

ESSEC II 2022

Le théorème limite central est l'un des théorèmes fondamentaux des probabilités : il occupe une place centrale aussi bien d'un point de vue théorique que des applications (notamment en statistiques). 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 $\mathbb{E}(X)$ son espérance et $\mathbb{V}(X)$ 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 $\mu = \mathbb{E}(X_i)$ leur espérance commune et $\sigma^2 = \mathbb{V}(X_i)$ leur variance commune avec $\sigma > 0$. Enfin, pour tout entier naturel $n \geq 1$, on définit les variables aléatoires :

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

Commentaire

L'énoncé original utilisait la notation $\text{Var}(X)$ plutôt que la notation $\mathbb{V}(X)$. Le programme opte quant à lui pour l'écriture $V(X)$. Les concepteurs ne souhaitent pas utiliser cette dernière notation pour éviter d'écrire, dans le cas où l'on considère une v.a.r. $V : V(V)$. On privilégiera cependant la notation $\mathbb{V}(X)$ dans ce corrigé pour mieux se conformer aux notations du programme de la filière ECE et rester lisible en cas de v.a.r. nommée V .

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

$$f_Z \text{ définie par : } f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

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

Commentaire

- Il convient de faire la distinction entre :
 - × f_Z , qui est une fonction de \mathbb{R} dans \mathbb{R} .
 - × $f_Z(x)$, image de x par la fonction f_Z (résultat de l'évaluation de la fonction f_Z au point x) qui est une quantité (un réel).

Si l'on souhaite présenter une fonction par le résultat de son évaluation en chacun de ses points, il serait préférable d'utiliser un quantificateur. Dans ce cas, on dira que f_Z est la fonction définie

$$\text{par : } \forall x \in \mathbb{R}, f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

- Rappelons qu'une fonction est un mécanisme d'association. Ainsi, f_Z est définie par :

$$\begin{aligned} f_Z &: \mathbb{R} \rightarrow \mathbb{R} \\ x &\mapsto f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

- Évidemment, cette remarque s'applique aussi à la définition de la fonction Φ .

a) Montrer que Φ est continue sur \mathbb{R} .

Démonstration.

La fonction Φ est la fonction de répartition de la v.a.r. Z qui est une v.a.r. à densité. Elle est donc continue sur \mathbb{R} .

La fonction Φ est continue sur \mathbb{R} .

□

b) Montrer que la fonction Φ est strictement croissante.

Démonstration.

La fonction $f_Z : x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ est continue sur \mathbb{R} .

La fonction Φ est une primitive de f_Z . Elle est donc de classe \mathcal{C}^1 sur \mathbb{R} et :

$$\forall x \in \mathbb{R}, \quad \Phi'(x) = f_Z(x) > 0$$

La fonction Φ est donc strictement croissante sur \mathbb{R} .

□

c) Montrer que la fonction Φ est une bijection de \mathbb{R} sur $]0, 1[$.

Démonstration.

La fonction φ est :

× continue sur \mathbb{R} d'après **1.a)**,

× strictement croissante sur \mathbb{R} d'après **1.b)**.

Ainsi, Φ réalise une bijection de $\mathbb{R} =]-\infty, +\infty[$ sur $\Phi(]-\infty, +\infty[)$.

$$\Phi(]-\infty, +\infty[) = \left] \lim_{x \rightarrow -\infty} \Phi(x), \lim_{x \rightarrow +\infty} \Phi(x) \right[=]0, 1[\quad (\text{car } \Phi \text{ est une fonction de répartition})$$

La fonction Φ réalise donc une bijection de \mathbb{R} sur $]0, 1[$.

□

d) Montrer que pour tout $x \in \mathbb{R}$, on a : $\Phi(-x) = 1 - \Phi(x)$.

Démonstration.

Soit $x \in \mathbb{R}$.

$$\Phi(-x) = \int_{-\infty}^{-x} f_Z(t) dt$$

Comme f_Z est une densité, l'intégrale $\int_{-\infty}^{+\infty} f_Z(t) dt$ est convergente (et vaut 1). Les intégrales

$\int_{-\infty}^{-x} f_Z(t) dt$ et $\int_{-\infty}^x f_Z(t) dt$ sont donc convergentes.

On effectue alors le changement de variable $u = -t$.

$$\left| \begin{array}{l} u = -t \\ \hookrightarrow du = -dt \quad \text{et} \quad dt = -du \\ \bullet t = -\infty \Rightarrow u = +\infty \\ \bullet t = -x \Rightarrow u = x \end{array} \right.$$

Ce changement de variable est valide car $\psi : u \mapsto -u$ est de classe \mathcal{C}^1 sur $[x, +\infty[$. On obtient :

$$\begin{aligned} \int_{-\infty}^{-x} f_Z(t) dt &= \int_x^{+\infty} f_Z(-u) du \\ &= \int_x^{+\infty} f_Z(u) du && \text{(car } f_Z \text{ est paire)} \\ &= \int_{-\infty}^{+\infty} f_Z(u) du - \int_{-\infty}^x f_Z(u) du \\ &= 1 - \Phi(x) && \text{(car } f_Z \text{ est une densité de probabilité)} \end{aligned}$$

$$\boxed{\forall x \in \mathbb{R}, \Phi(-x) = 1 - \Phi(x)}$$

□

2. a) Énoncer la loi faible des grands nombres pour la suite de variables aléatoires $(X_i)_{i \in \mathbb{N}^*}$.

Démonstration.

Les v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$:

- × sont indépendantes,
- × admettent toutes la même espérance μ ,
- × admettent toutes la même variance σ^2 .

$$\boxed{\text{Par la loi faible des grands nombres : } \forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}([\bar{X}_n - \mu] \geq \varepsilon) = 0.}$$

□

b) Rappeler les hypothèses du théorème limite central et en déduire que pour tout réel $x \in \mathbb{R}$, on a :

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

Démonstration.

• Les v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$ sont :

- × indépendantes,
- × de même loi,
- × de même espérance μ ,
- × et de même variance σ^2 non nulle.

Alors, par théorème central limite : $\bar{X}_n^* \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z$, où :

$$\bar{X}_n^* = \frac{\bar{X}_n - \mu}{\frac{\sigma}{n}} = \sqrt{n} \left(\frac{\bar{X}_n - \mu}{\sigma} \right) = Z_n$$

$$\boxed{\text{On en conclut : } Z_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z.}$$

• Soit $x \in \mathbb{R}$. On en déduit :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x]) = \mathbb{P}([Z \leq x]) = \Phi(x)$$

$$\boxed{\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x]) = \Phi(x)}$$

□

3. Donnons une première application du théorème central limite, pour un joueur de fléchettes. Au $i^{\text{ème}}$ 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 $\mathbb{E}(X_i)$.

Démonstration.

Soit $i \in \mathbb{N}^*$.

- La v.a.r. X_i admet une espérance car c'est une v.a.r. finie.
- De plus :

$$\begin{aligned} \mathbb{E}(X_i) &= 0 \times \mathbb{P}([X_i = 0]) + 2 \times \mathbb{P}([X_i = 2]) + 5 \times \mathbb{P}([X_i = 5]) + 10 \times \mathbb{P}([X_i = 10]) \\ &= 2 \times \frac{1}{2} + 5 \times \frac{1}{5} + 10 \times \frac{1}{10} \\ &= 3 \end{aligned}$$

Ainsi : $\forall i \in \mathbb{N}^*, \mathbb{E}(X_i) = 3.$

□

b) Calculer $\mathbb{V}(X_i)$.

Démonstration.

Soit $i \in \mathbb{N}^*$.

- La v.a.r. X_i admet une variance car c'est une v.a.r. finie.
- De plus :

$$\begin{aligned} \mathbb{E}(X_i^2) &= 0^2 \times \mathbb{P}([X_i = 0]) + 2^2 \times \mathbb{P}([X_i = 2]) + 5^2 \times \mathbb{P}([X_i = 5]) + 10^2 \times \mathbb{P}([X_i = 10]) \\ &= 2^2 \times \frac{1}{2} + 5^2 \times \frac{1}{5} + 10^2 \times \frac{1}{10} \\ &= 2 + 5 + 10 = 17 \end{aligned}$$

- Par formule de Koenig-Huygens :

$$\mathbb{V}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 17 - 3^2 = 17 - 9 = 8$$

Finalement : $\forall i \in \mathbb{N}^*, \mathbb{V}(X_i) = 8.$

□

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

$$f(x) = 0 \text{ si } x \in [0, \frac{1}{5}[, \quad f(x) = 2 \text{ si } x \in [\frac{1}{5}, \frac{7}{10}[, \quad f(x) = 5 \text{ si } x \in [\frac{7}{10}, \frac{9}{10}[, \quad f(x) = 10 \text{ si } x \in [\frac{9}{10}, 1]$$

(i) Soit U une variable aléatoire de loi uniforme sur $[0, 1]$. Montrer que $f(U)$ a même loi que X_i .

Démonstration.

- Tout d'abord :

$$\begin{aligned} (f(U))(\Omega) &= f(U(\Omega)) \\ &= f([0, 1]) \quad (\text{car } U \hookrightarrow \mathcal{U}([0, 1])) \\ &= \{0, 2, 5, 10\} \quad (\text{par définition de } f) \end{aligned}$$

Ainsi : $(f(U))(\Omega) = \{0, 2, 5, 10\}.$

• De plus :

× tout d'abord :

$$\begin{aligned}
 \mathbb{P}([f(U) = 0]) &= \mathbb{P}\left(\left[0 \leq U < \frac{1}{5}\right]\right) && \text{(par définition de } f) \\
 &= F_U\left(\frac{1}{5}\right) - F_U(0) && \text{(car } U \text{ est une} \\
 & && \text{v.a.r. à densité)} \\
 &= \frac{1}{5} - 0 && \text{(car } U \hookrightarrow \mathcal{U}([0, 1]) \text{ et} \\
 & && \frac{1}{5} \in [0, 1]) \\
 &= \frac{1}{5}
 \end{aligned}$$

× ensuite :

$$\begin{aligned}
 \mathbb{P}([f(U) = 2]) &= \mathbb{P}\left(\left[\frac{1}{5} \leq U < \frac{7}{10}\right]\right) && \text{(par définition de } f) \\
 &= F_U\left(\frac{7}{10}\right) - F_U\left(\frac{1}{5}\right) && \text{(car } U \text{ est une} \\
 & && \text{v.a.r. à densité)} \\
 &= \frac{7}{10} - \frac{1}{5} && \text{(car } U \hookrightarrow \mathcal{U}([0, 1]) \text{ et} \\
 & && \left(\frac{1}{5}, \frac{7}{10}\right) \in [0, 1]^2) \\
 &= \frac{5}{10} = \frac{1}{2}
 \end{aligned}$$

× de même :

$$\begin{aligned}
 \mathbb{P}([f(U) = 5]) &= \mathbb{P}\left(\left[\frac{7}{10} \leq U < \frac{9}{10}\right]\right) \\
 &= F_U\left(\frac{9}{10}\right) - F_U\left(\frac{7}{10}\right) \\
 &= \frac{9}{10} - \frac{7}{10} = \frac{1}{5}
 \end{aligned}$$

× enfin, comme $([f(U) = 0], [f(U) = 2], [f(U) = 5], [f(U) = 10])$ forme un système complet d'événements :

$$\begin{aligned}
 \mathbb{P}([f(U) = 10]) &= 1 - \left(\mathbb{P}([f(U) = 0]) + \mathbb{P}([f(U) = 2]) + \mathbb{P}([f(U) = 5])\right) \\
 &= 1 - \frac{1}{5} - \frac{1}{2} - \frac{1}{5} \\
 &= \frac{1}{10}
 \end{aligned}$$

On en déduit que la v.a.r. $f(U)$ suit la même loi que X .

□

- (ii) Compléter le programme **Scilab** suivant, qui permet de générer un nombre aléatoire de même loi que X_i . On rappelle que la fonction `rand()` simule une variable aléatoire de loi uniforme sur $[0, 1]$.

```

1  function x = X()
2      U = rand()
3      ...
4  endfunction

```

Démonstration.

On propose la fonction **Scilab** suivante :

```

1  function x = X()
2      U = rand()
3      if U < 1/5 then
4          x = 0
5      elseif U < 7/10 then
6          x = 2
7      elseif U < 9/10 then
8          x = 5
9      else
10         x = 10
11     end
12 endfunction

```

Détaillons les éléments de ce script.

• **Début de la fonction**

On commence par préciser la structure de la fonction :

- × cette fonction se nomme **X**,
- × elle ne prend pas de paramètre,
- × elle admet pour variable de sortie **x**.

```

1  function x = X()

```

• **Contenu de la fonction**

Les lignes 2 à 11 consistent à simuler la v.a.r. X , et de stocker le résultat dans la variable **x**. Comme la v.a.r. $f(U)$ suit la même loi que X , on simulera plutôt $f(U)$.

- × On commence donc par simuler la v.a.r. U de loi $\mathcal{U}([0, 1])$ à l'aide de `rand()` et on stocke le résultat obtenu dans la variable **U**.

```

2      U = rand()

```

- × On calcule ensuite $f(U)$. La fonction f étant définie par cas comme suit :

$$f : [0, 1] \rightarrow \mathbb{R}$$

$$u \mapsto \begin{cases} 0 & \text{si } u < \frac{1}{5} \\ 2 & \text{si } \frac{1}{5} \leq u < \frac{7}{10} \\ 5 & \text{si } \frac{7}{10} \leq u < \frac{9}{10} \\ 10 & \text{sinon} \end{cases}$$

Le calcul de $f(U)$ s'effectue donc à l'aide d'une structure conditionnelle.

```

3      if U < 1/5 then
4          x = 0
5      elseif U < 7/10 then
6          x = 2
7      elseif U < 9/10 then
8          x = 5
9      else
10         x = 10
11     end

```

Commentaire

- Afin de permettre une bonne compréhension des mécanismes en jeu, on a détaillé la réponse à cette question. Cependant, compléter correctement le programme **Scilab** démontre la bonne compréhension de la simulation demandée et permet certainement d'obtenir tous les points alloués à cette question.

On procédera de même dans les autres questions **Scilab**.

- Lors de l'écriture d'un programme informatique, on se soumet généralement à quelques règles de bonne conduite :

- (1) utilisation de commentaires indiquant le but de chaque fonction,
- (2) réflexion autour du découpage en sous-fonctions pouvant être réutilisées,
- (3) utilisation de noms explicites pour les fonctions et les variables,
- (4) indentation du code (utilisation correcte d'espaces et sauts de lignes).

Le but de ces règles est de produire un code lisible, intelligible et facilement modifiable à l'avenir. Ces règles sont globalement très bien respectées ici. L'introduction des variables informatiques x et U témoigne de cette volonté de produire du code intelligible.

- En revanche, le nom choisi pour la fonction, à savoir X semble peu pertinent. Il induit une confusion avec la variable informatique x qui contient la simulation de la v.a.r. X . Le nom `simulX` est bien plus adapté. C'est le nom choisi traditionnellement dans les énoncés d'ERICOME ou EML et il serait préférable qu'il soit également choisi dans les énoncés ESSEC.
- On pouvait également répondre à cette question en codant la fonction f à part. Ce n'était sans doute pas ce qui était attendu par l'énoncé au regard de la forme du script proposé, mais cela permettait tout aussi bien d'obtenir la totalité des points alloués à cette question. On aurait alors proposé les scripts suivants :

```

1  function y = f(x)
2      if x < 1/5 then
3          y = 0
4      elseif x < 7/10 then
5          y = 2
6      elseif x < 9/10 then
7          y = 5
8      else
9          y = 10
10     end
11 endfunction

```

```

1  function X = simulX()
2      U = rand()
3      X = f(U)
4  endfunction

```

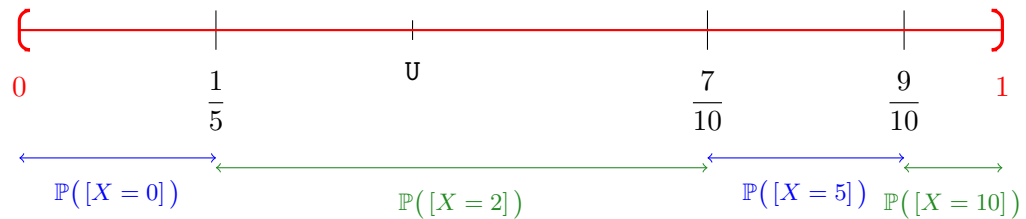
□

Commentaire

- L'objectif de cette question est de simuler une v.a.r. discrète (ici X). Ce type de simulation s'effectue toujours de la même façon :
 - 1) on commence par simuler une v.a.r. U de loi $\mathcal{U}([0, 1])$. Pour cela, on utilise l'instruction `rand()` et on stocke le résultat dans une variable U . Moins formellement, l'instruction `rand()` renvoie un réel choisi aléatoirement dans $[0, 1]$.
 - 2) on partitionne ensuite l'intervalle $[0, 1]$ en sous-intervalles de longueur $\mathbb{P}([X = x])$ où $x \in X(\Omega)$. (Notons que cela est toujours possible car $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements, donc : $\sum_{x \in X(\Omega)} \mathbb{P}([X = x]) = 1$)
 - 3) on affecte à la variable x la valeur x lorsque U appartient au sous-intervalle de longueur $\mathbb{P}([X = x])$.
- Détaillons ce procédé dans le cas de la v.a.r. X de l'énoncé.
 - 1) On simule une v.a.r. U de loi $\mathcal{U}([0, 1])$ et on stocke le résultat dans une variable U .

```
2      U = rand()
```

- 2) On partitionne l'intervalle $[0, 1]$ en sous-intervalles de longueur $\mathbb{P}([X = 0])$, $\mathbb{P}([X = 2])$, $\mathbb{P}([X = 5])$ et $\mathbb{P}([X = 10])$.



- 3) Quatre cas se présentent alors :

- si $U < \frac{1}{5}$, alors on affecte à x la valeur 0. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left(\left[0 \leq U < \frac{1}{5}\right]\right) = \mathbb{P}\left(\left[U < \frac{1}{5}\right]\right) = \frac{1}{5} = \mathbb{P}([X = 0])$$

```
3      if U < 1/5 then
4      x = 0
```

- si $\frac{1}{5} \leq U < \frac{7}{10}$, alors on affecte à x la valeur 2. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left(\left[\frac{1}{5} \leq U < \frac{7}{10}\right]\right) = \frac{7}{10} - \frac{1}{5} = \frac{1}{2} = \mathbb{P}([X = 2])$$

```
5      elseif U < 7/10 then
6      x = 2
```


Commentaire

- si $\frac{7}{10} \leq U < \frac{9}{10}$, alors on affecte à x la valeur 5. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left(\left[\frac{7}{10} \leq U < \frac{9}{10}\right]\right) = \frac{9}{10} - \frac{7}{10} = \frac{1}{5} = \mathbb{P}([X = 5])$$

```

7      elseif U < 9/10 then
8          x = 5

```

- si $\frac{9}{10} \leq U \leq 1$, alors on affecte à x la valeur 10. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left(\left[\frac{9}{10} \leq U \leq 1\right]\right) = 1 - \frac{9}{10} = \frac{1}{10} = \mathbb{P}([X = 10])$$

```

5      else
6          x = 10

```

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 .

Démonstration.

• On calcule :

$$\begin{aligned}
 Z_n &= \sqrt{n} \left(\frac{\bar{X}_n - \mu}{\sigma} \right) \\
 &= \sqrt{n} \left(\frac{\frac{1}{n} \sum_{i=1}^n X_i - \mu}{\sigma} \right) \\
 &= \sqrt{n} \left(\frac{\frac{1}{n} S_n - \mu}{\sigma} \right) \\
 &= \sqrt{n} \left(\frac{n \left(\frac{1}{n} S_n - \mu \right)}{n \sigma} \right) \\
 &= \sqrt{n} \left(\frac{S_n - n \mu}{n \sigma} \right) \\
 &= \frac{S_n - n \mu}{\sqrt{n} \sigma}
 \end{aligned}$$

• De plus :

× par définition de μ :

$$\mu = \mathbb{E}(X_i) = 3 \quad (\text{d'après 3.a})$$

× par définition de σ^2 :

$$\sigma^2 = \mathbb{V}(X_i) = 8 \quad (\text{d'après 3.b})$$

On obtient alors :

$$Z_n = \frac{S_n - 3n}{\sqrt{n}\sqrt{8}}$$

Finalement : $Z_n = \frac{S_n - 3n}{2\sqrt{2}n}$.

□

- 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 $6 \cdot 10^{-3}$.

Démonstration.

- Le score du joueur est la somme des points qu'il a marqué avec ces $n = 200$ fléchettes. Il correspond donc à la v.a.r. S_n . On cherche ainsi à évaluer :

$$\mathbb{P}([S_n \leq 500])$$

- Or, d'après la question 2.b), par le théorème central limite, pour tout $x \in \mathbb{R}$:

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

- De plus :

$$\begin{aligned} \mathbb{P}([Z_n \leq x]) &= \mathbb{P}\left(\left[\frac{S_n - 3n}{2\sqrt{2}n} \leq x\right]\right) && \text{(d'après la question précédente)} \\ &= \mathbb{P}([S_n - 3n \leq 2x\sqrt{2}n]) && \text{(car } 2\sqrt{2}n > 0) \\ &= \mathbb{P}([S_n \leq 3n + 2x\sqrt{2}n]) \\ &= \mathbb{P}([S_n \leq 3 \times 200 + 2x\sqrt{2 \times 200}]) && \text{(car } n = 200 \text{ dans cette question)} \\ &= \mathbb{P}([S_n \leq 600 + 2x \times 20]) \\ &= \mathbb{P}([S_n \leq 600 + 40x]) \end{aligned}$$

- Comme on cherche $\mathbb{P}([S_n \leq 500])$, on choisit x tel que : $600 + 40x = 500$.

$$600 + 40x = 500 \Leftrightarrow 40x = -100 \Leftrightarrow x = -\frac{100}{40} = -\frac{5}{2}$$

On obtient :

$$\begin{aligned} \mathbb{P}([S_n \leq 500]) &= \mathbb{P}\left(\left[Z_n \leq -\frac{5}{2}\right]\right) \\ &\approx \\ &\Phi\left(-\frac{5}{2}\right) \end{aligned}$$

Ainsi, pour $n = 200$, on a : $\mathbb{P}([S_n \leq 500]) \approx \Phi(-2, 5)$.

On en déduit que, pour $n = 200$, la probabilité que le score du joueur soit inférieur ou égal à 500 est approximativement $\Phi(-2, 5)$.

□

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 $k \in \{1, \dots, 2N - 1\}$, 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 :

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

Démonstration.

