+\infty}\frac{1}{2t^{\frac{3}{2}-r}}\,\mathrm{d}t$ est une intégrale de Riemann, au coefficient $\frac{1}{2}$ près, divergente.

(c). Il s'agit donc ici de résoudre l'équation $F_X(x)=\frac{1}{2}$. Compte-tenu de la défintion de F_X, si solutions il y a, ce ne peut être que sur l'intervalle $[1;+\infty[$. Par ailleurs, F_X étant continue strictement croissante sur l'intervalle $[1;+\infty[$, avec $F_X([1;+\infty[)=[0;1[$, l'équation $F_(x)=\frac{1}{2}$ y admet une unique solution d'après le théorème de la bijection. La résolution de l'équation $F_X(x)=\frac{1}{2}$, conduit à $x=4$, et donc $M=4$ qui est dans ce cas unique.

(d). Soit $k\in[\![1;n]\!]$. Puisque pour tout $x\in\mathbb{R}$, $f_{Y_k}(x)=$

$k\binom{n}{k} f_X(x) \left(F_X(x)\right)^{k-1} \left(1-F_X(x)\right)^{n-k}$, on en déduit que :

$$f_{Y_k}(x) = \begin{cases} 0 \text{ si } x<1 \\ k\binom{n}{k}\dfrac{1}{2x\sqrt{x}}\left(1-\dfrac{1}{\sqrt{x}}\right)^{k-1}\left(1-\left(1-\dfrac{1}{\sqrt{x}}\right)\right)^{n-k} & \text{si } x \geq 1 \end{cases}$$

ou encore après simplification de la deuxième expression :

$$f_{Y_k}(x) = \begin{cases} 0 \text{ si } x<1 \\ \dfrac{k}{2}\binom{n}{k}\left(\dfrac{1}{\sqrt{x}}\right)^{n-k+3}\left(1-\dfrac{1}{\sqrt{x}}\right)^{k-1} & \text{si } x \geq 1 \end{cases}$$

Par ailleurs, comme $\left(1-\dfrac{1}{\sqrt{x}}\right)^{k-1} \underset{x\to+\infty}{\longrightarrow} 1$, on en déduit que $f_{Y_k}(x) \underset{x\to+\infty}{\sim} \dfrac{k}{2}\binom{n}{k}\left(\dfrac{1}{\sqrt{x}}\right)^{n-k+3} = \dfrac{k}{2}\binom{n}{k}\left(\dfrac{1}{x}\right)^{\frac{n-k+3}{2}}$.

(3). (a). Soit $k \in [\![1;n-2]\!]$. On doit montrer que $\displaystyle\int_{-\infty}^{+\infty} t f_{Y_k}(t)\,\mathrm{d}t$ est convergente, et qu'en cas de convergence, $E(Y_k)$ est égale à la valeur de cette dernière.
Compte-tenu de l'expression de f_{Y_k}, il s'agit de montrer la convergence de $\displaystyle\int_1^{+\infty} t\frac{k}{2}\binom{n}{k}\left(\frac{1}{\sqrt{t}}\right)^{n-k+3}\left(1-\frac{1}{\sqrt{t}}\right)^{k-1}\mathrm{d}t$.
Or, pour tout $t \geq 1$, $t\dfrac{k}{2}\binom{n}{k}\left(\dfrac{1}{\sqrt{t}}\right)^{n-k+3}\left(1-\dfrac{1}{\sqrt{t}}\right)^{k-1} \geq 0$, et d'après la quetion précédente, on en déduit que :

$$t\frac{k}{2}\binom{n}{k}\left(\frac{1}{\sqrt{t}}\right)^{n-k+3}\left(1-\frac{1}{\sqrt{t}}\right)^{k-1} \underset{t\to+\infty}{\sim} \frac{k}{2}\binom{n}{k}\left(\frac{1}{t}\right)^{\frac{n-k+1}{2}}$$

Ainsi d'après le théorème d'équivalence pour les intégrales impropres de fonctions positives

$$\int_1^{+\infty} t\frac{k}{2}\binom{n}{k}\left(\frac{1}{\sqrt{t}}\right)^{n-k+3}\left(1-\frac{1}{\sqrt{t}}\right)^{k-1}\mathrm{d}t \qquad \text{et} \qquad \int_1^{+\infty}\frac{k}{2}\binom{n}{k}\left(\frac{1}{t}\right)^{\frac{n-k+1}{2}}\mathrm{d}t$$

sont de même nature, et comme $\dfrac{n-k+1}{2} > 1$, la deuxième est une intégrale (à une constante multiplicative près) de Riemann convergente. Par suite on en conclut que $\displaystyle\int_{-\infty}^{+\infty} t f_{Y_k}(t)\,\mathrm{d}t$ converge.

(b). Le changement de variable $t = \dfrac{1}{\sqrt{x}}$ étant de classe $\mathcal{C}^1$sur $[1;+\infty[$ et réalisant une bijection de l'intervalle $[1;+\infty[$ sur l'intervalle $]0;1]$, en étant décroissant, les intégrales

$$\int_1^{+\infty} t\frac{k}{2}\binom{n}{k}\left(\frac{1}{\sqrt{t}}\right)^{n-k+3}\left(1-\frac{1}{\sqrt{t}}\right)^{k-1}\mathrm{d}t \qquad \text{et} \qquad \int_0^1 k\binom{n}{k} t^{n-k-2}(1-t)^{j-1}\,\mathrm{d}t$$

sont de même nature, et en cas de convergence, elles sont égales. Or la question précédente, nous assure la converge de la première, et par suite, on en déduit la convergence de la seconde, et l'égalité de leurs valeurs.

(c). Pour $r \in \mathbb{N}^*$, on note $\mathcal{P}(r)$: « $\forall s \in \mathbb{N}^*, \quad I_{r,s} = \dfrac{(r-1)!(s-1)!}{(r+s-1)!}$ » et montrons par récurrence sur r que $\mathcal{P}(r)$ est vraie pour tout entier $r \geq 1$.

