

## CONCOURS BLANC ÉPREUVE II

### ECG2 MATHS APPLIQUÉES

Le théorème limite central est l'un des théorèmes fondamentaux en probabilités : il occupe une place centrale aussi bien du point de vue théorique que des applications (notamment en statistique). Le but de ce problème est d'explorer diverses applications de ce théorème. La première partie étudie quelques propriétés, applications simples, et généralisation du théorème limite central. La deuxième partie se concentre sur l'utilisation de ce théorème en statistique, en particulier dans le cadre de sondages électoraux. La troisième partie s'attelle à démontrer le théorème limite central, via une adaptation de la méthode de Lindeberg.

Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . Pour une variable aléatoire  $X$ , on notera  $E(X)$  son espérance et  $V(X)$  sa variance lorsqu'elles existent.

Pour tout le problème, on se donne une suite de variables aléatoires réelles  $(X_i)_{i \in \mathbb{N}^*}$  indépendantes et de même loi. On supposera qu'elles admettent un moment d'ordre deux et on notera  $\mu = E(X_i)$  leur espérance commune et  $\sigma^2 = V(X)$  leur variance commune avec  $\sigma > 0$ . Enfin, pour tout entier naturel  $n \geq 1$ , on définit les variables aléatoires

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad Z_n = \sqrt{n} \left( \frac{\bar{X}_n - \mu}{\sigma} \right).$$

Dans tout le sujet, on suppose déjà importées sous leur alias habituels les bibliothèques Python usuelles.

```
1 import numpy as np
2 import numpy.random as rd
3 import numpy.linalg as al
4 import matplotlib.pyplot as plt
```

#### PREMIÈRE PARTIE : AUTOUR DU THÉORÈME LIMITE CENTRAL

1. Soit  $Z$  une variable aléatoire de loi normale  $\mathcal{N}(0, 1)$ , c'est-à-dire que la loi de  $Z$  admet la densité  $f_Z$  définie par  $f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

On note  $\Phi$  la fonction de répartition de  $Z$  définie par :  $\Phi(x) = \int_{-\infty}^x f_Z(t) dt$ .

- Montrer que  $\Phi$  est continue sur  $\mathbb{R}$ .
  - Montrer que la fonction  $\Phi$  est strictement croissante sur  $\mathbb{R}$ .
  - Montrer que la fonction  $\Phi$  est une bijection de  $\mathbb{R}$  sur  $]0, 1[$ .
  - Montrer que pour tout  $x \in \mathbb{R}$  on a  $\Phi(-x) = 1 - \Phi(x)$ .
2. a. Énoncer la loi faible des grands nombres pour la suite de variables aléatoires  $(X_i)_{i \geq 1}$ .
- b. Rappeler les hypothèses du théorème limite central et en déduire que pour tout réel  $x \in \mathbb{R}$  on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n \leq x) = \Phi(x) \quad (*)$$

3. Donnons une première application du théorème limite central, pour un joueur de fléchettes. Au  $i$ -ème lancer de fléchette, le score est une variable aléatoire  $X_i$  qui prend ses valeurs dans  $\{0, 2, 5, 10\}$ . On suppose que les  $X_i$  sont indépendantes et de même loi donnée par :

$$\mathbb{P}(X_i = 0) = \frac{1}{5}, \quad \mathbb{P}(X_i = 2) = \frac{1}{2}, \quad \mathbb{P}(X_i = 5) = \frac{1}{5}, \quad \mathbb{P}(X_i = 10) = \frac{1}{10}.$$

- a. Calculer  $E(X_i)$ .

b. Calculer  $V(X_i)$ .

c. Soit  $f: [0, 1] \rightarrow \mathbb{R}$  la fonction définie de la manière suivante :

$$f(x) = \begin{cases} 0 & \text{si } x \in \left[0, \frac{1}{5}\right[ \\ 2 & \text{si } x \in \left[\frac{1}{5}, \frac{7}{10}\right[ \\ 5 & \text{si } x \in \left[\frac{7}{10}, \frac{9}{10}\right[ \\ 10 & \text{si } x \in \left[\frac{9}{10}, 1\right] \end{cases} .$$

