CCIP 2011 S
Eléments de correction

Première partie

1. a. Par indépendance de $X_1, \dots, X_n$, la fonction de répartition de $Y_n = \sup(X_1, \dots, X_n)$ est donnée par

$$F_{Y_n} : x \in \mathbb{R} \longmapsto \mathbb{P}(Y_n \leqslant x) = \mathbb{P}\Big(\bigcap_{k=1}^{n} [X_k \leqslant x]\Big) = \prod_{k=1}^{n} \mathbb{P}(X_k \leqslant x) = F_X(x)^n.$$

La fonction F_X étant continue sur $\mathbb{R}$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}$ privé d'un nombre fini de points, il en va de même de F_{Y_n} et la variable Y_n est donc à densité $f_{Y_n} : x \in \mathbb{R} \longmapsto n f_X(x) F_X(x)^{n-1}$.
On détermine de même la répartition de $Y_1 = \inf(X_1, \dots, X_n)$:

$$F_{Y_1} : x \in \mathbb{R} \longmapsto \mathbb{P}(Y_1 \leqslant x) = \mathbb{P}\Big(\bigcup_{k=1}^{n} [X_k \leqslant x]\Big) = 1 - \mathbb{P}\Big(\bigcap_{k=1}^{n} [X_k > x]\Big) = 1 - \big(1 - F_X(x)\big)^n$$

puis une densité $f_{Y_1} : x \in \mathbb{R} \longmapsto n f_X(x)\big(1 - F_X(x)\big)^{n-1}$.

b. Pour $x \in \mathbb{R}$ donné, chacune des variables $J_k(x)$, $1 \leqslant k \leqslant n$, suit la loi de Bernoulli de paramètre $\mathbb{P}\big(J_k(x) = 1\big) = \mathbb{P}(X_k \leqslant x) = F_X(x)$. Ces variables étant mutuellement indépendantes car $X_1, \dots, X_n$ le sont, leur somme $S_n(x)$ suit donc la loi binomiale $\mathcal{B}\big(n, F_X(x)\big)$.

c. L'événement $[Y_k \leqslant x]$ est réalisé si, et seulement si, au moins k parmi les n variables $X_1, \dots, X_n$ prennent une valeur inférieure ou égale à x, ce qui équivaut à $S_n(x) \geqslant k$.

d. Pour $x \in \mathbb{R}$, il vient :

$$F_{Y_k}(x) = \mathbb{P}(Y_k \leqslant x) = \mathbb{P}\big(S_n(x) \geqslant k\big) = \sum_{j=k}^{n} \mathbb{P}\big(S_n(x) = j\big) = \sum_{j=k}^{n} \binom{n}{j} F_X(x)^j \big(1 - F_X(x)\big)^{n-j}.$$

e. Les cas $k = 1$ et $k = n$ ayant déjà été traités, on se concentre sur le cas général $1 < k < n$. Comme en **a.**, la variable Y_k est à densité donnée, pour $x \in \mathbb{R}$, par

$$\begin{aligned} f_{Y_k}(x) &= \sum_{j=k}^{n} \binom{n}{j} \Big(j f_X(x) F_X(x)^{j-1} \big(1 - F_X(x)\big)^{n-j} - (n-j) f_X(x) F_X(x)^j \big(1 - F_X(x)\big)^{n-j-1} \Big) \\ &= \sum_{j=k}^{n} n \binom{n-1}{j-1} f_X(x) F_X(x)^{j-1} \big(1 - F_X(x)\big)^{n-j} \\ &\qquad\qquad - \sum_{j=k}^{n-1} n \binom{n-1}{n-j-1} f_X(x) F_X(x)^j \big(1 - F_X(x)\big)^{n-j-1} \end{aligned}$$

d'après la formule d'absorption, i.e.

$$\begin{aligned} f_{Y_k}(x) = \sum_{j=k-1}^{n-1} n \binom{n-1}{j} f_X(x) F_X(x)^j \big(1 - F_X(x)\big)^{n-j-1} \\ - \sum_{j=k}^{n-1} n \binom{n-1}{j} f_X(x) F_X(x)^j \big(1 - F_X(x)\big)^{n-j-1} \end{aligned}$$

par décalage d'indice dans la première somme, d'où le résultat :

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

f. Soient $k \in [\![1, n]\!]$ et $r \in \mathbb{N}^*$. Puisque F_X est à valeurs dans $[0, 1]$, on a

$$\forall x \in \mathbb{R}, \quad |x^r f_{Y_k}(x)| \leqslant k \binom{n}{k} |x|^r f_X(x)$$