— **Initialisation :** pour $r=1$, fixons $s \in \mathbb{N}^*$.
Par définition $I_{1,s} = \displaystyle\int_0^1 (1-t)^{s-1}\,\mathrm{d}t = \frac{1}{s}$ et on a bien $\dfrac{(1-1)!(s-1)!}{(1+s-1)!} = \dfrac{1}{s}$, et ce résultat étant vrai pour tout $s \geq 1$, on en déduit que $\mathcal{P}(1)$ est vraie.

— **Hérédité :** supposons que pour un entier $r \geq 1$, on ait la propriété $\mathcal{P}(r)$ et montrons que, sous cette hypothèse, on a la propriété $\mathcal{P}(r+1)$, à savoir :

$$\forall s \geq 1, \qquad I_{r+1,s} = \frac{r!(s-1)!}{(r+s)!}$$

Soit alors $s \geq 1$. Par définition, $I_{r+1} = \displaystyle\int_0^1 t^r(1-t)^{s-1}\,\mathrm{d}t$. On effectue alors l'intégration par parties :

$$\begin{array}{lcl} u(t) = t^r & \underset{\text{se dérive en}}{\rightsquigarrow} & u'(t) = rt^{r-1} \\ v(t) = -\dfrac{(1-t)^s}{s} & \underset{\text{se dérive en}}{\rightsquigarrow} & v'(t) = (1-t)^{s-1} \end{array}$$

où u et v sont de classe $\mathcal{C}^1$sur $[0;1]$, et il vient alors :

$$I_{r+1,s}(t) = \underbrace{\left[-t^r\frac{(1-t)^s}{s}\right]_0^1}_{0} + \frac{r}{s}\int_0^1 t^{r-1}(1-t)^s\,\mathrm{d}t = \frac{r}{s}I_{r,s+1}$$

Or l'hypothèse de récurrence appliquée pour r et $s+1$, donne $I_{r,s+1} = \dfrac{(r-1)!(s+1-1)!}{(r+s+1-1)!}$ et permet d'obtenir $I_{r+1,s} = \dfrac{r!(s-1)!}{r+s)!}$, qui donne donc $\mathcal{P}(r+1)$ puisque cela est vrai pour tout $s \geq 1$.

— La propriété $\mathcal{P}(r)$ étant vraie au rang $r=1$ et héréditaire, elle est donc vraie pour tout entier $r \geq 1$.

D'où $I_{r,s} = \dfrac{(r-1)!(s-1)!}{(r+s-1)!}$.

(d). On a directement d'après la question **I-(3)(b).**, $E(Y_k) = k\binom{n}{k}I_{n-k-1,k} = \dfrac{n(n-1)}{(n-k)(n-k-1)}$.

(e). Avec ces hypothèses sur n. Puisque $n = 2\ell+1$ avec $n \geq 5$, on en déduit d'après la question précédente que $\ell+1 \leq n-2$, et ainsi $3 \leq \ell+1 \leq n-2$. On en déduit donc que $E(Y_{\ell+1})$ existe et vaut $\dfrac{(2\ell+1)(2\ell)}{\ell(\ell-1)} = 4 + \dfrac{6}{\ell-1}$.
Par ailleurs, du fait que $E(Y_{\ell+1}) \underset{n\to+\infty}{\longrightarrow} 4$ puisque dans ce cas $\ell \underset{n\to+\infty}{\longrightarrow} +\infty$, et qui est la médiane théorique de X trouvée à la question **I(2)(c).**.

(4). (a). Soit $x \in \mathbb{R}$, alors $F_{Z_n \leq x} = \mathbb{P}\left(\dfrac{1}{n^2}Y_n \leq x\right) = \mathbb{P}\left(Y_n \leq n^2 x\right) = F_{Y_n}\left(n^2x\right)$. Comme :

$$F_{Y_n}\left(n^2x\right) = \left(F_X(n^2x)\right)^n = \begin{cases} \left(1-\dfrac{1}{\sqrt{n^2x}}\right)^n & \text{si } n^2x \in [1;+\infty[\\ 0 \text{ sinon} \end{cases}$$

il vient :

$$F_{Z_n}(x) = \begin{cases} \left(1-\dfrac{1}{n\sqrt{x}}\right)^n & \text{si } x \in \left[\dfrac{1}{n^2};+\infty\right[\\ 0 \text{ sinon} \end{cases}$$

(b). φ_Z est telle que $\lim\limits_{x\to 0^+}\varphi_Z(x) = 0 = \lim\limits_{x\to 0^-}\varphi_Z(x)$, donc φ_Z est continue sur $\mathbb{R}$, et de classe $\mathcal{C}^1$sur $\mathbb{R}$, sauf éventuellement en 0.

Reste à vérifier que φ_Z satisfait aux conditions pour être la fonction de répartition d'une variable aléatoire, en plus de la continuité à droite en tout point de $\mathbb{R}$ que l'on a déjà, à savoir :

— $\lim\limits_{x\to+\infty}\varphi_Z(x) = 1$ et $\lim\limits_{x\to-\infty}\varphi_Z(x) = 0$, ce qui est clairement le cas.

— φ_Z est croissante sur $\mathbb{R}_+^*$, car $\varphi_Z = f_1\circ f_2\circ f_3$ où $f_1 : x \longmapsto \mathrm{e}^{-x}$, $f_2 : x \longmapsto \dfrac{1}{x}$ qui sont décroissantes sur $\mathbb{R}_+^*$, et $f_3 : x \longmapsto \sqrt{x}$ qui est croissante sur $\mathbb{R}_+$.

Ainsi φ_Z est la fonction de répartition d'une variable aléatoire à densité.

(c). Puisque $\dfrac{1}{n^2} \underset{n\to+\infty}{\longrightarrow} 0$, il suffit de regarder, à x fixé dans $\mathbb{R}_+^*$, $\lim\limits_{n\to+\infty} F_{Z_n}(x)$ et de vérifier que $F_{Z_n}(x) \underset{n\to+\infty}{\longrightarrow} F_Z(x)$.