- (i) Soit  $U$  une variable aléatoire de loi uniforme sur  $[0, 1]$ . Montrer que  $f(U)$  a la même loi que  $X_i$ .
- (ii) Compléter le programme Python suivant, qui permet de générer un nombre aléatoire de même loi que  $X_i$ . On rappelle que la fonction `rd.rand()` simule une variable aléatoire uniforme sur  $[0, 1]$ .

```
1 def simul_X() :
2   U = rd.rand()
3   return ...
```

Après  $n$  lancers de fléchettes, le score du joueur est  $S_n = \sum_{i=1}^n X_i$ .

d. Exprimer  $Z_n$  en fonction de  $S_n$  et de  $n$ .

e. Un joueur lance  $n = 200$  fléchettes. En utilisant le théorème limite central, montrer que la probabilité que le score du joueur soit inférieur ou égal à 500 est approximativement  $\Phi(-2, 5)$ . Cette probabilité vaut environ  $6 \cdot 10^{-3}$ .

4. Soit  $N$  un entier naturel fixé supérieur ou égal à 1. La fonction  $\Phi: \mathbb{R} \rightarrow ]0, 1[$  étant strictement croissante et bijective, pour  $k \in \{1, 2, \dots, 2N - 1\}$ , on peut définir le réel  $x_k = \Phi^{-1}\left(\frac{k}{2N}\right)$ . On pose :  $x_0 = -\infty$  et  $x_{2N} = +\infty$ , avec par convention  $\Phi(x_0) = 0$  et  $\Phi(x_{2N}) = 1$ .

a. Montrer qu'il existe un  $n_0$  (qui dépend de  $N$ ) tel que pour tout  $n \geq n_0$ , on a :

$$\max_{k \in \{0, 1, \dots, 2N\}} |\mathbb{P}(Z_n \leq x_k) - \Phi(x_k)| \leq \frac{1}{2N}.$$

On divise l'ensemble des réels en intervalles  $I_k = ]x_{k-1}, x_k[$  pour  $k \in \{1, 2, \dots, 2N\}$ , avec par convention  $I_1 = ]-\infty, x_1[$  et  $I_{2N} = ]x_{2N-1}, +\infty[$ .

b. Soit  $k \in \{1, 2, \dots, 2N - 1\}$  et  $x$  un réel quelconque tel que  $x \in I_k$ . Soit  $n \geq n_0$ .

- (i) Montrer que l'on a  $\mathbb{P}(Z_n \leq x) - \Phi(x) \leq \mathbb{P}(Z_n \leq x_k) - \Phi(x_{k-1})$ .
- (ii) En déduire que  $\mathbb{P}(Z_n \leq x) - \Phi(x) \leq \Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{1}{N}$ .
- (iii) De même, montrer que l'on a  $\Phi(x) - \mathbb{P}(Z_n \leq x) \leq \frac{1}{N}$ .

c. En déduire que pour tout réel  $x$  et tout  $n \geq n_0$  on a  $|\mathbb{P}(Z_n \leq x) - \Phi(x)| \leq \frac{1}{N}$ .

d. Soit  $(M_n)_{n \geq 1}$  une suite de majorants des fonctions  $D_n$  définies pour  $x \in \mathbb{R}$  par

$$D_n(x) = \mathbb{P}(Z_n \leq x) - \Phi(x),$$

c'est-à-dire tels que  $|\mathbb{P}(Z_n \leq x) - \Phi(x)| \leq M_n$  pour tout  $x \in \mathbb{R}$  et tout entier  $n \geq 1$ . Montrer que l'on peut choisir la suite  $(M_n)_{n \geq 1}$  telle que

$$\lim_{n \rightarrow +\infty} M_n = 0. \quad (**)$$

Notons que  $(**)$  est une version plus forte que  $(*)$ .

5. Soit  $x$  un réel fixé.

a. Soit  $(x_n)_{n \in \mathbb{N}^*}$  une suite de réels telle que  $\lim_{n \rightarrow +\infty} x_n = x$ .

