1 Théorèmes limites

Exercice 1.

[Corrigé] ★★★☆

On jette 3600 fois un dé équilibré. Minorer la probabilité que le nombre d'apparitions du numéro 1 soit compris entre 480 et 720.

Exercice 2.

[Corrigé] ★★★☆

En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que :

$$\forall x > 0, \int_0^x e^{-\frac{t^2}{2}} dt \geq \sqrt{\frac{\pi}{2}} \left(1 - \frac{1}{x^2}\right).$$

Exercice 3.

[Corrigé] ★★★☆

Déterminer le nombre de lancers d'un dé équilibré permettant d'affirmer avec un risque inférieur à 0,05 que la fréquence d'apparition du 6 au cours des lancers diffère de $\frac{1}{6}$ d'au plus 0,01 en appliquant l'inégalité de Bienaymé-Tchebychev ou le théorème central limite.

Exercice 4.

[Corrigé] ★★★☆

Le papa de Monsieur Haïdar (aussi appelé Monsieur Haïdar) effectue deux fois par jour, cinq jours par semaine, et pendant 46 semaines, un trajet en voiture dont la durée est une variable aléatoire qui suit une loi d'espérance 45 minutes et d'écart-type 10 minutes. On suppose que les durées des trajets sont indépendantes.

Déterminer une valeur approchée de la probabilité que Monsieur Haïdar (père) passe au moins 350 heures dans sa voiture au cours de l'année. Même question avec un écart-type de 15 minutes.

Exercice 5.

[Corrigé] ★★★☆

Soit x un réel et soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes, de même loi, d'espérance m et d'écart-type σ . Pour tout $n \in \mathbb{N}^*$, on note $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$.

1. On suppose ici que $x < m$. On pose $\varepsilon = m - x$.

a. Montrer que, pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(\bar{X}_n \leq x) \leq \frac{\sigma^2}{n\varepsilon^2}$.

b. En déduire $\lim_{n \rightarrow +\infty} \mathbb{P}(\bar{X}_n \leq x)$.

2. On suppose ici que $x > m$. Déterminer $\lim_{n \rightarrow +\infty} \mathbb{P}(\bar{X}_n \leq x)$.

3. On suppose ici que $x = m$. Déterminer $\lim_{n \rightarrow +\infty} \mathbb{P}(\bar{X}_n \leq x)$.

Exercice 6.

[Corrigé] ★★★

soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes, de même loi de Poisson de paramètre 1. Pour tout $n \in \mathbb{N}^*$, on note $Y_n = \sum_{k=1}^n X_k$.

1. Déterminer $\lim_{n \rightarrow +\infty} \mathbb{P}(Y_n \leq n)$.

2. En déduire un équivalent simple de $\sum_{k=0}^n \frac{n^k}{k!}$ lorsque n tend vers $+\infty$.

Exercice 7.

[Corrigé] ★★★

Soit X une variable aléatoire réelle d'espérance m et de variance σ^2 . Soit $\alpha \in \mathbb{R}_+^*$.

1. Calculer $\mathbb{E}((X - m + \lambda)^2)$ pour tout $\lambda > 0$.

2. Démontrer que, pour tout $\lambda > 0$, $\mathbb{P}(X - m \geq \alpha) \leq \frac{\sigma^2 + \lambda^2}{\alpha^2 + 2\lambda\alpha + \lambda^2}$

3. En déduire que $\mathbb{P}(X - m \geq \alpha) \leq \frac{\sigma^2}{\alpha^2 + \sigma^2}$.

4. Démontrer que $\mathbb{P}(|X - m| \geq \alpha) \leq \frac{2\sigma^2}{\alpha^2 + \sigma^2}$. Pour quelles valeurs de α obtient-on une meilleure inégalité que celle de Bienaymé-Tchebychev ?

2 Statistiques

Exercice 8.

[Corrigé] ★☆☆

1. Écrire une fonction `moyenne_normale` qui simule n fois la loi normale $\mathcal{N}(m, \sigma^2)$ et qui renvoie la moyenne des valeurs obtenues.

2. En déduire une fonction `confiance` qui, étant donné un réel strictement positif ε , simule N fois la moyenne de n variables aléatoires de loi normale $\mathcal{N}(m, \sigma^2)$ et qui renvoie la proportion des moyennes dans l'intervalle $[m - \varepsilon, m + \varepsilon]$.

Exercice 9.

[Corrigé] ★★★☆

On suppose que le poids d'un nouveau-né est une variable aléatoire suivant une loi d'écart-type 0,5 kg.

Le poids moyen des 49 enfants nés en janvier 2018 dans un hôpital parisien est de 3,6 kg.