• D'après 2.b), pour tout $x \in \mathbb{R}$:

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

Autrement dit, par définition de la limite :

$$\forall \varepsilon > 0, \exists n_x \in \mathbb{N}, \forall n \geq n_x, \quad |\mathbb{P}([Z_n \leq x]) - \Phi(x)| \leq \varepsilon \quad (A)$$

• Soit $k \in \llbracket 0, 2N \rrbracket$. Trois cas se présentent alors :

× si $k = 0$, alors $x_0 = -\infty$. D'où, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} |\mathbb{P}([Z_n \leq x_0]) - \Phi(x_0)| &= |F_{Z_n}(x_0) - \Phi(x_0)| \\ &= |0 - 0| && \text{(par convention et car } F_{Z_n} \text{ est une} \\ & && \text{fonction de répartition, donc :} \\ & && \lim_{x \rightarrow -\infty} F_{Z_n}(x) = 0) \\ &= 0 \end{aligned}$$

Comme $\frac{1}{2N} > 0$, on obtient :

$$\forall n \in \mathbb{N}, \quad |\mathbb{P}([Z_n \leq x_0]) - \Phi(x_0)| \leq \frac{1}{2N}$$

Ainsi, en posant $n'_0 = 0$, il existe bien $n'_0 \in \mathbb{N}$ tel que :

$$\forall n \geq n'_0, \quad |\mathbb{P}([Z_n \leq x_{2N}]) - \Phi(x_{2N})| \leq \frac{1}{2N}$$

× si $k \in \llbracket 1, 2N - 1 \rrbracket$, alors on peut appliquer (A) en $x = x_k$. On obtient :

$$\forall \varepsilon > 0, \exists n'_k \in \mathbb{N}, \forall n \geq n'_k, \quad |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \varepsilon$$

En choisissant $\varepsilon = \frac{1}{2N} > 0$, il existe donc $n'_k \in \mathbb{N}$ tel que :

$$\forall n \geq n'_k, \quad |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \frac{1}{2N}$$

× si $k = 2N$, alors $x_{2N} = +\infty$. D'où, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} |\mathbb{P}([Z_n \leq x_{2N}]) - \Phi(x_{2N})| &= |F_{Z_n}(x_{2N}) - \Phi(x_{2N})| \\ &= |1 - 1| && \text{(par convention et car } F_{Z_n} \text{ est une} \\ & && \text{fonction de répartition, donc :} \\ & && \lim_{x \rightarrow +\infty} F_{Z_n}(x) = 1) \\ &= 0 \end{aligned}$$

Comme $\frac{1}{2N} > 0$, on obtient :

$$\forall n \in \mathbb{N}, \quad |\mathbb{P}([Z_n \leq x_{2N}]) - \Phi(x_{2N})| \leq \frac{1}{2N}$$

Ainsi, en posant $n'_{2N} = 0$, il existe bien $n'_{2N} \in \mathbb{N}$ tel que :

$$\forall n \geq n'_{2N}, \quad |\mathbb{P}([Z_n \leq x_{2N}]) - \Phi(x_{2N})| \leq \frac{1}{2N}$$

Enfinement, pour tout $k \in \llbracket 0, 2N \rrbracket$, il existe $n'_k \in \mathbb{N}$ tel que :

$$\forall n \geq n'_k, \quad |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \frac{1}{2N}.$$

- On pose alors : $n_0 = \max_{j \in \llbracket 0, 2N \rrbracket} (n'_j)$. Vérifions :

$$\forall n \geq n_0, \forall k \in \llbracket 0, 2N \rrbracket, \quad |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \frac{1}{2N}$$

Soit $n \geq n_0$. Soit $k \in \llbracket 0, 2N \rrbracket$.

$$n_0 = \max_{j \in \llbracket 0, 2N \rrbracket} (n'_j) \geq n'_k$$

Ainsi, d'après le point précédent :

$$|\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \frac{1}{2N}$$

Cette propriété est vérifiée pour tout élément k de $\llbracket 0, 2N \rrbracket$. On obtient ainsi $2N+1$ ($= \text{Card}(\llbracket 0, 2N \rrbracket)$) inégalités. On obtient en particulier celle dont le membre de gauche est le plus grand, à savoir :

$$\max_{k \in \llbracket 0, 2N \rrbracket} (|\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)|) \leq \frac{1}{2N}$$

Enfinement, on a démontré qu'il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$:

$$\max_{k \in \llbracket 0, 2N \rrbracket} (|\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)|) \leq \frac{1}{2N}$$

Commentaire

- La difficulté d'un sujet se mesure en grande partie à la manière dont chaque question est découpée en sous-questions. Moins il y a de sous-questions, plus le candidat doit prendre des initiatives et plus le sujet est difficile. Ainsi, un même thème peut amener à un traitement différent lorsqu'il est abordé dans un sujet du TOP3 ou du TOP5.
- Dans les sujets, on distingue grossièrement trois types de questions :
 - (1) des questions abordables qui sont traitées par un grand nombre de candidats. Il peut s'agir de questions de cours ou de questions classiques (celles qui reviennent chaque année aux concours).
 - (2) des questions plus difficiles qui permettent de bien classer les candidats. Celles-ci ont un rôle fort pour le classement des candidats car sont abordées avec plus ou moins de succès.
 - (3) des questions très difficiles qui ne sont bien traitées par presque aucun candidat.

Dans les sujets du TOP5, on trouve essentiellement des questions de type 1) et 2). Dans les sujets du TOP3, on trouve les trois types de questions, avec un pourcentage élevé de questions de type 2). C'est d'ailleurs essentiellement sur ces questions que se font les différences et en aucun cas les questions de type 3).

Commentaire

- Ces dernières années, la taille des sujets a eu tendance à grossir pour s'établir à près de 60 questions par énoncé. Pour finir un sujet, un candidat ne dispose donc que d'environ 4 minutes pour traiter chaque question. Il ne faut pas s'inquiéter pour autant car les barèmes permettent de rebattre les cartes. En effet, un candidat qui obtient 50% des points d'un sujet aura une très bonne note (de 16 à 20 en fonction des années et des épreuves).

Cette considération sur la taille des sujets et la distinction précédente permettent d'éclairer sur la stratégie à adopter lors des concours :

- il est essentiel de savoir repérer la difficulté d'une question. Ce n'est pas chose aisée car cela requiert d'avoir du recul.
- il est essentiel de savoir traiter la majorité des questions de type (1).
- il est important de réussir à traiter correctement des questions de type (2). C'est sur le bon traitement de ces questions que se joue le classement.
- pour les questions de type (3) (ou les questions de type (2) les plus difficiles), il faut garder en tête que plus une question est difficile, plus le correcteur est indulgent avec les candidats qui s'y aventurent. On peut donc aborder ces questions avec l'idée que tout ce qu'un candidat écrit de juste sera retenu en sa faveur (quelques points accordés). Mais tenter de traiter ces questions en entier est une perte de temps préjudiciable : le nombre de points alloués ne sera certainement pas à hauteur du temps investi.

En résumé, il faut aborder en priorité les questions de type (1) et (2) et les traiter en entier. Il ne faut surtout pas abandonner une question que l'on sait traiter. L'important est de maximiser le nombre de questions que l'on traite entièrement. En revanche, il ne faut pas hésiter à passer ou ne pas traiter en entier (on signalera alors au correcteur qu'on ne sait pas comment conclure) les questions les plus difficiles.

- La question 4.a) du sujet est clairement de type (3) car elle nécessite une prise d'initiatives beaucoup trop importante. Pour que cette question puisse être abordée avec profit par certains candidats, il aurait été préférable de la découper en sous-questions. En particulier, on aurait pu demander en amont la démonstration de la proposition suivante :

$$\exists n_k \in \mathbb{N}, \forall n \geq n_k, \quad |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \frac{1}{2N} \quad \square$$

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

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

(i) Montrer que l'on a : $\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \mathbb{P}([Z_n \leq x_k]) - \Phi(x_{k-1})$.

Démonstration.

On sait : $x \in I_k$. Autrement dit :

$$x_{k-1} < x \leq x_k$$

- D'une part, la fonction F_{Z_n} (fonction de répartition de Z_n) est croissante sur \mathbb{R} car c'est une fonction de répartition. Ainsi :

$$F_{Z_n}(x_{k-1}) \leq F_{Z_n}(x) \leq F_{Z_n}(x_k)$$

Commentaire

Notons que cette relation est bien valable pour $k = 1$ et $k = 2N$ avec la convention $F_{Z_n}(x_0) = 0$ et $F_{Z_n}(x_{2N}) = 1$. En effet, comme F_{Z_n} est une fonction de répartition, pour tout $x \in \mathbb{R} : 0 \leq F_{Z_n}(x) \leq 1$. Ainsi : $F_{Z_n}(x_0) \leq F_{Z_n}(x)$ et $F_{Z_n}(x) \leq F_{Z_n}(x_{2N})$.

D'où :

$$\mathbb{P}([Z_n \leq x_{k-1}]) \leq \mathbb{P}([Z_n \leq x]) \leq \mathbb{P}([Z_n \leq x_k])$$

- D'autre part, la fonction Φ est (strictement) croissante sur \mathbb{R} d'après **1.b**. D'où :

$$\Phi(x_{k-1}) \leq \Phi(x) \leq \Phi(x_k)$$

(notons une nouvelle fois que cet encadrement est valide pour $k = 1$ et $k = 2N$ avec la convention $\Phi(x_0) = 0$ et $\Phi(x_{2N}) = 1$)

D'où :

$$-\Phi(x_{k-1}) \geq -\Phi(x) \geq -\Phi(x_k)$$

On en déduit : $\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \mathbb{P}([Z_n \leq x_k]) - \Phi(x_{k-1})$.

□

(ii) En déduire : $\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{1}{N}$.

Démonstration.

- D'après la question précédente :

$$\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \mathbb{P}([Z_n \leq x_k]) - \Phi(x_{k-1})$$

- On cherche alors à majorer $\mathbb{P}([Z_n \leq x_k])$.

× Comme $n \geq n_0$, d'après **4.a** :

$$\max_{j \in [0, 2N]} \left(|\mathbb{P}([Z_n \leq x_j]) - \Phi(x_j)| \right) \leq \frac{1}{2N}$$

Or :

$$|\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \max_{j \in [0, 2N]} \left(|\mathbb{P}([Z_n \leq x_j]) - \Phi(x_j)| \right)$$

Donc, par transitivité :

$$|\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)| \leq \max_{j \in [0, 2N]} \left(|\mathbb{P}([Z_n \leq x_j]) - \Phi(x_j)| \right) \leq \frac{1}{2N}$$

× De plus, par définition de $|\cdot|$:

$$\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k) \leq |\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k)|$$

D'où :

$$\mathbb{P}([Z_n \leq x_k]) - \Phi(x_k) \leq \frac{1}{2N}$$

Ainsi :

$$\mathbb{P}([Z_n \leq x_k]) \leq \Phi(x_k) + \frac{1}{2N}$$

- On en déduit :

$$\begin{aligned} \mathbb{P}([Z_n \leq x]) - \Phi(x) &\leq \mathbb{P}([Z_n \leq x_k]) - \Phi(x_{k-1}) \\ &\leq \Phi(x_k) + \frac{1}{2N} - \Phi(x_{k-1}) \quad (\text{d'après le point précédent}) \end{aligned}$$

On obtient bien : $\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N}$.

- Il reste à démontrer : $\Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{1}{N}$.

× Par définition de x_k :

$$x_k = \Phi^{-1}\left(\frac{k}{2N}\right) \quad \text{donc} \quad \Phi(x_k) = \frac{k}{2N}$$

De même : $\Phi(x_{k-1}) = \frac{k-1}{2N}$.

× On obtient :

$$\Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{k}{2N} - \frac{k-1}{2N} + \frac{1}{2N} = \frac{1}{2N} + \frac{1}{2N} = \frac{2}{2N} = \frac{1}{N}$$

$$\boxed{\Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{1}{N}}$$

□

(iii) De même, montrer que l'on a : $\Phi(x) - \mathbb{P}([Z_n \leq x]) \leq \frac{1}{N}$.

Démonstration.

- On reprend les encadrements de 4.b)(i).

× D'une part :

$$\begin{aligned} \mathbb{P}([Z_n \leq x_{k-1}]) &\leq \mathbb{P}([Z_n \leq x]) \leq \mathbb{P}([Z_n \leq x_k]) \\ \text{donc} \quad -\mathbb{P}([Z_n \leq x_{k-1}]) &\geq -\mathbb{P}([Z_n \leq x]) \geq -\mathbb{P}([Z_n \leq x_k]) \end{aligned}$$

× D'autre part :

$$\Phi(x_{k-1}) \leq \Phi(x) \leq \Phi(x_k)$$

On obtient :

$$\Phi(x) - \mathbb{P}([Z_n \leq x]) \leq \Phi(x_k) - \mathbb{P}([Z_n \leq x_{k-1}])$$

- On raisonnant comme en question précédente, on a :

$$|\mathbb{P}([Z_n \leq x_{k-1}]) - \Phi(x_{k-1})| \leq \max_{j \in [0, 2N]} (|\mathbb{P}([Z_n \leq x_j]) - \Phi(x_j)|) \leq \frac{1}{2N}$$

Par définition de $|\cdot|$:

$$\begin{aligned} -\left(\mathbb{P}([Z_n \leq x_{k-1}]) - \Phi(x_{k-1})\right) &\leq |\mathbb{P}([Z_n \leq x_{k-1}]) - \Phi(x_{k-1})| \\ &\stackrel{\text{II}}{\leq} \Phi(x_{k-1}) - \mathbb{P}([Z_n \leq x_{k-1}]) \end{aligned}$$

D'où :

$$\Phi(x_{k-1}) - \mathbb{P}([Z_n \leq x_{k-1}]) \leq \frac{1}{2N}$$

Ainsi :

$$-\mathbb{P}([Z_n \leq x_{k-1}]) \leq -\Phi(x_{k-1}) + \frac{1}{2N}$$

- On en déduit :

$$\begin{aligned} \Phi(x) - \mathbb{P}([Z_n \leq x]) &\leq \Phi(x_k) - \mathbb{P}([Z_n \leq x_{k-1}]) \\ &\leq \Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} \quad (\text{d'après le point précédent}) \end{aligned}$$

Or, d'après la question précédente :

$$\Phi(x_k) - \Phi(x_{k-1}) + \frac{1}{2N} = \frac{1}{N}$$

Finalement : $\Phi(x) - \mathbb{P}([Z_n \leq x]) \leq \frac{1}{N}$.

□

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

Démonstration.

Soit $x \in \mathbb{R}$. Soit $n \geq n_0$.

- Comme $I_1, \dots, I_{2N-1}, I_{2N}$ forment une partition de \mathbb{R} , alors il existe $k \in \llbracket 1, 2N \rrbracket$ tel que : $x \in I_k$. Alors :

× d'après **4.b(ii)** :

$$\mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \frac{1}{N}$$

× d'après **4.b(iii)** :

$$\Phi(x) - \mathbb{P}([Z_n \leq x]) \leq \frac{1}{N}$$

$$\text{donc } -(\Phi(x) - \mathbb{P}([Z_n \leq x])) \geq -\frac{1}{N}$$

$$\text{d'où } \mathbb{P}([Z_n \leq x]) - \Phi(x) \geq -\frac{1}{N}$$

- On en déduit :

$$-\frac{1}{N} \leq \mathbb{P}([Z_n \leq x]) - \Phi(x) \leq \frac{1}{N}$$

Ainsi :

$$|\mathbb{P}([Z_n \leq x]) - \Phi(x)| \leq \frac{1}{N}$$

Finalement : $\forall x \in \mathbb{R}, \forall n \geq n_0, |\mathbb{P}([Z_n \leq x]) - \Phi(x)| \leq \frac{1}{N}$.

□

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 $x \in \mathbb{R}$, pour 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 **(*)**.

Démonstration.

Soit $x \in \mathbb{R}$.

- On souhaite construire une suite (M_n) telle que :

(i) $\forall n \in \mathbb{N}^*, |D_n(x)| \leq M_n,$

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

Pour cela, on utilise le résultat de la question précédente :

$$\forall N \in \mathbb{N}^*, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, |D_n(x)| \leq \frac{1}{N}$$

L'idée est d'appliquer cette proposition à $N = 1, N = 2, \dots, N = i, \dots$

× Pour $N = 1$, il existe $n_1 \in \mathbb{N}$ tel que : $\forall n \geq n_1, |D_n(x)| \leq 1$.

Remarquons que l'on peut choisir $n_1 = 1$. En effet, pour tout $n \geq 1$:

- d'une part : $0 \leq \mathbb{P}([Z_n \leq x]) \leq 1$
- d'autre part : $0 \leq \Phi(x) \leq 1$ (car Φ est une fonction de répartition)

$$\text{donc : } 0 \geq -\Phi(x) \geq -1$$

Ainsi :

$$\begin{aligned} -1 \leq \mathbb{P}([Z_n \leq x]) - \Phi(x) &\leq 1 \\ &\parallel \\ &D_n(x) \end{aligned}$$

D'où : $|D_n(x)| \leq 1$.

Finalement, on choisit $n_1 = 1$ et on a :

$$\forall n \geq n_1, |D_n(x)| \leq 1 \quad (1)$$

× Pour $N = 2$, il existe $p_2 \in \mathbb{N}^*$ tel que : $\forall n \geq p_2, |D_n(x)| \leq \frac{1}{2}$.

Deux cas se présentent alors :

- si $p_2 > n_1$, alors on pose $n_2 = p_2$. On a donc : $\forall n \geq n_2, |D_n(x)| \leq \frac{1}{2}$.
- si $p_2 = n_1$, alors on pose $n_2 = n_1 + 1$. Soit $n \geq n_2$. Alors : $n \geq n_2 > n_1 \geq p_2$.
Ainsi : $\forall n \geq n_2, |D_n(x)| \leq \frac{1}{2}$.

Finalement, dans tous les cas, on peut choisir $n_2 > n_1$ tel que :

$$\forall n \geq n_2, |D_n(x)| \leq \frac{1}{2} \quad (2)$$

On note alors : $M_{n_1} = M_{n_1+1} = M_{n_1+2} = \dots = M_{n_2-1} = 1$.

À ce stade de la démonstration, on sait que pour tout $n \in \mathbb{N}$:

- ▶ si $n \in \llbracket n_1, n_2 - 1 \rrbracket$, $|D_n(x)| \leq M_n$ (d'après (1)).
- ▶ si $n \geq n_2$, $|D_n(x)| \leq \frac{1}{2}$.

× Pour $N = 3$, il existe $p_3 \in \mathbb{N}^*$ tel que : $\forall n \geq p_3, |D_n(x)| \leq \frac{1}{3}$.

Deux cas se présentent alors :

- si $p_3 > n_2$, alors on pose $n_3 = p_3$. On a donc : $\forall n \geq n_3, |D_n(x)| \leq \frac{1}{3}$.
- si $p_3 \leq n_2$, alors on pose $n_3 = n_2 + 1$. Soit $n \geq n_3$. Alors : $n \geq n_3 > n_2 \geq p_3$.
Ainsi : $\forall n \geq n_3, |D_n(x)| \leq \frac{1}{3}$.

Finalement, dans tous les cas, on peut choisir $n_3 > n_2$ tel que :

$$\forall n \geq n_3, |D_n(x)| \leq \frac{1}{3} \quad (3)$$

On note alors : $M_{n_2} = M_{n_2+1} = \dots = M_{n_3-1} = \frac{1}{2}$.

À ce stade de la démonstration, on sait que pour tout $n \in \mathbb{N}$:

▶ si $n \in \llbracket n_1, n_2 - 1 \rrbracket$, $|D_n(x)| \leq M_n$ (d'après (1)).

▶ si $n \in \llbracket n_2, n_3 - 1 \rrbracket$, $|D_n(x)| \leq M_n$ (d'après (2)).

▶ si $n \geq n_3$, $|D_n(x)| \leq \frac{1}{3}$.

× ...

× Pour $N = i$, il existe $p_i \in \mathbb{N}^*$ tel que : $\forall n \geq p_i$, $|D_n(x)| \leq \frac{1}{i}$.

Deux cas se présentent alors :

- si $p_i > n_{i-1}$, alors on pose $n_i = p_i$. On a donc : $\forall n \geq n_i$, $|D_n(x)| \leq \frac{1}{i}$.

- si $p_i \leq n_{i-1}$, alors on pose $n_i = n_{i-1} + 1$. Soit $n \geq n_i$. Alors : $n \geq n_i > n_{i-1} \geq p_i$.

Ainsi : $\forall n \geq n_i$, $|D_n(x)| \leq \frac{1}{i}$

Finalement, dans tous les cas, on peut choisir $n_i > n_{i-1}$ tel que :

$$\forall n \geq n_i, \quad |D_n(x)| \leq \frac{1}{i} \quad (i)$$

On note alors : $M_{n_{i-1}} = M_{n_{i-1}+1} = \dots = M_{n_i-1} = \frac{1}{i-1}$.

À ce stade de la démonstration, on sait que pour tout $n \in \mathbb{N}$:

▶ si $n \in \llbracket n_1, n_i - 1 \rrbracket$, $|D_n(x)| \leq M_n$ (d'après (1), (2), \dots , $(i-1)$).

▶ si $n \geq n_i$, $|D_n(x)| \leq \frac{1}{i}$.

× ...

- On construit ainsi une suite $(n_i)_{i \in \mathbb{N}^*}$:

× strictement croissante. En effet, pour tout $i \in \mathbb{N}^*$, on a choisi : $n_{i+1} > n_i$.

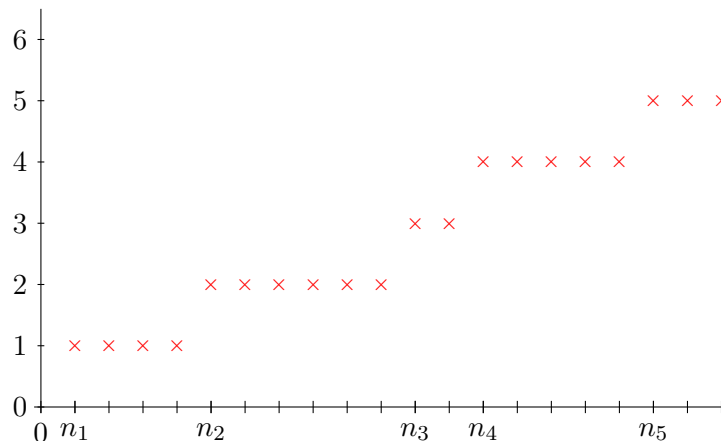
× telle que : $\forall n \in \llbracket n_i, n_{i+1} - 1 \rrbracket$, $|D_n(x)| \leq M_n$.

Formellement, on a défini la suite $(M_n)_{n \in \mathbb{N}^*}$ de la manière suivante :

$$\forall n \in \mathbb{N}^*, \quad M_n = \frac{1}{h_n}$$

où (h_n) est la suite définie par : $\forall n \in \mathbb{N}^*$, $h_n = i$ si $n \in \llbracket n_i, n_{i+1} - 1 \rrbracket$.

L'allure de la représentation graphique de (h_n) est donc :



(ici $n_1 = 1$, $n_2 = 5$, $n_3 = 11$, $n_4 = 13$ et $n_5 = 18$)

Commentaire

Notons qu'il était nécessaire de construire la suite $(n_i)_{i \in \mathbb{N}^*}$ pour qu'elle soit strictement croissante. En effet, si ce n'est pas le cas, alors certains des intervalles $\llbracket n_i, n_{i+1} - 1 \rrbracket$ ne seraient pas bien définis, et donc la suite (M_n) non plus.

- Il reste à démontrer qu'avec la suite (M_n) ainsi définie, on obtient bien les propositions (i) et (ii).

(i) On souhaite démontrer : $\forall n \in \mathbb{N}^*, |D_n(x)| \leq M_n$.

Soit $n \in \mathbb{N}^*$. Alors il existe $i \in \mathbb{N}^*$ tel que : $n \in \llbracket n_i, n_{i+1} - 1 \rrbracket$. En particulier : $n \geq n_i$.

Par définition de n_i , on obtient :

$$|D_n(x)| \leq \frac{1}{i}$$

Or, par définition de M_n et h_n , comme $n \in \llbracket n_i, n_{i+1} - 1 \rrbracket$, alors : $h_n = i$. Et donc : $M_n = \frac{1}{i}$.

Ainsi :

$$|D_n(x)| \leq M_n$$

Avec la suite (M_n) définie plus haut : $\forall n \in \mathbb{N}^*, |D_n(x)| \leq M_n$.

(ii) On souhaite démontrer : $\lim_{n \rightarrow +\infty} M_n = 0$.

Pour cela, on va démontrer : $\lim_{n \rightarrow +\infty} h_n = +\infty$ (ce qui semble être cohérent avec l'allure de la représentation graphique de (h_n)). Autrement dit, on veut montrer :

$$\forall A \in \mathbb{R}_+, \exists p \in \mathbb{N}, \forall n \geq p, h_n \geq A$$

Soit $A \in \mathbb{R}_+$. On note : $p = n_{\lceil A \rceil}$. Alors, par définition de (h_n) :

$$\forall n \geq p, h_n \geq \lceil A \rceil$$

Or, par définition de la partie entière par excès : $\lceil A \rceil \geq A$. D'où :

$$\forall n \geq p, h_n \geq A$$

On a bien démontré : $\forall A \in \mathbb{R}_+, \exists p \in \mathbb{N}, \forall n \geq p, h_n \geq A$. Autrement dit : $\lim_{n \rightarrow +\infty} h_n = +\infty$.

On en déduit : $\lim_{n \rightarrow +\infty} M_n = 0$.

Commentaire

- Si chaque étape de la démonstration est à portée d'un bon élève de classe ECE, la prise d'initiative est beaucoup trop importante pour espérer qu'un élève en vienne à bout. Cette question n'a donc pas le rôle discriminant qu'ont généralement les questions de concours : classer les élèves selon qu'ils ont traité de manière satisfaisante ou non la question. La présence d'une telle question permet de comprendre la stratégie à adopter lors des concours :

- × il est essentiel de savoir repérer les questions les plus difficiles. Elles permettent de discriminer les candidats puisqu'il faut avoir du recul pour juger du niveau d'une question.
- × il faut aborder ces questions en ayant en tête que le correcteur sera plus indulgent pour les candidats qui s'y aventurent. Cependant, il ne faut pas perdre du temps à essayer de les traiter jusqu'au bout : le nombre de points alloués ne sera certainement pas à la hauteur du temps investi pour traiter une telle question.