- (i) Montrer que  $\lim_{n \rightarrow +\infty} \Phi(x_n) = \Phi(x)$ .
- (ii) En appliquant le résultat (\*\*) de la question 4.d, montrer que
 
$$\lim_{n \rightarrow +\infty} |\mathbb{P}(Z_n \leq x_n) - \Phi(x_n)| = 0.$$
- (iii) En conclure que l'on a  $\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n \leq x_n) = \Phi(x)$ .
- b. (i) Montrer que pour tout  $n \geq 1$  on a  $\mathbb{P}(Z_n \leq x - \frac{1}{n}) \leq \mathbb{P}(Z_n < x) \leq \mathbb{P}(Z_n \leq x)$ .
- (ii) En déduire que  $\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n < x) = \Phi(x)$ .
- c. Montrer que, pour tous réels  $a, b$  qui vérifient  $a < b$ , on a
 
$$\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n \in [a, b]) = \Phi(b) - \Phi(a).$$

DEUXIÈME PARTIE : APPLICATIONS EN STATISTIQUES

Dans toute cette partie, on suppose que les variables aléatoires  $X_i$  sont indépendantes et de loi de Bernoulli de paramètre  $p$ , où  $p \in ]0, 1[$ . On se servira de ces variables aléatoires pour modéliser une élection entre les candidats A et B :  $X_i = 1$  si la  $i$ -ème personne vote pour le candidat A et  $X_i = 0$  si la  $i$ -ème personne vote pour le candidat B. Le paramètre  $p$  représente la proportion des voix qu'obtient le candidat A.

Les sondages cherchent à estimer le paramètre  $p$  inconnu pour anticiper le résultat de l'élection. On sélectionne  $n$  personnes (dans la population totale) et on note comme précédemment  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ , qui correspond à la proportion des personnes (parmi celles sélectionnées) qui votent pour le candidat A.

- 6. a. Démontrer que  $E(X_i) = p$  et que  $V(X_i) = p(1 - p)$ .
- b. On notera dans la suite  $\sigma = \sqrt{p(1 - p)}$ . Montrer que  $\sigma \leq \frac{1}{2}$ .
- c. Montrer que  $E(\bar{X}_n) = p$ .
- d. Montrer que  $V(\bar{X}_n) = \frac{1}{n}\sigma^2$ .
- 7. On peut construire un intervalle de confiance pour  $p$  en appliquant le théorème limite central.

a. Montrer que pour tout  $a > 0$  on a

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{\sqrt{n}}{\sigma}(\bar{X}_n - p) \in [-a, a]\right) = \Phi(a) - \Phi(-a).$$

b. En déduire que

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(p \in \left[\bar{X}_n - a \frac{\sigma}{\sqrt{n}}, \bar{X}_n + a \frac{\sigma}{\sqrt{n}}\right]\right) = 2\Phi(a) - 1.$$

c. Une table des valeurs de  $\Phi$  donne  $\Phi(1,96) \approx 0,975$ . En déduire que pour  $n$  grand, le paramètre  $p$  a approximativement 95% de chances d'appartenir à l'intervalle  $\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}}\right]$ .

Un problème est que  $\sigma$  dépend de  $p$ , donc l'intervalle ci-dessus dépend encore de  $p$ , qui est inconnu.

d. Montrer que pour  $n$  grand, le paramètre  $p$  a approximativement plus de 95% de chances d'appartenir à l'intervalle  $\left[\bar{X}_n - \frac{0,98}{\sqrt{n}}, \bar{X}_n + \frac{0,98}{\sqrt{n}}\right]$ .

Une autre solution est d'utiliser les observations de  $X_1, \dots, X_n$  pour estimer  $\sigma$ .

- 8. On pose, pour  $n \geq 1$ ,  $V_n = \bar{X}_n(1 - \bar{X}_n) + \frac{1}{n}$ . Soit  $\varepsilon > 0$ .
- a. Montrer que  $V_n - \sigma^2 - \frac{1}{n} = (\bar{X}_n - p)(1 - \bar{X}_n - p)$ .
- b. En déduire que  $|V_n - \sigma^2| \leq 2|\bar{X}_n - p| + \frac{1}{n}$ .
- c. Montrer que  $\mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}(|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n})$ .