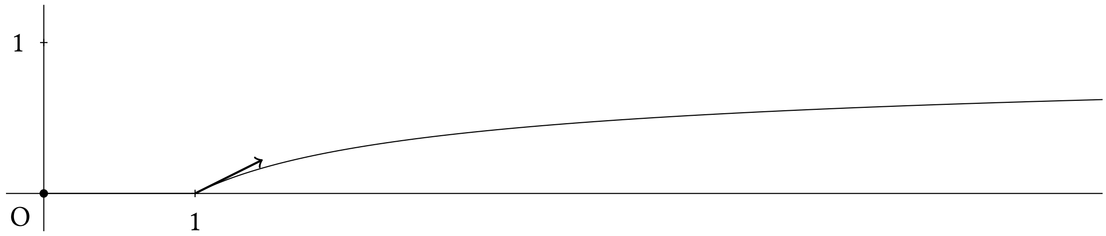
où l'intégrale $\int_{-\infty}^{+\infty} |x|^r f_X(x)\,\mathrm{d}x$ converge puisque X admet un moment d'ordre r par hypothèse. On en déduit que l'intégrale $\int_{-\infty}^{+\infty} x^r f_{Y_k}(x)\,\mathrm{d}x$ converge absolument, ce qui signifie que la variable Y_k admet un moment d'ordre r.

2. L'énoncé ne demande pas de vérifier que F_X est une fonction de répartition.

a. La fonction F_X est continue sur $\mathbb{R}$ (même en 0) et de classe $\mathscr{C}^1$ sur $\mathbb{R} \setminus \{1\}$. La variable X est donc à densité

$$f_X : x \in \mathbb{R} \longmapsto \begin{cases} 0 & \text{si } x < 1 \\ \frac{1}{2x^{3/2}} & \text{si } x \geqslant 1 \end{cases}.$$

On obtient pour F_X le graphe ci-dessous :



Puisque la restriction de F_X à $[1, +\infty[$ est de classe $\mathscr{C}^1$, la pente de la demi-tangente à droite au point d'abscisse 1 est donnée par

$$\lim_{\substack{x \to 1 \\ x > 1}} F_X'(x) = \lim_{\substack{x \to 1 \\ x > 1}} f_X(x) = \frac{1}{2}.$$

b. Pour $r \geqslant 1$:

$$x^r f_X(x) = \frac{1}{2x^{3/2-r}} \geqslant 0, \quad x \to +\infty$$

où $\frac{3}{2} - r \leqslant \frac{3}{2} - 1 < 1$ et l'intégrale de Riemann $\int_{-\infty}^{+\infty} x^r f_X(x)\,\mathrm{d}x$ est donc divergente. La variable X n'admet donc aucun moment.

c. Pour $x \in \mathbb{R}$,

$$F_X(x) = \frac{1}{2} \quad \Longleftrightarrow \quad \begin{cases} x \geqslant 1 \\ 1 - \frac{1}{\sqrt{x}} = \frac{1}{2} \end{cases} \quad \Longleftrightarrow \quad x = 4.$$

La variable X admet donc une unique médiane théorique $M = 4$.

d. D'après **1.e.**,

$$\forall x \in \mathbb{R}, \quad f_{Y_k}(x) = \begin{cases} 0 & \text{si } x < 1 \\ \frac{k}{2}\binom{n}{k}\left(1 - \frac{1}{\sqrt{x}}\right)^{k-1} \frac{1}{x^{(n-k+3)/2}} & \text{si } x \geqslant 1 \end{cases}.$$

En particulier,

$$f_{Y_k}(x) \sim \frac{k}{2}\binom{n}{k}\frac{1}{x^{(n-k+3)/2}}, \quad x \to +\infty.$$

3. a. Pour $k \in [\![1, n-2]\!]$,

$$x f_{Y_k}(x) \sim \frac{k}{2}\binom{n}{k}\frac{1}{x^{(n-k+1)/2}} \geqslant 0, \quad x \to +\infty$$

d'où l'on déduit la convergence absolue de l'intégrale $\int_{-\infty}^{+\infty} x f_{Y_k}(x)\,\mathrm{d}x$ par comparaison aux intégrales de Riemann puisque $\frac{n-k+1}{2} \geqslant \frac{3}{2} > 1$ étant donné que $n - k \geqslant 2$. La variable Y_k admet donc une espérance.

b. Le changement de variable $t \longmapsto x = \frac{1}{t^2}$, de classe $\mathscr{C}^1$ et strictement décroissant de $]0, 1]$ sur $[1, +\infty[$, donne pour $k \in [\![1, n-2]\!]$:

$$\begin{aligned} \mathbb{E}(Y_k) &= \frac{k}{2}\binom{n}{k}\int_1^{+\infty}\left(1 - \frac{1}{\sqrt{x}}\right)^{k-1}\frac{\mathrm{d}x}{x^{(n-k+1)/2}} = \frac{k}{2}\binom{n}{k}\int_0^1 (1-t)^{k-1} t^{n-k+1}\frac{2\,\mathrm{d}t}{t^3} \\ &= k\binom{n}{k}\int_0^1 (1-t)^{k-1} t^{n-k-2}\,\mathrm{d}t. \end{aligned}$$

c. Par intégration par parties sur le segment $[0, 1]$, il vient pour $r \geqslant 1$ et $s \geqslant 2$:

$$I_{r,s} = \int_0^1 t^{r-1}(1-t)^{s-1}\,\mathrm{d}t = \left[\frac{t^r}{r}(1-t)^{s-1}\right]_0^1 + \frac{s-1}{r}\int_0^1 t^r(1-t)^{s-2}\,\mathrm{d}t = \frac{s-1}{r} I_{r+1,s-1}.$$