1. Déterminer un intervalle de confiance à 95% pour le poids moyen d'un nouveau-né dans cet hôpital.

2. Quel serait le niveau de confiance d'un intervalle de longueur 0,1 kg centré en 3,6 pour ce poids moyen ?

Exercice 10.[\[Corrigé\]](#) ★★☆☆

Un appareil de télécommunication reçoit un signal stocké à chaque unité de temps dans une suite de variable $(X_n)_{n \in \mathbb{N}^*}$. Cet appareil doit détecter un signal effectif en le différenciant du bruit ambiant.

On suppose que le bruit est une suite de variables aléatoires suivant la loi normale centrée réduite et on suppose que pour un signal effectif, la moyenne n'est pas nulle.

Aujourd'hui, on a observé une suite de 40 valeurs $x_1, \dots, x_{40}$ de variance 1 et de moyenne empirique 0,6. S'agit-il de bruit ?

Construire un test pour répondre à cette question.

Exercice 11.[\[Corrigé\]](#) ★★★

On utilise une nouvelle variété de pommes de terre dans une exploitation agricole. Le rendement de l'ancienne variété était de 41,5 tonnes à l'hectare. La nouvelle est cultivée sur 100 hectares, avec un rendement moyen de 45 tonnes à l'hectare et un écart-type de 11,25 tonnes.

Faut-il, au vu de ce rendement, favoriser la culture de cette nouvelle variété ?

Construire un test pour répondre à cette question.

Exercice 12.[\[Corrigé\]](#) ★★★

Lors de la construction d'un collège accueillant 500 élèves, il est prévu la construction d'une cantine comprenant deux salles, chacune disposant de N places. On fait l'hypothèse que les élèves qui mangent à la cantine choisissent au hasard et de façon équiprobable l'une des deux salles, indépendamment les uns des autres.

On note X la variable aléatoire correspondant au nombre d'élèves choisissant la première salle.

1. Montrer que tous les étudiants trouvent une place si, et seulement si, $500 - N \leq X \leq N$.

2. Justifier qu'on peut considérer que la variable $Y = \frac{X - 250}{\sqrt{125}}$ suit une loi normale.

3. En déduire qu'il faut et il suffit que $N \geq 279$ pour que la probabilité que chaque élève trouve une place dans la salle qu'il a choisie soit supérieure à 0,99. *On rappelle que $\Phi(2,58) \approx 0,995$ où Φ désigne la fonction de répartition de la loi $\mathcal{N}(0,1)$.*

Exercice 13.[\[Corrigé\]](#) ★★☆☆

Soit $\theta \in \left[0, \frac{1}{10}\right]$. Dans un lot de 400 bouteilles de lait, choisies au hasard dans un stock important, on évalue le nombre de bouteilles de lait contaminées par une bactérie donnée.

Dans le stock complet, la probabilité de contamination d'une bouteille est θ .

1. On choisit de tester une à une les bouteilles, de manière indépendante. On note X_k la variable aléatoire égale à 1 si la k -ème bouteille est contaminée, 0 sinon.

a. Justifier que $Z = \frac{1}{400} \sum_{k=1}^{400} X_k$ admet pour espérance θ .

b. Déterminer $\mathbb{V}(Z)$ et vérifier que $\mathbb{V}(Z) \leq \frac{1}{1600}$.

c. Montrer que $\mathbb{P}(|Z - \theta| \geq 0,05) \leq 0,25$. Que peut-on en déduire ?

2. On regroupe désormais les bouteilles en 40 lots de 10. On regroupe une partie du contenu de chaque bouteille d'un lot, dont on teste la contamination.

- Si le lot est contaminé, on teste une à une toutes les bouteilles du lot.

- Si le lot n'est pas contaminé, on considère que les dix bouteilles du lot ne le sont pas.

On note Y_n la variable aléatoire de Bernoulli qui vaut 1 si le lot n est contaminé. On note T le nombre de tests effectués par cette méthode.

a. Déterminer $\mathbb{E}(Y_n)$ pour tout $n \in \llbracket 1, 40 \rrbracket$.

b. Justifier que $T = 40 + 10 \sum_{k=1}^{40} Y_k$. En déduire $\mathbb{E}(T)$.

Exercice 14.[\[Corrigé\]](#) ★★★

On veut étudier la proportion p de gens qui vont au cinéma chaque mois. On prend un échantillon de taille $n = 100$. Soit N le nombre de personnes dans l'échantillon qui vont au cinéma chaque mois.

1. Quelle est la loi de N ? Par quelle loi peut-on l'approcher ? Sous quelles conditions ?

Par quelle loi peut-on approcher la loi de $F = \frac{N}{n}$?

2. On observe une proportion $f = 0,1$ de personnes allant au cinéma chaque mois. Déterminer un intervalle de confiance asymptotique pour p au niveau 95%.