Or on a $\left(1-\dfrac{1}{n\sqrt{x}}\right)^n = \exp\left(n\ln\left(1-\dfrac{1}{n\sqrt{x}}\right)\right)$. Comme $\dfrac{1}{n\sqrt{x}} \underset{n\to+\infty}{\longrightarrow} 0$, on en déduit que $\ln\left(1-\dfrac{1}{n\sqrt{x}}\right) \underset{n\to+\infty}{\sim} -\dfrac{1}{n\sqrt{x}}$ et par suite que $n\ln\left(1-\dfrac{1}{n\sqrt{x}}\right) \underset{n\to+\infty}{\sim} -\dfrac{1}{\sqrt{x}}$.

Par conséquent $\lim\limits_{n\to+\infty} n\ln\left(1-\dfrac{1}{n\sqrt{x}}\right) = -\dfrac{1}{\sqrt{x}}$, et par composition des limites par la fonction exp qui est continue sur $\mathbb{R}$, on en déduit que $\lim\limits_{n\to+\infty}\exp\left(n\ln\left(1-\dfrac{1}{n\sqrt{x}}\right)\right) = \exp\left(-\dfrac{1}{\sqrt{x}}\right) = F_Z(x)$.

Partie II - Existence et unicité d'un estimateur optimal

(1). Les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes de même loi $\mathcal{N}(\theta,1)$, donc par théorème $X_1+\ldots+X_n \hookrightarrow \mathcal{N}(n\theta, n)$. Par suite, $\overline{X_n} = \dfrac{1}{n}(X_1+\ldots+X_n) \hookrightarrow \mathcal{N}\left(\theta, \dfrac{1}{n}\right)$.

On en déduit donc que $\overline{X_n}$ est un estimateur sans biais de θ, et grâce à l'inégalité de Bienaymé-Tchebychev puisque $\overline{X_n}$ admet un moment d'ordre 2, on en déduit que $\left(\overline{X_n}\right)_{n\geq 1}$ converge en probabilité vers θ puisque $v\left(\overline{X_n}\right) \underset{n\to+\infty}{\longrightarrow} 0$, et on a donc un estimateur convergent.

(2). (a). La fonction Φ étant continue strictement croissante sur $\mathbb{R}$ et à valeurs dans $]0;1[$, elle réalise ainsi une bijection de $\mathbb{R}$ sur $]0;1[$, et par suite Φ^{-1} existe.

(b). Puisque $(X_1, \ldots, X_n)$ est un $n-$échantillon i.i.d. de $X \hookrightarrow \mathcal{N}(\theta,1)$, on sait qu'un intervalle de confiance au niveau de confiance $1-\alpha$ est $\left[\overline{X_n} - t_\alpha\dfrac{1}{\sqrt{n}} ; \overline{X_n} + t_\alpha\dfrac{1}{\sqrt{n}}\right]$ où t_α est tel que $\Phi(t_\alpha) = \dfrac{2-\alpha}{2}$.

On en déduit alors que $t_\alpha = -\Phi^{-1}\left(\dfrac{\alpha}{2}\right) < 0$, et on obtient pour intervalle de confiance :

$$\left[\overline{X_n} + \frac{\Phi^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{n}} ; \overline{X_n} - \frac{\Phi^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{n}}\right]$$

qui est bien un intervalle de confiance de θ au niveau de confiance $1-\alpha$ centré en $\overline{X_n}$, et on a bien $\mu(\alpha) = -\dfrac{\Phi^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{n}}$.

(c). Puisque $\mu(\beta) = b\mu(\alpha)$, on en déduit que $\dfrac{\Phi^{-1}\left(\frac{\beta}{2}\right)}{\sqrt{n}} = -b\dfrac{\Phi^{-1}\left(\frac{\alpha}{2}\right)}{\sqrt{n}}$, et donc que $\Phi^{-1}\left(\dfrac{\beta}{2}\right) = b\Phi^{-1}\left(\dfrac{\alpha}{2}\right)$, et en composant par Φ, on en déduit que $\dfrac{\beta}{2} = \Phi\left(b\Phi^{-1}\left(\dfrac{\alpha}{2}\right)\right)$ et donc que $\beta = 2\Phi\left(b\Phi^{-1}\left(\dfrac{\alpha}{2}\right)\right)$.

Puisque $0<\alpha<1$, on a $0 < \dfrac{\alpha}{2} < \dfrac{1}{2}$ et donc $\Phi^{-1}\left(\dfrac{\alpha}{2}\right) < 0$ par croissance Φ^{-1} et parce que $\Phi(0) = \dfrac{1}{2}$. Par suite, on a $b\Phi^{-1}\left(\dfrac{\alpha}{2}\right) > \Phi^{-1}\left(\dfrac{\alpha}{2}\right)$ car $0<b<1$, et donc $\beta > 2\Phi\left(\Phi^{-1}\left(\dfrac{\alpha}{2}\right)\right) = 2\times\dfrac{\alpha}{2} = \alpha$ et donc $\beta > \alpha$.

Le fait que $\mu(\beta) = b\mu(\alpha)$ avec $0<b<1$ signifie que l'on souhaite réduire la largeur de l'intervalle de confiance, ce qui ne peut se faire qu'en diminuant le niveau de confiance.

(3). (a). $\overline{X_n}$ est un estimateur sans biais de θ, qui possède une variance, et donc $\overline{X_n} \in \mathcal{E}(\theta)$ qui n'est donc pas vide.

(b). Soit $U_n \in \mathcal{E}(\theta)$. On va montrer que $v(U_n) \geq (v\overline{X_n})$.

On a $\mathrm{Cov}\left(\overline{X_n}, U_n - \overline{X_n}\right) = 0$ d'après l'énoncé, et donc par bilinéarité de la variance, on en déduit que $\mathrm{Cov}\left(\overline{X_n}, U_n\right) = \mathrm{Cov}\left(\overline{X_n}, \overline{X_n}\right) = v\left(\overline{X_n}\right)$. D'où $\mathrm{Cov}\left(\overline{X_n}, U_n\right) = v\left(\overline{X_n}\right)$.