On en déduit par récurrence sur $s \geqslant 1$ (noter l'ordre des quantificateurs : le prédicat de récurrence au rang s spécifie que la formule est valable pour tout r) que :

$$\forall s \geqslant 1, \quad \forall r \geqslant 1, \quad I_{r,s} = \frac{(s-1)!}{r(r+1)\cdots(r+s-2)} I_{r+s-1,1} = \frac{(s-1)!(r-1)!}{(r+s-1)!}.$$

d. D'après **b.** et **c.**,

$$\forall k \in [\![1, n-2]\!]\,,\quad \mathbb{E}(\mathrm{Y}_k) = k\binom{n}{k}\mathrm{I}_{n-k-1,k} = \frac{n(n-1)}{(n-k)(n-k-1)}.$$

e. De $n \geqslant 5$, on déduit que $\ell + 1 \leqslant n-2$ si bien que :

$$\mathbb{E}(\mathrm{Y}_{\ell+1}) = \frac{2\ell(2\ell+1)}{\ell(\ell-1)} = 4 + \frac{6}{\ell-1}.$$

On constate que pour n assez grand, la médiane empirique prend en moyenne des valeurs proches de la médiane théorique.

4. a. Pour $n \in \mathbb{N}^*$,

$$\forall x \in \mathbb{R},\quad \mathrm{F}_{\mathrm{Z}_n}(x) = \mathbb{P}(\mathrm{Y}_n \leqslant n^2 x) = \mathrm{F}_\mathrm{X}(n^2x)^n = \begin{cases} 0 & \text{si } x < \frac{1}{n^2} \\ \left(1 - \frac{1}{n\sqrt{x}}\right)^n & \text{si } x \geqslant \frac{1}{n^2} \end{cases}.$$

b. La fonction φ_Z est continue sur $\mathbb{R}$ (même en 0), de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$, croissante, de limites 0 en $-\infty$ et 1 en $+\infty$, c'est donc la fonction de répartition d'une variable aléatoire Z à densité.

c. Pour $x \leqslant 0$,

$$\forall n \in \mathbb{N}^*,\quad \mathrm{F}_{\mathrm{Z}_n}(x) = 0 \xrightarrow[n\to\infty]{} 0 = \varphi_\mathrm{Z}(x),$$

et pour $x > 0$ fixé,

$$\begin{aligned}\forall n \geqslant \frac{1}{\sqrt{x}},\quad \mathrm{F}_{\mathrm{Z}_n}(x) &= \left(1 - \frac{1}{n\sqrt{x}}\right)^n = \exp\left[n\left(-\frac{1}{n\sqrt{x}} + o\left(\frac{1}{n}\right)\right)\right] \\ &= \exp\left(-\frac{1}{\sqrt{x}} + o(1)\right) \xrightarrow[n\to\infty]{} \exp\left(-\frac{1}{\sqrt{x}}\right) = \varphi_\mathrm{Z}(x),\end{aligned} \qquad n \to \infty$$

ce qui établit la convergence en loi de $(\mathrm{Z}_n)_{n\geqslant 1}$ vers Z.

Deuxième partie

5. Les variables $\mathrm{X}_1, \ldots, \mathrm{X}_n$ étant indépendantes, leur somme suit la loi normale $\mathcal{N}(n\theta, n)$ et $\overline{\mathrm{X}}_n$ suit donc la loi $\mathcal{N}(\theta, \frac{1}{n})$.
La variable $\overline{\mathrm{X}}_n$ étant fonction de $\mathrm{X}_1, \ldots, \mathrm{X}_n$, c'est un estimateur de θ, sans biais puisque $\mathbb{E}(\overline{\mathrm{X}}_n) = \theta$. De plus son risque quadratique $r(\overline{\mathrm{X}}_n) = \mathbb{V}(\overline{\mathrm{X}}_n) = \frac{1}{n}$ tend vers 0 lorsque $n \to \infty$, d'où l'on déduit que $\overline{\mathrm{X}}_n$ est un estimateur convergent de θ.