Corrigé de l'exercice 1. [Énoncé]

On note X le nombre d'apparitions du numéro 1. La variable aléatoire X correspond au nombre de succès ("obtenir un 1") lors de la répétition de 3600 expériences indépendantes de Bernoulli de paramètre $\frac{1}{6}$; elle suit donc la loi binomiale de paramètres $n = 3600$ et $p = \frac{1}{6}$. Puisque X admet une variance, on peut appliquer la formule de Bienaymé-Tchebychev :

$$\forall \varepsilon > 0, \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

En choisissant $\varepsilon = 121$, on trouve :

$$\mathbb{P}(|X - 600| \geq 121) \leq \frac{500}{121^2}.$$

Par passage au complémentaire, on trouve :

$$\mathbb{P}(|X - 600| < 121) \geq 1 - \frac{500}{121^2}.$$

Ainsi :

$$\mathbb{P}(|X - 600| \leq 120) \geq 1 - \frac{500}{121^2}.$$

i.e.

$$\mathbb{P}(480 \leq X \leq 720) \geq 1 - \frac{500}{121^2} \approx 0,966$$

Seconde rédaction (plus rapide) pour ceux qui ont plus d'intuition :

$$\begin{aligned} \mathbb{P}(480 \leq X \leq 720) &= \mathbb{P}(|X - 600| \leq 120) \\ &= \mathbb{P}(|X - 600| < 121) \\ &= 1 - \mathbb{P}(|X - \mathbb{E}(X)| \geq 121). \end{aligned}$$

En appliquant l'inégalité de Bienaymé-Tchebychev (X admet un moment d'ordre 2), on trouve :

$$\mathbb{P}(480 \leq X \leq 720) \geq 1 - \frac{\mathbb{V}(X)}{121^2} \geq 1 - \frac{500}{121^2} \approx 0,966$$

Corrigé de l'exercice 2. [Énoncé]

Soit x un réel strictement positif. et soit X une variable aléatoire suivant la loi normale centrée réduite. Puisque X admet une variance, on peut appliquer l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq x) \leq \frac{\mathbb{V}(X)}{x^2},$$

i.e. (les inégalités larges deviennent strictes puisque X est à densité) :

$$1 - \mathbb{P}(-x \leq X \leq x) \leq \frac{1}{x^2},$$

ou encore :

$$\mathbb{P}(-x \leq X \leq x) \geq 1 - \frac{1}{x^2},$$

Par parité d'une densité de la loi normale centrée réduite, on trouve :

$$\int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \mathbb{P}(0 \leq X \leq x) = \frac{1}{2} \mathbb{P}(-x \leq X \leq x) \geq \frac{1}{2} \left(1 - \frac{1}{x^2}\right).$$

En multipliant cette inégalité par $\sqrt{2\pi}$, on obtient l'inégalité souhaitée :

$$\forall x > 0, \int_0^x e^{-\frac{t^2}{2}} dt \geq \sqrt{\frac{\pi}{2}} \left(1 - \frac{1}{x^2}\right).$$

Corrigé de l'exercice 3. [Énoncé]

Notons, pour tout $n \in \mathbb{N}^*$, X_n la variable de Bernoulli égale à 1 si, et seulement si, on obtient 6 au n -ème lancer. On note alors, $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ pour tout $k \in \mathbb{N}^*$.

On cherche donc $n \in \mathbb{N}^*$ tel que $\mathbb{P}\left(\left|\overline{X}_n - \frac{1}{6}\right| \geq 0,01\right) \leq 0,05$.

- Puisque $\mathbb{E}(\overline{X}_n) = \frac{1}{6}$ (par linéarité de l'espérance) et $\mathbb{V}(\overline{X}_n) = \frac{5}{36n}$ (par indépendance des lancers), on trouve via l'inégalité de Bienaymé-Tchebychev :

$$\forall n \in \mathbb{N}^*, \mathbb{P}\left(\left|\overline{X}_n - \frac{1}{6}\right| \geq 0,01\right) \leq \frac{5}{0,01^2 \times 36n}.$$

Or :

$$\frac{5}{0,01^2 \times 36n} \leq 0,05 \Leftrightarrow n \geq \frac{10^6}{36} \Leftrightarrow n \geq 27778 \text{ (car } n \in \mathbb{N}^* \text{)}.$$

Pour un nombre de lancers supérieur ou égal à 27778, la fréquence d'apparition du 6 au cours des lancers diffère de $\frac{1}{6}$ d'au plus 0,01, avec un risque inférieur à 0,05.