On a nécessairement $v(U_n) \neq 0$ car sinon, $\mathrm{Cov}\left(\overline{X_n}, U_n\right) = 0$ implique $v\left(\overline{X_n}\right) = \dfrac{1}{n} = 0$, ce qui est impossible. Ainsi, on a :

$$\frac{v\left(\overline{X_n}\right)}{v(U_n)} = \frac{\left(v\left(\overline{X_n}\right)\right)^2}{v\left(\overline{X_n}\right)v(U_n)} = \frac{\left(\mathrm{Cov}\left(\overline{X_n}, U_n\right)\right)^2}{v\left(\overline{X_n}\right)v(U_n)} = \left(\rho_{\overline{X_n},U_n}\right)^2$$

où $\rho_{\overline{X_n},U_n}$ est le coefficient de corrélation linéaire du couple de variables aléatoires $\left(\overline{X_n}, U_n\right)$, dont on sait qu'il appartient, en valeur absolue, à l'intervalle $[0;1]$. Par suite, on en déduit que $0 \leq \dfrac{v\left(\overline{X_n}\right)}{v(U_n)} \leq 1$, et donc que $v(U_n) \geq (v\overline{X_n})$.

(c). $A_n(\lambda)$ est la somme de trois variables aléatoires admettant une variance, donc admet une variance, ainsi qu'une espérance puisqu'ayant un moment d'ordre 2. Par linéarité de l'espérance, on tire $E(A_n\lambda) = \theta$, donc $A_n(\lambda) \in \mathcal{E}(\theta)$.

Pour la variance de $A_n(\lambda)$, du fait qu'il n'y a pas indépendance des différentes variables aléatoires on a :

$$\begin{aligned} v(A_n(\lambda)) &= v(Z_n) + v(\lambda(U_n - Z_n)) + 2\mathrm{Cov}\left(Z_n, \lambda(U_n - Z_n)\right) \\ &= v(Z_n) + \lambda^2 v(U_n - Z_n) + 2\lambda\mathrm{Cov}\left(Z_n, U_n - Z_n\right) \end{aligned}$$

On en déduit, pour tout $\lambda \in \mathbb{R}$:

$\underbrace{v(A_n(\lambda)) - v(Z_n)}_{\geq 0 \text{ car } Z_n \text{ est optimal}} = \lambda^2 v(U_n - Z_n) + 2\lambda\mathrm{Cov}\left(Z_n, U_n - Z_n\right)$ d'où $\lambda^2 v(U_n - Z_n) + 2\lambda\mathrm{Cov}\left(Z_n, U_n - Z\right.$

Ainsi, pour tout $\lambda \in \mathbb{R}$: $\lambda\left(\lambda v(U_n - Z_n) + 2\mathrm{Cov}\left(Z_n, U_n - Z_n\right)\right) \geq 0$, c'est à dire, que λ et $\lambda v(U_n - Z_n) + 2\mathrm{Cov}\left(Z_n, U_n - Z_n\right)$ sont de même signe. Par suite :

— lorsque $\lambda \longrightarrow 0^+$, on en déduit que $\underbrace{\lambda v(U_n - Z_n)}_{\geq 0} + 2\mathrm{Cov}\left(Z_n, U_n - Z_n\right) \underset{\lambda \to 0^+}{\longrightarrow}$ $2\mathrm{Cov}\left(Z_n, U_n - Z_n\right) \geq 0$;

— lorsque $\lambda \longrightarrow 0^-$, on en déduit que $\underbrace{\lambda v(U_n - Z_n)}_{\leq 0} + 2\mathrm{Cov}\left(Z_n, U_n - Z_n\right) \underset{\lambda \to 0^-}{\longrightarrow}$ $2\mathrm{Cov}\left(Z_n, U_n - Z_n\right) \leq 0$;

et donc que $2\mathrm{Cov}\left(Z_n, U_n - Z_n\right) = 0$ soit $\mathrm{Cov}\left(Z_n, U_n - Z_n\right) = 0$.

(d). Il suffit par exemple de montrer que $v\left(\overline{X_n} - Z_n\right) = 0$ pour s'assurer que $\overline{X_n} - Z_n$ est presque sûrement constante.
Comme $v\left(\overline{X_n} - Z_n\right) = \mathrm{Cov}\left(\overline{X_n} - Z_n, \overline{X_n} - Z_n\right)$ il vient :

$$\mathrm{Cov}\left(\overline{X_n} - Z_n, \overline{X_n} - Z_n\right) = \mathrm{Cov}\left(\overline{X_n}, \overline{X_n} - Z_n\right) - \mathrm{Cov}\left(Z_n, \overline{X_n} - Z_n\right)$$

De $\mathrm{Cov}\left(Z_n, \overline{X_n} - Z_n\right) = 0$ et $\mathrm{Cov}\left(\overline{X_n}, Z_n - \overline{X_n}\right) = 0$, il vient $\mathrm{Cov}\left(\overline{X_n}, \overline{X_n} - Z_n\right) - \mathrm{Cov}\left(Z_n, \overline{X_n} - Z_n\right) = 0$, et ainsi $\mathrm{Cov}\left(\overline{X_n} - Z_n, \overline{X_n} - Z_n\right) = 0$ et par suite $\overline{X_n} - Z_n$ est presque sûrement constante, et comme $\overline{X_n}$ et Z_n ont même espérance, cette constante est nécessairement nulle.
On peut en conclure que seul $\overline{X_n}$ est optimal dans $\mathcal{E}(\theta)$.

(4). (a). La fonction de répartition d'une loi normale étant strictement croissante, on en déduit l'existence et l'unicité d'un $x \in \mathbb{R}$ tel que $F_X(x) = \dfrac{1}{2}$.

Par suite, puisque $X \hookrightarrow \mathcal{N}(\theta, 1)$, la centrée réduite $X^* = \dfrac{X - \theta}{1} = X - \theta$ associée à X suit quant à elle une loi $\mathcal{N}(0, 1)$. Ainsi :

$$F_X(x) = \frac{1}{2} \quad \Leftrightarrow \quad \mathbb{P}\left(X^* \leq x - \theta\right) = \frac{1}{2} \quad \Leftrightarrow \quad \Phi(x - \theta) = \frac{1}{2} \quad \Leftrightarrow \quad \Phi(x - \theta) = \Phi(0)$$

et ainsi, $x = \theta$ puisque Φ réalise une bijection de $\mathbb{R}$ sur $]0; 1[$.
Ainsi, M existe et est unique, et vaut θ.