6. L'énoncé ne précise pas qu'il ne considère que des intervalles de confiance dont la longueur est fixe (alors qu'elle pourrait être aléatoire).

a. La fonction Φ est continue et strictement croissante sur $\mathbb{R}$, de limites 0 en $-\infty$ et 1 en $+\infty$. Elle réalise donc une bijection de $\mathbb{R}$ sur $]0, 1[$.

b. Il s'agit de montrer que pour $\mu(\alpha) = -\frac{\Phi^{-1}(\alpha/2)}{\sqrt{n}} > 0$ (car $\frac{\alpha}{2} < \frac{1}{2} = \Phi(0)$),

$$\mathbb{P}\big(\theta \in \big[\overline{\mathrm{X}}_n - \mu(\alpha), \overline{\mathrm{X}}_n + \mu(\alpha)\big]\big) \geqslant 1 - \alpha.$$

(La même propriété est alors automatiquement vérifiée pour toute marge d'erreur supérieure à $\mu(\alpha)$...). Or, en notant $t = \Phi^{-1}(\frac{\alpha}{2}) < 0$,

$$\mathbb{P}\big(\theta \in \big[\overline{\mathrm{X}}_n - \mu(\alpha), \overline{\mathrm{X}}_n + \mu(\alpha)\big]\big) = \mathbb{P}\big(\big|\overline{\mathrm{X}}_n - \theta\big| \leqslant \mu(\alpha)\big) = \mathbb{P}\big(\sqrt{n}\,\big|\overline{\mathrm{X}}_n - \theta\big| \leqslant -t\big)$$

où $\sqrt{n}(\overline{\mathrm{X}}_n - \theta) = \overline{\mathrm{X}}_n^*$ suit la loi normale centrée réduite, si bien que

$$\begin{aligned}\mathbb{P}\big(\theta \in \big[\overline{\mathrm{X}}_n - \mu(\alpha), \overline{\mathrm{X}}_n + \mu(\alpha)\big]\big) &= \mathbb{P}(t \leqslant \overline{\mathrm{X}}_n^* \leqslant -t) = \mathbb{P}(\overline{\mathrm{X}}_n^* \leqslant -t) - \mathbb{P}(\overline{\mathrm{X}}_n^* < t) \\ &= \Phi(-t) - \Phi(t) = 1 - 2\Phi(t) = 1 - \alpha.\end{aligned}$$

c. Il vient :

$$\mu(\beta) = b\mu(\alpha) \iff \Phi^{-1}\left(\frac{\beta}{2}\right) = b\Phi^{-1}\left(\frac{\alpha}{2}\right) \iff \beta = 2\Phi\left(b\Phi^{-1}\left(\frac{\alpha}{2}\right)\right).$$

Par stricte croissance de Φ et à partir des inégalités $b < 1$ et $\Phi^{-1}(\frac{\alpha}{2}) < 0$, on obtient $\beta > 2\Phi\big(\Phi^{-1}(\frac{\alpha}{2})\big) = \alpha$: en réduisant l'intervalle de confiance ($\mu(\beta) < \mu(\alpha)$), on augmente donc le niveau de risque ($\beta > \alpha$), rien de surprenant...

7. a. L'ensemble $\mathcal{E}_\theta$ contient $\overline{X}_n$ d'après la question **5.**.

b. Pour $U_n \in \mathcal{E}_\theta$,

$$\mathbb{V}(U_n) = \mathbb{V}\big(\overline{X}_n + (U_n - \overline{X}_n)\big) = \mathbb{V}(\overline{X}_n) + 2\operatorname{cov}(\overline{X}_n, U_n - \overline{X}_n) + \mathbb{V}(U_n - \overline{X}_n) \geqslant \mathbb{V}(\overline{X}_n) \qquad (1)$$

car $\operatorname{cov}(\overline{X}_n, U_n - \overline{X}_n) = 0$ et $\mathbb{V}(U_n - \overline{X}_n) \geqslant 0$. L'estimateur $\overline{X}_n$ est donc optimal dans $\mathcal{E}_\theta$.

c. La variable $A_n(\lambda)$ est fonction de Z_n et de U_n donc de $X_1, \ldots, X_n$. De plus, $\mathbb{E}\big(A_n(\lambda)\big) = \mathbb{E}(Z_n) + \lambda\big(\mathbb{E}(U_n) - \mathbb{E}(Z_n)\big) = \theta$ car Z_n et U_n appartiennent à $\mathcal{E}_\theta$. Enfin,

$$\mathbb{V}\big(A_n(\lambda)\big) = \mathbb{V}(Z_n) + 2\lambda\operatorname{cov}(Z_n, U_n - Z_n) + \lambda^2\,\mathbb{V}(U_n - Z_n)$$

d'où, puisque Z_n est optimal dans $\mathcal{E}_\theta$,

$$\forall \lambda \in \mathbb{R}, \quad \lambda^2\,\mathbb{V}(U_n - Z_n) + 2\lambda\operatorname{cov}(Z_n, U_n - Z_n) \geqslant 0.$$

Cela implique $\operatorname{cov}(Z_n, U_n - Z_n) = 0$ sans quoi le membre de gauche, équivalent à $2\lambda\operatorname{cov}(Z_n, U_n - Z_n)$ lorsque $\lambda \to 0$, changerait de signe au voisinage de 0.

d. Si Z_n est un estimateur optimal dans $\mathcal{E}_\theta$ alors $\mathbb{V}(Z_n) \leqslant \mathbb{V}(\overline{X}_n)$ et donc $\mathbb{V}(Z_n) = \mathbb{V}(\overline{X}_n)$ puisque $\overline{X}_n$ est lui-même optimal. La relation (1) appliquée à $U_n = Z_n$ donne donc $\mathbb{V}(Z_n - \overline{X}_n) = 0$ i.e. $Z_n = \overline{X}_n$ presque sûrement. Ceci justifie (sans utiliser **c.**...) l'unicité de l'estimateur optimal dans $\mathcal{E}_\theta$, l'existence ayant été établie en **b.**.