- Puisque les variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ sont indépendantes, de même loi, d'espérance $\frac{1}{6}$ et de variance $\frac{5}{36}$, on peut approcher la loi de la variable aléatoire $Z_n = \frac{\overline{X}_n - \frac{1}{6}}{\sqrt{\frac{5}{36n}}}$

par la loi normale centrée réduite. Ainsi, en notant Φ la fonction de répartition de la loi normale centrée réduite :

$$\begin{aligned} \mathbb{P}\left(\left|\overline{X}_n - \frac{1}{6}\right| \geq 0,01\right) &= 1 - \mathbb{P}\left(\left|\overline{X}_n - \frac{1}{6}\right| < 0,01\right) \\ &= 1 - \mathbb{P}\left(-0,01 \times \sqrt{\frac{36n}{5}} < Z_n < 0,01 \times \sqrt{\frac{36n}{5}}\right) \\ &= 1 - \left(\Phi\left(0,01 \times \sqrt{\frac{36n}{5}}\right) - \Phi\left(-0,01 \times \sqrt{\frac{36n}{5}}\right)\right) \\ &= 1 - \left(2\Phi\left(0,01 \times \sqrt{\frac{36n}{5}}\right) - 1\right) \\ &= 2\left(1 - \Phi\left(0,01 \times \sqrt{\frac{36n}{5}}\right)\right). \end{aligned}$$

Or :

$$\begin{aligned} 2\left(1 - \Phi\left(0,01 \times \sqrt{\frac{36n}{5}}\right)\right) \leq 0,05 &\Leftrightarrow \Phi\left(0,01 \times \sqrt{\frac{36n}{5}}\right) \geq 0,975 \\ &\Leftrightarrow 0,01 \times \sqrt{\frac{36n}{5}} \geq 1,96 \\ &\Leftrightarrow n \geq \frac{5}{36} \times 196^2 \approx 5335,5. \end{aligned}$$

Pour un nombre de lancers supérieur ou égal à 5336, la fréquence d'apparition du 6 au cours des lancers diffère de $\frac{1}{6}$ d'au plus 0,01, avec un risque inférieur à 0,05.

Corrigé de l'exercice 4. [Énoncé]

On dénombre 460 trajets annuels. On note alors $X_1, \dots, X_{460}$ les variables aléatoires égales aux temps de trajets successifs (en minutes) de Monsieur Haïdar. On note $\overline{X} = \frac{1}{460} \sum_{k=1}^{460} X_k$ la moyenne empirique des variables $(X_k)_{1 \leq k \leq 460}$. On cherche une valeur approchée de $\mathbb{P}(460\overline{X} \geq 350 \times 60)$.

1. On sait que, pour tout $k \in \llbracket 1, 460 \rrbracket$, $\mathbb{E}(X_k) = 45$ et $\sigma(X_k) = 10$.

Les variables aléatoires $(X_k)_{1 \leq k \leq 460}$ étant indépendantes identiquement distribuées et admettant une espérance et une variance non nulle, on peut appliquer le théorème central

limite : on peut approcher la loi de la variable $\frac{\overline{X} - 45}{\frac{10}{\sqrt{460}}}$ par la loi normale centrée réduite.

Grâce à la table de la loi normale centrée réduite (on note Φ sa fonction de répartition), on trouve :

$$\mathbb{P}(460\overline{X} \geq 350 \times 60) = \mathbb{P}\left(\frac{\overline{X} - 45}{\frac{10}{\sqrt{460}}} \geq \frac{\frac{2100}{46} - 45}{\frac{10}{\sqrt{460}}}\right) \approx 1 - \Phi(1,4) \approx 0,0808.$$

2. On sait que, pour tout $k \in \llbracket 1, 460 \rrbracket$, $\mathbb{E}(X_k) = 45$ et $\sigma(X_k) = 15$. Par le même raisonnement, on trouve :

$$\mathbb{P}(460\overline{X} \geq 350 \times 60) = \mathbb{P}\left(\frac{\overline{X} - 45}{\frac{15}{\sqrt{460}}} \geq \frac{\frac{2100}{46} - 45}{\frac{15}{\sqrt{460}}}\right) \approx 1 - \Phi(0,93) \approx 0,1762.$$

Corrigé de l'exercice 5. [Énoncé]

1. On suppose ici que $x < m$. On pose $\varepsilon = m - x$.

a. Soit $n \in \mathbb{N}^*$. Puisque $[\overline{X}_n - m \leq -\varepsilon] \subset [|\overline{X}_n - m| \geq \varepsilon]$, on a :

$$\mathbb{P}(\overline{X}_n \leq x) = \mathbb{P}(\overline{X}_n - m \leq -\varepsilon) \leq \mathbb{P}(|\overline{X}_n - m| \geq \varepsilon)$$

La moyenne empirique $\overline{X}_n$ admet un moment d'ordre 2 : $\mathbb{E}(\overline{X}_n) = m$ et $\mathbb{V}(\overline{X}_n) = \frac{\sigma^2}{n}$. On peut donc appliquer l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}(\overline{X}_n \leq x) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

b. Puisque $\lim_{n \rightarrow +\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$, et :

$$\forall n \in \mathbb{N}^*, 0 \leq \mathbb{P}(\overline{X}_n \leq x) \leq \frac{\sigma^2}{n\varepsilon^2},$$

on en déduit que $\lim_{n \rightarrow +\infty} \mathbb{P}(\overline{X}_n \leq x)$ par encadrement.