(b). Puisque pour tout $x \in \mathbb{R}$, $f_X(x) = \dfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\theta)^2}{2}}$, on en déduit directement que $f_X(M) = f_X(\theta) = \dfrac{1}{\sqrt{2\pi}}$.
Ensuite, pour tout $x \in \mathbb{R}$:

$$F_X(2M - x) = \mathbb{P}(X \leq 2M - x) = \mathbb{P}(X \leq 2\theta - x) = \mathbb{P}(X - \theta \leq \theta - x) = \Phi(\theta - x)$$
$$\text{et} \qquad 1 - F_X(x) = 1 - \mathbb{P}(X \leq x) = 1 - \mathbb{P}(X - \theta \leq x - \theta) = 1 - \Phi(x - \theta) = \Phi(\theta - x)$$

et ainsi $F_X(2M - x) = 1 - F_X(x)$.
En dérivant cette dernière expression, ce que l'on peut faire pour tout $x \in \mathbb{R}$ puisque Φ est de classe $\mathcal{C}^1$sur $\mathbb{R}$, on en déduit que :

$$\forall x \in \mathbb{R}, \qquad f_X(2M - x) = f_X(x)$$

(c). D'après **I(1)(f).**, puisque X admet un moment d'ordre 1, Y_k admet aussi un moment d'ordre 1.
Montrer que $E(Y_k - M) = E(M - Y_{n-k+1})$ revient à montrer que $E(Y_k) = 2M - E(Y_{n-k+1})$

Puisque Y_k admet une espérance, $\displaystyle\int_{-\infty}^{+\infty} t f_{Y_k}(t)\,\mathrm{d}t$ est convergente et il reste à en calculer sa valeur, ainsi que celle de $\displaystyle\int_{-\infty}^{+\infty} t f_{Y_{n-k+1}}(t)\,\mathrm{d}t$, ou par exemple de trouver une relation entre la densité de Y_k et celle de Y_{n-k+1}.

Soit alors $x \in \mathbb{R}$. On a :

$$\begin{aligned}
f_{Y_k}(x) &= k\binom{n}{k} \left(F_X(x)\right)^{k-1} \left(1 - F_X(x)\right)^{n-k} f_X(x) \\
&= k\binom{n}{k} \left(1 - F_X(2M - x)\right)^{k-1} \left(1 - (1 - F_X(2M - x))\right)^{n-k} f_X(2M - x) \\
&= k\binom{n}{k} \left(F_X(2M - x)\right)^{n-k} \left(1 - F_X(2M - x)\right)^{k-1} f_X(2M - x) \\
&= \underbrace{k\binom{n}{k}}_{(n-k+1)\binom{n}{n-k+1}} \left(F_X(2M - x)\right)^{(n-k+1)-1} \left(1 - F_X(2M - x)\right)^{n-(n-k+1)} f_X(2M - x) \\
&= f_{Y_{n-k+1}}(2M - x)
\end{aligned}$$

Le changement de variables $u = 2M - t$ étant décroissant, bijectif de $\mathbb{R}$ sur $\mathbb{R}$ et de classe $\mathcal{C}^1$, on en déduit que $\displaystyle\int_{-\infty}^{+\infty} t f_{Y_k}(t)\,\mathrm{d}t$ et $\displaystyle\int_{-\infty}^{+\infty} (2M - u) f_{Y_{n-k+1}}(2M - u)\,\mathrm{d}u$ sont de même nature, et égales puisqu'il y a convergence.
Pour tout $u \in \mathbb{R}$, on a $(2M - u) f_{Y_{n-k+1}}(u) = 2M f_{Y_{n-k+1}}(u) - u f_{Y_{n-k+1}}(u)$.

Or $\displaystyle\int_{-\infty}^{+\infty} f_{Y_{n-k+1}}(u)\,\mathrm{d}u$ converge et vaut 1, et $\displaystyle\int_{-\infty}^{+\infty} u f_{Y_{n-k+1}}(u)\,\mathrm{d}u$ converge et vaut $E(Y_{n-k+1})$, donc $\displaystyle\int_{-\infty}^{+\infty} (2M - u) f_{Y_{n-k+1}}(2M - u)\,\mathrm{d}u$ converge par somme de deux intégrales convergentes, et vaut $2M - E(Y_{n-k+1})$, d'où le résultat.

(d). De la relation précédente, on tire $E(Y_{\ell+1} - M) = E(M - Y_{n-(\ell+1)+1})$, c'est à dire $E(Y_{\ell+1}) - M = M - E(Y_{\ell+1})$ puisque $n = 2\ell + 1$ et donc $E(Y_{\ell+1}) = M$.
Puisque $Y_{\ell+1} \in \mathcal{E}(\theta)$ et que $Y_{\ell+1}$ admet une variance puisque X en admet une d'après **I-(3)(f).**, il vient $v(Y_{\ell+1}) \geq v\left(\overline{X_n}\right) = \dfrac{1}{n}$.

Partie III - Résultats asymptotiques

(1). (a). e^{sJ} est une variable aléatoire discrète finie, donc son espérance existe et vaut d'après le théorème de transfert :

$$E\left(\mathrm{e}^{sJ}\right) = \mathrm{e}^{s \times 0}\,\mathbb{P}(J = 0) + \mathrm{e}^{s \times 1}\,\mathbb{P}(J = 1) = 1 - p + p\mathrm{e}^{s}$$

d'où la valeur de $L_j(s)$.

(b). D'après le théorème de transfert $E\left(\mathrm{e}^{sT}\right)$ existe, si et seulement si $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{st} \varphi(t)\,\mathrm{d}t$ converge absolument.

Or, pour tout $t \in \mathbb{R}$, $\mathrm{e}^{st}\varphi(t) = \ldots = \mathrm{e}^{\frac{1}{2}s^2} \dfrac{1}{1 \times \sqrt{2\pi}} \mathrm{e}^{-\frac{(t-s)^2}{2 \times 1^2}}$.