Remarque. Il aurait été moins artificiel d'admettre que $\overline{X}_n$ est un estimateur optimal dans $\mathcal{E}_\theta$... On s'appuie alors sur le résultat démontré en **c.** pour établir l'unicité d'un tel estimateur.

8. a. La variable $X - \theta$ suivant la loi normale centrée réduite, on a pour $x \in \mathbb{R}$:

$$\mathbb{P}(X \leqslant x) = \frac{1}{2} \iff \Phi(x - \theta) = \Phi(0) \iff x = \theta$$

par injectivité de Φ. Ainsi la variable X admet $M = \theta$ pour unique médiane théorique.

b. On considérera que la variable X admet pour densité

$$f_X : x \in \mathbb{R} \longmapsto \frac{1}{\sqrt{2\pi}}\mathbf{e}^{-(x-\theta)^2/2}.$$

Il vient alors immédiatement $f_X(M) = \frac{1}{\sqrt{2\pi}}$ et, pour tout $x \in \mathbb{R}$, $f_X(2M - x) = f_X(x)$ (cette propriété traduisant la symétrie du graphe de f_X par rapport à la droite d'équation $x = M$).

Sachant que $X - \theta$ suit la loi normale centrée réduite, la variable X admet pour fonction de répartition $F_X : x \in \mathbb{R} \longmapsto \Phi(x - \theta)$ qui vérifie, d'après les propriétés de Φ :

$$\forall x \in \mathbb{R}, \quad F_X(2M - x) = \Phi(M - x) = 1 - \Phi(x - M) = 1 - F_X(x).$$

Remarque. La relation sur f_X peut se déduire par dérivation de celle sur F_X.

c. Pour $k \in [\![1, n]\!]$, la formule établie en **1.e.** fait apparaître, à la lumière des résultats obtenus en **b.**, que :

$$\begin{aligned}\forall x \in \mathbb{R}, \quad f_{Y_k}(2M - x) &= n\binom{n-1}{k-1}F_X(2M - x)^{k-1}\big(1 - F_X(2M - x)\big)^{n-k}f_X(2M - x)\\ &= n\binom{n-1}{n-k}\big(1 - F_X(x)\big)^{k-1}F_X(x)^{n-k}f_X(x) = f_{Y_{n-k+1}}(x)\end{aligned}$$

d'où l'on déduit, en appliquant le changement de variable affine $y = 2M - x$ (les espérances existent d'après **1.f.**), que :

$$\begin{aligned}\mathbb{E}(M - Y_k) &= \int_{-\infty}^{+\infty}(M - x)f_{Y_k}(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty}(y - M)f_{Y_k}(2M - y)\,\mathrm{d}y\\ &= \int_{-\infty}^{+\infty}(y - M)f_{Y_{n-k+1}}(y)\,\mathrm{d}y = \mathbb{E}(Y_{n-k+1} - M).\end{aligned}$$

Remarque. Ce résultat aurait pu être justifié plus directement. En effet, à partir de la question **b.**, on justifie que les variables $X'_j = 2M - X_j$, $1 \leqslant j \leqslant n$, ont même loi que X et sont par ailleurs indépendantes. Les variables $Y'_k = 2M - Y_{n-k+1}$, $1 \leqslant k \leqslant n$, donnent alors le réarrangement croissant des valeurs prises par les variables $X'_1, \ldots, X'_n$. Dans ces conditions, $Y'_k = 2M - Y_{n-k+1}$ a même loi que Y_k, si bien que $\mathbb{E}(2M - Y_{n-k+1}) = \mathbb{E}(Y_k)$.

d. La relation obtenue en **c.** appliquée à $k = \ell+1$ donne $\mathbb{E}(\mathrm{M}-\mathrm{Y}_{\ell+1}) = \mathbb{E}(\mathrm{Y}_{\ell+1}-\mathrm{M})$ i.e. $\mathbb{E}(\mathrm{Y}_{\ell+1}) = \mathrm{M}$. Ainsi la médiane empirique $\mathrm{Y}_{\ell+1}$ est un estimateur sans biais de $\mathrm{M} = \theta$ si bien, d'après **7.b.**, que $\mathbb{V}(\mathrm{Y}_{\ell+1}) \geqslant \mathbb{V}(\overline{\mathrm{X}}_n) = \frac{1}{n}\mathbb{V}(\mathrm{X}) = \frac{1}{n}$. Il en ressort que la moyenne empirique est un meilleur estimateur de la médiane théorique M que la médiane empirique.