Il ne faut donc pas hésiter à passer les questions les plus difficiles et aller chercher les points où ils sont, à savoir sur les questions plus abordables du sujet.

- Précisons par ailleurs qu'un tel niveau de détail n'est pas attendu aux concours. Donner l'idée de la construction de la suite (M_n) et signaler qu'elle tend vers 0 par construction est bien suffisant pour obtenir tous les points. □

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) Démontrer : $\lim_{n \rightarrow +\infty} \Phi(x_n) = \Phi(x)$.

Démonstration.

D'après l'énoncé : $\lim_{n \rightarrow +\infty} x_n = x$.

Or, d'après la question 1.a), la fonction Φ est continue sur \mathbb{R} .

$$\text{On en déduit : } \lim_{n \rightarrow +\infty} \Phi(x_n) = \Phi(x).$$

Commentaire

La question précédente était particulièrement difficile. Celle-ci, en revanche, est particulièrement simple. On en conclut qu'il n'y a pas forcément dans tout le sujet de croissance linéaire de la difficulté. Ainsi, passer une question ne doit pas empêcher de tenter d'aborder la suivante. □

(ii) En appliquant le résultat (**) de la question 4.d), démontrer :

$$\lim_{n \rightarrow +\infty} |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| = 0$$

Démonstration.

Soit $n \in \mathbb{N}^*$. D'après 4.d) :

$$\forall x \in \mathbb{R}, \quad |\mathbb{P}([Z_n \leq x]) - \Phi(x)| \leq M_n$$

En particulier, en appliquant l'inégalité précédente à $x = x_n$, on obtient :

$$|\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| \leq M_n$$

D'où :

$$0 \leq |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| \leq M_n$$

Or, toujours d'après 4.d) : $\lim_{n \rightarrow +\infty} M_n = 0$.

$$\text{Par théorème d'encadrement, on en déduit : } \lim_{n \rightarrow +\infty} |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| = 0. \quad \square$$

(iii) En conclure que l'on a : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x_n]) = \Phi(x)$.

Démonstration.

• Pour démontrer : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x_n]) = \Phi(x)$, on montrera :

$$|\mathbb{P}([Z_n \leq x_n]) - \Phi(x)| \xrightarrow{n \rightarrow +\infty} 0$$

• Soit $n \in \mathbb{N}^*$.

$$\begin{aligned} |\mathbb{P}([Z_n \leq x_n]) - \Phi(x)| &= \left| \left(\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n) \right) + \left(\Phi(x_n) - \Phi(x) \right) \right| \\ &\leq |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| + |\Phi(x_n) - \Phi(x)| \quad (\text{par inégalité triangulaire}) \end{aligned}$$

Ainsi :

$$0 \leq |\mathbb{P}([Z_n \leq x_n]) - \Phi(x)| \leq |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| + |\Phi(x_n) - \Phi(x)|$$

Or :

$$\times \text{ d'après } \mathbf{5.a)(ii)} : \lim_{n \rightarrow +\infty} |\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| = 0,$$

$$\times \text{ d'après } \mathbf{5.a)(i)} : \lim_{n \rightarrow +\infty} \Phi(x_n) = \Phi(x). \text{ D'où : } \lim_{n \rightarrow +\infty} |\Phi(x_n) - \Phi(x)| = 0.$$

Ainsi :

$$\lim_{n \rightarrow +\infty} \left(|\mathbb{P}([Z_n \leq x_n]) - \Phi(x_n)| + |\Phi(x_n) - \Phi(x)| \right) = 0$$

Par théorème d'encadrement, on obtient :

$$|\mathbb{P}([Z_n \leq x_n]) - \Phi(x)| \xrightarrow{n \rightarrow +\infty} 0$$

Finalement : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x_n]) = \Phi(x).$

□

b) (i) Montrer que pour tout $n \geq 1$, on a : $\mathbb{P}\left(\left[Z_n \leq x - \frac{1}{n}\right]\right) \leq \mathbb{P}([Z_n < x]) \leq \mathbb{P}([Z_n \leq x]).$

Démonstration.

Soit $n \in \mathbb{N}^*$.

• Démontrons :

$$\left[Z_n \leq x - \frac{1}{n}\right] \subset [Z_n < x] \subset [Z_n \leq x]$$

Soit $\omega \in \left[Z_n \leq x - \frac{1}{n}\right]$, alors :

$$Z_n(\omega) \leq x - \frac{1}{n} < x$$

Ainsi : $Z_n(\omega) < x$. D'où : $\omega \in [Z_n < x]$.

Donc : $\left[Z_n \leq x - \frac{1}{n}\right] \subset [Z_n < x].$

De plus :

$$[Z_n \leq x] = [Z_n < x] \cup [Z_n = x]$$

Ainsi : $[Z_n < x] \subset [Z_n \leq x].$

• On a obtenu :

$$\left[Z_n \leq x - \frac{1}{n}\right] \subset [Z_n < x] \subset [Z_n \leq x]$$

Par croissance de \mathbb{P} , on en déduit :

$$\mathbb{P}\left(\left[Z_n \leq x - \frac{1}{n}\right]\right) \leq \mathbb{P}([Z_n < x]) \leq \mathbb{P}([Z_n \leq x])$$

$\forall n \in \mathbb{N}^*, \mathbb{P}\left(\left[Z_n \leq x - \frac{1}{n}\right]\right) \leq \mathbb{P}([Z_n < x]) \leq \mathbb{P}([Z_n \leq x])$

□

(ii) En déduire : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n < x]) = \Phi(x)$.

Démonstration.

- On définit la suite (x_n) par :

$$\forall n \in \mathbb{N}^*, \quad x_n = x - \frac{1}{n}$$

Alors : $\lim_{n \rightarrow +\infty} x_n = x$.

- On a démontré en question précédente, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}([Z_n \leq x_n]) \leq \mathbb{P}([Z_n < x]) \leq \mathbb{P}([Z_n \leq x])$$

Or :

× comme $\lim_{n \rightarrow +\infty} x_n = x$, d'après la question **5.a)(iii)** : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x_n]) = \Phi(x)$.

× d'après **2.b)** : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq x]) = \Phi(x)$.

Par théorème d'encadrement : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n < x]) = \Phi(x)$.

□

c) Montrer que, pour tous réels a, b qui vérifient $a < b$, on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \in [a, b]]) = \Phi(b) - \Phi(a)$$

Démonstration.

Soit $(a, b) \in \mathbb{R}^2$ tel que : $a < b$.

- Soit $n \in \mathbb{N}^*$.

$$\begin{aligned} \mathbb{P}([Z_n \in [a, b]]) &= \mathbb{P}([a \leq Z_n \leq b]) \\ &= \mathbb{P}([Z_n \leq b]) - \mathbb{P}([Z_n < a]) \end{aligned}$$

En effet :

$$[Z_n \leq b] = [Z_n < a] \cup [a \leq Z_n \leq b]$$

Les événements $[Z_n < a]$ et $[a \leq Z_n \leq b]$ étant incompatibles, on obtient :

$$\mathbb{P}([Z_n \leq b]) = \mathbb{P}([Z_n < a]) + \mathbb{P}([a \leq Z_n \leq b])$$

- Or :

× d'après la question précédente appliquée à $x = a$: $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n < a]) = \Phi(a)$.

× d'après la question **2.b)** appliquée à $x = b$: $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq b]) = \Phi(b)$.

On en déduit :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \in [a, b]]) = \Phi(b) - \Phi(a)$$

Pour tout $(a, b) \in \mathbb{R}^2$ tels que $a < b$: $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \in [a, b]]) = \Phi(b) - \Phi(a)$.

□

Deuxième partie : Applications en statistique

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 $i^{\text{ème}}$ personne vote pour le candidat A et $X_i = 0$ si la $i^{\text{ème}}$ 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 .

Commentaire

- Dans un énoncé de probabilités, on manipule différents niveaux d'objets.
 - 1) Au premier niveau, on trouve l'expérience aléatoire considérée.
On note Ω l'univers des possibles : c'est l'**ensemble** des résultats possibles (appelés aussi issues) de l'expérience. Ici, Ω est un ensemble fixé au début du problème par l'énoncé.
 - 2) Au deuxième niveau, on trouve les événements : un événement A n'est rien d'autre qu'un **ensemble** qui regroupe certaines issues de l'expérience. Ainsi : $A \subset \Omega$.
On peut écrire : $A = \{\omega \in \Omega \mid \omega \in A\}$.
Lorsque $\omega \in A$, on dit que ω **réalise** l'événement A .
 - 3) Au troisième niveau, on trouve les v.a.r. . Ce sont des **applications** particulières :
 - elles prennent comme argument un résultat possible de l'expérience et renvoient une valeur réelle. Par exemple, avec une réalisation $\omega \in \Omega$, la v.a.r. X_i peut prendre la valeur 0 ou la valeur 1.
En effet, d'après l'énoncé, X_i prend la valeur 1 si la $i^{\text{ème}}$ personne vote pour le candidat A et 0 sinon.
 - elles sont des machines à créer des événements. Par exemple, $[X_i = 1]$ est un événement.
Il regroupe **toutes** les réalisations ω tels que : $X_i(\omega) = 1$.
Autrement dit : $[X_i = 1] = \{\omega \in \Omega \mid X_i(\omega) = 1\} \subset \Omega$.
 Ce deuxième point nous replonge au deuxième niveau. Ainsi, pour comprendre le chapitre sur les v.a.r. , il est essentiel de maîtriser celui sur les probabilités générales.
- Les notations $X_i = 1$ ou $X_i = 0$ de l'énoncé sont malvenues.
Écrire $X_i = 1$ signifie : $\forall \omega \in \Omega, X_i(\omega) = 1$. Il n'en est rien. Il faut alors traduire rigoureusement l'énoncé pour comprendre la définition de la v.a.r. X_i : si la $i^{\text{ème}}$ personne vote pour le candidat A , alors la v.a.r. X_i prend la valeur 1 (on dit aussi que l'événement $[X_i = 1]$ est réalisé). Mais en aucun cas, il ne faut écrire $X_i = 1$ car cela marque une confusion d'objets.

6. a) Démontrer : $\mathbb{E}(X_i) = p$ et $\mathbb{V}(X_i) = p(1 - p)$.

Démonstration.

Soit $i \in \mathbb{N}^*$.

- La v.a.r. X_i admet une variance (et donc une espérance) car c'est une v.a.r. finie.
- Tout d'abord :

$$\begin{aligned} \mathbb{E}(X_i) &= \cancel{0 \times \mathbb{P}([X_i = 0])} + 1 \times \mathbb{P}([X_i = 1]) \\ &= p \quad (\text{car } X_i \hookrightarrow \mathcal{B}(p)) \end{aligned}$$

$$\boxed{\mathbb{E}(X_i) = p}$$

- Ensuite :

$$\mathbb{E}(X_i^2) = 0^2 \times \mathbb{P}([X_i = 0]) + 1^2 \times \mathbb{P}([X_i = 1]) = 1 \times p = p$$

Par formule de Koenig-Huygens :

$$\mathbb{V}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = p - p^2 = p(1 - p)$$

$$\mathbb{V}(X_i) = p(1 - p)$$

Commentaire

Il faut prêter attention à la formulation de l'énoncé. Bien que l'espérance et la variance d'une v.a.r. de loi $\mathcal{B}(p)$ soient connues de tout candidat ECE, il s'agit bien ici de **retrouver** ces valeurs à l'aide d'une démonstration. Une réponse lapidaire du type « $X_i \hookrightarrow \mathcal{B}(p)$ donc $\mathbb{E}(X_i) = p$ et $\mathbb{V}(X_i) = p(1 - p)$ » ne peut permettre d'obtenir des points à cette question. \square

- b) On notera dans la suite $\sigma = \sqrt{p(1 - p)}$. Démontrer : $\sigma \leq \frac{1}{2}$.

Démonstration.

On remarque :

$$\begin{aligned} \sigma \leq \frac{1}{2} &\Leftrightarrow \sqrt{p(1 - p)} \leq \frac{1}{2} \\ &\Leftrightarrow p(1 - p) \leq \frac{1}{4} \quad (\text{car la fonction } x \mapsto x^2 \text{ est} \\ &\quad \text{strictement croissante sur } \mathbb{R}_+) \\ &\Leftrightarrow p - p^2 \leq \frac{1}{4} \\ &\Leftrightarrow 0 \leq p^2 - p + \frac{1}{4} \\ &\Leftrightarrow 0 \leq \left(p - \frac{1}{2}\right)^2 \end{aligned}$$

Cette dernière assertion est vérifiée.

Grâce au raisonnement par équivalence, la première inégalité l'est également.

$$\text{On en déduit : } \sigma \leq \frac{1}{2}.$$

Commentaire

La majoration classique « $p(1 - p) \leq \frac{1}{4}$ » (équivalente à celle demandée par l'énoncé par stricte croissance de la fonction carrée sur \mathbb{R}_+) peut aussi se démontrer grâce à l'étude de la fonction $f : p \mapsto p(1 - p)$ sur $[0, 1]$.

- La fonction f est une fonction polynomiale, donc dérivable sur $[0, 1]$ (et même sur \mathbb{R}), et :

$$\forall p \in [0, 1], \quad f'(p) = 1 - 2p$$

- On obtient alors le tableau de variations suivant qui permet d'obtenir l'inégalité souhaitée :

p	0	$\frac{1}{2}$	1
Signe de $f'(p)$		+	-
Variations de f		\nearrow	\searrow
	0	$\frac{1}{4}$	0

\square

c) Démontrer : $\mathbb{E}(\bar{X}_n) = p$.

Démonstration.

- La v.a.r. \bar{X}_n admet une espérance en tant que combinaison linéaire de v.a.r. qui en admettent une.
- De plus :

$$\begin{aligned}\mathbb{E}(\bar{X}_n) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) \quad (\text{par linéarité de l'espérance}) \\ &= \frac{1}{n} \sum_{i=1}^n p \quad (\text{car, d'après 6.a) : } \\ &\quad \forall i \in \llbracket 1, n \rrbracket, \mathbb{E}(X_i) = p \\ &= \frac{1}{n} \times n p\end{aligned}$$

Finalement : $\mathbb{E}(\bar{X}_n) = p$.

□

d) Démontrer : $\mathbb{V}(\bar{X}_n) = \frac{1}{n} \sigma^2$.

Démonstration.

- La v.a.r. \bar{X}_n admet une variance en tant que combinaison linéaire de v.a.r. qui en admettent une.
- De plus :

$$\begin{aligned}\mathbb{V}(\bar{X}_n) &= \mathbb{V}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \mathbb{V}\left(\sum_{i=1}^n X_i\right) \quad (\text{par propriété de la variance}) \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) \quad (\text{car } X_1, \dots, X_n \text{ sont} \\ &\quad \text{mutuellement indépendantes}) \\ &= \frac{1}{n^2} \sum_{i=1}^n p(1-p) \quad (\text{car, d'après 6.a) : } \\ &\quad \forall i \in \llbracket 1, n \rrbracket, \mathbb{V}(X_i) = p(1-p)) \\ &= \frac{1}{n^2} \times n p(1-p)\end{aligned}$$

Finalement : $\mathbb{V}(\bar{X}_n) = \frac{p(1-p)}{n} = \frac{\sigma^2}{n}$.

□

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 :

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

Démonstration.

- D'après 5.c), pour tout $(a, b) \in \mathbb{R}^2$ tels que $a < b$, on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \in [a, b]]) = \Phi(b) - \Phi(a)$$

Soit $a > 0$. Alors : $-a < a$. On obtient donc :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \in [-a, a]]) = \Phi(a) - \Phi(-a)$$

- De plus, d'après l'énoncé :

$$Z_n = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}$$

Or :

- × d'une part, μ est l'espérance commune des v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$.

Ainsi, d'après **6.a**) : $\mu = p$.

- × d'autre part, σ^2 est la variance commune des v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$.

Ainsi, d'après **6.a**) : $\sigma^2 = p(1-p)$ (et donc : $\sigma = \sqrt{p(1-p)}$).

- On en déduit : $Z_n = \sqrt{n} \frac{\bar{X}_n - p}{\sigma}$. D'où :

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

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

□

- b)** En déduire :

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

Démonstration.

- D'après la question précédente :

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

- D'une part, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} \mathbb{P}\left(\left[\frac{\sqrt{n}}{\sigma} (\bar{X}_n - p) \in [-a, a]\right]\right) &= \mathbb{P}\left(\left[-a \leq \frac{\sqrt{n}}{\sigma} (\bar{X}_n - p) \leq a\right]\right) \\ &= \mathbb{P}\left(\left[-a \frac{\sigma}{\sqrt{n}} \leq \bar{X}_n - p \leq a \frac{\sigma}{\sqrt{n}}\right]\right) \quad (\text{car } \frac{\sigma}{\sqrt{n}} > 0) \\ &= \mathbb{P}\left(\left[a \frac{\sigma}{\sqrt{n}} \geq p - \bar{X}_n \geq -a \frac{\sigma}{\sqrt{n}}\right]\right) \\ &= \mathbb{P}\left(\left[\bar{X}_n + a \frac{\sigma}{\sqrt{n}} \geq p \geq \bar{X}_n - a \frac{\sigma}{\sqrt{n}}\right]\right) \\ &= \mathbb{P}\left(\left[p \in \left[\bar{X}_n - a \frac{\sigma}{\sqrt{n}}, \bar{X}_n + a \frac{\sigma}{\sqrt{n}}\right]\right]\right) \end{aligned}$$

- D'autre part, d'après **1.d**) :

$$\Phi(a) - \Phi(-a) = \Phi(a) - (1 - \Phi(a)) = 2\Phi(a) - 1$$

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

□

- c) Une table des valeurs de Φ 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 $\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right]$.

Démonstration.

D'après la question précédente, pour n grand :

$$\begin{aligned} \mathbb{P} \left(\left[p \in \left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right] \right) &\approx 2\Phi(1,96) - 1 \\ &\approx 2 \times 0,975 - 1 = 0,95 \end{aligned}$$

Ainsi, 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 σ 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 $\left[\bar{X}_n - \frac{0,98}{\sqrt{n}}, \bar{X}_n + \frac{0,98}{\sqrt{n}} \right]$.

Démonstration.

- Soit $n \in \mathbb{N}^*$.

$$\mathbb{P} \left(\left[p \in \left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right] \right) = \mathbb{P} \left(\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right)$$

- Or, d'après la question **6.b)** :

$$\sigma \leq \frac{1}{2} \quad \text{et} \quad -\sigma \geq -\frac{1}{2}$$

On en déduit :

$$\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \subset \left[\bar{X}_n - 1,96 \frac{1}{2\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{1}{2\sqrt{n}} \right]$$

- Ainsi, par croissance de \mathbb{P} :

$$\mathbb{P} \left(\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right) \leq \mathbb{P} \left(\left[\bar{X}_n - 1,96 \frac{1}{2\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{1}{2\sqrt{n}} \right] \right)$$

$$\mathbb{P} \left(\left[p \in \left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right] \right) \leq \mathbb{P} \left(\left[p \in \left[\bar{X}_n - \frac{0,98}{\sqrt{n}}, \bar{X}_n + \frac{0,98}{\sqrt{n}} \right] \right] \right)$$

On en conclut :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right) \leq \lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \in \left[\bar{X}_n - \frac{0,98}{\sqrt{n}}, \bar{X}_n + \frac{0,98}{\sqrt{n}} \right] \right] \right)$$

- De plus, d'après la question précédente :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[\bar{X}_n - 1,96 \frac{\sigma}{\sqrt{n}} \leq p \leq \bar{X}_n + 1,96 \frac{\sigma}{\sqrt{n}} \right] \right) = 2\Phi(1,96) - 1 \approx 0,95$$

On en déduit :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \in \left[\bar{X}_n - \frac{0,98}{\sqrt{n}}, \bar{X}_n + \frac{0,98}{\sqrt{n}} \right] \right] \right) \geq 2\Phi(1,96) - 1 \approx 0,95$$

Ainsi, 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 X_1, \dots, X_n pour estimer σ .

8. On pose, pour $n \geq 1$: $V_n = \bar{X}_n (1 - \bar{X}_n) + \frac{1}{n}$. Soit $\varepsilon > 0$.

a) Démontrer : $V_n - \sigma^2 - \frac{1}{n} = (\bar{X}_n - p)(1 - \bar{X}_n - p)$.

Démonstration.

Soit $n \in \mathbb{N}^*$.

• D'une part :

$$V_n - \sigma^2 - \frac{1}{n} = \bar{X}_n (1 - \bar{X}_n) + \frac{1}{n} - \sigma^2 - \frac{1}{n} = \bar{X}_n - \bar{X}_n^2 - p + p^2$$

• D'autre part :

$$(\bar{X}_n - p)(1 - \bar{X}_n - p) = \bar{X}_n - \bar{X}_n^2 - p\bar{X}_n - p + p\bar{X}_n + p^2 = \bar{X}_n - \bar{X}_n^2 - p + p^2$$

$$\text{Finalement : } \forall n \in \mathbb{N}^*, V_n - \sigma^2 - \frac{1}{n} = (\bar{X}_n - p)(1 - \bar{X}_n - p).$$

□

b) En déduire : $|V_n - \sigma^2| \leq 2|\bar{X}_n - p| + \frac{1}{n}$.

Démonstration.

Soit $n \in \mathbb{N}^*$.

• Tout d'abord :

$$\begin{aligned} |V_n - \sigma^2| &= \left| (\bar{X}_n - p)(1 - p - \bar{X}_n) + \frac{1}{n} \right| \quad (\text{d'après la question précédente}) \\ &\leq \left| (\bar{X}_n - p)(1 - p - \bar{X}_n) \right| + \frac{1}{n} \quad (\text{par inégalité triangulaire}) \\ &\leq |\bar{X}_n - p| |1 - p - \bar{X}_n| + \frac{1}{n} \end{aligned}$$

• De plus, par inégalité triangulaire :

$$|1 - p - \bar{X}_n| \leq |1 - p| + |\bar{X}_n|$$

Or :

× d'une part, comme $p \in]0, 1[$, on obtient : $0 < 1 - p < 1$. D'où : $|1 - p| < 1$.

× d'autre part : $\forall i \in \mathbb{N}^*, X_i(\Omega) = \{0, 1\}$. Donc :

$$\forall i \in \mathbb{N}^*, \quad 0 \leq X_i \leq 1$$

En sommant ces encadrements pour $i \in \llbracket 1, n \rrbracket$, on obtient :

$$\begin{aligned} 0 &\leq \sum_{i=1}^n X_i \leq \sum_{i=1}^n 1 \\ \text{donc } 0 &\leq \sum_{i=1}^n X_i \leq n \\ \text{d'où } 0 &\leq \frac{1}{n} \sum_{i=1}^n X_i \leq 1 \quad (\text{car : } \frac{1}{n} > 0) \\ \text{ainsi } 0 &\leq \bar{X}_n \leq 1 \end{aligned}$$

On en déduit : $|\bar{X}_n| \leq 1$.

On obtient alors, par transitivité :

$$|1 - p - \bar{X}_n| \leq |1 - p| + |\bar{X}_n| \leq 2$$

- Comme $|\bar{X}_n - p| \geq 0$, on en déduit :

$$|\bar{X}_n - p| |1 - p - \bar{X}_n| \leq 2 |\bar{X}_n - p|$$

Alors, par transitivité :

$$|V_n - \sigma^2| \leq |\bar{X}_n - p| |1 - p - \bar{X}_n| + \frac{1}{n} \leq 2 |\bar{X}_n - p| + \frac{1}{n}$$

$$\forall n \in \mathbb{N}^*, |V_n - \sigma^2| \leq 2 |\bar{X}_n - p| + \frac{1}{n}$$

□

c) Démontrer : $\mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}\left(|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right)$.

Démonstration.

Soit $n \in \mathbb{N}^*$.

- D'après la question précédente :

$$|V_n - \sigma^2| \leq 2 |\bar{X}_n - p| + \frac{1}{n}$$

Donc :

$$[|V_n - \sigma^2| > \varepsilon] \subset \left[2 |\bar{X}_n - p| + \frac{1}{n} > \varepsilon\right]$$

Commentaire

Détaillons l'obtention de cette inclusion.

Soit $\omega \in [|V_n - \sigma^2| > \varepsilon]$. Alors : $|V_n(\omega) - \sigma^2| > \varepsilon$.

Ainsi, par transitivité :

$$\varepsilon < |V_n(\omega) - \sigma^2| \leq 2 |\bar{X}_n(\omega) - p| + \frac{1}{n}$$

On en déduit : $2 |\bar{X}_n(\omega) - p| + \frac{1}{n} > \varepsilon$. D'où : $\omega \in \left[2 |\bar{X}_n - p| + \frac{1}{n} > \varepsilon\right]$.