Comme $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{1 \times \sqrt{2\pi}} \mathrm{e}^{-\frac{(t-s)^2}{2 \times 1^2}}\,\mathrm{d}t$ converge et vaut 1 car il s'agit de la densité d'une variable aléatoire qui suit une loi $\mathcal{N}(s, 1)$, on en déduit que $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{st}\varphi(t)\,\mathrm{d}t$ converge absolument.

(c). Par suite $E\left(\mathrm{e}^{sT}\right)$ existe et vaut $\mathrm{e}^{\frac{s^2}{2}}$.

(d). En écrivant $\mathrm{e}^{s(\sigma T + \theta)} = \mathrm{e}^{s\theta} \times \mathrm{e}^{s\theta T}$, il vient que $E\left(\mathrm{e}^{s(\sigma T + \theta)}\right)$ est existe et vaut $\mathrm{e}^{s\theta} \times E\left(\mathrm{e}^{s\theta t}\right) = \mathrm{e}^{s\theta} \times \mathrm{e}^{\frac{(s\sigma)^2}{2}} = \exp\left(\sigma^2 \dfrac{s^2}{2} + \theta s\right)$.

(2). (a). Il va montrer que $\lim\limits_{n\to+\infty} \dfrac{k(n)-\frac{n}{2}}{\sqrt{n}} = 0$ pour conclure à $k(n)-\dfrac{n}{2} = o\left(\sqrt{n}\right)$.

Puisque $k(n) = \left\lfloor \dfrac{n}{2} \right\rfloor + 1$, on a $\left\lfloor \dfrac{n}{2} \right\rfloor = k(n)-1$, donc par définition de la partie entière, il vient $k(n)-1 \leq \dfrac{n}{2} < k(n)$. Par suite, on en déduit que $0 < k(n)-\dfrac{n}{2} \leq 1$, et en divisant par $\sqrt{n}$, il vient $0 < \dfrac{k(n)-\frac{n}{2}}{\sqrt{n}} \leq \dfrac{1}{\sqrt{n}} \underset{n\to+\infty}{\longrightarrow} 0$. D'où le résultat.

(b). F_X étant la fonction de répartition de la loi normale $\mathcal{N}(\theta,1)$, on sait qu'elle est de classe $\mathcal{C}^2$ (au moins) sur $\mathbb{R}$. Par suite, pour tout x au voisinage de $M=\theta$, on a :

$$F_X(x) = F_X(M) + F_X'(M)(x-M) + \frac{1}{2}F_X''(M)(x-M)^2 + o\left((x-M)^2\right)$$

On a, pour tout $x \in \mathbb{R}$, $F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X-\theta \leq x-\theta) = \Phi(x-\theta)$ puisque $X-\theta \hookrightarrow \mathcal{N}(0,1)$. Ainsi :

— Comme $M=\theta$, $F_X(M) = \dfrac{1}{2}$.

— Pour tout $x \in \mathbb{R}$, $F_X'(x) = \varphi(x-\theta)$ et $F_X''(x) = \varphi''(x-\theta)$.

Ainsi, pour tout x au voisinage de θ,

$$F_X(x) = \frac{1}{2} + \varphi(\theta-\theta)(x-\theta) + \frac{1}{2}\varphi''(\theta-\theta)(x-\theta)^2 + o\left((x-\theta)^2\right)$$

soit finalement : $F_X(x) = \dfrac{1}{2} + \dfrac{1}{2}(x-\theta) + o\left((x-\theta)^2\right)$ puisque $\varphi(0) = \dfrac{1}{\sqrt{2\pi}}$ et $\varphi'(0) = 0$.

Ainsi, en appliquant cette égalité à $y_n = M + \dfrac{x}{\sqrt{n}} \underset{n\to+\infty}{\longrightarrow} \theta$ qui est donc dans un voisinage de θ à partir d'un certain rang, on en déduit que :

$$q_n = \frac{1}{2} + \frac{x}{\sqrt{2\pi n}} + o\left(\frac{x^2}{n}\right)$$

et comme $o\left(\dfrac{x^2}{n}\right) = o\left(\dfrac{1}{\sqrt{n}}\right)$, on en déduit que :

$$q_n = \frac{1}{2} + \frac{x}{\sqrt{2\pi n}} + o\left(\frac{1}{\sqrt{n}}\right)$$

(c). De $k(n) = \dfrac{n}{2} + o\left(\sqrt{n}\right)$, on en déduit que $\dfrac{1}{\sqrt{n}}k(n) = \dfrac{\sqrt{n}}{2} + o(1)$, et de $q_n = \dfrac{1}{2} + \dfrac{x}{\sqrt{2\pi n}} + o\left(\dfrac{1}{\sqrt{n}}\right)$, on en déduit que $\sqrt{n}q_n = \dfrac{\sqrt{n}}{2} + \dfrac{x}{\sqrt{2\pi}} + o(1)$, donc par somme, on en déduit que :

$$u_n = -\frac{x}{\sqrt{2\pi}} + o(1) \underset{n\to+\infty}{\longrightarrow} -\frac{x}{\sqrt{n}}$$

d'où la convergence et la limite de la suite $(u_n)_{n\geq 1}$.

(3). (a). Soit $s \in \mathbb{R}$. W_n est une variable aléatoire discrète finie puisque $S_n(x)$ suit une loi binomiale, et donc son espérance existe. Par suite, d'après le théorème de transfert pour les variables aléatoires finie, $E\left(\mathrm{e}^{sW_n}\right)$ existe et on a :

$$\begin{aligned}
E\left(\mathrm{e}^{sW_n}\right) &= E\left(\mathrm{e}^{\frac{s}{\sqrt{n}}(S_n(y_n)-nq_n)}\right) \\
&= E\left(\mathrm{e}^{-s\sqrt{n}q_n} \times \mathrm{e}^{\frac{s}{\sqrt{n}}S_n(y_n)}\right) \\
&= \mathrm{e}^{s\sqrt{n}q_n} \times E\left(\mathrm{e}^{\frac{s}{\sqrt{n}}\sum\limits_{k=1}^{n} J_k(y_n)}\right) \\
&= \mathrm{e}^{-s\sqrt{n}q_n} E\left(\prod_{k=1}^{n} \mathrm{e}^{\frac{s}{\sqrt{n}}J_k(y_n)}\right) \\
&\underset{\substack{\textbf{indepédance des}\\ J_1(y_n),\dots,J_n(y_k)}}{=} \mathrm{e}^{-s\sqrt{n}q_n} \prod_{k=1}^{n} E\left(\mathrm{e}^{\frac{s}{\sqrt{n}}J_k(y_n)}\right)
\end{aligned}$$