- d. Montrer que pour  $n$  assez grand  $\mathbb{P}(|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}) \leq \mathbb{P}(|\bar{X}_n - p| > \frac{\varepsilon}{4})$ .  
 e. Conclure que pour tout  $\varepsilon > 0$  on a  $\lim_{n \rightarrow +\infty} \mathbb{P}(|V_n - \sigma^2| > \varepsilon) = 0$ .

9. On pose maintenant  $W_n = \frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - p) = \frac{\sigma}{\sqrt{V_n}}Z_n$ . On se fixe un réel  $x$ .

a. Soit  $\varepsilon > 0$ .

(i) Montrer que  $\mathbb{P}(W_n \leq x) \leq \mathbb{P}(Z_n \leq (1 + \varepsilon)x) + \mathbb{P}\left(\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right)$ .

(ii) Montrer que  $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right) = 0$ .

(iii) Montrer que  $\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n \leq (1 + \varepsilon)x) = \Phi((1 + \varepsilon)x)$ .

(iv) En déduire qu'il existe un entier  $n_\varepsilon$  tel que pour tout  $n \geq n_\varepsilon$  on a

$$\mathbb{P}(W_n \leq x) \leq \Phi((1 + \varepsilon)x) + \varepsilon.$$

b. On admettra que, de manière symétrique, pour tout  $\varepsilon > 0$  il existe un  $n_\varepsilon$  tel que pour tout  $n \geq n_\varepsilon$  on a  $\mathbb{P}(W_n \leq x) \geq \Phi((1 - \varepsilon)x) - \varepsilon$ . En conclure que l'on a  $\lim_{n \rightarrow +\infty} \mathbb{P}(W_n \leq x) = \Phi(x)$ .

10. a. Montrer que pour tout  $x \in \mathbb{R}$  on a  $\lim_{n \rightarrow +\infty} \mathbb{P}\left(p \geq \bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}\right) = \Phi(x)$ .

b. Le candidat A remporte effectivement l'élection si on a  $p \geq \frac{1}{2}$ . Une semaine avant l'élection, un sondage réalisé auprès de  $n = 1000$  personnes donne pour  $\bar{X}_n$  la valeur 0,52 (et donc pour  $V_n$  la valeur 0,2506).

Montrer que la probabilité que le candidat A remporte l'élection est approximativement

$$\Phi\left(\frac{\sqrt{1000}}{\sqrt{0.2506}}\left(0,52 - \frac{1}{2}\right)\right).$$

On a  $\frac{\sqrt{1000}}{\sqrt{0.2506}}\left(0,52 - \frac{1}{2}\right) \approx 1,27$  et une table donne  $\Phi(1,27) \approx 0,9$ . Il y a donc environ 1 chance sur 10 d'avoir  $p < \frac{1}{2}$ , c'est-à-dire que le candidat B remporte l'élection.

11. Lors des dernières élections, on s'est rendu compte que les électeurs pouvaient mentir lors du sondage (ou bien tout simplement changer d'avis entre le sondage et l'élection) : avec une probabilité  $q$  déterminée, un électeur votera pour B alors qu'il avait déclaré qu'il voterait pour A.

La réponse enregistrée par l'institut de sondage est  $Y_i = X_i + (1 - X_i)T_i$ , où  $T_i = 1$  si la  $i$ -ème personne change d'avis en faveur de B et  $T_i = 0$  sinon. On suppose que les variables  $T_i$  sont indépendantes, de loi de Bernoulli de paramètre  $q$  et qu'elles sont indépendantes des  $X_i$ .

a. Montrer que  $Y_i$  est une variable aléatoire de Bernoulli dont on déterminera le paramètre  $r$ .

On définit  $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$ , qui est la mesure effectuée par le sondage et on pose  $U_n = \bar{Y}_n(1 - \bar{Y}_n) + \frac{1}{n}$ .

De la même manière que dans la question 9, on admet qu'on a, pour tout  $x \in \mathbb{R}$

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{\sqrt{n}}{\sqrt{U_n}}(\bar{Y}_n - r) \leq x\right) = \Phi(x).$$

b. Montrer que  $\lim_{n \rightarrow +\infty} \mathbb{P}\left(p \geq \frac{1}{1-q}\left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q\right)\right) = \Phi(x)$ .

c. On suppose que le sondage sur  $n$  personnes a donné à  $\bar{Y}_n$  la valeur  $\bar{y}_n$ , et donc pour  $U_n$  la valeur  $u_n = \bar{y}_n(1 - \bar{y}_n) + \frac{1}{n}$ .