Troisième partie

9. a. Le théorème de transfert pour une variable finie donne :

$$\forall s \in \mathbb{R}, \quad \mathrm{L}_\mathrm{J}(s) = \mathbb{E}(\mathbf{e}^{s\mathrm{J}}) = 1 - p + p\mathbf{e}^s.$$

b. Toujours par transfert, les intégrales étant absolument convergentes par référence aux densités gaussiennes,

$$\begin{aligned}\forall s \in \mathbb{R}, \quad \mathrm{L}_\mathrm{T}(s) = \mathbb{E}(\mathbf{e}^{s\mathrm{T}}) &= \int_{-\infty}^{+\infty} \mathbf{e}^{st} f_\mathrm{T}(t)\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathbf{e}^{st-t^2/2}\,\mathrm{d}t \\ &= \frac{\mathbf{e}^{s^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathbf{e}^{-(t-s)^2/2}\,\mathrm{d}t = \mathbf{e}^{s^2/2}.\end{aligned}$$

c. Pour $(\theta,\sigma) \in \mathbb{R}\times\mathbb{R}_+^*$, on a donc :

$$\forall s \in \mathbb{R}, \quad \mathrm{L}_{\sigma\mathrm{T}+\theta}(s) = \mathbb{E}(\mathbf{e}^{s(\sigma\mathrm{T}+\theta)}) = \mathbf{e}^{s\theta}\,\mathbb{E}(\mathbf{e}^{(s\sigma)\mathrm{T}}) = \exp\Big(\sigma^2\frac{s^2}{2} + \theta s\Big).$$

10. a. Par construction, $0 \leqslant \lfloor \frac{n}{2}\rfloor + 1 - \frac{n}{2} \leqslant 1$ pour tout $n \in \mathbb{N}$. Ainsi $\lfloor \frac{n}{2}\rfloor + 1 - \frac{n}{2}$ est borné donc négligeable devant $\sqrt{n}$ lorsque $n \to \infty$, si bien que

$$k(n) = \Big\lfloor\frac{n}{2}\Big\rfloor + 1 = \frac{n}{2} + \Big(\Big\lfloor\frac{n}{2}\Big\rfloor + 1 - \frac{n}{2}\Big) = \frac{n}{2} + o(\sqrt{n}), \quad n \to \infty.$$

b. La fonction F_X étant de classe $\mathscr{C}^2$ sur $\mathbb{R}$, on peut lui appliquer la formule de Taylor-Young à l'ordre 1 au voisinage de M :

$$\mathrm{F}_\mathrm{X}(y) = \mathrm{F}_\mathrm{X}(\mathrm{M}) + \mathrm{F}_\mathrm{X}'(\mathrm{M})(y-\mathrm{M}) + o(y-\mathrm{M}) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}(y-\mathrm{M}) + o(y-\mathrm{M}), \quad y \to \mathrm{M}$$

d'où, puisque $y_n = \mathrm{M} + \frac{x}{\sqrt{n}} \to \mathrm{M}$,

$$q_n = \mathrm{F}_\mathrm{X}(y_n) = \frac{1}{2} + \frac{x}{\sqrt{2\pi n}} + o\Big(\frac{1}{\sqrt{n}}\Big), \quad n \to \infty.$$

c. De **a.** et **b.**, on déduit :

$$u_n = \frac{1}{\sqrt{n}}\big(k(n) - nq_n\big) = \frac{1}{\sqrt{n}}\Big(-\frac{x}{\sqrt{2\pi}}\sqrt{n} + o(\sqrt{n})\Big) = -\frac{x}{\sqrt{2\pi}} + o(1) \to -\frac{x}{\sqrt{2\pi}}, \quad n \to \infty.$$

11. a. Pour $s \in \mathbb{R}$, les variables $\mathrm{S}_n(y_n)$ donc W_n et $\mathbf{e}^{s\mathrm{W}_n}$ sont finies et admettent ainsi une espérance donnée par :

$$\mathrm{L}_{\mathrm{W}_n}(s) = \mathbb{E}\Big[\exp\Big(\frac{s}{\sqrt{n}}\sum_{k=1}^n \mathrm{J}_k(y_n) - s\sqrt{n}q_n\Big)\Big] = \mathbf{e}^{-s\sqrt{n}q_n}\,\mathbb{E}\Big[\prod_{k=1}^n \mathbf{e}^{s\mathrm{J}_k(y_n)/\sqrt{n}}\Big]$$

i.e., puisque les variables $\mathrm{J}_1(y_n),\dots,\mathrm{J}_n(y_n)$ sont indépendantes et identiquement distribuées de loi $\mathcal{B}(q_n)$,

$$\mathrm{L}_{\mathrm{W}_n}(s) = \mathbf{e}^{-s\sqrt{n}q_n}\prod_{k=1}^n \mathbb{E}(\mathbf{e}^{s\mathrm{J}_k(y_n)/\sqrt{n}}) = \mathbf{e}^{-s\sqrt{n}q_n}\Big[\mathrm{L}_{\mathrm{J}_1(y_n)}\Big(\frac{s}{\sqrt{n}}\Big)\Big]^n = \mathbf{e}^{-s\sqrt{n}q_n}\big(1 - q_n + q_n\mathbf{e}^{s/\sqrt{n}}\big)^n.$$

b. Pour $s \in \mathbb{R}$ donné (considéré comme une constante !), on effectue un développement asymptotique [1]