Finalement : $[|V_n - \sigma^2| > \varepsilon] \subset \left[2 |\bar{X}_n - p| + \frac{1}{n} > \varepsilon\right]$.

Par croissance de \mathbb{P} , on obtient :

$$\mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}\left(\left[2|\bar{X}_n - p| + \frac{1}{n} > \varepsilon\right]\right)$$

• Or :

$$\begin{aligned}\mathbb{P}\left(\left[2|\bar{X}_n - p| + \frac{1}{n}\right]\right) &= \mathbb{P}\left(\left[2|\bar{X}_n - p| > \varepsilon - \frac{1}{n}\right]\right) \\ &= \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right)\end{aligned}$$

Finalement : $\mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right)$.

□

d) Démontrer, pour n assez grand : $\mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right)$.

Démonstration.

• On sait : $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$. Ainsi, pour tout $\varepsilon' > 0$, il existe $n_0 \in \mathbb{N}^*$ tel que : $\forall n \geq n_0, \left|\frac{1}{n} - 0\right| < \varepsilon'$.

Autrement dit : $\forall n \geq n_0, \frac{1}{n} < \varepsilon'$.

• Soit $n \geq n_0$. Alors : $-\frac{1}{n} > -\varepsilon'$. D'où :

$$\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right] \subset \left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{\varepsilon'}{2}\right]$$

Commentaire

Détaillons l'obtention de cette inclusion.

Soit $\omega \in \left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]$. Alors : $|\bar{X}_n(\omega) - p| > \frac{\varepsilon}{2} - \frac{1}{2n}$.

Ainsi, par transitivité :

$$|\bar{X}_n(\omega) - p| > \frac{\varepsilon}{2} - \frac{1}{2n} > \frac{\varepsilon}{2} - \frac{\varepsilon'}{2}$$

On en déduit : $|\bar{X}_n(\omega) - p| > \frac{\varepsilon}{2} - \frac{\varepsilon'}{2}$. D'où : $\omega \in \left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{\varepsilon'}{2}\right]$.

Finalement : $\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right] \subset \left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{\varepsilon'}{2}\right]$.

• On choisit alors ε' de sorte que : $\frac{\varepsilon}{2} - \frac{\varepsilon'}{2} = \frac{\varepsilon}{4}$.

$$\frac{\varepsilon}{2} - \frac{\varepsilon'}{2} = \frac{\varepsilon}{4} \Leftrightarrow \frac{\varepsilon}{2} - \frac{\varepsilon}{4} = \frac{\varepsilon'}{2} \Leftrightarrow \frac{\varepsilon}{4} = \frac{\varepsilon'}{2} \Leftrightarrow \frac{\varepsilon}{2} = \varepsilon'$$

On choisit donc : $\varepsilon' = \frac{\varepsilon}{2} > 0$. Ainsi :

$$\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right] \subset \left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]$$

- Par croissance de \mathbb{P} :

$$\mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right)$$

Finalement, pour n assez grand ($n \geq n_0$) : $\mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right)$. □

- e) Conclure que pour tout $\varepsilon > 0$, on a : $\lim_{n \rightarrow +\infty} \mathbb{P}(|V_n - \sigma^2| > \varepsilon) = 0$.

Démonstration.

Soit $n \in \mathbb{N}^*$ tel que : $n \geq n_0$.

- D'après **8.c)** et **8.d)**, par transitivité :

$$\mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{2} - \frac{1}{2n}\right]\right) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right)$$

- Les v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$:

× sont indépendantes,

× admettent toutes la même espérance p d'après **6.a)**,

× admettent toutes la même variance $p(1-p)$ d'après **6.a)**.

D'après la loi faible des grands nombres :

$$\forall \varepsilon' > 0, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(|\bar{X}_n - p| > \varepsilon') = 0$$

En appliquant cette propriété à $\varepsilon' = \frac{\varepsilon}{4} > 0$, on obtient :

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right) = 0$$

- On sait donc :

× d'une part :

$$0 \leq \mathbb{P}(|V_n - \sigma^2| > \varepsilon) \leq \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right)$$

× d'autre part : $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[|\bar{X}_n - p| > \frac{\varepsilon}{4}\right]\right) = 0$

Alors, par théorème d'encadrement : $\lim_{n \rightarrow +\infty} \mathbb{P}(|V_n - \sigma^2| > \varepsilon) = 0$ □

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

Démonstration.

Soit $n \in \mathbb{N}^*$.

- Tout d'abord, par définition de W_n :

$$\begin{aligned} \mathbb{P}([W_n \leq x]) &= \mathbb{P}\left(\left[\frac{\sigma}{\sqrt{V_n}} Z_n \leq x\right]\right) \\ &= \mathbb{P}\left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x\right]\right) \quad (\text{car } \frac{\sqrt{V_n}}{\sigma} \text{ est une v.a.r.} \\ &\quad \text{strictement positive}) \end{aligned}$$

- On cherche alors à faire apparaître l'événement $\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right]$. Pour cela, on remarque que la famille $\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right], \left[\frac{\sqrt{V_n}}{\sigma} \leq 1 + \varepsilon \right] \right)$ est un système complet d'événements. Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned} & \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \right) \\ &= \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \leq 1 + \varepsilon \right] \right) \end{aligned}$$

Étudions chacune de ces probabilités **dans le cas $x > 0$** .

× D'une part :

$$\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \subset \left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right]$$

$$\text{Ainsi : } \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right).$$

× D'autre part, comme $x > 0$, alors :

$$\left[\frac{\sqrt{V_n}}{\sigma} \leq 1 + \varepsilon \right] = \left[\frac{\sqrt{V_n}}{\sigma} x \leq (1 + \varepsilon) x \right]$$

Donc :

$$\begin{aligned} \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \leq 1 + \varepsilon \right] &= \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} x \leq (1 + \varepsilon) x \right] \\ &= \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \leq (1 + \varepsilon) x \right] \end{aligned}$$

Or :

$$\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \leq (1 + \varepsilon) x \right] \subset [Z_n \leq (1 + \varepsilon) x]$$

Ainsi, par croissance de \mathbb{P} :

$$\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \leq (1 + \varepsilon) x \right] \right) \leq \mathbb{P}([Z_n \leq (1 + \varepsilon) x])$$

$$\text{Donc : } \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \leq 1 + \varepsilon \right] \right) \leq \mathbb{P}([Z_n \leq (1 + \varepsilon) x]).$$

On en déduit :

$$\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \right) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq (1 + \varepsilon) x])$$

C'est-à-dire :

$$\mathbb{P}([W_n \leq x]) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq (1 + \varepsilon) x])$$

$$\forall x > 0, \mathbb{P}([W_n \leq x]) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq (1 + \varepsilon) x])$$

- Démontrons maintenant que l'inégalité de l'énoncé est vraie pour $x = 0$.

× D'une part :

$$\begin{aligned} \mathbb{P}([W_n \leq 0]) &= \mathbb{P} \left(\left[\frac{\sigma}{\sqrt{V_n}} Z_n \leq 0 \right] \right) \\ &= \mathbb{P}([Z_n \leq 0]) \quad \left(\text{car } \frac{\sqrt{V_n}}{\sigma} \text{ est une v.a.r.} \right. \\ &\quad \left. \text{strictement positive} \right) \end{aligned}$$

× D'autre part :

$$\begin{aligned} \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq (1 + \varepsilon) \times 0]) &= \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq 0]) \\ &\geq \mathbb{P}([Z_n \leq 0]) \end{aligned}$$

On en déduit : $\mathbb{P}([W_n \leq 0]) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) + \mathbb{P}([Z_n \leq (1 + \varepsilon) \times 0])$.

Commentaire

- Notons que cette démonstration ne permet pas d'aboutir à l'inégalité de l'énoncé dans le cas $x < 0$. Pour tout $x < 0$, on peut néanmoins démontrer l'inégalité suivante, qui permet bien d'obtenir un résultat similaire à celui de la question **9.a)(i)** (nous y reviendrons) :

$$\mathbb{P}([W_n \leq x]) \leq \mathbb{P}([Z_n \leq (1 - \varepsilon)x]) + \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right)$$

- Démontrons cette assertion.

Tout d'abord, comme dans le 1^{er} point de la démonstration :

$$\mathbb{P}([W_n \leq x]) = \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \right)$$

La famille $\left(\left[\frac{\sqrt{V_n}}{\sigma} \geq 1 - \varepsilon \right], \left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right)$ forme un système complet d'événements. Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned} &\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \right) \\ &= \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \geq 1 - \varepsilon \right] \right) + \mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right) \end{aligned}$$

Étudions chacune de ces probabilités.

× D'une part :

$$\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \subset \left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right]$$

Ainsi :

$$\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right) \leq \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right)$$

Commentaire

× D'autre part, comme $x < 0$, alors :

$$\left[\frac{\sqrt{V_n}}{\sigma} \geq 1 - \varepsilon \right] = \left[\frac{\sqrt{V_n}}{\sigma} x \leq (1 - \varepsilon) x \right]$$

Donc :

$$\begin{aligned} \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \geq 1 - \varepsilon \right] &= \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} x \leq (1 - \varepsilon) x \right] \\ &= \left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \leq (1 - \varepsilon) x \right] \end{aligned}$$

Or :

$$\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \leq (1 - \varepsilon) x \right] \subset [Z_n \leq (1 - \varepsilon) x]$$

Ainsi :

$$\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \cap \left[\frac{\sqrt{V_n}}{\sigma} \leq 1 - \varepsilon \right] \right) \leq \mathbb{P}([Z_n \leq (1 - \varepsilon) x])$$

On en déduit :

$$\mathbb{P} \left(\left[Z_n \leq \frac{\sqrt{V_n}}{\sigma} x \right] \right) \leq \mathbb{P}([Z_n \leq (1 - \varepsilon) x]) + \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right)$$

C'est-à-dire :

$$\mathbb{P}([W_n \leq x]) \leq \mathbb{P}([Z_n \leq (1 - \varepsilon) x]) + \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon \right] \right)$$

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

Démonstration.

• Soit $n \in \mathbb{N}^*$.

$$\begin{aligned} \mathbb{P} \left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon \right] \right) &= \mathbb{P} \left(\left[\frac{V_n}{\sigma^2} > (1 + \varepsilon)^2 \right] \right) && \text{(par stricte croissance de la} \\ & && \text{fonction } x \mapsto x^2 \text{ sur } \mathbb{R}_+) \\ &= \mathbb{P}([V_n > \sigma^2(1 + 2\varepsilon + \varepsilon^2)]) && \text{(car } \sigma^2 > 0) \\ &= \mathbb{P}([V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]) \end{aligned}$$

• On rappelle qu'on a démontré en question **8.e**) :

$$\forall \varepsilon' > 0, \quad \lim_{n \rightarrow +\infty} \mathbb{P}([|V_n - \sigma^2| > \varepsilon']) = 0$$

On va donc chercher à utiliser ce résultat pour $\varepsilon' = \sigma^2(2\varepsilon + \varepsilon^2) > 0$.

× Tout d'abord :

$$[|V_n - \sigma^2| > \sigma^2(2\varepsilon + \varepsilon^2)] = [V_n - \sigma^2 < -\sigma^2(2\varepsilon + \varepsilon^2)] \cup [V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]$$

Ainsi :

$$[V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)] \subset [V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]$$

Donc :

$$\mathbb{P}([V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]) \leq \mathbb{P}([|V_n - \sigma^2| > \sigma^2(2\varepsilon + \varepsilon^2)])$$

× On en déduit :

$$0 \leq \mathbb{P}([V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]) \leq \mathbb{P}(|V_n - \sigma^2| > \sigma^2(2\varepsilon + \varepsilon^2))$$

Or, comme $\sigma^2(2\varepsilon + \varepsilon^2) > 0$, d'après **8.e** :

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|V_n - \sigma^2| > \sigma^2(2\varepsilon + \varepsilon^2)) = 0$$

Par théorème d'encadrement, on obtient :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]) = 0$$

On en conclut : $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) = 0.$

Commentaire

- Revenons sur l'égalité :

$$[|V_n - \sigma^2| > \sigma^2(2\varepsilon + \varepsilon^2)] = [V_n - \sigma^2 < -\sigma^2(2\varepsilon + \varepsilon^2)] \cup [V_n - \sigma^2 > \sigma^2(2\varepsilon + \varepsilon^2)]$$

On rappelle que, pour tout $c \in \mathbb{R}_+$ et tout $x \in \mathbb{R}$:

$$|x| \leq c \Leftrightarrow -c \leq x \leq c \Leftrightarrow (-c \leq x \text{ ET } x \leq c)$$

Par contraposée :

$$|x| > c \Leftrightarrow (-c > x \text{ OU } x > c) \Leftrightarrow (x < -c \text{ OU } x > c)$$

En appliquant ce résultat à $x = V_n - \sigma^2$ et $c = \sigma^2(2\varepsilon + \varepsilon^2)$, on obtient bien l'égalité entre événements souhaitée.

- On peut démontrer avec un raisonnement similaire, pour tout $\varepsilon \in]0, 1[$:

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon\right]\right) = 0$$

× Soit $n \in \mathbb{N}^*$.

$$\begin{aligned} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon\right]\right) &= \mathbb{P}\left(\left[\frac{V_n}{\sigma^2} < (1 - \varepsilon)^2\right]\right) && \text{(par stricte croissance de la} \\ & && \text{fonction } x \mapsto x^2 \text{ sur } \mathbb{R}_+) \\ &= \mathbb{P}([V_n < \sigma^2(1 - 2\varepsilon + \varepsilon^2)]) && \text{(car } \sigma^2 > 0) \\ &= \mathbb{P}([V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)]) \end{aligned}$$

× Ensuite :

$$[|V_n - \sigma^2| > \sigma^2(2\varepsilon - \varepsilon^2)] = [V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)] \cup [V_n - \sigma^2 > \sigma^2(2\varepsilon - \varepsilon^2)]$$

Ainsi :

$$[V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)] \subset [|V_n - \sigma^2| > \sigma^2(2\varepsilon - \varepsilon^2)]$$

Donc :

$$\mathbb{P}([V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)]) \leq \mathbb{P}(|V_n - \sigma^2| > \sigma^2(2\varepsilon - \varepsilon^2))$$

Commentaire

× On en déduit :

$$0 \leq \mathbb{P}([V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)]) \leq \mathbb{P}(|V_n - \sigma^2| > \sigma^2(2\varepsilon - \varepsilon^2))$$

Or, comme $\varepsilon \in]0, 1[$, alors : $\sigma^2(2\varepsilon - \varepsilon^2) > 0$. On obtient alors d'après **8.e** :

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|V_n - \sigma^2| > \sigma^2(2\varepsilon - \varepsilon^2)) = 0$$

Par théorème d'encadrement, on obtient :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([V_n - \sigma^2 < -\sigma^2(2\varepsilon - \varepsilon^2)]) = 0$$

On en conclut : $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon\right]\right) = 0$.

(iv) Démontrer : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) = \Phi((1 + \varepsilon)x)$.

Démonstration.

Comme $(1 + \varepsilon)x \in \mathbb{R}$, d'après la question **2.b** :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) = \Phi((1 + \varepsilon)x)$$

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) = \Phi((1 + \varepsilon)x)$$

Commentaire

Remarquons que, dans le cas $x < 0$, toujours d'après la question **2.b** :

$$\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 - \varepsilon)x]) = \Phi((1 - \varepsilon)x)$$

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

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

Démonstration.

• Soit $n \in \mathbb{N}^*$. D'après **9.a)(i)** :

$$\mathbb{P}([W_n \leq x]) \leq \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) + \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right)$$

• De plus, d'après **9.a)(ii)** : $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) = 0$.

Pour tout $\varepsilon' > 0$, il existe donc $n_1 \in \mathbb{N}$ tel que, pour tout $n \geq n_1$:

$$\left| \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) - 0 \right| \leq \varepsilon'$$

En particulier :

$$\mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) \leq \varepsilon'$$

- D'après **9.a)(iii)** : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) = \Phi((1 + \varepsilon)x)$.

Pour tout $\varepsilon' > 0$, il existe donc $n_2 \in \mathbb{N}$ tel que, pour tout $n \geq n_2$:

$$|\mathbb{P}([Z_n \leq (1 + \varepsilon)x]) - \Phi((1 + \varepsilon)x)| \leq \varepsilon'$$

En particulier :

$$\mathbb{P}([Z_n \leq (1 + \varepsilon)x]) - \Phi((1 + \varepsilon)x) \leq \varepsilon'$$

D'où :

$$\mathbb{P}([Z_n \leq (1 + \varepsilon)x]) \leq \Phi((1 + \varepsilon)x) + \varepsilon'$$

- Soit $\varepsilon' > 0$. On pose $n_{\varepsilon'} = \max(n_1, n_2)$. Alors, pour tout $n \geq n_{\varepsilon'}$, par transitivité :

$$\begin{aligned} \mathbb{P}([W_n \leq x]) &\leq \mathbb{P}([Z_n \leq (1 + \varepsilon)x]) + \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) \\ &\leq \left(\Phi((1 + \varepsilon)x) + \varepsilon'\right) + \varepsilon' \end{aligned}$$

- Ceci étant valable pour tout $\varepsilon' > 0$, pour obtenir l'inégalité de l'énoncé, on cherche à choisir ce réel tel que : $2\varepsilon' = \varepsilon$.

On choisit donc : $\varepsilon' = \frac{\varepsilon}{2} > 0$.

On conclut alors qu'il existe $n_\varepsilon \in \mathbb{N}$ tel que : $\forall n \geq n_\varepsilon, \mathbb{P}([W_n \leq x]) \leq \Phi((1 + \varepsilon)x) + \varepsilon$.

Commentaire

On peut une nouvelle fois proposer une inégalité similaire dans le cas $x < 0$.

Rappelons qu'on a démontré dans la remarque de la question **9.a)(i)** :

$$\mathbb{P}([W_n \leq x]) \leq \mathbb{P}([Z_n \leq (1 - \varepsilon)x]) + \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon\right]\right)$$

Or :

- d'après la remarque de **9.a)(ii)** : $\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} < 1 - \varepsilon\right]\right) = 0$,
- d'après la remarque de la question **9.a)(iii)** : $\lim_{n \rightarrow +\infty} \mathbb{P}([Z_n \leq (1 - \varepsilon)x]) = \Phi((1 - \varepsilon)x)$

Avec un raisonnement similaire à celui de cette question, il existe alors $n'_\varepsilon \in \mathbb{N}$ tel que :

$$\forall n \geq n'_\varepsilon, \quad \mathbb{P}([W_n \leq x]) \leq \Phi((1 - \varepsilon)x) + \varepsilon$$

- b) On admettra que, de manière symétrique, pour tout $\varepsilon > 0$, il existe n_ε 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 : $\lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$.

Commentaire

- L'inégalité admise par l'énoncé est directement obtenue avec le raisonnement de la question précédente. En effet, pour tout $\varepsilon' > 0$, on sait :

× qu'il existe $n_1 \in \mathbb{N}$ tel que, pour tout $n \geq n_1$:

$$\left| \mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) - 0 \right| \leq \varepsilon'$$

En particulier :

$$\mathbb{P}\left(\left[\frac{\sqrt{V_n}}{\sigma} > 1 + \varepsilon\right]\right) \geq -\varepsilon'$$

Commentaire

× qu'il existe $n_2 \in \mathbb{N}$ tel que, pour tout $n \geq n_2$:

$$|\mathbb{P}([Z_n \leq (1 + \varepsilon)x]) - \Phi((1 + \varepsilon)x)| \leq \varepsilon'$$

En particulier :

$$\mathbb{P}([Z_n \leq (1 + \varepsilon)x]) \geq \Phi((1 + \varepsilon)x) - \varepsilon'$$

En utilisant la question **9.a)(i)** et en choisissant toujours $\varepsilon' = \frac{\varepsilon}{2}$, on obtient bien qu'il existe $n_\varepsilon = \max(n_1, n_2) \in \mathbb{N}$ tel que :

$$\forall n \geq n_\varepsilon \quad \mathbb{P}([W_n \leq x]) \geq \Phi((1 + \varepsilon)x) - \varepsilon$$

De même, si $x < 0$, on obtient qu'il existe $n'_\varepsilon \in \mathbb{N}$ tel que :

$$\forall n \geq n'_\varepsilon \quad \mathbb{P}([W_n \leq x]) \geq \Phi((1 - \varepsilon)x) - \varepsilon$$

- La formulation de cette question est un peu malheureuse. En écrivant « pour tout $\varepsilon > 0$, il existe n_ε tel que... », la quantification universelle en ε sous-entend que le rang n_ε introduit dans cette question **9.b)** est éventuellement différent de celui considéré en question précédente (ce qui n'est pas le cas d'après la démonstration que l'on vient d'effectuer). L'énoncé choisit cependant de conserver la même notation pour le rang introduit en question **9.a)**. Rédiger rigoureusement cette question ne demandait pourtant pas beaucoup de modifications :

- × soit l'énoncé considère que le rang à introduire en question **9.b)** est différent de celui de **9.a)**. On aurait pu alors écrire : « On admettra que, de manière symétrique, pour tout $\varepsilon > 0$, il existe N_ε tel que, pour tout $n \geq N_\varepsilon$... ».
- × soit l'énoncé admet que ce sont les mêmes rangs. On écrirait alors : « On admettra que, de manière symétrique, pour tout $n \geq n_{\varepsilon, \dots}$ ».

Démonstration.

Soit $\varepsilon > 0$.

- D'après la question précédente et la propriété admise dans cette question, il existe $n_\varepsilon \in \mathbb{N}$ tel que pour tout $n \geq n_\varepsilon$:

$$\Phi((1 + \varepsilon)x) - \varepsilon \leq \mathbb{P}([W_n \leq x]) \leq \Phi((1 + \varepsilon)x) + \varepsilon$$

- Par ailleurs, on souhaite démontrer : $\lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$. Autrement dit, pour tout $\varepsilon' > 0$, il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$:

$$|\mathbb{P}([W_n \leq x]) - \Phi(x)| \leq \varepsilon'$$

C'est-à-dire :

$$-\varepsilon' \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq \varepsilon'$$

- Soit $n \geq n_\varepsilon$.

$$\Phi((1 + \varepsilon)x) - \Phi(x) - \varepsilon \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq \Phi((1 + \varepsilon)x) - \Phi(x) + \varepsilon \quad (\star)$$

On cherche donc à encadrer $\Phi((1 + \varepsilon)x) - \Phi(x)$. Pour cela, nous allons appliquer l'inégalité des accroissements finis à la fonction Φ .

- La fonction Φ est :
 - × dérivable sur \mathbb{R} (il a été démontré en **1.b**) que la fonction Φ est de classe \mathcal{C}^1 sur \mathbb{R}).
 - × telle que : $\forall t \in \mathbb{R}, 0 \leq \Phi'(t) \leq \frac{1}{\sqrt{2\pi}}$. En effet, pour tout $t \in \mathbb{R}$:

$$\Phi'(t) = f_Z(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$$\text{Or : } 0 \leq e^{-\frac{t^2}{2}} \leq 1. \text{ D'où : } 0 \leq \Phi'(t) \leq \frac{1}{\sqrt{2\pi}}.$$

Par inégalité des accroissements finis :

$$\forall (y, z) \in \mathbb{R}^2, \quad 0 \times (z - y) \leq \Phi(z) - \Phi(y) \leq \frac{1}{\sqrt{2\pi}} \times (z - y)$$

En particulier, pour $z = (1 + \varepsilon)x$ et $y = x$, on obtient :

$$0 \leq \Phi((1 + \varepsilon)x) - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} ((1 + \varepsilon)x - x)$$

D'où :

$$0 \leq \Phi((1 + \varepsilon)x) - \Phi(x) \leq \frac{x}{\sqrt{2\pi}} \varepsilon$$

- On en déduit, par transitivité (cf (\star)) :

$$-\varepsilon \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq \frac{x}{\sqrt{2\pi}} \varepsilon + \varepsilon = \left(1 + \frac{x}{\sqrt{2\pi}}\right) \varepsilon$$

En notant $c = 1 + \frac{x}{\sqrt{2\pi}}$, on obtient finalement :

$$-c\varepsilon \leq -\varepsilon \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq c\varepsilon$$

On a donc démontré :

$$\forall \varepsilon > 0, \exists n_\varepsilon \in \mathbb{N}, \forall n \geq n_\varepsilon, \quad |\mathbb{P}([W_n \leq x])| \leq c\varepsilon$$

Ainsi :

$$\forall \varepsilon' > 0, \exists n_{\varepsilon'} \in \mathbb{N}, \forall n \geq n_{\varepsilon'}, \quad |\mathbb{P}([W_n \leq x])| \leq \varepsilon'$$

Par définition de la limite, on en déduit : $\lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$.