2. On raisonne de la même manière qu'aux deux questions précédentes. Remarquons que $-\varepsilon > 0$ et $[|\overline{X}_n - m| < -\varepsilon] \subset [\overline{X}_n - m \leq -\varepsilon]$. Ainsi :

$$\mathbb{P}(\overline{X}_n \leq x) = \mathbb{P}(\overline{X}_n - m \leq -\varepsilon) \geq \mathbb{P}(|\overline{X}_n - m| < -\varepsilon) = 1 - \mathbb{P}(|\overline{X}_n - m| \geq -\varepsilon).$$

En appliquant l'inégalité de Bienaymé-Tchebychev, on trouve :

$$1 \geq \mathbb{P}(\overline{X_n} \leq x) \geq 1 - \frac{\sigma^2}{n(-\varepsilon)^2}.$$

On conclut par encadrement que $\lim_{n \rightarrow +\infty} \mathbb{P}(\overline{X_n} \leq x) = 1$.

3. Notons Φ la fonction de répartition de la loi normale centrée réduite.

On peut appliquer le théorème central limite puisque les variables $(X_n)_{n \in \mathbb{N}^*}$ sont indépendamment distribuées et indépendantes, et admettent un moment d'ordre 2 :

$$\mathbb{P}(\overline{X_n} \leq x) = \mathbb{P}(\overline{X_n} - m \leq 0) = \mathbb{P}\left(\frac{\overline{X_n} - m}{\frac{\sigma}{\sqrt{n}}} \leq 0\right) \xrightarrow{n \rightarrow +\infty} \Phi(0) = \frac{1}{2}.$$

On en déduit que $\lim_{n \rightarrow +\infty} \mathbb{P}(\overline{X_n} \leq x) = \frac{1}{2}$.

Corrigé de l'exercice 6. [Énoncé]

1. Soit $n \in \mathbb{N}^*$. Remarquons que $\frac{Y_n}{n}$ est la moyenne empirique des variables aléatoires $(X_k)_{1 \leq k \leq n}$. Notons Φ la fonction de répartition de la loi normale centrée réduite.

On peut appliquer le théorème central limite puisque les variables $(X_n)_{n \in \mathbb{N}^*}$ sont indépendamment distribuées et indépendantes, et admettent une espérance et une variance (égales à 1) :

$$\mathbb{P}(Y_n \leq n) = \mathbb{P}\left(\frac{Y_n}{n} - 1 \leq 0\right) = \mathbb{P}\left(\frac{Y_n - 1}{\frac{1}{\sqrt{n}}} \leq 0\right) \xrightarrow{n \rightarrow +\infty} \Phi(0) = \frac{1}{2}.$$

On en déduit que $\lim_{n \rightarrow +\infty} P(Y_n \leq n) = \frac{1}{2}$.

2. On sait que la variable Y_n suit la loi de Poisson de paramètre n en tant que somme de variable aléatoires indépendantes suivant des lois de Poisson. Ainsi :

$$\mathbb{P}(Y_n \leq n) = \sum_{k=0}^n \mathbb{P}(Y_n = k) = \sum_{k=0}^n \frac{n^k}{k!} e^{-n} = e^{-n} \sum_{k=0}^n \frac{n^k}{k!}$$

D'après la question précédente, $\mathbb{P}(Y_n \leq n) \underset{n \rightarrow +\infty}{\sim} \frac{1}{2}$ donc :

$$\sum_{k=0}^n \frac{n^k}{k!} \underset{n \rightarrow +\infty}{\sim} \frac{e^n}{2}.$$

Corrigé de l'exercice 7. [Énoncé]

1. Soit $\lambda > 0$. La variable aléatoire X admet une variance donc $(X - m + \lambda)$ admet un moment d'ordre 2 et :

$$\mathbb{E}((X - m + \lambda)^2) = \mathbb{E}((X - m)^2) + 2\lambda\mathbb{E}(X - m) + \lambda^2 = \sigma^2 + \lambda^2.$$

2. Soit $\lambda > 0$. La variable aléatoire $(X - m + \lambda)^2$ est positive et admet une espérance, donc, d'après l'inégalité de Markov, on trouve :

$$\begin{aligned} \mathbb{P}(X - m \geq \alpha) &= \mathbb{P}(X - m + \lambda \geq \alpha + \lambda) \\ &\leq \mathbb{P}((X - m + \lambda)^2 \geq (\alpha + \lambda)^2) \\ &\leq \frac{\sigma^2 + \lambda^2}{(\alpha + \lambda)^2} = \frac{\sigma^2 + \lambda^2}{\alpha^2 + 2\lambda\alpha + \lambda^2}. \end{aligned}$$