Montrer que la probabilité que le candidat A remporte effectivement l'élection, c'est-à-dire que  $p \geq \frac{1}{2}$ , vaut approximativement  $\Phi\left(\frac{\sqrt{n}}{\sqrt{u_n}}\left(\bar{y}_n - \frac{1}{2}(1 + q)\right)\right)$ .

d. Prenons les mêmes données que plus haut : le sondage de  $n = 1000$  personnes donne  $\bar{y}_n = 0,52$ . Si  $q = 0,04$ , montrer que la probabilité que le candidat A remporte effectivement l'élection n'est plus que de  $1/2$ .

TROISIÈME PARTIE : DÉMONSTRATION DU THÉORÈME LIMITE CENTRAL

Dans cette partie, on suppose que les  $X_i$  admettent un moment d'ordre 3, donc  $E(X)$ ,  $E(X^2)$  et  $E(X^3)$  existent. Sous cette condition, on va démontrer le théorème limite central, c'est-à-dire le résultat (\*).

On supposera aussi pour simplifier que  $\mu = E(X) = 0$  et  $\sigma^2 = V(X) = 1$ , de sorte que  $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ .

Soit  $x$  un réel fixé pour toute la suite de cette partie.

On rappelle que  $\mathcal{C}^k$  désigne l'ensemble des fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  qui sont  $k$  fois dérivable et de dérivée  $k$ -ème continue sur  $\mathbb{R}$ . Pour  $f$  une fonction bornée sur  $\mathbb{R}$ , on notera  $M_f$  un majorant de  $|f|$ , c'est-à-dire un réel tel que  $|f(t)| \leq M_f$  pour tout  $t \in \mathbb{R}$ .

12. a. (i) Montrer, grâce à des intégrations par parties successives, que  $\int_0^1 u^3(1-u)^3 du = \frac{1}{20} \int_0^1 u^6 du$ .

(ii) En déduire que  $\int_0^1 u^3(1-u)^3 du = \frac{1}{140}$ .

On définit la fonction  $h: \mathbb{R} \rightarrow \mathbb{R}$  de la manière suivante : pour  $z \in \mathbb{R}$ , on pose

$$h(z) = \begin{cases} 0 & \text{si } z \leq 0 \\ 140 \int_0^z u^3(1-u)^3 du & \text{si } z \in ]0, 1[ \\ 1 & \text{si } z \geq 1 \end{cases}$$

b. Montrer que  $h$  est continue sur  $\mathbb{R}$ .

c. Montrer que pour tout  $z \in \mathbb{R}$ , on a  $0 \leq h(z) \leq 1$ .

On admettra que  $h'$ ,  $h''$  et  $h'''$  sont aussi continues et bornées sur  $\mathbb{R}$ .

13. On pose  $a_n = n^{-1/12}$  et  $g_n(z) = 1 - h\left(\frac{1}{a_n}(z-x)\right)$ .

a. (i) Montrer que  $g_n$  est continue sur  $\mathbb{R}$  et que pour tout  $z \in \mathbb{R}$  on a  $0 \leq g_n(z) \leq 1$ .

On admettra que  $g'_n$ ,  $g''_n$  et  $g'''_n$  sont aussi continues et bornées sur  $\mathbb{R}$ .

(ii) Montrer que l'on peut choisir un majorant  $M_{g'''_n}$  de  $g'''_n$  tel que  $M_{g'''_n} \leq n^{1/4} M_{h''''}$ .

b. (i) Montrer que  $g_n(z) = 1$  si  $z \leq x$  et  $g_n(z) = 0$  si  $z > x + a_n$ .

Pour un événement  $A$ , on définit la variable aléatoire  $\mathbb{1}_A$  par  $\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$ .