Commentaire

- Détaillons l'obtention de l'assertion :

$$\forall \varepsilon' > 0, \exists n_{\varepsilon'} \in \mathbb{N}, \forall n \geq n_{\varepsilon'}, \quad |\mathbb{P}([W_n \leq x])| \leq \varepsilon'$$

Soit $\varepsilon' > 0$. On a démontré :

$$\forall \varepsilon > 0, \exists n_\varepsilon \in \mathbb{N}, \forall n \geq n_\varepsilon, \quad |\mathbb{P}([W_n \leq x])| \leq c\varepsilon$$

Cette propriété étant vérifiée pour tout $\varepsilon > 0$, elle l'est en particulier pour : $\varepsilon = \frac{\varepsilon'}{c}$.

On en déduit qu'il existe $n_{\varepsilon'}$ tel que, pour tout $n \geq n_{\varepsilon'}$:

$$|\mathbb{P}([W_n \leq x])| \leq c \frac{\varepsilon'}{c} = \varepsilon'$$

Commentaire

- Cette limite est aussi valide pour $x < 0$. En effet, en reprenant l'inégalité des accroissements finis pour $z = (1 - \varepsilon)x$ et $y = x$, on obtient :

$$0 \leq \Phi((1 - \varepsilon)x) - \Phi(x) \leq -\frac{x}{\sqrt{2\pi}} \varepsilon$$

D'où :

$$-\varepsilon \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq -\frac{x}{\sqrt{2\pi}} \varepsilon + \varepsilon = \left(1 - \frac{x}{\sqrt{2\pi}}\right) \varepsilon$$

En notant $c = 1$, on obtient :

$$-c\varepsilon \leq \mathbb{P}([W_n \leq x]) - \Phi(x) \leq c\varepsilon$$

Cela permet une nouvelle fois de conclure : $\lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$.

Notons que l'on a démontré : $\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$. D'où :

$$W_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z$$

- Revenons un instant sur les questions 7. à 9.
 - × En question 7., on obtient un intervalle contenant p avec probabilité 95% à l'aide de la convergence en loi de Z_n vers Z . Cependant, la v.a.r. $Z_n = \frac{\sqrt{n}}{\sqrt{\sigma^2}} (\bar{X}_n - p)$ s'exprime à l'aide du réel σ^2 qui dépend de p (en effet : $\sigma^2 = p(1-p)$). L'intervalle obtenu n'est donc pas un intervalle de confiance de p .
 - × L'idée derrière les questions 8. et 9. est de remplacer σ^2 par un estimateur convergent de σ^2 . Il s'agit de l'estimateur V_n . Cet estimateur de σ^2 est bien convergent d'après 8.e).
On obtient alors une nouvelle suite de v.a.r. (W_n) dont on cherche à montrer qu'elle possède la même propriété que la suite de v.a.r. (Z_n) : $W_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z$ (c'est l'objet de la question 9.b)). La question 10 va alors permettre d'en déduire un intervalle de confiance de p .

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

Démonstration.

Soit $x \in \mathbb{R}$.

- Soit $n \in \mathbb{N}^*$. Par définition de W_n :

$$\begin{aligned} \mathbb{P}([W_n \leq x]) &= \mathbb{P}\left(\left[\frac{\sqrt{n}}{\sqrt{V_n}} (\bar{X}_n - p) \leq x\right]\right) \\ &= \mathbb{P}\left(\left[\bar{X}_n - p \leq x \frac{\sqrt{V_n}}{\sqrt{n}}\right]\right) \quad (\text{car } \frac{\sqrt{V_n}}{\sqrt{n}} \text{ est une v.a.r.} \\ &\quad \text{strictement positive)} \\ &= \mathbb{P}\left(\left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}} \leq p\right]\right) \end{aligned}$$

- Or, d'après **9.b**) : $\lim_{n \rightarrow +\infty} \mathbb{P}([W_n \leq x]) = \Phi(x)$.

$$\text{On en déduit : } \forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}} \leq p\right]\right) = \Phi(x).$$

□

- 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 pour 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.

Commentaire

- Lors d'un sondage, le mathématicien se place à l'instant précédant les interrogations des individus. Ces interrogations sont alors considérées comme des expériences aléatoires.

× On définit alors naturellement les **variables aléatoires** X_1, \dots, X_n, \dots (on devrait parler plus rigoureusement des v.a.r. de la suite $(X_i)_{i \in \mathbb{N}^*}$) introduites dans l'énoncé qui prennent la valeur 1 si le $i^{\text{ème}}$ individu interrogé vote pour le candidat A et prend la valeur 0 sinon.

× Les résultats obtenus après le sondage, définissant une suite $(x_i)_{i \in \mathbb{N}^*}$, sont alors les réalisations des variables aléatoires de la suite $(X_i)_{i \in \mathbb{N}^*}$. Les **réels** de la suite $(x_i)_{i \in \mathbb{N}^*}$ sont les données statistiques du problème.

On prendra garde à bien utiliser :

- × des majuscules lorsqu'on souhaite parler de variables aléatoires,
- × des minuscules lorsqu'on souhaite parler de leurs réalisations.

Ainsi, il aurait sans doute été préférable d'opter dans cette question pour une formulation similaire à celle de la question **11.c**) : « lors d'un sondage réalisé auprès de $n = 1000$ personnes, la variable aléatoire \bar{X}_n a pris la valeur $\bar{x}_n = 0,52$ et V_n la valeur $v_n = 0,2506$ ».

- On a démontré en question précédente :

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}} \leq p\right]\right) = \Phi(x)$$

Autrement dit :

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

Ou encore, puisque $p \in]0, 1[$:

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[p \in \left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}, 1\right]\right]\right) = \Phi(x)$$

Commentaire

- S'il y a effectivement une probabilité $\Phi(x)$ pour que p appartienne à l'intervalle aléatoire construit sur les X_i , cette probabilité est de 0 ou 1 concernant l'intervalle réalisé $\left[\bar{x}_n - x \frac{\sqrt{v_n}}{\sqrt{n}}, 1 \right]$. En effet, p appartient ou pas à un tel intervalle déterministe, il n'y a pas d'autre issue. On pourrait déterminer ce fait en interrogeant tous les membres du panel (mais cela prendrait du temps). En pratique, on ne le fait pas et on espère avoir eu de la chance lors de la réalisation de l'intervalle de confiance.
Pour éclairer cette distinction, pensez à un tirage du Loto : étant donnée une grille, avant le tirage, il lui est associé une (faible) probabilité de gain du gros lot ; après le tirage, la grille est soit gagnante, soit perdante. Évidemment, à parier, on parierait plutôt sur le fait de ne pas avoir eu le gros lot.
- La question posée par l'énoncé n'a donc ici pas de sens. Si l'on se place après que le sondage a été effectué, on doit considérer l'intervalle réalisé $\left[\bar{x}_n - x \frac{\sqrt{v_n}}{\sqrt{n}}, 1 \right]$ et non l'intervalle de confiance obtenu en question précédente. L'assertion « $\frac{1}{2} \leq \bar{x}_n - x \frac{\sqrt{v_n}}{\sqrt{n}}$ » est vraie (donc de probabilité 1) ou fausse (donc de probabilité 0). On ne peut dire qu'elle est de probabilité $\Phi\left(\frac{\sqrt{1000}}{\sqrt{0,2506}}\left(0,52 - \frac{1}{2}\right)\right)$.

Démonstration.

- D'après la question précédente :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}} \leq p\right]\right) = \Phi(x)$$

Or on remarque, pour tout $x \in \mathbb{R}$, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} \mathbb{P}\left(\left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}} \leq p\right]\right) &= \mathbb{P}\left(\left[p \in \left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}, +\infty\right]\right]\right) \\ &= \mathbb{P}\left(\left[p \in \left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}, 1\right]\right]\right) \quad (\text{car } p \in]0, 1[) \end{aligned}$$

On obtient :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left[p \in \left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}, 1\right]\right]\right) = \Phi(x)$$

On en déduit :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}\left(\left[p \in \left[\bar{X}_{1000} - x \frac{\sqrt{V_{1000}}}{\sqrt{1000}}, 1\right]\right]\right) \approx \Phi(x)$$

Si on note $x_0 = \frac{\sqrt{1000}}{\sqrt{0,2506}}\left(0,52 - \frac{1}{2}\right)$, alors :

$$\mathbb{P}\left(\left[p \in \left[\bar{X}_{1000} - x_0 \frac{\sqrt{V_{1000}}}{\sqrt{1000}}, 1\right]\right]\right) \approx \Phi(x_0)$$

$$\mathbb{P}\left(\left[p \in \left[\bar{X}_{1000} - \frac{\sqrt{1000}}{\sqrt{0,2506}}\left(0,52 - \frac{1}{2}\right) \frac{\sqrt{V_{1000}}}{\sqrt{1000}}, 1\right]\right]\right) \approx \Phi\left(\frac{\sqrt{1000}}{\sqrt{0,2506}}\left(0,52 - \frac{1}{2}\right)\right)$$

- Le réel x_0 est choisi afin que l'intervalle réalisé $\left[\bar{x}_n - x_0 \frac{\sqrt{v_n}}{\sqrt{n}}, 1 \right]$ soit égal à $\left[\frac{1}{2}, 1 \right]$. Or :

$$\begin{aligned} \left[\bar{x}_n - x_0 \frac{\sqrt{v_n}}{\sqrt{n}}, 1 \right] = \left[\frac{1}{2}, 1 \right] &\Leftrightarrow \frac{1}{2} = \bar{x}_n - x_0 \frac{\sqrt{v_n}}{\sqrt{n}} \\ &\Leftrightarrow x_0 \frac{\sqrt{v_n}}{\sqrt{n}} = \bar{x}_n - \frac{1}{2} \\ &\Leftrightarrow x_0 = \frac{\sqrt{n}}{\sqrt{v_n}} \left(\bar{x}_n - \frac{1}{2} \right) \end{aligned}$$

Ainsi, en choisissant $x_0 = \frac{\sqrt{1000}}{\sqrt{0,2506}} \left(0,52 - \frac{1}{2} \right)$, la réalisation de l'intervalle de confiance trouvé en question précédente est $\left[\frac{1}{2}, 1 \right]$.

Commentaire

Comme expliqué en remarque précédente, cette question **10.b)** n'a pas de sens. La démarche usuelle en statistiques est en fait la suivante :

1) on établit le résultat :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \in \left[\bar{X}_n - x \frac{\sqrt{V_n}}{\sqrt{n}}, 1 \right] \right] \right) = \Phi(x)$$

2) on choisit ensuite x de sorte que $\Phi(x) = 95\%$. Notons u ce réel (remarquons que ce réel existe et est unique d'après la question **1.c)**). On obtient alors :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \in \left[\bar{X}_n - u \frac{\sqrt{V_n}}{\sqrt{n}}, 1 \right] \right] \right) = 95\%$$

L'intervalle $\left[\bar{X}_n - u \frac{\sqrt{V_n}}{\sqrt{n}}, 1 \right]$ est donc un intervalle de confiance asymptotique de p au niveau de confiance 95%.

3) Une fois le sondage effectué, on obtient une réalisation de l'intervalle de confiance précédent : $\left[\bar{x}_n - u \frac{\sqrt{v_n}}{\sqrt{n}}, 1 \right]$. Si cet intervalle n'est pas inclus dans $\left[\frac{1}{2}, 1 \right]$, alors on rejette l'hypothèse que le candidat A remporte l'élection. □

11. Lors des dernières élections, on s'est rendu compte que les électeurs pouvaient mentir lors du sondage (ou bien 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 $i^{\text{ème}}$ personne change d'avis en faveur de B et $T_i = 0$ sinon. On suppose que les variables aléatoires 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 .

Démonstration.

Soit $i \in \mathbb{N}^*$.

- Tout d'abord : $Y_i(\Omega) \subset \{0, 1\}$.

En effet, pour tout $\omega \in \Omega$, deux cas se présentent :

× si $X_i(\omega) = 1$, alors :

$$Y_i(\omega) = X_i(\omega) + (1 - X_i(\omega))T_i(\omega) = 1 + (1 - 1)T_i(\omega) = 1$$

× si $X_i(\omega) = 0$, alors :

$$Y_i(\omega) = X_i(\omega) + (1 - X_i(\omega)) T_i(\omega) = 0 + (1 - 0) T_i(\omega) = T_i(\omega)$$

Or : $T_i(\Omega) = \{0, 1\}$. D'où : $Y_i(\omega) \in \{0, 1\}$.

$$Y_i(\Omega) \subset \{0, 1\}$$

- Déterminons $\mathbb{P}([Y_i = 1])$.

La famille $([X_i = 1], [X_i = 0])$ forme un système complet d'événements.

Ainsi, d'après la formule des probabilités totales :

$$\mathbb{P}([Y_i = 1]) = \mathbb{P}([X_i = 1] \cap [Y_i = 1]) + \mathbb{P}([X_i = 0] \cap [Y_i = 1])$$

× D'une part :

$$\begin{aligned} \mathbb{P}([X_i = 1] \cap [Y_i = 1]) &= \mathbb{P}([X_i = 1] \cap [X_i + (1 - X_i) T_i = 1]) \quad (\text{par définition de } Y_i) \\ &= \mathbb{P}([X_i = 1] \cap [1 + (1 - 1) T_i = 1]) \\ &= \mathbb{P}([X_i = 1] \cap [1 = 1]) \\ &= \mathbb{P}([X_i = 1] \cap \Omega) \\ &= \mathbb{P}([X_i = 1]) \\ &= p \quad (\text{car } X_i \hookrightarrow \mathcal{B}(p)) \end{aligned}$$

× D'autre part :

$$\begin{aligned} \mathbb{P}([X_i = 0] \cap [Y_i = 1]) &= \mathbb{P}([X_i = 0] \cap [X_i + (1 - X_i) T_i = 1]) \quad (\text{par définition de } Y_i) \\ &= \mathbb{P}([X_i = 0] \cap [0 + (1 - 0) T_i = 1]) \\ &= \mathbb{P}([X_i = 0] \cap [T_i = 1]) \\ &= \mathbb{P}([X_i = 0]) \times \mathbb{P}([T_i = 1]) \quad (\text{car les v.a.r. } X_i \text{ et } T_i \\ &\hspace{15em} \text{sont indépendantes}) \\ &= (1 - p) \times q \quad (\text{car } X_i \hookrightarrow \mathcal{B}(p) \text{ et } \\ &\hspace{15em} T_i \hookrightarrow \mathcal{B}(q)) \end{aligned}$$

Finalement :

$$\mathbb{P}([Y_i = 1]) = p + (1 - p)q = p + q - pq = p(1 - q) + q$$

$$\text{On en déduit : } Y_i \hookrightarrow \mathcal{B}(r), \text{ où } r = p(1 - q) + q. \quad \square$$

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 $x \in \mathbb{R}$:

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

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

Démonstration.

Soit $x \in \mathbb{R}$.

- Soit $n \in \mathbb{N}^*$.

$$\begin{aligned} \mathbb{P} \left(\left[\frac{\sqrt{n}}{\sqrt{U_n}} (\bar{Y}_n - r) \leq x \right] \right) &= \mathbb{P} \left(\left[\bar{Y}_n - r \leq x \frac{\sqrt{U_n}}{\sqrt{n}} \right] \right) && \text{(car la v.a.r. } \frac{\sqrt{U_n}}{\sqrt{n}} \text{ est} \\ & && \text{strictement positive)} \\ &= \mathbb{P} \left(\left[\bar{Y}_n - (p(1-q) + q) \leq x \frac{\sqrt{U_n}}{\sqrt{n}} \right] \right) && \text{(d'après la question} \\ & && \text{précédente)} \\ &= \mathbb{P} \left(\left[\bar{Y}_n - q - x \frac{\sqrt{U_n}}{\sqrt{n}} \leq p(1-q) \right] \right) \\ &= \mathbb{P} \left(\left[\frac{1}{1-q} \left(\bar{Y}_n - q - x \frac{\sqrt{U_n}}{\sqrt{n}} \right) \leq p \right] \right) && \text{(car } 1-q > 0) \end{aligned}$$

- Or, d'après la proposition admise juste avant cette question **11.b**), on a :

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

On en déduit : $\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \geq \frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right) \right] \right) = \Phi(x)$.

□

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 $\Phi \left(\frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right) \right)$.

Commentaire

Cette question n'a une nouvelle fois pas de sens, pour les mêmes raisons que la question **10.b**). On fournit, comme en **10.b**), des éléments qui auraient pu convenir à une formulation correcte de la question.

Démonstration.

- D'après la question précédente :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \geq \frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right) \right] \right) = \Phi(x)$$

Or on remarque, pour tout $x \in \mathbb{R}$, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} \mathbb{P} \left(\left[p \geq \frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right) \right] \right) &= \mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), +\infty \right] \right] \right) \\ &= \mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) \end{aligned}$$

où la dernière égalité est obtenue car $p \in]0, 1[$.

On obtient :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) = \Phi(x)$$

On en déduit, pour n suffisamment grand :

$$\forall x \in \mathbb{R}, \quad \mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - x \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) \approx \Phi(x)$$

Si on note $y_0 = \frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right)$, alors :

$$\mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - y_0 \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) \approx \Phi(y_0)$$

$$\mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - y_0 \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) \approx \Phi \left(\frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right) \right)$$

- Le réel y_0 est choisi afin que l'intervalle réalisé $\left[\frac{1}{1-q} \left(\bar{y}_n - y_0 \frac{\sqrt{u_n}}{\sqrt{n}} - q \right), 1 \right]$ soit égal à $\left[\frac{1}{2}, 1 \right]$. Or :

$$\begin{aligned} \left[\frac{1}{1-q} \left(\bar{y}_n - y_0 \frac{\sqrt{u_n}}{\sqrt{n}} - q \right), 1 \right] &= \left[\frac{1}{2}, 1 \right] &\Leftrightarrow \frac{1}{2} &= \frac{1}{1-q} \left(\bar{y}_n - y_0 \frac{\sqrt{u_n}}{\sqrt{n}} - q \right) \\ &&\Leftrightarrow \frac{1}{2} (1-q) &= \bar{y}_n - y_0 \frac{\sqrt{u_n}}{\sqrt{n}} - q \\ &&\Leftrightarrow y_0 \frac{\sqrt{u_n}}{\sqrt{n}} &= \bar{y}_n - \frac{1}{2} (1-q) - q \\ &&\Leftrightarrow y_0 \frac{\sqrt{u_n}}{\sqrt{n}} &= \bar{y}_n - \frac{1}{2} (1+q) \\ &&\Leftrightarrow y_0 &= \frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right) \end{aligned}$$

Ainsi, en choisissant $y_0 = \frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right)$, la réalisation de l'intervalle de confiance trouvé en question précédente est $\left[\frac{1}{2}, 1 \right]$. \square

- 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 $\frac{1}{2}$.

Commentaire

Cette question n'a une nouvelle fois pas de sens, pour les mêmes raisons que les questions **10.b)** et **11.c)**. On fournit ici des éléments qui auraient pu convenir à une formulation correcte de la question.

Démonstration.

- D'après la question précédente :

$$\begin{aligned}
 y_0 &= \frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right) \\
 &= \frac{\sqrt{n}}{\sqrt{u_n}} \left(0,52 - \frac{1}{2} (1+0,04) \right) \quad (\text{d'après les données de l'énoncé}) \\
 &= \frac{\sqrt{n}}{\sqrt{u_n}} \left(0,52 - \frac{1}{2} \times 1,04 \right) \\
 &= \frac{\sqrt{n}}{\sqrt{u_n}} (0,52 - 0,52) \\
 &= 0
 \end{aligned}$$

- Or, d'après la question précédente :

$$\mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - y_0 \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) \approx \Phi \left(\frac{\sqrt{n}}{\sqrt{u_n}} \left(\bar{y}_n - \frac{1}{2} (1+q) \right) \right) \approx \Phi(y_0)$$

Or : $\Phi(y_0) = \Phi(0) = \frac{1}{2}$.

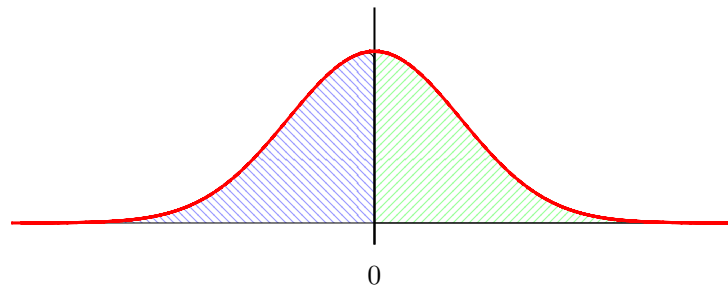
On en déduit : $\mathbb{P} \left(\left[p \in \left[\frac{1}{1-q} \left(\bar{Y}_n - y_0 \frac{\sqrt{U_n}}{\sqrt{n}} - q \right), 1 \right] \right] \right) = \frac{1}{2}$.

Commentaire

- La valeur de $\Phi(0)$ provient de la parité de f_Z et peut se lire graphiquement. En effet, comme Φ est la fonction de répartition de la v.a.r. Z de densité f_Z , on a, pour tout $x \in \mathbb{R}$:

$$\Phi(x) = \int_{-\infty}^x f_Z(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Ainsi, $\Phi(x)$ est l'aire sous la courbe représentative de f_Z entre les abscisses $-\infty$ et x .



= $\Phi(0)$ et = $1 - \Phi(0)$

Par parité de f_Z , on lit bien : $\Phi(0) = \frac{1}{2}$.

- Dans cet énoncé, on démontre en question **1.d**) : $\forall x \in \mathbb{R}, \quad \Phi(-x) = 1 - \Phi(x)$. Cette relation permet d'obtenir rapidement la valeur de $\Phi(0)$. En effet : $\Phi(0) = 1 - \Phi(0)$.

Donc : $2\Phi(0) = 1$. D'où : $\Phi(0) = \frac{1}{2}$.

□

Troisième partie : Démonstration du théorème limite central

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

On supposera aussi pour simplifier :

$$\mu = \mathbb{E}(X_i) = 0 \quad \text{et} \quad \sigma^2 = \mathbb{V}(X_i) = 1$$

de sorte que $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$.

Soit x un réel fixé dans 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 dérivables et de dérivée $k^{\text{ème}}$ 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 : $\int_0^1 u^3 (1-u)^3 du = \frac{1}{20} \int_0^1 u^6 du$.

Démonstration.

- On procède par intégrations par parties (IPP).

$$\left| \begin{array}{ll} f(u) = (1-u)^3 & f'(u) = 3(1-u)^2(-1) \\ g'(u) = u^3 & g(u) = \frac{1}{4} u^4 \end{array} \right.$$

Cette IPP est valide car les fonctions f et g sont de classe \mathcal{C}^1 sur le SEGMENT $[0, 1]$.

On obtient :

$$\begin{aligned} \int_0^1 u^3 (1-u)^3 du &= \left[\frac{1}{4} u^4 (1-u)^3 \right]_0^1 - \int_0^1 3 \frac{1}{4} u^4 (1-u)^2 (-1) du \\ &= \frac{1}{4} [u^4 (1-u)^3]_0^1 + \frac{3}{4} \int_0^1 u^4 (1-u)^2 du \\ &= \frac{1}{4} (\cancel{1^4 (1-1)^3} - \cancel{0^4 (1-0)^3}) + \frac{3}{4} \int_0^1 u^4 (1-u)^2 du \end{aligned}$$

$$\boxed{\int_0^1 u^3 (1-u)^3 du = \frac{3}{4} \int_0^1 u^4 (1-u)^2 du}$$

- On réalise alors une nouvelle IPP en suivant la même idée.