3. L'étude de la fonction $f : \lambda \mapsto \frac{\sigma^2 + \lambda^2}{(\alpha + \lambda)^2}$ montre que f admet un minimum global en

$$\lambda = \frac{\sigma^2}{\alpha}, \text{ égal à } f\left(\frac{\sigma^2}{\alpha}\right) = \frac{\sigma^2}{\alpha^2 + \sigma^2}.$$

Puisque $\mathbb{P}(X - m \geq \alpha) \leq f(\lambda)$ pour tout $\lambda > 0$, on en déduit bien que $P(X - m \geq \alpha) \leq \frac{\sigma^2}{\alpha^2 + \sigma^2}$.

4. On peut montrer de la même manière que $\mathbb{P}(m - X \geq \alpha) \leq \frac{\sigma^2}{\alpha^2 + \sigma^2}$. On en déduit que :

$$\mathbb{P}(|X - m| \geq \alpha) = \mathbb{P}(X - m \geq \alpha) + \mathbb{P}(m - X \geq \alpha) \leq \frac{2\sigma^2}{\alpha^2 + \sigma^2}.$$

Cette majoration est meilleure que celle de l'inégalité que celle de Bienaymé-Tchebychev si, et seulement si $\frac{2\sigma^2}{\alpha^2 + \sigma^2} < \frac{\sigma^2}{\alpha^2}$, i.e. $\alpha < \sigma$.

Corrigé de l'exercice 8. [Énoncé]

```
1. 

---


import random as rd

def moyenne_normale(m, sigma, n):
    s = 0
    for k in range(n):
        s += rd.gauss(mu, sigma)
    return s/n


---


```

```

import random as rd

def confiance(m, sigma, n, N, e):
    nb = 0
    for k in range(N):
        if m - e <= moyenne_normale(m, sigma, n) <= m + e:
            nb += 1
    return nb/N

```

Exercice 9.[\[Corrigé\]](#) ★★☆

2. Le cours nous donne directement la réponse ; en notant $\bar{x}_{49} = 3,6$, $s_{49} = 0,5$, $\alpha = 0,05$ et $u_{1-\frac{\alpha}{2}} = 1,96$ (le quantile d'ordre 97,5% de la loi normale centrée réduite), un intervalle de confiance à 95% pour le poids moyen d'un nouveau-né dans cet hôpital est donné par :

$$\left[\bar{x}_{49} - u_{1-\frac{\alpha}{2}} \frac{s_{49}}{\sqrt{49}}, \bar{x}_{49} + u_{1-\frac{\alpha}{2}} \frac{s_{49}}{\sqrt{49}} \right] = [3,46; 3,74]$$

2. Notons m le poids moyen théorique des enfants nés dans cet hôpital, $\bar{X}_{49}$ la moyenne empirique des poids de 49 nouveaux-nés et Φ la fonction de répartition de la loi normale $\mathcal{N}(0,1)$.

$$\begin{aligned} \mathbb{P}(\bar{X}_{49} - 0,05 \leq m \leq \bar{X}_{49} + 0,05) &= \mathbb{P}\left(-\frac{0,05}{\frac{s_{49}}{\sqrt{49}}} \leq \frac{m - \bar{X}_{49}}{\frac{s_{49}}{\sqrt{49}}} \leq \frac{0,05}{\frac{s_{49}}{\sqrt{49}}}\right) \\ &= \Phi\left(\frac{0,05}{\frac{s_{49}}{\sqrt{49}}}\right) - \Phi\left(-\frac{0,05}{\frac{s_{49}}{\sqrt{49}}}\right) \\ &= 2\Phi\left(\frac{0,05}{\frac{s_{49}}{\sqrt{49}}}\right) - 1 \\ &= 2\Phi(0,7) - 1 \\ &= 0,516 \end{aligned}$$

Le niveau de confiance d'un intervalle de longueur 0,1 kg centré en 3,6 pour ce poids moyen serait alors de 51,6%.

Corrigé de l'exercice 10. [\[Énoncé\]](#)

On souhaite savoir si l'échantillon des 40 valeurs reçues est un bruit. On construit pour cela un test de conformité à la moyenne (d'un bruit).

En notant μ l'espérance (ou moyenne théorique) du signal reçu aujourd'hui, l'hypothèse nulle est $(H_0) : \mu = 0$ et l'hypothèse alternative est $\mu \neq 0$.