1. En faisant attention à la gestion des restes : par exemple,
- ➢ puisque (q_n) converge vers $\frac{1}{2} \neq 0$, un terme en $o(\frac{q_n}{n^k})$ est tout simplement un $o(\frac{1}{n^k})$;
- ➢ après avoir utilisé le développement à l'ordre 2 de exp, un développement de ln à l'ordre 1 serait insuffisant car il produirait un développement asymptotique du crochet à la précision $o(\frac{1}{\sqrt{n}})$.

lorsque $n \to \infty$, sachant la suite (q_n) convergente vers $\frac{1}{2} \neq 0$:

$$\begin{aligned} n\ln\left(1-q_n+q_n\mathbf{e}^{s/\sqrt{n}}\right) &= n\ln\left[1-q_n+q_n\left(1+\frac{s}{\sqrt{n}}+\frac{s^2}{2n}+o\left(\frac{1}{n}\right)\right)\right] \\ &= n\ln\left[1+s\frac{q_n}{\sqrt{n}}+s^2\frac{q_n}{2n}+o\left(\frac{1}{n}\right)\right] = n\left[s\frac{q_n}{\sqrt{n}}+s^2\frac{q_n}{2n}+o\left(\frac{1}{n}\right)-\frac{s^2}{2}\frac{q_n^2}{n}+o\left(\frac{q_n^2}{n}\right)\right] \\ &= s\sqrt{n}q_n+\frac{s^2}{2}(q_n-q_n^2)+o(1) = s\sqrt{n}q_n+\frac{s^2}{2}\left(\frac{1}{4}+o(1)\right)+o(1) = s\sqrt{n}q_n+\frac{s^2}{8}+o(1) \end{aligned}$$

d'où

$$\ln\left(\mathrm{L}_{\mathrm{W}_n}(s)\right) = -s\sqrt{n}q_n+n\ln\left(1-q_n+q_n\mathbf{e}^{s/\sqrt{n}}\right) = \frac{s^2}{8}+o(1) \to \frac{s^2}{8}, \quad n\to\infty$$

et, par continuité de l'exponentielle,

$$\mathrm{L}_{\mathrm{W}_n}(s) \xrightarrow[n\to\infty]{} \mathbf{e}^{s^2/8} = \mathrm{L}_{\mathrm{T}/2}(s).$$

12. Pour $x=0$, $\mathrm{S}_n(y_n)=\mathrm{S}_n(\mathrm{M})$ suit la loi $\mathcal{B}(n,\frac{1}{2})$ d'après **1.b.** et

$$2\mathrm{W}_n = \frac{1}{\sqrt{\frac{n}{4}}}\left(\mathrm{S}_n(\mathrm{M})-\frac{n}{2}\right) = \mathrm{S}_n(\mathrm{M})^* = \bar{\mathrm{J}}_n(\mathrm{M})^*$$

est la variable aléatoire centrée réduite associée à la moyenne empirique

$$\bar{\mathrm{J}}_n(\mathrm{M}) = \frac{\mathrm{S}_n(\mathrm{M})}{n} = \frac{1}{n}\sum_{k=1}^{n}\mathrm{J}_k(\mathrm{M}).$$

Puisque les variables $\mathrm{J}_n(\mathrm{M})$, $n\in\mathbb{N}^*$, sont indépendantes et identiquement distribuées admettant une variance, le théorème limite central assure que la suite $\left(\bar{\mathrm{J}}_n(\mathrm{M})^*\right)_{n\in\mathbb{N}^*}$ converge en loi vers la loi normale centrée réduite :

$$\forall x\in\mathbb{R}, \quad \mathbb{P}(2\mathrm{W}_n \leqslant x) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathrm{T}\leqslant x).$$

Il en résulte immédiatement que

$$\forall x\in\mathbb{R}, \quad \mathbb{P}(\mathrm{W}_n\leqslant x) = \mathbb{P}(2\mathrm{W}_n\leqslant 2x) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathrm{T}\leqslant 2x) = \mathbb{P}\left(\frac{\mathrm{T}}{2}\leqslant x\right),$$

i.e. que $(\mathrm{W}_n)_{n\in\mathbb{N}^*}$ converge en loi vers $\frac{1}{2}\mathrm{T}$.