En posant : $f(u) = (1-u)^2$ et $g'(u) = u^4$, on obtient de même :

$$\begin{aligned} \int_0^1 u^4 (1-u)^2 du &= \left[\frac{1}{5} u^5 (1-u)^2 \right]_0^1 - \int_0^1 \frac{1}{5} u^5 \times 2(1-u)(-1) du \\ &= \frac{1}{5} [\cancel{u^5 (1-u)^2}]_0^1 + \frac{2}{5} \int_0^1 u^5 (1-u) du \end{aligned}$$

$$\boxed{\int_0^1 u^4 (1-u)^2 du = \frac{2}{5} \int_0^1 u^5 (1-u) du}$$

- On réalise alors une nouvelle IPP en posant : $f(u) = (1-u)$ et $g'(u) = u^5$.

$$\boxed{\text{On obtient alors : } \int_0^1 u^5 (1-u) du = \frac{1}{6} \int_0^1 u^6 du.}$$

- Finalement :

$$\begin{aligned} \int_0^1 u^3 (1-u)^3 du &= \frac{3}{4} \int_0^1 u^4 (1-u)^2 du \\ &= \frac{3}{4} \times \frac{2}{5} \int_0^1 u^5 (1-u) du \\ &= \frac{\cancel{3}}{4} \times \frac{\cancel{2}}{5} \times \frac{1}{\cancel{6}} \int_0^1 u^6 du = \frac{1}{4 \times 5} \int_0^1 u^6 du \end{aligned}$$

On a bien : $\int_0^1 u^3 (1-u)^3 du = \frac{1}{20} \int_0^1 u^6 du.$

Commentaire

- Cette question exige de faire 3 intégrations par parties successives. S'il convient de présenter de manière très rigoureuse la première, cela suffit pour démontrer la bonne acquisition de la méthode. Ainsi, les 2 IPP suivantes n'ont pas à être détaillées tout autant.
- On peut généraliser ce résultat. En effet, pour tout $(i, j) \in \mathbb{N} \times \mathbb{N}^*$:

$$\int_0^1 u^i (1-u)^j du = \frac{j}{i+1} \int_0^1 u^{i+1} (1-u)^{j-1} du$$

□

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

Démonstration.

$$\begin{aligned} \int_0^1 u^3 (1-u)^3 du &= \frac{1}{20} \int_0^1 u^6 du \quad (\text{d'après la question précédente}) \\ &= \frac{1}{20} \left[\frac{1}{7} u^7 \right]_0^1 \\ &= \frac{1}{20} \frac{1}{7} (1^7 - 0^7) \end{aligned}$$

On a bien : $\int_0^1 u^3 (1-u)^3 du = \frac{1}{140}.$

Commentaire

Ce sujet, comme on l'a dit précédemment, alterne parfaitement le niveau de difficulté des questions. Lorsque les sujets sont aussi bien construits, il est fortement conseillé de ne laisser aucune partie de côté : chacune d'entre elles contient des questions abordables. Ajoutons que le début de chaque nouvelle partie (voire sous-partie) est souvent l'occasion d'une entrée en matière douce (comme c'est le cas ici), et il est donc primordial de les aborder.

□

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} .

Démonstration.

- La fonction $f : u \mapsto u^3 (1-u)^3$ est continue sur $]0, 1[$.

Elle admet donc une primitive F de classe \mathcal{C}^1 sur $]0, 1[$. On a alors, pour tout $z \in]0, 1[$:

$$h(z) = 140 \int_0^z u^3 (1-u)^3 du = 140 [F(u)]_0^z = 140 F(z) - 140 F(0)$$

Ainsi, la fonction h est continue (et même de classe \mathcal{C}^1) sur $]0, 1[$ car F l'est.

Commentaire

- On peut aussi rédiger en se servant du fait que la fonction h est (à multiplication par 140 près), la primitive de f sur $]0, 1[$ qui s'annule au point 0. Ainsi, h est de classe \mathcal{C}^1 sur $]0, 1[$ et : $\forall z \in]0, 1[, h'(z) = 140 z^3 (1-z)^3$.
- L'intérêt de la démonstration précédente est qu'elle est plus générale et peut donc être adaptée à tous les cas particuliers. Imaginons par exemple une fonction h définie par :

$$\forall z \in]0, 1[, \frac{1}{140} h(z) = \int_0^{z^2} f(t) dt = [F(t)]_0^{z^2} = F(z^2) - F(0)$$

La fonction $\frac{1}{140} h$ N'EST PAS une primitive de f .

L'expression ci-dessus permet toutefois de conclure que h est de classe \mathcal{C}^1 sur $]0, 1[$ comme composée de $z \mapsto z^2$ par F toutes les deux de classe \mathcal{C}^1 sur les intervalles adéquats.

De plus :

$$\forall z \in]0, 1[, \frac{1}{140} h'(z) = 2z \times F'(z^2) = 2z \times f(z^2)$$

- Par ailleurs, la fonction h est :
 - × continue sur $] -\infty, 0[$ car elle est constante sur cet intervalle.
 - × continue sur $]1, +\infty[$ car elle est constante sur cet intervalle.

- Enfin :

$$\times \lim_{z \rightarrow 0^-} h(z) = \lim_{z \rightarrow 0} 0 = 0$$

$$\text{et } \lim_{z \rightarrow 0^+} h(z) = \lim_{z \rightarrow 0} 140 \int_0^z u^3 (1-u)^3 du = 140 \int_0^0 u^3 (1-u)^3 du = 0.$$

$$\lim_{z \rightarrow 0^-} h(z) = h(0) = \lim_{z \rightarrow 0^+} h(z)$$

$$\times \lim_{z \rightarrow 1^-} h(z) = \lim_{z \rightarrow 1} 140 \int_0^z u^3 (1-u)^3 du = 140 \int_0^1 u^3 (1-u)^3 du = 140 \frac{1}{140} = 1$$

$$\text{et } \lim_{z \rightarrow 1^+} h(z) = \lim_{z \rightarrow 1} 1 = 1.$$

$$\lim_{z \rightarrow 1^-} h(z) = h(1) = \lim_{z \rightarrow 1^+} h(z)$$

Ainsi, la fonction h est bien continue sur \mathbb{R} .

Commentaire

Rappelons que la continuité (et de manière plus générale la régularité) d'une fonction se démontre sur les intervalles **OUVERTS**. On complète ensuite si nécessaire par une étude des limites aux points restants. □

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

Démonstration.

- D'après la question précédente, la fonction h est de classe \mathcal{C}^1 sur $]0, 1[$.
De plus, pour tout $z \in]0, 1[$:

$$h'(z) = 140 z^3 (1-z)^3 > 0 \quad (\text{car } z > 0 \text{ et } 1-z > 0 \text{ car } z < 1)$$

Ainsi la fonction h est strictement croissante sur $]0, 1[$.

On en déduit que pour tout $(a, b, z) \in]0, 1[^3$ tel que $a \leq z \leq b$:

$$h(a) \leq h(z) \leq h(b)$$

On en déduit :

$$\begin{array}{ccc} \lim_{a \rightarrow 0} h(a) & \leq & \lim_{a \rightarrow 0} h(z) \leq \lim_{a \rightarrow 0} h(b) \\ \parallel & & \parallel \\ h(0) & & h(b) \end{array}$$

et enfin :

$$\begin{array}{ccc} \lim_{b \rightarrow 1} h(0) & \leq & \lim_{b \rightarrow 1} h(z) \leq \lim_{b \rightarrow 1} h(b) \\ \parallel & & \parallel \\ 0 = h(0) & & h(1) = 1 \end{array}$$

- Par définition de h , on sait de plus :
 - × $\forall z \in]-\infty, 0]$, $h(z) = 0 \in [0, 1]$.
 - × $\forall z \in [1, +\infty)$, $h(z) = 1 \in [0, 1]$.

On a bien : $\forall z \in [0, 1], 0 \leq h(z) \leq 1$.

□

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

13. On pose $a_n = n^{-\frac{1}{12}}$ 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 : $0 \leq g_n(z) \leq 1$.

Démonstration.

- La fonction g_n est continue sur \mathbb{R} car elle est la transformée affine $g_n = 1 - h \circ u_n$ où h et $u_n : z \mapsto \frac{1}{a_n}(z-x)$ sont deux fonctions continues sur \mathbb{R} .
- Soit $z \in \mathbb{R}$. Comme $u_n(z) = \frac{1}{a_n}(z-x) \in \mathbb{R}$, alors, d'après la question précédente :

$$\begin{array}{ccc} 0 & \leq & h(u_n(z)) \leq 1 \\ \text{donc} & 0 & \geq -h(u_n(z)) \geq -1 \\ \text{donc} & 1 & \geq 1 - h(u_n(z)) \geq 0 \end{array}$$

Finalement : $\forall z \in \mathbb{R}, g_n(z) = 1 - h(u_n(z)) \in [0, 1]$.

Commentaire

- Dans le sujet, on admet que les fonctions h' , h'' et h''' sont continues sur \mathbb{R} . Cela sous-entend que h est trois fois dérivable sur \mathbb{R} . Mieux, comme de plus h''' est supposée continue sur \mathbb{R} , alors h est de classe \mathcal{C}^3 sur \mathbb{R} . On l'admettra par la suite.
- On peut aussi remarquer que la variable z de la définition de $g_n(z)$ n'est pas quantifiée. Il faut donc présumer que l'on définit ici une fonction g_n en précisant sa valeur en tout point $z \in \mathbb{R}$. Ce manque de quantification est une critique assez générale que l'on peut faire sur le sujet.
- On utilise ici l'argument que la composée de deux fonctions continues sur \mathbb{R} est une fonction continue sur \mathbb{R} . Attention cependant à cet argument, la composée de deux fonctions continues sur un intervalle I n'est pas forcément une fonction continue sur I . Par exemple, la fonction :

$$g : z \mapsto \ln(-z)$$

est la composée $g = g_2 \circ g_1$ où :

- × $g_1 : z \mapsto -z$ est continue sur $]0, +\infty[$,
- × $g_2 : z \mapsto \ln(z)$ est continue sur $]0, +\infty[$.

Cependant, g n'est PAS continue sur $]0, +\infty[$. Le problème provient du fait que la fonction g_1 est à valeurs dans $] - \infty, 0[$. C'est sur ce dernier intervalle que la fonction g_2 devrait être continue. Or, elle n'est même pas définie sur cet intervalle.

- Profitons-en pour rappeler le schéma de démonstration détaillé pour la régularité d'une fonction qui s'écrit comme une composée $g = g_2 \circ g_1$. La fonction g est [type de régularité] sur l'intervalle I car elle est **la composée** $g = g_2 \circ g_1$ où :
 - × g_1 est :
 - [type de régularité] sur I .
 - telle que $g_1(I) \subset J$
 - × g_2 est [type de régularité] sur J .
- Comme on l'a vu au début de la remarque, lorsque $g_1(I) \neq I$, affirmer que « g est [type de régularité] sur I car est la composée de deux fonctions [type de régularité] sur I » est inexact. Si on ne souhaite pas détailler le schéma précédent, il faudra alors a minima affirmer :

« g est [type de régularité] sur I car est la composée de deux fonctions
[type de régularité] sur des intervalles adéquats »

- Il est pertinent de s'interroger sur le niveau de détail qu'il faut apporter aux questions portant sur la régularité de fonctions. Il est conseillé :
 - × de rédiger très proprement la régularité d'une fonction pour les questions que l'on traite en premier. On démontre ainsi au correcteur sa capacité à rédiger ce type de questions.
 - × de rédiger très proprement la régularité lorsqu'il s'agit du cœur de la question (« Démontrer que la fonction est continue / de classe \mathcal{C}^1 sur ... »).

Dans les autres cas, on pourra se contenter d'écrire que la fonction f est continue / dérivable / de classe \mathcal{C}^1 sur l'intervalle J car elle est la somme / le produit / le quotient dont le dénominateur ne s'annule pas sur J , de fonctions continues / dérivables / de classe \mathcal{C}^1 sur l'intervalle J . □

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 : $M_{g'''_n} \leq n^{\frac{1}{4}} M_{h'''}$.

Démonstration.

- Par hypothèse admise de l'énoncé, la fonction g_n est de classe \mathcal{C}^3 sur \mathbb{R} (elle est dérivable trois fois sur \mathbb{R} et g'''_n est continue sur \mathbb{R}).

Soit $z \in \mathbb{R}$. Avec les notations de la question précédente :

$$g_n(z) = 1 - h(u_n(z))$$

$$\begin{aligned} \text{donc } g'_n(z) &= -h'(u_n(z)) \times u'_n(z) \\ &= -h'(u_n(z)) \times \frac{1}{a_n} = -\frac{1}{a_n} h'(u_n(z)) \end{aligned}$$

$$\begin{aligned} \text{puis } g''_n(z) &= -\frac{1}{a_n} h''(u_n(z)) \times u'_n(z) \\ &= -\frac{1}{a_n} h''(u_n(z)) \times \frac{1}{a_n} = -\left(\frac{1}{a_n}\right)^2 h''(u_n(z)) \end{aligned}$$

$$\begin{aligned} \text{enfin } g'''_n(z) &= -\left(\frac{1}{a_n}\right)^2 h'''(u_n(z)) \times u'_n(z) \\ &= -\left(\frac{1}{a_n}\right)^2 h'''(u_n(z)) \times \frac{1}{a_n} = -\left(\frac{1}{a_n}\right)^3 h'''(u_n(z)) \end{aligned}$$

- On en déduit :

$$|g'''_n(z)| = \left| -\left(\frac{1}{a_n}\right)^3 h'''(u_n(z)) \right| = |-1| \times \left| \left(\frac{1}{a_n}\right)^3 \right| \times |h'''(u_n(z))| = \left| \left(\frac{1}{a_n}\right)^3 \right| \times |h'''(u_n(z))|$$

Et :

$$\left| \left(\frac{1}{a_n}\right)^3 \right| = \left| \left(\frac{1}{n^{-\frac{1}{12}}}\right)^3 \right| = \left| \left(n^{\frac{1}{12}}\right)^3 \right| = \left| n^{\frac{1}{4}} \right| = n^{\frac{1}{4}}$$

- Finalement :

$$|g'''_n(z)| = n^{\frac{1}{4}} |h'''(u_n(z))| \leq n^{\frac{1}{4}} M_{h'''} \quad (\text{où } M_{h'''} \text{ est un majorant de } |h'''| \text{ sur } \mathbb{R})$$

- On note alors $M_{g'''_n} = n^{\frac{1}{4}} M_{h'''}$. On a démontré que pour tout $z \in \mathbb{R}$:

$$g_n(z) \leq |g'''_n(z)| \leq M_{g'''_n} = n^{\frac{1}{4}} M_{h'''}$$

On a donc bien démontré l'existence d'un majorant de g'''_n sur \mathbb{R} qui vérifie $M_{g'''_n} \leq n^{\frac{1}{4}} M_{h'''}$.

Commentaire

- Rappelons tout d'abord que si une fonction g possède un majorant M sur \mathbb{R} , alors elle en possède une infinité. En effet, tout réel A plus grand que M est aussi un majorant de g . En effet, avec ces notations, on a, pour tout $z \in \mathbb{R}$:

$$g(z) \leq M \leq A$$

- La formulation de la question n'est pas très rigoureuse. Il faut comprendre que la propriété est vérifiée pour n'importe quel majorant de $|h'''|$. Il y a un « quel que soit » caché dans cette question et il aurait été plus rigoureux de l'expliciter. On peut par ailleurs s'interroger sur la pertinence de la notation $M_{g_n''}$ pour représenter un majorant de g_n'' alors que cette notation est introduite dans l'énoncé pour désigner un majorant de $|g_n'''|$. D'ailleurs, la démonstration consiste à exhiber un majorant de $|g_n'''|$ (qui est donc en particulier un majorant de g_n''). \square

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

Démonstration.

Soit $z \in \mathbb{R}$. Deux cas se présentent.

- Si $z \leq x$ alors : $z - x \leq 0$ et donc $\frac{1}{a_n}(z - x) \leq 0$.

On en déduit :

$$g_n(z) = 1 - h\left(\frac{1}{a_n}(z - x)\right) = 1 - 0 \quad (\text{car pour tout } u \leq 0, h(u) = 0)$$

$$\text{Pour tout } z \leq x, g_n(z) = 1.$$

- Si $z > x + a_n$ alors : $z - x > a_n$ et donc $\frac{1}{a_n}(z - x) > 1$.

On en déduit :

$$g_n(z) = 1 - h\left(\frac{1}{a_n}(z - x)\right) = 1 - 1 \quad (\text{car pour tout } u > 1, h(u) = 1)$$

$$\text{Pour tout } z > x + a_n, g_n(z) = 0. \quad \square$$

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

Commentaire

- La notion de variable indicatrice n'est pas au programme ECE. Cependant, c'est un objet incontournable dans les sujets du TOP3. Il est donc fortement conseillé de l'avoir manipulé au préalable. C'est d'autant plus primordial que les énoncés ne s'apesantissent généralement pas sur cette notion et se contentent d'en donner la définition. Profitons-en pour exposer quelques propriétés usuelles.

× Loi de $\mathbb{1}_A$.

- Par définition de $\mathbb{1}_A$, cette v.a.r. ne prend comme valeur que 0 et 1.

Donc $\mathbb{1}_A(\Omega) = \{0, 1\}$.

- Soit $\omega \in \Omega$.

$$\omega \in [\mathbb{1}_A = 1] \Leftrightarrow \mathbb{1}_A(\omega) = 1 \Leftrightarrow \omega \in A$$

D'où : $[\mathbb{1}_A = 1] = A$. Ainsi : $\mathbb{P}([\mathbb{1}_A = 1]) = \mathbb{P}(A)$.

On en déduit : $\mathbb{1}_A \hookrightarrow \mathcal{B}(\mathbb{P}(A))$.

× En particulier : $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$ et $\mathbb{V}(\mathbb{1}_A) = \mathbb{P}(A)(1 - \mathbb{P}(A))$.

Commentaire

- Il peut aussi être utile de savoir démontrer les propriétés suivantes.
Soient B et C deux événements.

$$\mathbb{1}_{B \cap C} = \mathbb{1}_B \times \mathbb{1}_C$$

et

$$\mathbb{1}_B + \mathbb{1}_{\bar{B}} = 1$$

On peut se référer au sujet **ESSEC-II 2018** pour la démonstration de ces deux propriétés.

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

Démonstration.

Soit X une v.a.r.

- Démontrons tout d'abord l'inégalité de gauche. Il s'agit de démontrer :

$$\forall \omega \in \Omega, \mathbb{1}_{[X \leq x]}(\omega) \leq g_n(X(\omega))$$

Soit $\omega \in \Omega$. Deux cas se présentent.

× Si $X(\omega) \leq x$ alors $\omega \in [X \leq x]$ et donc, par définition : $\mathbb{1}_{[X \leq x]}(\omega) = 1$.

Par ailleurs, d'après la question précédente : $g_n(X(\omega)) = 1$ car $X(\omega) \leq x$.

$$\text{Si } X(\omega) \leq x \text{ alors : } \mathbb{1}_{[X \leq x]}(\omega) \leq g_n(X(\omega)).$$

× Si $X(\omega) > x$ alors $\omega \notin [X \leq x]$ et donc, par définition : $\mathbb{1}_{[X \leq x]}(\omega) = 0$.

Or, d'après la question **13.a(i)**, g_n est à valeurs dans $[0, 1]$. En particulier, on a donc :

$$g_n(X(\omega)) \geq 0$$

$$\text{Si } X(\omega) > x \text{ alors : } \mathbb{1}_{[X \leq x]}(\omega) \leq g_n(X(\omega)).$$

$$\text{Finalement : } \forall \omega \in \Omega, \mathbb{1}_{[X \leq x]}(\omega) \leq g_n(X(\omega)).$$

- On raisonne de même pour l'inégalité de gauche. Il s'agit de démontrer :

$$\forall \omega \in \Omega, g_n(X(\omega)) \leq \mathbb{1}_{[X \leq x + a_n]}(\omega)$$

Soit $\omega \in \Omega$. Deux cas se présentent.

× Si $X(\omega) \leq x + a_n$ alors $\omega \in [X \leq x + a_n]$ et donc, par définition : $\mathbb{1}_{[X \leq x + a_n]}(\omega) = 1$.

Or, d'après la question **13.a(i)**, g_n est à valeurs dans $[0, 1]$. En particulier, on a donc :

$$g_n(X(\omega)) \leq 1$$

$$\text{Si } X(\omega) \leq x + a_n \text{ alors : } g_n(X(\omega)) \leq \mathbb{1}_{[X \leq x + a_n]}(\omega).$$

× Si $X(\omega) > x + a_n$ alors $\omega \notin [X \leq x + a_n]$ et donc, par définition : $\mathbb{1}_{[X \leq x + a_n]}(\omega) = 0$.

Par ailleurs, d'après la question précédente : $g_n(X(\omega)) = 0$ car $X(\omega) > x + a_n$.

$$\text{Si } X(\omega) > x \text{ alors : } g_n(X(\omega)) \leq \mathbb{1}_{[X \leq x + a_n]}(\omega).$$

$$\text{Finalement : } \forall \omega \in \Omega, \mathbb{1}_{[X \leq x]}(\omega) \leq g_n(X(\omega)) \leq \mathbb{1}_{[X \leq x + a_n]}(\omega).$$

Commentaire

Une variable indicatrice étant une fonction (de Ω dans \mathbb{R}) **définie par cas**, il est naturel de procéder par disjonction de cas pour les démonstrations des résultats qui utilisent cet objet. \square

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

Démonstration.

- Tout d'abord, en utilisant la question précédente pour la v.a.r. $X = Z_n$, on obtient :

$$\mathbb{1}_{[Z_n \leq x]} \leq g_n(Z_n) \leq \mathbb{1}_{[Z_n \leq x + a_n]}$$

On en déduit, par croissance de l'espérance, en ne considérant que l'inégalité de gauche :

$$\mathbb{E}(\mathbb{1}_{[Z_n \leq x]}) \leq \mathbb{E}(g_n(Z_n))$$

Or, comme $\mathbb{1}_{[Z_n \leq x]}(\Omega) = \{0, 1\}$, on a :

$$\begin{aligned} \mathbb{E}(\mathbb{1}_{[Z_n \leq x]}) &= 1 \times \mathbb{P}(\mathbb{1}_{[Z_n \leq x]} = 1) + 0 \times \mathbb{P}(\mathbb{1}_{[Z_n \leq x]} = 0) \\ &= \mathbb{P}([Z_n \leq x]) \end{aligned}$$

Ce dernier résultat provient de l'égalité : $[\mathbb{1}_{[Z_n \leq x]} = 1] = [Z_n \leq x]$.

En effet, pour tout $\omega \in \Omega$:

$$\mathbb{1}_{[Z_n \leq x]}(\omega) = 1 \Leftrightarrow Z_n(\omega) \leq x$$

$$\text{Finalement, on a bien : } \mathbb{P}([Z_n \leq x]) \leq \mathbb{E}(g_n(Z_n)).$$

- On procède de même en considérant maintenant la v.a.r. $X = Z_n + a_n$.
D'après la question précédente :

$$\mathbb{1}_{[Z_n + a_n \leq x]} \leq g_n(Z_n + a_n) \leq \mathbb{1}_{[Z_n \leq x + a_n]}$$

On en déduit, par croissance de l'espérance (en ne considérant que l'inégalité de droite) :

$$\begin{aligned} \mathbb{E}(g_n(Z_n + a_n)) &\leq \mathbb{E}(\mathbb{1}_{[Z_n \leq x]}) \\ &= \\ &= \mathbb{P}([Z_n \leq x]) \end{aligned}$$

$$\text{On a bien : } \mathbb{E}(\mathbb{1}_{[Z_n + a_n \leq x]}) \leq \mathbb{P}([Z_n + a_n \leq x]).$$

Commentaire

- L'énoncé original comporte une erreur. Il était en effet demandé de démontrer :

$$\mathbb{E}(g_n(Z_n - a_n)) \leq \mathbb{P}([Z_n \leq x]) \leq \mathbb{E}(g_n(Z_n))$$

Il est globalement rare que les sujets contiennent des erreurs. Malheureusement, malgré la relecture soignée des concepteurs, il peut arriver que certaines coquilles subsistent. Un candidat repérant une coquille peut le signaler sur sa copie. Attention cependant au faux positif : signaler qu'on a repéré une coquille alors qu'il n'y en a pas fait plutôt mauvais effet. Il est à noter que, quand la coquille est avérée, la question sort généralement du barème.

Commentaire

- Il manque une étape essentielle dans la question, à savoir démontrer l'existence des espérances considérées. Comme on considère une transformée de la v.a.r. Z_n , cela passe a priori par le théorème de transfert. C'est en effet l'argument clé dans ce cas. Au vu de la définition de la fonction g_n , il est peu probable que les concepteurs exigent une rédaction en ce sens. C'est certainement plutôt un argument de domination qui était attendu. Malheureusement, ce résultat n'est pas au programme ECE. Détaillons-le afin de comprendre comment on aurait pu traiter plus rigoureusement cette question.
- Le « théorème de domination » est un résultat assez classique en probabilités. Il est présent dans le programme ECS mais absent du programme ECE. Dans le programme officiel ECS, il s'énonce comme suit.

Soient U et V deux variables aléatoires réelles.