On fixe le risque $\alpha = 5\%$. Le quantile $u_{1-\frac{\alpha}{2}}$ de la loi normale centrée réduite est alors 1,96. Notons $\bar{x}_{40} = 0,6$ et $s_{40} = 1$. L'intervalle de confiance de la moyenne observé au risque α est :

$$\left[\bar{x}_{40} - 1,96 \frac{s_{40}}{\sqrt{40}}, \bar{x}_{40} + 1,96 \frac{s_{40}}{\sqrt{40}} \right] \approx [0,29; 0,91].$$

Puisque 0 (l'espérance théorique d'un bruit) n'appartient à cet intervalle de confiance, on rejette l'hypothèse (H_0) avec un risque égal à 5%.

On peut donc affirmer qu'il ne s'agit pas d'un bruit, avec une probabilité d'erreur inférieure ou égale à 5%.

Corrigé de l'exercice 11. [\[Énoncé\]](#)

On cherche à tester $(H_0) : m = 41,5$ contre l'hypothèse $(H_1) : m \neq 41,5$.

On considère pour cela la statistique de test $Z = \frac{\bar{X}_n - m}{\frac{S_n}{\sqrt{n}}}$ où :

- $\bar{X}_n$ désigne la moyenne empirique des $n = 100$ hectares cultivés avec la nouvelle variété et S_n leur écart-type empirique.
- m désigne le rendement théorique (espérance) de la nouvelle variété.

Le théorème central limite assure qu'on peut approcher la loi de Z par la loi normale centrée réduite. On sait que $\mathbb{P}(-1,96 \leq Z \leq 1,96) \approx 0,95$ (d'après une table de la loi normale centrée réduite. Or on obtient $z_{obs} = 3,11 \notin [-1,96, 1,96]$. On rejette donc l'hypothèse (H_0) : la nouvelle variété a un rendement significativement différent de l'ancienne. On privilégiera donc la nouvelle variété.

Corrigé de l'exercice 12. [\[Énoncé\]](#)

1. Tous les étudiants (au nombre de X) ayant choisi la première salle trouvent une place si, et seulement si, $X \leq N$. Tous les étudiants (au nombre de $500 - X$) ayant choisi la seconde salle trouvent une place si, et seulement si, $500 - X \leq N$, i.e. $500 - N \leq X$. Ainsi, tous les étudiants trouvent une place si, et seulement si, $500 - N \leq X \leq N$.

2. La variable aléatoire X suit la loi binomiale de paramètres $n = 500$ et $p = \frac{1}{2}$. En effet, X est le nombre de succès ("choisir la première salle") lors de la répétition de 500 expériences indépendantes de Bernoulli de probabilité de succès $\frac{1}{2}$ (les choix des salles sont équiprobables). On en déduit immédiatement que $\mathbb{E}(X) = np = 250$ et $\sigma(X) = \sqrt{np(1-p)} = \sqrt{125}$. Puisque $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, on en déduit,

d'après le théorème de Moivre-Laplace, qu'on peut approcher la loi de X par la loi normale d'espérance 250 et d'écart-type $\sqrt{125}$. Sous cette approximation, $Y = \frac{X - 250}{\sqrt{125}} = X^*$ suit la loi normale centrée réduite.

3. Remarquons que :

$$\begin{aligned} \mathbb{P}(500 - N \leq X \leq N) &= \mathbb{P}\left(\frac{250 - N}{\sqrt{125}} \leq \frac{X - 250}{\sqrt{125}} \leq \frac{N - 250}{\sqrt{125}}\right) \\ &= \mathbb{P}\left(-\frac{N - 250}{\sqrt{125}} \leq Y \leq \frac{N - 250}{\sqrt{125}}\right) \\ &= \Phi\left(\frac{N - 250}{\sqrt{125}}\right) - \Phi\left(-\frac{N - 250}{\sqrt{125}}\right) \\ &= 2\Phi\left(\frac{N - 250}{\sqrt{125}}\right) - 1. \end{aligned}$$

Ainsi :

$$\begin{aligned} \mathbb{P}(500 - N \leq X \leq N) \geq 0,99 &\Leftrightarrow 2\Phi\left(\frac{N - 250}{\sqrt{125}}\right) - 1 \geq 0,99 \\ &\Leftrightarrow \Phi\left(\frac{N - 250}{\sqrt{125}}\right) \geq 0,995 \\ &\Leftrightarrow \frac{N - 250}{\sqrt{125}} \geq 2,58 \\ &\Leftrightarrow N \geq 2,58\sqrt{125} + 250 \approx 278,8. \end{aligned}$$

Puisque le nombre de places est entier, il faut et il suffit que $N \geq 279$ pour que la probabilité que chaque élève trouve une place dans la salle qu'il a choisie soit supérieure à 0,99.

Corrigé de l'exercice 13. [\[Énoncé\]](#)

1. a. Pour tout $k \in \llbracket 1, 400 \rrbracket$, la variable aléatoire X_k suit la loi de Bernoulli de paramètre θ par définition, donc $\mathbb{E}(X_k) = \theta$.