Or pour tout $k \in [\![1;n]\!]$, $E\left(\mathrm{e}^{\frac{s}{\sqrt{n}}J_k(y_n)}\right) = 1 - q_n + q_n\mathrm{e}^{\frac{s}{\sqrt{n}}}$ d'après **III(1)(a)**, ce qui donnera bien :

$$E\left(\mathrm{e}^{sW_n}\right) = \mathrm{e}^{-s\sqrt{n}q_n}\left(1 + q_n\mathrm{e}^{\frac{s}{\sqrt{n}}} - q_n\right)^n$$

(b). Puisque $E\left(\mathrm{e}^{\frac{s}{\sqrt{n}}J_k(y_n)}\right) = 1 - q_n + q_n\mathrm{e}^{\frac{s}{\sqrt{n}}}$ et que $\mathrm{e}^{\frac{s}{\sqrt{n}}J_k(y_n)}$ est une variable aléatoire à valeurs strictement positives son espérance l'est aussi, la quantité $\ln\left(L_{W_n}(x)\right)$ est bien définie. Par suite, par propriété du logarithme, on en déduit :

$$\ln\left(L_{W_n}(s)\right) = -s\sqrt{n}q_n + n\ln\left(1 + q_n\mathrm{e}^{\frac{s}{\sqrt{n}}} - q_n\right)$$

Or d'après **III(2)(b)**, on a montré que $q_n = \dfrac{1}{2} + \dfrac{x}{\sqrt{2\pi n}} + o\left(\dfrac{1}{\sqrt{n}}\right)$ mais aussi que $q_n = \dfrac{1}{2} + \dfrac{x}{\sqrt{2\pi n}} + o\left(\dfrac{1}{n}\right)$, d'où $s\sqrt{n}q_n = -\dfrac{s\sqrt{n}}{2} - \dfrac{s\sqrt{n}}{\sqrt{2\pi}} + o(1)$.

Par ailleurs, $\mathrm{e}^{\frac{s}{\sqrt{n}}} = 1 + \dfrac{s}{\sqrt{n}} + \dfrac{1}{2}\left(\dfrac{s}{\sqrt{n}}\right)^2 + o\left(\dfrac{1}{n}\right)$.

Le produit des deux développements limités donne alors :

$$q_n\mathrm{e}^{\frac{s}{\sqrt{n}}} = \frac{1}{2} + \frac{s}{2\sqrt{n}} + \frac{1}{4}\frac{s^2}{n} + \frac{x}{\sqrt{2\pi n}} + \frac{xs}{\sqrt{2\pi n}} + o\left(\frac{1}{n}\right)$$

$$\text{et} \qquad 1 + q_n\mathrm{e}^{\frac{s}{\sqrt{n}}} - q_n = 1 + \frac{s}{2\sqrt{n}} + \frac{s^2}{4n} + \frac{sx}{\sqrt{2\pi n}} + o\left(\frac{1}{n}\right)$$

Comme $\dfrac{s}{2\sqrt{n}} + \dfrac{s^2}{4n} + \dfrac{sx}{\sqrt{2\pi n}} + o\left(\dfrac{1}{n}\right) \underset{n\to+\infty}{\longrightarrow} 0$, on en déduit par composition des développements limités que :

$$\ln\left(1 + q_n\mathrm{e}^{\frac{1}{\sqrt{n}}} - q_n\right) = \frac{1}{2\sqrt{n}} + \frac{s^2}{8n} + \frac{sx}{\sqrt{2\pi n}} + o\left(\frac{1}{n}\right)$$

$$\text{et finalement} \qquad n\ln\left(1 + q_n\mathrm{e}^{\frac{1}{\sqrt{n}}} - q_n\right) = \frac{\sqrt{n}}{2} + \frac{s^2}{8} + \frac{sx}{\sqrt{2\pi}} + o(1)$$

Finalement on obtient : $\ln\left(L_{W_n}(s)\right) = \dfrac{s^2}{8} + o(1) \underset{n\to+\infty}{\longrightarrow} \dfrac{s^2}{8}$, et par continuité de la fonction exponentielle, on en déduit que $\lim\limits_{n\to+\infty} L_{W_n}(s) = \mathrm{e}^{\frac{s^2}{8}}$ et on a bien ainsi $\lim\limits_{n\to+\infty} L_{W_n}(s) = L_{\frac{T}{2}}(s)$ puisque $L_{\frac{T}{2}} = L_{\frac{1}{2}T+0}(s) = \mathrm{e}^{\left(\frac{1}{2}\right)^2\frac{s^2}{2}+0\times s} = \mathrm{e}^{\frac{s^2}{8}}$.

(4). On suppose que $x=0$. Alors $W_n = \dfrac{1}{\sqrt{n}}\left(S_n(\theta) - \dfrac{n}{2}\right)$ avec $S_n(\theta) \hookrightarrow \mathcal{B}\left(n;\, \frac{1}{2}\right)$ et donc d'espérance $\dfrac{n}{2}$ et de variance $\dfrac{n}{4}$.
Le théorème de la limite centré appliqué à la suite de variables aléatoires de même loi (de Bernoulli de paramètre $\frac{1}{2}$) $(J_n(\theta))_{n\geq 1}$ indépendantes de même espérance $\dfrac{1}{2}$ et de variance $\dfrac{1}{4}$, permet d'affirmer que la variable aléatoire $S_n^*(\theta) = \dfrac{S_n(\theta) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$ converge en loi vers une variable aléatoire T qui suit une loi normale $\mathcal{N}(0,1)$.
En remarquant alors que $W_n = \dfrac{1}{2} S_n^*(\theta)$, il vient :

$$\forall \omega \in \mathbb{R}, \qquad \lim_{n\to+\infty} \mathbb{P}(W_n \leq \omega) = \Phi(2\omega) = \mathbb{P}(T \leq 2\omega) = \mathbb{P}\left(\frac{T}{2} \leq \omega\right)$$

ce qui est bien la traduction de W_n converge en loi vers $\dfrac{T}{2}$.