(ii) Montrer que pour tout variable aléatoire  $X$  on a  $\mathbb{1}_{[X \leq x]} \leq g_n(X) \leq \mathbb{1}_{[X \leq x + a_n]}$ .

c. Montrer que l'on a  $E(g_n(Z_n + a_n)) \leq \mathbb{P}(Z_n \leq x) \leq E(g_n(Z_n))$ .

Il suffit alors de montrer que l'on a

$$\lim_{n \rightarrow +\infty} E(g_n(Z_n + a_n)) = \lim_{n \rightarrow +\infty} E(g_n(Z_n)) = \Phi(x).$$

On va se concentrer sur la dernière égalité.

14. Soit  $g$  une fonction de classe  $\mathcal{C}^3$  sur  $\mathbb{R}$ .

a. Par des intégrations par parties successives, montrer que pour tout  $z, u \in \mathbb{R}$  fixés, on a

$$\frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) t = -\frac{1}{2} u^2 g''(z) - u g'(z) + g(z+u) - g(z).$$

b. En déduire que pour tous  $u, v, z$  réels, on a

$$g(z+u) - g(z+v) = g'(z)(u-v) + \frac{1}{2} g''(z)(u^2 - v^2) + R(z, u, v)$$

$$\text{où } R(z, u, v) = \frac{1}{2} \int_z^{z+u} (z+u-t) g'''(t) t - \frac{1}{2} \int_z^{z+v} (z+v-t) g'''(t) t.$$

c. Montrer que si  $g'''$  est bornée alors pour tous réels  $z, u, v$  on a  $|R(z, u, v)| \leq \frac{1}{6} M_{g'''} (|u|^3 + |v|^3)$ .

15. Soient  $(Y_i)_{i \geq 1}$  une suite de variables aléatoires de loi normale  $\mathcal{N}(0, 1)$ . On suppose les variables  $Y_i$  indépendantes entre elles et indépendantes des variables  $X_j$ .

a. Soit  $n$  un entier naturel supérieur ou égal à 2.

(i) Justifier que la loi de  $\sum_{i=1}^n Y_i$  est une loi normale.

(ii) On pose  $T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ . Déterminer la loi de  $T_n$ .

b. Pour  $k \in \{2, 3, \dots, n-1\}$ , on pose  $W_k = \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{k-1} + X_{k+1} + \dots + X_n)$ ,

avec  $W_1 = \frac{1}{\sqrt{n}} \sum_{i=2}^n X_i$  et  $W_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} Y_i$ .

(i) Montrer que pour tout  $k \in \{1, 2, \dots, n-1\}$ ,  $W_k + \frac{1}{\sqrt{n}} Y_k = W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1}$ .

(ii) En déduire que  $g_n(Z_n) - g_n(T_n) = \sum_{k=1}^n \left( g_n \left( W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left( W_k + \frac{1}{\sqrt{n}} Y_k \right) \right)$ .

On va maintenant calculer l'espérance de chacun des termes de la somme.

c. (i) Montrer que  $E((X_k - Y_k)g'_n(W_k)) = 0$ .

(ii) Montrer que  $E((X_k^2 - Y_k^2)g''_n(W_k)) = 0$ .

(iii) En déduire que  $E \left( g_n \left( W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left( W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) = E \left( R \left( W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right)$ .

d. Montrer que l'on a  $|E(g_n(Z_n)) - E(g_n(T_n))| \leq \sum_{k=1}^n E \left( \left| R \left( W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right| \right)$ .

e. Conclure que l'on a

$$|E(g_n(Z_n)) - E(g_n(T_n))| \leq \frac{1}{6\sqrt{n}} M_{g'''} \left( E(|X_1|^3) + E(|Y_1|^3) \right).$$

f. En déduire que l'on a  $\lim_{n \rightarrow +\infty} |E(g_n(Z_n)) - E(g_n(T_n))| = 0$ .

16. a. Montrer que l'on a  $\mathbb{P}(T_n \leq x) \leq E(g_n(T_n)) \leq \mathbb{P}(T_n \leq x + a_n)$ .

b. En déduire que  $\lim_{n \rightarrow +\infty} E(g_n(T_n)) = \Phi(x)$ .

c. Conclure que l'on a  $\lim_{n \rightarrow +\infty} E(g_n(T_n)) = \Phi(x)$ .