La moyenne empirique de $(X_k)_{1 \leq k \leq 400}$ est donc un estimateur sans biais de θ .

b. La variable aléatoire Z est finie (car $X_1, \dots, X_{400}$ le sont) ; elle admet donc une variance.

Par propriétés de la variance et indépendances des variables $(X_k)_{1 \leq k \leq 400}$, on trouve :

$$\mathbb{V}(Z) = \frac{1}{400^2} \sum_{k=1}^{400} \mathbb{V}(X_k) = \frac{\theta(1-\theta)}{400}$$

L'étude de la fonction $\varphi : x \mapsto x(1-x)$ sur $[0, 1]$ montre qu'elle atteint un maximum en $\frac{1}{2}$, égal à $\frac{1}{4}$. On en déduit que $\mathbb{V}(Z) \leq \frac{1}{1600}$.

c. La variable Z admet une variance donc, d'après l'inégalité de Bienaymé-Tchebychev, on a :

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq 0,05) \leq \frac{\mathbb{V}(Z)}{0,05^2},$$

c'est-à-dire :

$$\mathbb{P}(|Z - \theta| \geq 0,05) \leq \frac{1}{1600 \times 0,05^2} = 0,25.$$

On en déduit un intervalle de confiance de θ au niveau 75% : $[Z - 0,05; Z + 0,05]$.

2. a. Soit $n \in \llbracket 1, 40 \rrbracket$. La variable aléatoire Y_n suit une loi de Bernoulli puisqu'elle ne prend que 0 et 1 comme valeur.

Par indépendance (supposée) des contaminations de bouteilles d'un même lot, $\mathbb{P}(Y_n = 0) = (1 - \theta)^{10}$. On en déduit donc que $Y_n \hookrightarrow \mathcal{B}(1 - (1 - \theta)^{10})$ et ainsi que $\mathbb{E}(Y_n) = 1 - (1 - \theta)^{10}$.

b. Soit $n \in \llbracket 1, 40 \rrbracket$. Pour le lot n , on fait un test de contamination, puis 0 si le test est négatif, et 10 tests dans le cas contraire. On fait donc $1 + 10Y_n$ pour le lot n . Ainsi :

$$T = \sum_{k=1}^{40} (1 + 10Y_k) = 40 + 10 \sum_{k=1}^{40} Y_k.$$

Par linéarité de l'espérance, T admet une espérance, égale à :

$$\mathbb{E}(T) = 40 + 10 \sum_{k=1}^{40} \mathbb{E}(Y_k) = 440 - 400(1 - \theta)^{10}.$$

Corrigé de l'exercice 14. [\[Énoncé\]](#)

1. La variable aléatoire N est le nombre de succès ("la personne sondée va au cinéma") lors de la répétition de n expériences de Bernoulli (qu'on peut supposer indépendantes) de même paramètre p .

La variable aléatoire N suit donc la loi binomiale de paramètres n et p .

Le théorème de Moivre-Laplace assure donc qu'on peut approcher la loi de N par la loi normale d'espérance np et de variance $np(1-p)$. On considère souvent que cette approximation est valable dès que $n \geq 30$, $np > 5$ et $n(1-p) > 5$. On peut alors approcher la loi de $F = \frac{N}{n}$ par la loi normale d'espérance $\mathbb{E}(F) = \mathbb{E}\left(\frac{N}{n}\right) = p$ et de variance

$$\mathbb{V}(F) = \mathbb{V}\left(\frac{N}{n}\right) = \frac{\mathbb{V}(N)}{n^2} = \frac{p(1-p)}{n}.$$

2. D'après la question précédente et une table de la loi normale, on a :

$$\mathbb{P} \left(-1,96 \leq \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}} \leq 1,96 \right) = 0,95.$$

Ainsi :

$$\mathbb{P} \left(F - 1,96\sqrt{\frac{p(1-p)}{n}} \leq p \leq F + 1,96\sqrt{\frac{p(1-p)}{n}} \right) = 0,95.$$

Puisque $p(1-p) \leq \frac{1}{4}$ (ce qu'on peut prouver en étudiant la fonction $x \mapsto x(1-x)$), on en déduit que :

$$\mathbb{P} \left(p \in \left[F - 1,96\sqrt{\frac{p(1-p)}{n}}, F + 1,96\sqrt{\frac{p(1-p)}{n}} \right] \right) \leq \mathbb{P} \left(p \in \left[F - \frac{1,96}{2\sqrt{n}}, F + \frac{1,96}{2\sqrt{n}} \right] \right)$$

On en déduit que l'intervalle

$$\left[f - \frac{1,96}{2\sqrt{n}}, f + \frac{1,96}{2\sqrt{n}} \right] = [0,002; 0,198]$$

est un intervalle de confiance asymptotique pour p à un niveau de confiance supérieur ou égal à 95%.