On suppose :

- × $0 \leq |U| \leq V$ (ce qui signifie : $\forall \omega \in \Omega, 0 \leq |U(\omega)| \leq V(\omega)$)
- × la v.a.r. Y admet une espérance

Alors la v.a.r. U admet une espérance et :

$$|\mathbb{E}(U)| \leq \mathbb{E}(V)$$

- Notons que, d'après la question **13.a)(i)**, on a, pour toute v.a.r. X et pour tout $\omega \in \Omega$:

$$0 \leq g_n(X(\omega)) \leq 1$$

On est dans le cadre d'application du théorème de domination en considérant : $U = g_n(X)$ et $Y = 1$. On en conclut ainsi, que pour toute v.a.r. X , la v.a.r. $g_n(X)$ admet une espérance. De plus :

$$|\mathbb{E}(g_n(X))| \leq \mathbb{E}(1) = 1$$

Au passage : $|\mathbb{E}(g_n(X))| = \mathbb{E}(g_n(X))$ car $g_n(X)$ est à valeurs positives.

- Il est à noter que le théorème de domination peut s'utiliser aussi bien :
 - × pour des v.a.r. discrètes,
 - × pour des v.a.r. à densité,
 - × pour des v.a.r. qui ne sont ni discrètes ni à densité.
- Profitons-en pour rappeler brièvement que l'ensemble des v.a.r. discrètes et l'ensemble des v.a.r. à densité sont disjoints mais que ces deux ensembles ne forment pas une partition de l'ensemble des v.a.r. . Ainsi, une v.a.r. ne peut être à la fois discrète et à densité mais peut n'être ni discrète ni à densité. Les v.a.r. qui vérifient cette dernière propriété sont un peu délicates à étudier au sein du programme EC. En particulier, si on sait parfaitement définir l'espérance d'une v.a.r. discrète ou à densité, on ne sait pas, dans le cadre du programme, définir l'espérance d'une v.a.r. ni discrète ni à densité. On joue alors un peu les apprentis sorciers : on peut, sous les hypothèses du théorème de domination, démontrer qu'une v.a.r. X ni discrète, ni à densité, admet une espérance $\mathbb{E}(X)$ (et même conclure $|\mathbb{E}(X)| \leq \mathbb{E}(Y)$) sans pour autant savoir calculer $\mathbb{E}(X)$.
- En réalité, il existe une théorie permettant d'unifier, sous une même écriture, les définitions de l'espérance dans le cas des v.a.r. discrètes, à densité ou quelconques. Cette théorie n'est pas à notre portée en ECE / ECS mais permet de justifier l'utilisation de certains théorèmes du programme dans le cas de v.a.r. quelconques.

□

Il suffit alors de montrer que l'on a :

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

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 :

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

Démonstration.

Soit $z \in \mathbb{R}$ et soit $u \in \mathbb{R}$.

- On procède par intégrations par parties (IPP).

$$\left| \begin{array}{ll} f(t) = (z+u-t)^2 & f'(t) = 2(z+u-t)(-1) \\ h'(t) = g'''(t) & h(t) = g''(t) \end{array} \right.$$

Cette IPP est valide car :

- × la fonction f est de classe \mathcal{C}^1 sur le SEGMENT $[z, z+u]$ en tant que fonction polynomiale.
- × la fonction $h = g''$ est de classe \mathcal{C}^1 sur le SEGMENT $[z, z+u]$ car g''' est continue sur ce segment car elle est de classe \mathcal{C}^3 sur \mathbb{R} tout entier.

On obtient :

$$\begin{aligned} \int_z^{z+u} (z+u-t)^2 g'''(t) dt &= [(z+u-t)^2 g''(t)]_z^{z+u} - 2 \int_z^{z+u} (z+u-t)(-1) g''(t) dt \\ &= \left(\cancel{((z+u)-(z+u))^2} g''(z+u) - ((z+u)-z)^2 g''(z) \right) \\ &\quad + 2 \int_z^{z+u} (z+u-t) g''(t) dt \\ &= -u^2 g''(z) + 2 \int_z^{z+u} (z+u-t) g''(t) dt \end{aligned}$$

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

- On réalise alors une nouvelle IPP en suivant la même idée.
En posant : $f(t) = z+u-t$ et $h'(t) = g''(t)$, on obtient de même :

$$\begin{aligned} \int_z^{z+u} (z+u-t) g''(t) dt &= [(z+u-t) g'(t)]_z^{z+u} - \int_z^{z+u} (-1) g'(t) dt \\ &= \left(\cancel{((z+u)-(z+u))} g'(z+u) - ((z+u)-z) g'(z) \right) + \int_z^{z+u} g'(t) dt \\ &= -u g''(z) + [g(t)]_z^{z+u} \end{aligned}$$

$$\boxed{\int_z^{z+u} (z+u-t) g''(t) dt = -u g''(z) + (g(z+u) - g(z))}$$

- Finalement, en combinant ces deux résultats, on obtient bien le résultat annoncé.

$$\forall z \in \mathbb{R}, \forall u \in \mathbb{R}, \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt = -\frac{1}{2} u^2 g''(z) - u g'(z) + g(z+u) - g(z) \quad \square$$

- b) En déduire que pour tous u, v, z réels, on a :

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

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

Démonstration.

Soit $(u, v, z) \in \mathbb{R}^3$.

D'après la question précédente, en réordonnant :

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

$$\text{et aussi } g(z+v) = v g'(z) + \frac{1}{2} v^2 g''(z) + g(z) + \frac{1}{2} \int_z^{z+v} (z+v-t)^2 g'''(t) dt$$

On obtient alors, en soustrayant terme à terme :

$$\begin{aligned} g(z+u) - g(z+v) &= u g'(z) - v g'(z) \\ &+ \frac{1}{2} u^2 g''(z) - \frac{1}{2} v^2 g''(z) \\ &+ \cancel{g(z)} - \cancel{g(z)} \\ &+ \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt - \frac{1}{2} \int_z^{z+v} (z+v-t)^2 g'''(t) dt \\ &= g'(z)(u-v) + \frac{1}{2} g''(z)(u^2 - v^2) + R(z, u, v) \end{aligned}$$

$$\forall (u, v, z) \in \mathbb{R}^3, g(z+u) - g(z+v) = g'(z)(u-v) + \frac{1}{2} g''(z)(u^2 - v^2) + R(z, u, v) \quad \square$$

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

Démonstration.

Supposons que g''' est bornée.

On en déduit que $|g'''|$ est majorée et on note $M_{g'''}$ un majorant de $|g'''|$.

- Par définition :

$$\begin{aligned} &|R(z, u, v)| \\ &= \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt - \frac{1}{2} \int_z^{z+v} (z+v-t)^2 g'''(t) dt \right| \\ &\leq \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| + \left| \frac{1}{2} \int_z^{z+v} (z+v-t)^2 g'''(t) dt \right| \quad (\text{par inégalité triangulaire}) \end{aligned}$$

- Considérons tout d'abord le cas où $u \geq 0$:

$$\begin{aligned}
 & \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \\
 &= \frac{1}{2} \left| \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \\
 &\leq \frac{1}{2} \int_z^{z+u} |(z+u-t)^2 g'''(t)| dt && \text{(par inégalité triangulaire,} \\
 &&& \text{les bornes d'intégration} \\
 &&& \text{étant dans l'ordre croissant)} \\
 &= \frac{1}{2} \int_z^{z+u} |(z+u-t)^2| |g'''(t)| dt \\
 &= \frac{1}{2} \int_z^{z+u} (z+u-t)^2 |g'''(t)| dt && \text{(car un carré est toujours positif)} \\
 &\leq \frac{1}{2} \int_z^{z+u} (z+u-t)^2 M_{g'''} dt
 \end{aligned}$$

Il convient de s'arrêter sur cette dernière inégalité. Pour tout $t \in [z, z+u]$:

$$\begin{aligned}
 |g'''(t)| &\leq M_{g'''} \\
 \text{donc } (z+u-t)^2 |g'''(t)| &\leq (z+u-t)^2 M_{g'''} \quad \text{(car } (z+u-t)^2 \geq 0)
 \end{aligned}$$

Ainsi, par croissance de l'intégrale, les bornes étant dans l'ordre croissant ($z+u \geq z$) :

$$\int_z^{z+u} (z+u-t)^2 |g'''(t)| dt \leq \int_z^{z+u} (z+u-t)^2 M_{g'''} dt = M_{g'''} \int_z^{z+u} (z+u-t)^2 dt$$

Enfin :

$$\begin{aligned}
 \int_z^{z+u} (z+u-t)^2 dt &= \left[\frac{-1}{3} (z+u-t)^3 \right]_z^{z+u} \\
 &= -\frac{1}{3} \left(\left((z+u) - (z+u) \right)^3 - \left((z+u) - z \right)^3 \right) \\
 &= -\frac{1}{3} (-u^3) = \frac{u^3}{3}
 \end{aligned}$$

Finalement :

$$\left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \leq \frac{1}{2} \int_z^{z+u} (z+u-t)^2 M_{g'''} dt = \frac{1}{2} M_{g'''} \frac{u^3}{3}$$

$$\boxed{\forall u \geq 0, \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \leq \frac{1}{2} M_{g'''} \frac{u^3}{3}}$$

- Il reste à traiter le cas $u < 0$. On effectue une démonstration similaire, en remarquant :

$$\begin{aligned} \left| \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| &= \left| - \int_{z+u}^z (z+u-t)^2 g'''(t) dt \right| \\ &= |-1| \left| \int_{z+u}^z (z+u-t)^2 g'''(t) dt \right| \\ &= \left| \int_{z+u}^z (z+u-t)^2 g'''(t) dt \right| \end{aligned}$$

Les bornes étant dans l'ordre croissant, en appliquant l'inégalité triangulaire, on obtient :

$$\begin{aligned} \left| \int_{z+u}^z (z+u-t)^2 g'''(t) dt \right| &\leq \int_{z+u}^z (z+u-t)^2 M_{g'''} dt \\ &= M_{g'''} \left[-\frac{1}{3} (z+u-t)^3 \right]_{z+u}^z \\ &= -\frac{1}{3} M_{g'''} \left(((z+u)-z)^3 - \cancel{((z+u)-(z+u))^3} \right) \end{aligned}$$

En combinant ces résultats, on obtient :

$$\forall u < 0, \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \leq -\frac{1}{2} M_{g'''} \frac{u^3}{3}.$$

- On remarque alors : $|u|^3 = \begin{cases} u^3 & \text{si } u \geq 0 \\ (-u)^3 = -u^3 & \text{si } u < 0 \end{cases}$.

Cela permet de synthétiser les deux cas précédents :

$$\forall u \in \mathbb{R}, \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| \leq \frac{1}{6} M_{g'''} |u|^3.$$

- On reprend alors le début de la démonstration :

$$\begin{aligned} |R(z, u, v)| &= \left| \frac{1}{2} \int_z^{z+u} (z+u-t)^2 g'''(t) dt \right| + \left| \frac{1}{2} \int_z^{z+v} (z+v-t)^2 g'''(t) dt \right| \\ &\leq \frac{1}{6} M_{g'''} |u|^3 + \frac{1}{6} M_{g'''} |v|^3 \end{aligned}$$

$$\text{On a bien : } \forall (u, v, z) \in \mathbb{R}^3, |R(z, u, v)| \leq \frac{1}{6} M_{g'''} (|u|^3 + |v|^3).$$

□

15. Soit $(Y_i)_{i \geq 1}$ 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.

Démonstration.

Les v.a.r. de la famille $(Y_i)_{i \in [1, n]}$:

× sont indépendantes.

× suivent toutes la loi normale $\mathcal{N}(0, 1)$.

On en déduit, par stabilité par somme des lois normales : $\sum_{i=1}^n Y_i \hookrightarrow \mathcal{N}(0 + \dots + 0, 1 + \dots + 1)$.

$$\sum_{i=1}^n Y_i \hookrightarrow \mathcal{N}(0, n)$$

Commentaire

- L'énoncé demande de « Justifier » le résultat. Cette terminologie est souvent associée à des démonstrations courtes et éventuellement moins formelles.
- La formulation de l'énoncé renseigne sur la méthode à choisir :
 - × si l'énoncé demande de déterminer la loi d'une v.a.r. , il convient de déterminer la fonction de répartition de cette v.a.r.
 - × si l'énoncé demande de « Justifier » qu'une v.a.r. suit une loi donnée, il faut s'orienter vers une propriété de stabilité des lois usuelles ou vers un résultat précédemment démontré dans l'énoncé.

□

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

Démonstration.

• Dans la question précédente, on a démontré : $\sum_{i=1}^n Y_i \hookrightarrow \mathcal{N}(0, n)$.

Comme $T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ est une transformée affine de la v.a.r. $\sum_{i=1}^n Y_i$ qui suit une loi normale, on en conclut que T_n suit une loi normale $\mathcal{N}(m, s^2)$. On détermine alors les paramètres (m, s^2) de cette loi normale en déterminant espérance et variance de T_n .

• Comme T_n suit une loi normale alors T_n admet une espérance et une variance. Tout d'abord :

$$\begin{aligned} m &= \mathbb{E}(T_n) \\ &= \mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i\right) \\ &= \frac{1}{\sqrt{n}} \mathbb{E}\left(\sum_{i=1}^n Y_i\right) && \text{(par linéarité de l'espérance)} \\ &= \frac{1}{\sqrt{n}} \times 0 && \text{(d'après la question précédente)} \\ &= 0 \end{aligned}$$

- De plus :

$$\begin{aligned}
 s^2 &= \mathbb{V}(T_n) \\
 &= \mathbb{V}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i\right) \\
 &= \left(\frac{1}{\sqrt{n}}\right)^2 \mathbb{V}\left(\sum_{i=1}^n Y_i\right) \quad (\text{par propriété de la variance}) \\
 &= \frac{1}{n} n \quad (\text{d'après la question précédente}) \\
 &= 1
 \end{aligned}$$

Finalemment : $T_n \hookrightarrow \mathcal{N}(0, 1)$.

Commentaire

Rappelons que les lois normales sont stables par transformation affine. De manière générale, cela s'écrit, pour tout $(a, b, m) \in \mathbb{R}^3$ et tout $\sigma \geq 0$:

$$X \hookrightarrow \mathcal{N}(m, \sigma^2) \Leftrightarrow aX + b \hookrightarrow \mathcal{N}(am + b, a^2 \sigma^2)$$

□

b) Pour $k \in \{2, \dots, n-1\}$, 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 $k \in \{1, 2, \dots, n-1\}$: $W_k + \frac{1}{\sqrt{n}} Y_k = W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1}$.

Démonstration.

- Soit $k \in \{2, \dots, n-2\}$.

$$\begin{aligned}
 W_k + \frac{1}{\sqrt{n}} Y_k &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{k-1} + X_{k+1} + \dots + X_n) + \frac{1}{\sqrt{n}} Y_k \quad (\text{par définition avec } k \in \llbracket 2, n-1 \rrbracket) \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{k-1}) + \frac{1}{\sqrt{n}} Y_k + \frac{1}{\sqrt{n}} (X_{k+1} + \dots + X_n) \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{k-1} + Y_k) + \frac{1}{\sqrt{n}} (X_{k+1} + \dots + X_n) \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_k) + \frac{1}{\sqrt{n}} (X_{k+1} + X_{k+2} + \dots + X_n) \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_k) + \frac{1}{\sqrt{n}} (X_{k+2} + \dots + X_n) + \frac{1}{\sqrt{n}} X_{k+1} \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_k + X_{k+2} + \dots + X_n) + \frac{1}{\sqrt{n}} X_{k+1} \\
 &= W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1} \quad (\text{par définition avec } k+1 \in \llbracket 2, n-1 \rrbracket)
 \end{aligned}$$

$$\forall k \in \llbracket 2, n-2 \rrbracket, W_k + \frac{1}{\sqrt{n}} Y_k = W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1}$$

- Traitons maintenant le cas $k = 1$.

$$\begin{aligned}
 W_1 + \frac{1}{\sqrt{n}} Y_1 &= \frac{1}{\sqrt{n}} \sum_{i=2}^n X_i + \frac{1}{\sqrt{n}} Y_1 && (\text{par définition de } W_1) \\
 &= \frac{1}{\sqrt{n}} (Y_1 + X_3 + \dots + X_n) + \frac{1}{\sqrt{n}} X_2 \\
 &= W_2 + \frac{1}{\sqrt{n}} X_2
 \end{aligned}$$

$$W_1 + \frac{1}{\sqrt{n}} Y_1 = W_2 + \frac{1}{\sqrt{n}} X_2$$

- Traitons enfin le cas $k = n - 1$.

$$\begin{aligned}
 W_{n-1} + \frac{1}{\sqrt{n}} Y_{n-1} &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{n-2} + X_n) + \frac{1}{\sqrt{n}} Y_{n-1} \\
 &= \frac{1}{\sqrt{n}} (Y_1 + \dots + Y_{n-2} + Y_{n-1}) + \frac{1}{\sqrt{n}} X_n \\
 &= W_n + \frac{1}{\sqrt{n}} X_n && (\text{par définition de } W_n)
 \end{aligned}$$

$$W_{n-1} + \frac{1}{\sqrt{n}} Y_{n-1} = W_n + \frac{1}{\sqrt{n}} X_n$$

$$\text{Finalement, on a bien : } \forall k \in \llbracket 1, n-1 \rrbracket, W_k + \frac{1}{\sqrt{n}} Y_k = W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1}. \quad \square$$

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

Démonstration.

- Tout d'abord :

$$\begin{aligned}
 &\sum_{k=1}^n \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) \\
 &= \sum_{k=1}^{n-1} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) \\
 &\quad + g_n \left(W_n + \frac{1}{\sqrt{n}} X_n \right) - g_n \left(W_n + \frac{1}{\sqrt{n}} Y_n \right) \\
 &= \sum_{k=1}^{n-1} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_{k+1} + \frac{1}{\sqrt{n}} X_{k+1} \right) \right) && (\text{d'après la question précédente}) \\
 &\quad + g_n \left(W_n + \frac{1}{\sqrt{n}} X_n \right) - g_n \left(W_n + \frac{1}{\sqrt{n}} Y_n \right) \\
 &= g_n \left(W_1 + \frac{1}{\sqrt{n}} X_1 \right) - g_n \left(W_n + \frac{1}{\sqrt{n}} X_n \right) && (\text{par sommation télescopique}) \\
 &\quad + g_n \left(W_n + \frac{1}{\sqrt{n}} X_n \right) - g_n \left(W_n + \frac{1}{\sqrt{n}} Y_n \right)
 \end{aligned}$$

- Or :

$$W_1 + \frac{1}{\sqrt{n}} X_1 = \left(\frac{1}{\sqrt{n}} \sum_{i=2}^n X_i \right) + \frac{1}{\sqrt{n}} X_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i = Z_n$$

et :

$$W_n + \frac{1}{\sqrt{n}} Y_n = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} Y_i \right) + \frac{1}{\sqrt{n}} Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i = T_n$$

Finalemment, on a bien :

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

Commentaire

- L'énoncé original comporte une erreur. Il était en effet demandé de démontrer :

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

(on considèrerait une somme $\sum_{k=1}^{n-1} \dots$ en lieu et place de $\sum_{k=1}^n \dots$)

- Étant donnée cette coquille, il est probable que tout candidat présentant la sommation télescopique se voie attribuer l'ensemble des points alloués à la question. On retiendra que tout ce qui est juste est retenu pour le candidat. On en profite pour rappeler qu'il est possible de ne traiter qu'une partie de la question. Il est alors primordial de signaler au correcteur que la piste suivie ne semble pas aboutir. □

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

c) (i) Montrer : $\mathbb{E} \left((X_k - Y_k) g'_n(W_k) \right) = 0$.

Démonstration.

Soit $k \in \llbracket 1, n \rrbracket$.

- Par hypothèse, les v.a.r. $Y_1, \dots, Y_{k-1}, Y_k, X_k, X_{k+1}, \dots, X_n$ sont indépendantes. On en déduit, par le lemme des coalitions, que toute transformée des v.a.r. X_k et Y_k (notamment la v.a.r. $X_k - Y_k$) est indépendante de toute transformée des v.a.r. $Y_1, \dots, Y_{k-1}, X_{k+1}, \dots, X_n$ (notamment la v.a.r. $g'_n(W_k)$).
- On en déduit (sous réserve de l'existence des espérances) :

$$\begin{aligned} \mathbb{E} \left((X_k - Y_k) g'_n(W_k) \right) &= \mathbb{E}(X_k - Y_k) \mathbb{E} \left(g'_n(W_k) \right) && \text{(par indépendance des} \\ &&& \text{v.a.r. } X_k - Y_k \text{ et } g'_n(W_k)) \\ &= \left(\mathbb{E}(X_k) - \mathbb{E}(Y_k) \right) \mathbb{E} \left(g'_n(W_k) \right) && \text{(par linéarité de l'espérance)} \\ &= (0 - 0) \mathbb{E} \left(g'_n(W_k) \right) && \text{(par hypothèse sur } X_k \\ &&& \text{et car } Y_k \hookrightarrow \mathcal{N}(0, 1)) \\ &= 0 \end{aligned}$$

$$\text{On a bien : } \mathbb{E} \left((X_k - Y_k) g'_n(W_k) \right) = 0$$

Commentaire

- On insiste une nouvelle fois sur le manque de quantification de certaines variables de l'énoncé. Il est très étonnant que l'ensemble d'appartenance de k ne soit pas spécifié dans cette question alors qu'il l'est dans les questions précédentes et notamment la question **15.b(i)**.
- Comme en question **13.c)**, il manque une étape essentielle de la question, à savoir démontrer l'existence de l'espérance considérée. Comme les v.a.r. $X_k - Y_k$ et $g'_n(W_k)$ sont indépendantes, il suffit de démontrer que chacune des ces deux v.a.r. sont indépendantes. Or :
 - × la v.a.r. $X_k - Y_k$ admet une espérance en tant que différence de deux v.a.r. qui admettent une espérance.
 - × la v.a.r. $g'_n(W_k)$ admet une espérance car la fonction g'_n est bornée. Pour comprendre cet argument, il faut se reporter au théorème de domination, résultat détaillé dans la remarque de la question **13.c)**.

□

(ii) Montrer : $\mathbb{E}\left((X_k^2 - Y_k^2) g''_n(W_k)\right) = 0$.

Démonstration.

Soit $k \in \llbracket 1, n \rrbracket$.

- × En utilisant les arguments développés en question précédente, on démontre que les v.a.r. $X_k^2 - Y_k^2$ et $g''_n(W_k)$ sont indépendantes.
- × On en déduit (sous réserve de l'existence des espérances) :

$$\begin{aligned}
 \mathbb{E}\left((X_k^2 - Y_k^2) g''_n(W_k)\right) &= \mathbb{E}(X_k^2 - Y_k^2) \mathbb{E}\left(g''_n(W_k)\right) && \text{(par indépendance des} \\
 & && \text{v.a.r. } X_k^2 - Y_k^2 \text{ et } g''_n(W_k)) \\
 &= \left(\mathbb{E}(X_k^2) - \mathbb{E}(Y_k^2)\right) \mathbb{E}\left(g''_n(W_k)\right) && \text{(par linéarité de l'espérance)} \\
 &= (1 - 1) \mathbb{E}\left(g''_n(W_k)\right) = 0
 \end{aligned}$$

En effet, par hypothèse de l'énoncé, $\mathbb{V}(X_k) = 1$. On en déduit, par la formule de Kœnig-Huygens :

$$\begin{aligned}
 \mathbb{E}(X_k^2) &= \mathbb{V}(X_k) + (\mathbb{E}(X_k))^2 \\
 &= 1 + 0^2 = 1
 \end{aligned}$$

On en déduit de même : $\mathbb{E}(Y_k^2)$ car, comme $Y_k \leftrightarrow \mathcal{N}(0, 1)$, $\mathbb{V}(Y_k) = 1$ et $(\mathbb{E}(Y_k))^2 = 0^2 = 0$.

On a bien : $\mathbb{E}\left((X_k - Y_k) g'_n(W_k)\right) = 0$

Commentaire

La remarque précédente s'applique :

- × la formulation de cette question fait apparaître un manque de quantificateur.
- × la démonstration de l'existence de l'espérance de la v.a.r. $g'_n(W_k)$ peut être obtenue facilement par le théorème de domination puisque la fonction g'_n est bornée.

□

$$(iii) \text{ En d\u00e9duire : } \mathbb{E} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) = \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right).$$

D\u00e9monstration.

- Soit $\omega \in \Omega$. On applique la question **14.b**), en consid\u00e9rant :

\(\times\) $g = g_n$ qui est une fonction de classe \mathcal{C}^3 sur \mathbb{R} .

\(\times\) $z = W_k(\omega) \in \mathbb{R}$.