(5). (a). D'après **I(1)(c)**, on a $\left[Y_{k(n)} \leq y_n\right] = [S_n(y_n) \geq k(n)]$. Mais aussi :

$$[Y_{k(n)} \leq y_k] = \left[Y_{k(n)} \leq \theta + \frac{x}{\sqrt{n}}\right] = \left[\sqrt{n}\,(Y_{k(n)} - \theta) \leq x\right]$$

d'où $[S_n(y_n) \geq k(n)] = \left[Y_{k(n)} \leq \theta + \dfrac{x}{\sqrt{n}}\right]$.
Par ailleurs :

$$[S_n(y_n) \geq k(n)] = [S_n(y_n) - nq_n \geq k(n) - nq_n] = \left[\frac{S_n(y-n) - nq_n}{\sqrt{n}} \geq \frac{k(n) - nq_n}{\sqrt{n}}\right] = [W_n \geq u_n]$$

(b). A finaliser

(c). On sait que $\mathbb{P}\left(\sqrt{n}\,(Y_{k(n)} - M) \leq x\right) = \mathbb{P}(W_n \geq u_n)$ pour tout $x \in \mathbb{R}$ et $n \in \mathbb{R}$. D'où en utilisant le résultat de la question précédente, $\lim\limits_{n\to+\infty} \mathbb{P}\left(\sqrt{n}\,(Y_{k(n)} - M) \leq x\right) = \mathbb{P}\left(T\sqrt{\dfrac{\pi}{2}} \leq x\right) = \mathbb{P}\left(T \leq \sqrt{\dfrac{\pi}{2}}x\right)$.
Or $\mathbb{P}\left(\sqrt{n}\,(Y_{k(n)} - M) \leq x\right) = \mathbb{P}\left(\sqrt{\dfrac{2n}{\pi}}\,(Y_{k(n)} - M) \leq \sqrt{\dfrac{2}{\pi}}x\right)$, donc :

$$\mathbb{P}\left(\sqrt{\frac{2n}{\pi}}\,(Y_{k(n)} - M) \leq \sqrt{\frac{2}{\pi}}x\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}\left(T \leq \sqrt{\frac{2}{\pi}}x\right)$$

et comme $x \longmapsto \sqrt{\dfrac{2}{\pi}}x$ est une bijection de $\mathbb{R}$ sur $\mathbb{R}$, on en déduit que :

$$\forall x \in \mathbb{R}, \qquad \mathbb{P}\left(\sqrt{\frac{2n}{\pi}}\,(Y_{k(n)} - M) \leq x\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(T \leq x)$$

d'où $\left(\sqrt{\dfrac{2n}{\pi}}\,(Y_{k(n)} - M)\right)_{n\geq 1}$ converge en loi vers T.

(6). (a). Puisque $n = 2\ell + 1$, on en déduit que $k(n) = \left\lfloor \dfrac{n}{2} \right\rfloor + 1 = \left\lfloor \ell + \dfrac{1}{2} \right\rfloor + 1 = \ell + 1$ car $\ell \in \mathbb{N}$.

(b). On sait que $E(Y_{\ell+1}) = \theta$ d'après **II(4)(d).**, et par suite $\lim\limits_{n\to+\infty} E(Y_{k(n)}) = \theta$.

(c). Puisque X possède un moment d'ordre 2, il en est de même de Y_k et ainsi $Y_{k(n)}$ admet une variance.
Or par définition de la variance, on a $v(Y_{k(n)}) = E\left((Y_{k(n)} - M)^2\right)$ car $E(Y_{\ell+1}) = \theta = M$, et comme on a que :

$$E\left(\sqrt{\frac{2n}{\pi}}\,(Y_{k(n)} - M)^2\right) \underset{n\to+\infty}{\longrightarrow} \frac{2n}{\pi} v(Y_{k(n)})$$

et que cette dernière limite est admise égale à $E(T^2) = 1$ car $T \hookrightarrow \mathcal{N}(0,1)$, il vient $\lim\limits_{n\to+\infty} \dfrac{2n}{\pi} v(Y_{k(n)}) = 1$ et par suite $v(Y_{k(n)}) \underset{n\to+\infty}{\sim} \dfrac{\pi}{2n}$.

(d). On sait que $Y_{k(n)} \in \mathcal{E}(\theta)$, donc $\mathrm{Cov}\left(\overline{X_n},\, Y_{k(n)} - \overline{X_n}\right) = 0$ et $\mathrm{Cov}\left(\overline{X_n},\, Y_{k(n)}\right) = \mathrm{Cov}\left(\overline{X_n},\, \overline{X_n}\right)$ c'est à dire $\mathrm{Cov}\left(\overline{X_n},\, Y_{k(n)}\right) = \dfrac{1}{n}$.
Par suite, le coefficient de corrélation linéaire du couple $\left(\overline{X_n}, Y_{k(n)}\right)$ vaut :

$$\rho_n = \frac{\mathrm{Cov}\left(\overline{X_n},\, Y_{k(n)}\right)}{\sqrt{v\left(\overline{X_n}\right)}\sqrt{v(Y_{k(n)})}} = \frac{1}{\sqrt{n}\sqrt{v(Y_{k(n)})}} \underset{n\to+\infty}{\sim} \frac{1}{\sqrt{n}}\sqrt{\frac{2n}{\pi}} \underset{n\to+\infty}{\longrightarrow} \sqrt{\frac{2}{\pi}}$$

---

a. Il est à noter que la notion de moment d'ordre r n'est au programme. . .sauf pour $r=1$ et $r=2$ dans le cas des variables aléatoires à densité. Cela aurait donc mérité de la part du concepteur du sujet, un (r)appel de la définition.