13. a. Pour $n\in\mathbb{N}^*$ et $x\in\mathbb{R}$, on a par définition et d'après **1.c.** :

$$\left[\sqrt{n}(\mathrm{Y}_{k(n)}-\mathrm{M})\leqslant x\right] = \left[\mathrm{Y}_{k(n)}\leqslant y_n\right] = \left[\mathrm{S}_n(y_n)\geqslant k(n)\right] = [\mathrm{W}_n\geqslant u_n].$$

b. On a (en partie) établi en **11.** la convergence en loi de (W_n) vers $\frac{1}{2}\mathrm{T}$. Par ailleurs, la suite (u_n) converge vers $\ell = -\frac{x}{\sqrt{2\pi}}$ et cette convergence a aussi lieu en probabilité : en effet, pour $\varepsilon>0$ donné, on a $|u_n-\ell|<\varepsilon$ à partir d'un certain rang et la suite de terme général $\mathbb{P}(|u_n-\ell|\geqslant\varepsilon)$, $n\in\mathbb{N}^*$, nulle à partir d'un certain rang, converge donc vers 0.
Dans ces conditions, le lemme de Slutsky assure la convergence en loi de la suite (W_n-u_n) vers $\frac{1}{2}\mathrm{T}+\frac{x}{\sqrt{2\pi}}$. En composant par la fonction continue $z\longmapsto -z$, on en déduit que la suite $(u_n-\mathrm{W}_n)$ converge en loi vers $-\frac{1}{2}\mathrm{T}-\frac{x}{\sqrt{2\pi}}$:

$$\forall z\in\mathbb{R}, \quad \mathbb{P}(u_n-\mathrm{W}_n\leqslant z) \xrightarrow[n\to\infty]{} \mathbb{P}\left(-\frac{\mathrm{T}}{2}-\frac{x}{\sqrt{2\pi}}\leqslant z\right).$$

Pour $z=0$, on obtient en particulier :

$$\mathbb{P}(\mathrm{W}_n\geqslant u_n) \xrightarrow[n\to\infty]{} \mathbb{P}\left(-\frac{\mathrm{T}}{2}-\frac{x}{\sqrt{2\pi}}\leqslant 0\right) = 1-\Phi\left(-x\sqrt{\frac{2}{\pi}}\right) = \Phi\left(x\sqrt{\frac{2}{\pi}}\right) = \mathbb{P}\left(\sqrt{\frac{\pi}{2}}\mathrm{T}\leqslant x\right).$$

c. D'après **a.** et **b.**,

$$\begin{aligned} \forall x\in\mathbb{R}, \quad \mathbb{P}\left(\sqrt{\frac{2n}{\pi}}(\mathrm{Y}_{k(n)}-\mathrm{M})\leqslant x\right) &= \mathbb{P}\left(\sqrt{n}(\mathrm{Y}_{k(n)}-\mathrm{M})\leqslant x\sqrt{\frac{\pi}{2}}\right) \\ &\xrightarrow[n\to\infty]{} \mathbb{P}\left(\sqrt{\frac{\pi}{2}}\mathrm{T}\leqslant x\sqrt{\frac{\pi}{2}}\right) = \mathbb{P}(\mathrm{T}\leqslant x), \end{aligned}$$

ce qui établit la convergence en loi de la suite de terme général $\sqrt{\frac{2n}{\pi}}(\mathrm{Y}_{k(n)}-\mathrm{M})$, $n\in\mathbb{N}^*$, vers la loi normale centrée réduite.

14. a. On a $k(n) = \ell + 1$.

b. D'après **8.d.**, $\mathbb{E}(\mathrm{Y}_{k(n)}) = \mathbb{E}(\mathrm{Y}_{\ell+1}) = \mathrm{M} \to \mathrm{M}$ lorsque $n \to \infty$.

c. D'après **1.f.**, $\mathrm{Y}_{k(n)}$ admet une variance. Le résultat admis donne :

$$\mathbb{V}(\mathrm{Y}_{k(n)}) = \mathbb{E}\big[\big(\mathrm{Y}_{k(n)} - \mathrm{M}\big)^2\big] = \frac{\pi}{2n}\,\mathbb{E}\Big[\Big(\sqrt{\frac{2n}{\pi}}\big(\mathrm{Y}_{k(n)} - \mathrm{M}\big)\Big)^2\Big] \sim \frac{\pi}{2n}\,\mathbb{E}(\mathrm{T}^2) = \frac{\pi}{2n}, \quad n \to \infty.$$

d. D'après le résultat admis en **7.**,

$$\mathrm{cov}(\mathrm{Y}_{k(n)}, \overline{\mathrm{X}}_n) = \mathrm{cov}(\overline{\mathrm{X}}_n, \overline{\mathrm{X}}_n) + \mathrm{cov}(\mathrm{Y}_{k(n)} - \overline{\mathrm{X}}_n, \overline{\mathrm{X}}_n) = \mathbb{V}(\overline{\mathrm{X}}_n)$$

d'où, d'après **5.** et **c.**,

$$\varrho_n = \frac{\mathrm{cov}(\mathrm{Y}_{k(n)}, \overline{\mathrm{X}}_n)}{\sigma(\mathrm{Y}_{k(n)})\sigma(\overline{\mathrm{X}}_n)} = \frac{\sigma(\overline{\mathrm{X}}_n)}{\sigma(\mathrm{Y}_{k(n)})} \sim \sqrt{\frac{2}{\pi}} \to \sqrt{\frac{2}{\pi}}, \quad n \to \infty.$$