\(\times\) $u = \frac{1}{\sqrt{n}} X_k(\omega) \in \mathbb{R}$.

\(\times\) $v = \frac{1}{\sqrt{n}} Y_k(\omega) \in \mathbb{R}$.

On obtient :

$$\begin{aligned} & g_n \left(W_k(\omega) + \frac{1}{\sqrt{n}} X_k(\omega) \right) - g_n \left(W_k(\omega) + \frac{1}{\sqrt{n}} Y_k(\omega) \right) \\ &= g'_n(W_k(\omega)) \left(\frac{1}{\sqrt{n}} X_k(\omega) - \frac{1}{\sqrt{n}} Y_k(\omega) \right) + \frac{1}{2} g''_n(W_k(\omega)) \left(\left(\frac{1}{\sqrt{n}} X_k(\omega) \right)^2 - \left(\frac{1}{\sqrt{n}} Y_k(\omega) \right)^2 \right) \\ &+ R \left(W_k(\omega), \frac{1}{\sqrt{n}} X_k(\omega), \frac{1}{\sqrt{n}} Y_k(\omega) \right) \end{aligned}$$

Autrement dit, l'\u00e9galit\u00e9 entre v.a.r. suivante est v\u00e9rifi\u00e9e :

$$\begin{aligned} & g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \\ &= \frac{1}{\sqrt{n}} g'_n(W_k) (X_k - Y_k) + \frac{1}{2} g''_n(W_k) \frac{1}{n} (X_k^2 - Y_k^2) + R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \end{aligned}$$

- Toutes les v.a.r. en jeu admettant une esp\u00e9rance, on en d\u00e9duit, par lin\u00e9arit\u00e9 de l'esp\u00e9rance :

$$\begin{aligned} & \mathbb{E} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) \\ &= \frac{1}{\sqrt{n}} \mathbb{E} \left(g'_n(W_k) (X_k - Y_k) \right) + \frac{1}{2n} \mathbb{E} \left(g''_n(W_k) (X_k^2 - Y_k^2) \right) + \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right) \\ &= 0 + 0 + \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right) \end{aligned}$$

Enfinement : $\mathbb{E} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) = \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right).$

Commentaire

Dans cette question encore, on ne d\u00e9montre pas l'existence des esp\u00e9rances. Cela a \u00e9t\u00e9 dit pr\u00e9c\u00e9demment, comme la fonction g_n est born\u00e9e, on peut en d\u00e9duire, via le th\u00e9or\u00e8me de domination, que pour toute v.a.r. X , la transform\u00e9e $g_n(X)$ admet une esp\u00e9rance. En isolant la v.a.r. $R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right)$ dans l'\u00e9galit\u00e9 de la d\u00e9monstration, on d\u00e9montre que cette v.a.r. s'\u00e9crit comme somme de v.a.r. qui admettent une esp\u00e9rance. Elle admet donc une esp\u00e9rance, ce qui permet de conclure que toutes les v.a.r. pr\u00e9sentes en admettent une. \square

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

Démonstration.

- D'après la question **15.b)(ii)** :

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

- Toutes les v.a.r. en jeu admettant une espérance, on en déduit par linéarité de l'espérance :

$$\begin{aligned} \mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n)) &= \sum_{k=1}^n \mathbb{E} \left(g_n \left(W_k + \frac{1}{\sqrt{n}} X_k \right) - g_n \left(W_k + \frac{1}{\sqrt{n}} Y_k \right) \right) \\ &= \sum_{k=1}^n \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right) \quad (\text{d'après la question précédente}) \end{aligned}$$

- On en déduit :

$$\begin{aligned} |\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))| &= \left| \sum_{k=1}^n \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right) \right| \\ &\leq \sum_{k=1}^n \left| \mathbb{E} \left(R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right) \right| \quad (\text{par inégalité triangulaire}) \\ &\leq \sum_{k=1}^n \mathbb{E} \left(\left| R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right| \right) \quad (\text{par inégalité triangulaire}) \end{aligned}$$

$$\boxed{|\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))| \leq \sum_{k=1}^n \mathbb{E} \left(\left| R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right| \right)}$$

Commentaire

- On se sert dans cette question de la propriété suivante :

Soit X une variable aléatoire réelle.
 On suppose que X admet une espérance.
 Alors $|X|$ admet une espérance et : $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

Il s'agit d'une propriété usuelle en probabilités mais elle n'apparaît pas dans le programme ECE. Son utilisation requiert donc une démonstration. Cette démonstration est accessible dans le cas où X est une v.a.r. discrète (ce que l'on détaille plus loin dans la remarque) ainsi que dans le cas où X est une v.a.r. à densité. Le cas des v.a.r. quelconques n'est pas accessible puisque le programme ne permet pas de définir (sauf cas particuliers) l'espérance d'une v.a.r. ni discrète ni à densité. L'esprit du programme est d'exposer les propriétés de l'espérance dans le cas des v.a.r. discrètes / à densité puis d'admettre « que les propriétés opératoires usuelles de l'espérance et de la variance se généralisent aux variables aléatoires quelconques ».

Commentaire

- On constate une nouvelle fois dans cet énoncé qu'il nous manque un résultat permettant de répondre de manière complète à une question. L'énoncé étant déjà (démessurément) long, on peut comprendre la réticence du concepteur à ajouter de nouvelles questions ou à admettre des propriétés qui ne sont pas au programme. Cet argument est assez peu convaincant car les énoncés ne devraient contenir que des questions qui peuvent être traitées dans le cadre du programme. Si cela oblige à densifier l'énoncé, il faut alors essayer de supprimer d'autres éléments afin que le résultat ait encore du sens pour une épreuve d'une durée de 4h.
- Passons maintenant à la démonstration de cette propriété dans le cas des v.a.r. discrètes. On suppose que X admet une espérance. Deux cas se présentent alors.
 - × Si X est une v.a.r. finie alors X prend un nombre fini $m \in \mathbb{N}^*$ de valeurs distinctes. On peut écrire son ensemble image sous la forme :

$$X(\Omega) = \{x_1, \dots, x_m\}$$

La v.a.r. $|X|$ est elle aussi finie ($X(\Omega) = \{|x_i| \mid i \in \llbracket 1, m \rrbracket\}$) et admet donc une espérance. On a alors :

$$\begin{aligned} |\mathbb{E}(X)| &= \left| \sum_{i=1}^m x_i \mathbb{P}([X = x_i]) \right| \\ &\leq \sum_{i=1}^m |x_i \mathbb{P}([X = x_i])| && \text{(par inégalité triangulaire)} \\ &= \sum_{i=1}^m |x_i| \mathbb{P}([X = x_i]) \\ &= \sum_{i=1}^m |x_i| \mathbb{P}([X = x_i]) \\ &= \mathbb{E}(|X|) && \text{(par théorème de transfert)} \end{aligned}$$

- × Si X n'est pas une v.a.r. finie alors elle prend un nombre infini dénombrable de valeurs distinctes. On peut écrire son ensemble image sous la forme :

$$X(\Omega) = \{x_i \mid i \in \mathbb{N}^*\}$$

On a fait l'hypothèse que X admet une espérance. Par définition, on en conclut que la série $\sum_{i \in \mathbb{N}^*} x_i \mathbb{P}([X = x_i])$ est absolument convergente. On peut alors appliquer l'inégalité triangulaire comme dans le cas précédent. Plus précisément :

$$\begin{aligned} |\mathbb{E}(X)| &= \left| \sum_{i=1}^{+\infty} x_i \mathbb{P}([X = x_i]) \right| \leq \sum_{i=1}^{+\infty} |x_i \mathbb{P}([X = x_i])| \\ &= \sum_{i=1}^{+\infty} |x_i| \mathbb{P}([X = x_i]) = \mathbb{E}(|X|) \end{aligned}$$

La démonstration est très similaire dans le cas des v.a.r. à densité : on utilise encore le fait que la valeur absolue d'une somme est majorée par la somme des valeurs absolues. La différence majeure est que dans le cas des v.a.r. à densité, on travaille sur des intégrales (c'est-à-dire des sommes dans le monde continu) et plus des sommes dénombrables de réels. \square

e) Conclure que l'on a :

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

Démonstration.

- Soit $\omega \in \Omega$. On applique la question 14.c), en considérant :
 - × $g = g_n$ qui est une fonction telle que g_n''' est bornée (hypothèse admise dans l'énoncé).
 - × $z = W_k(\omega) \in \mathbb{R}$.
 - × $u = \frac{1}{\sqrt{n}} X_k(\omega) \in \mathbb{R}$.
 - × $v = \frac{1}{\sqrt{n}} Y_k(\omega) \in \mathbb{R}$.

On obtient :

$$\left| R\left(W_k(\omega), \frac{1}{\sqrt{n}} X_k(\omega), \frac{1}{\sqrt{n}} Y_k(\omega)\right) \right| \leq \frac{1}{6} M_{g_n'''} \left(\left| \frac{1}{\sqrt{n}} X_k(\omega) \right|^3 + \left| \frac{1}{\sqrt{n}} Y_k(\omega) \right|^3 \right)$$

Autrement dit, l'inégalité suivante entre v.a.r. est vérifiée :

$$\begin{aligned} \left| R\left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k\right) \right| &\leq \frac{1}{6} M_{g_n'''} \left(\left| \frac{1}{\sqrt{n}} X_k \right|^3 + \left| \frac{1}{\sqrt{n}} Y_k \right|^3 \right) \\ &= \frac{1}{6} M_{g_n'''} \left(\left| \frac{1}{\sqrt{n}} \right|^3 |X_k|^3 + \left| \frac{1}{\sqrt{n}} \right|^3 |Y_k|^3 \right) \\ &= \frac{1}{6} \frac{1}{n\sqrt{n}} M_{g_n'''} \left(|X_k|^3 + |Y_k|^3 \right) \quad (\text{en effet, on a : } (\sqrt{n})^3 = \sqrt{n} \sqrt{n} \sqrt{n} = n \sqrt{n}) \end{aligned}$$

- Par hypothèse de l'énoncé, la v.a.r. X_k admet un moment d'ordre 3. Cela démontre que la v.a.r. $|X_k|$ admet elle aussi un moment d'ordre 3.

De plus, comme la v.a.r. Y_k suit la loi normale $\mathcal{N}(0, 1)$, elle admet un moment à tout ordre et donc en particulier à l'ordre 3. On en déduit que la v.a.r. $|Y_k|$ admet elle aussi un moment d'ordre 3.

Toutes les v.a.r. en jeu admettant une espérance, on en déduit, par croissance de l'espérance :

$$\begin{aligned} \mathbb{E} \left(\left| R\left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k\right) \right| \right) &\leq \mathbb{E} \left(\frac{1}{6} \frac{1}{n\sqrt{n}} M_{g_n'''} \left(|X_k|^3 + |Y_k|^3 \right) \right) \\ &= \frac{1}{6} \frac{1}{n\sqrt{n}} M_{g_n'''} \left(\mathbb{E}(|X_k|^3) + \mathbb{E}(|Y_k|^3) \right) \quad (\text{par linéarité de l'espérance}) \\ &= \frac{1}{6} \frac{1}{n\sqrt{n}} M_{g_n'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) \end{aligned}$$

Cette dernière égalité provient du fait que toutes les v.a.r. de la famille $(X_i)_{i \in [1, n]}$ ont même loi. De même, toutes les v.a.r. de la famille $(Y_i)_{i \in [1, n]}$ ont même loi ($\mathcal{N}(0, 1)$ en l'occurrence).

- On obtient alors, d'après la question précédente :

$$\begin{aligned}
 & \left| \mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n)) \right| \\
 & \leq \sum_{k=1}^n \mathbb{E} \left(\left| R \left(W_k, \frac{1}{\sqrt{n}} X_k, \frac{1}{\sqrt{n}} Y_k \right) \right| \right) && \text{(d'après le début de la question)} \\
 & \leq \sum_{k=1}^n \left(\frac{1}{6} \frac{1}{n\sqrt{n}} M_{g'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) \right) \\
 & = \cancel{n} \left(\frac{1}{6} \frac{1}{\cancel{n}\sqrt{n}} M_{g'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) \right) && \text{(car le contenu de la somme ne dépend pas de la variable de sommation } k \text{)}
 \end{aligned}$$

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

Commentaire

- Insistons tout d'abord sur la propriété suivante :

$$\text{La v.a.r. } U \text{ admet un moment d'ordre } r \in \mathbb{N} \iff \text{La v.a.r. } |U| \text{ admet un moment d'ordre } r \in \mathbb{N}$$

La démonstration de ce résultat est directe car l'existence de l'espérance d'une v.a.r. nécessite une hypothèse de convergence absolue.

- Dans cette question, on rappelle qu'une v.a.r. qui suit la loi $\mathcal{N}(0, 1)$ admet des moments à tout ordre. Ce résultat n'étant pas au programme ECE, on le détaille ci-dessous.
- Commençons par rappeler qu'une densité de Y_k est : $f_{Y_k} : t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$.

Soit $r \in \mathbb{N}$. La v.a.r. Y_k admet un moment à l'ordre r si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t^r f_{Y_k}(t) dt$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour ce type de moment. Par un argument de parité (détaillé plus bas), il suffit de démontrer la convergence de l'intégrale $\int_0^{+\infty} t^r f_{Y_k}(t) dt$.

La fonction $t \mapsto t^r f_{Y_k}(t)$ est continue sur $[0, +\infty[$.

L'intégrale étudiée est donc impropre seulement en $+\infty$. Or :

$$\times \forall t \in [0, +\infty[, e^{-t} \geq 0.$$

$$\times t^r f_{Y_k}(t) = \underset{t \rightarrow +\infty}{o} (e^{-t}) \text{ car : } \frac{t^r f_{Y_k}(t)}{e^{-t}} = \frac{1}{\sqrt{2\pi}} t^r e^{-\frac{1}{2}t^2} e^t = \frac{1}{\sqrt{2\pi}} \frac{t^r}{e^t} e^{2t - \frac{1}{2}t^2} \xrightarrow{t \rightarrow +\infty} 0.$$

$$\blacktriangleright \frac{t^r}{e^t} \xrightarrow{t \rightarrow +\infty} 0 \text{ par croissances comparées.}$$

Plus précisément :

$$\blacktriangleright e^{2t - \frac{1}{2}t^2} \xrightarrow{t \rightarrow +\infty} 0 \text{ par composition de limite.}$$

$$\times \text{L'intégrale } \int_0^{+\infty} e^{-t} dt \text{ est convergente en tant que moment d'ordre } 0 \text{ d'une v.a.r. de loi } \mathcal{E}(1).$$

Par critère de négligeabilité d'intégrales généralisées de fonctions continues positives, l'intégrale $\int_0^{+\infty} t^r f_{Y_k}(t) dt$ est convergente.

Commentaire

- Deux cas se présentent alors.

× Si r est pair alors l'intégrande $t \mapsto t^r f_{Y_k}(t)$ est pair. On a alors :

$$\int_{-\infty}^0 t^r f_{Y_k}(t) dt = \int_0^{+\infty} t^r f_{Y_k}(t) dt$$

On en conclut alors que Y_k admet un moment d'ordre r et que celui-ci vaut :

$$\mathbb{E}(Y_k^r) = \int_{-\infty}^0 t^r f_{Y_k}(t) dt + \int_0^{+\infty} t^r f_{Y_k}(t) dt = 2 \int_0^{+\infty} t^r f_{Y_k}(t) dt$$

× Si r est impair alors l'intégrande $t \mapsto t^r f_{Y_k}(t)$ est impair. On a alors :

$$\int_{-\infty}^0 t^r f_{Y_k}(t) dt = - \int_0^{+\infty} t^r f_{Y_k}(t) dt$$

On en conclut alors que Y_k admet un moment d'ordre r et que celui-ci vaut :

$$\mathbb{E}(Y_k^r) = \int_{-\infty}^0 t^r f_{Y_k}(t) dt + \int_0^{+\infty} t^r f_{Y_k}(t) dt = 0$$

- Dans ce qui précède, on a établi un critère de négligeabilité à l'aide de la fonction $t \mapsto e^{-t}$. On pouvait utiliser ce même théorème en utilisant la fonction $t \mapsto \frac{1}{t^2}$. Il est d'ailleurs fréquent de procéder ainsi lorsque l'intégrande apparaît comme suffisamment petit en $+\infty$ pour que l'intégrale associée soit convergente. Cette dernière méthode est un peu plus longue à établir car elle fait apparaître une intégrale de Riemann impropre en $+\infty$, ce qui nécessite de « s'écarter de 0 ». Plus précisément, on rédige en deux temps :

× l'intégrande $t \mapsto t^r f_{Y_k}(t)$ est continu sur le **SEGMENT** $[0, 1]$ donc l'intégrale $\int_0^1 t^r f_{Y_k}(t) dt$ est bien définie.

× l'intégrale $\int_1^{+\infty} t^r f_{Y_k}(t) dt$ est convergente car l'intégrande est négligeable en $+\infty$ devant $t \mapsto \frac{1}{t^2}$, et car l'intégrale $\int_1^{+\infty} \frac{1}{t^2} dt$ est convergente en tant qu'intégrale de Riemann impropre en $+\infty$ et d'exposant 2 (> 1).

Ces deux points permettent de conclure que l'intégrale $\int_0^{+\infty} t^r f_{Y_k}(t) dt$ est convergente.

- De manière générale, lorsque l'intégrande présente un terme de la forme « $e^{-g(t)}$ », il peut être intéressant de penser à utiliser la comparaison à une intégrande de la forme $t \mapsto e^{-\alpha t}$ où $\alpha > 0$ est un réel correctement choisi. Cela permet de ne pas avoir à effectuer le découpage d'intégrale auquel on est contraint lors d'une comparaison avec une intégrale de Riemann. \square

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

Démonstration.

• Tout d'abord :

$$\begin{aligned} |\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))| &\leq \frac{1}{6\sqrt{n}} M_{g_n'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) && \text{(d'après la question précédente)} \\ &\leq \frac{1}{6\sqrt{n}} n^{\frac{1}{4}} M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) && \text{(d'après la question 13.a)(ii)} \\ &= \frac{1}{6n^{\frac{1}{4}}} M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) && \text{(car } \frac{n^{\frac{1}{4}}}{\sqrt{n}} = \frac{n^{\frac{1}{4}}}{n^{\frac{1}{2}}} = \frac{1}{n^{\frac{1}{2}-\frac{1}{4}}}) \end{aligned}$$

Finalement : $0 \leq |\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))| \leq \frac{1}{6n^{\frac{1}{4}}} M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right)$.

• Or :

$$\times 0 \xrightarrow[n \rightarrow +\infty]{} 0.$$

$$\times \frac{1}{6n^{\frac{1}{4}}} M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) \xrightarrow[n \rightarrow +\infty]{} 0 \quad \text{car :}$$

$$\blacktriangleright \frac{1}{6n^{\frac{1}{4}}} \xrightarrow[n \rightarrow +\infty]{} 0$$

$$\blacktriangleright M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right) \xrightarrow[n \rightarrow +\infty]{} M_{h'''} \left(\mathbb{E}(|X_1|^3) + \mathbb{E}(|Y_1|^3) \right)$$

On en conclut, par théorème d'encadrement : $\lim_{n \rightarrow +\infty} |\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))| = 0$. □

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

Démonstration.

• En appliquant la question 13.b)(ii) à la v.a.r. $X = T_n$, on obtient :

$$\mathbf{1}_{[T_n \leq x]} \leq g_n(T_n) \leq \mathbf{1}_{[T_n \leq x + a_n]}$$

Par croissance de l'espérance, on obtient alors :

$$\begin{array}{ccc} \mathbb{E}(\mathbf{1}_{[T_n \leq x]}) & \leq & \mathbb{E}(g_n(T_n)) \leq & \mathbb{E}(\mathbf{1}_{[T_n \leq x + a_n]}) \\ \parallel & & & \parallel \\ \mathbb{P}([T_n \leq x]) & & & \mathbb{P}([T_n \leq x + a_n]) \end{array}$$

On a bien : $\mathbb{P}([T_n \leq x]) \leq \mathbb{E}(g_n(T_n)) \leq \mathbb{P}([T_n \leq x + a_n])$. □

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

Démonstration.

• D'après la question 15.a)(ii), $T_n \hookrightarrow \mathcal{N}(0, 1)$. On en déduit :

$$\mathbb{P}([T_n \leq x]) = F_{T_n}(x) = \Phi(x) \quad \text{et} \quad \mathbb{P}([T_n \leq x + a_n]) = F_{T_n}(x + a_n) = \Phi(x + a_n)$$

- Cela permet de réécrire le résultat de la question précédente comme suit :

$$\begin{aligned} & \Phi(x) \leq \mathbb{E}(g_n(T_n)) \leq \Phi(x + a_n) \\ \text{Or :} & \\ \times \Phi(x) & \xrightarrow[n \rightarrow +\infty]{} \Phi(x). \\ \times \Phi(x + a_n) & \xrightarrow[n \rightarrow +\infty]{} \Phi(x + 0) = \Phi(x) \text{ par composition de limite car } \Phi \text{ est continue au point } x. \end{aligned}$$

On en conclut, par théorème d'encadrement : $\lim_{n \rightarrow +\infty} \mathbb{E}(g_n(T_n)) = 0$. □

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

Démonstration.

- Tout d'abord :

$$\begin{aligned} & \left| \mathbb{E}(g_n(Z_n)) - \Phi(x) \right| \\ = & \left| (\mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n))) + (\mathbb{E}(g_n(T_n)) - \Phi(x)) \right| \\ = & \left| \mathbb{E}(g_n(Z_n)) - \mathbb{E}(g_n(T_n)) \right| + \left| \mathbb{E}(g_n(T_n)) - \Phi(x) \right| \\ \xrightarrow[n \rightarrow +\infty]{} & \quad 0 \quad + \quad 0 \quad \quad \quad (d'après les questions \mathbf{15.f} et \mathbf{16.b}) \end{aligned}$$

$$\lim_{n \rightarrow +\infty} \left| \mathbb{E}(g_n(Z_n)) - \Phi(x) \right| = 0$$

- On en conclut : $\lim_{n \rightarrow +\infty} \mathbb{E}(g_n(Z_n)) - \Phi(x) = 0$.

Autrement dit, on a bien : $\lim_{n \rightarrow +\infty} \mathbb{E}(g_n(Z_n)) = \Phi(x)$.

Commentaire

- Profitons de cette dernière question pour faire un point sur cette **Partie 3**. Elle est relativement technique et contient plusieurs questions calculatoires. C'est le cas notamment des questions **12.a**) et des questions **14.a**) et **14.b**) qui consistent à manipuler des intégrales. Il peut être intéressant de repérer ces questions d'analyse qui peuvent être traitées indépendamment de tout le reste de l'énoncé.
- Cette partie contient des questions de difficultés variées. On ne peut que se féliciter que le découpage en sous-questions (comme dans le reste du sujet) soit très bien pensé ce qui permet de rendre la plupart des questions abordables (même si elles n'ont pas vocation à être toutes abordées). La principale critique que l'on peut émettre est la longueur de cette partie. Elle contient à elle seule 25 questions alors que les deux précédentes parties en contenaient déjà 48. Cela laisse un peu plus de 3 minutes pour réfléchir et rédiger chaque question. Il est évident qu'un tel sujet n'est pas conçu dans l'idée qu'un candidat puisse aborder toutes les questions. Avec une telle avalanche de questions, un candidat souhaitant se démarquer devra « faire du volume », c'est-à-dire traiter un grand nombre de questions. En optant pour un sujet aussi dense, les concepteurs courent le risque d'avoir à corriger des copies un peu bâclées, qui insistent sur les calculs et qui laissent de côté les aspects les plus théoriques. La **Partie 3** penche d'ailleurs parfois un peu dans ce sens puisque l'existence de la plupart des espérances des v.a.r. ne semble pas devoir être traitée. En agissant ainsi, il est à craindre que des candidats qui optent pour des stratégies visant à rédiger peu et sans trop se soucier des objets considérés puissent obtenir de très bons résultats.

Commentaire

- Finissons maintenant sur l'aspect mathématique de cette partie.
En question **13.c)**, on a démontré :

$$\mathbb{E}(g_n(Z_n + a_n)) \leq \mathbb{P}([Z_n \leq x]) \leq \mathbb{E}(g_n(Z_n))$$

L'énoncé précise alors qu'il s'agit maintenant de démontrer :

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

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

Si on y parvient, on peut alors conclure, par théorème d'encadrement :

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

ce qui est précisément le résultat du théorème central limite.

Notons que l'on démontre **uniquement** la deuxième limite. La première est laissé en suspens, ce qui a du sens au vu de la longueur de cette partie.

□