⧫ Corrigés des exercices du chapitre 16

Exercice 1. On lance 3600 fois un dé équilibré. En utilisant l'inégalité de Bienaymé-Tchebychef, minorer la probabilité que le nombre d'apparitions du nombre 1 soit compris entre 480 et 720,

Solution. La variable aléatoire X égale au nombre d'apparition de 1 lors des 3600 lancers suit une loi binomiale de paramètres 3600 et $\frac{1}{6}$ donc $\mathbf{E}(X) = 3600 \times \frac{1}{6} = 600$ et $\mathbf{V}(X) = 3600 \times \frac{1}{6} \times \frac{5}{6} = 500$.

Or,

$$\mathbf{P}(480 \leqslant X \leqslant 720) = \mathbf{P}(-120 \leqslant X - 600 \leqslant 120) = \mathbf{P}(|X - 600| \leqslant 120)$$

donc $\mathbf{P}(480 \leqslant X \leqslant 720) = 1 - \mathbf{P}(|X - \mathbf{E}(X)| > 120)$. Comme X est une variable discrète, on en déduit que $\mathbf{P}(480 \leqslant X \leqslant 720) = 1 - \mathbf{P}(|X - \mathbf{E}(X)| \geqslant 121)$ et, d'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbf{P}(|X - \mathbf{E}(X)| \geqslant 121) \leqslant \frac{\mathbf{V}(X)}{121^2} = \frac{500}{121^2}$$

donc

$$\boxed{\mathbf{P}(480 \leqslant X \leqslant 720) \geqslant 1 - \frac{500}{121^2} = \frac{14141}{14641}}.$$

Exercice 2. Soit X une variable aléatoire telle que e^X admet une espérance. Soit $t \in \mathbb{R}$.

1. Justifier que $\{X \geqslant t\} = \{\mathrm{e}^X \geqslant \mathrm{e}^t\}$.
2. En déduire, en utilisant l'inégalité de Markov, que $\mathbf{P}(X \geqslant t) \leqslant \mathbf{E}(e^{X-t})$.

Solution.

1. Comme la fonction *exp* est croissante sur $\mathbb{R}$, pour tout réel x, $x \geqslant t$ est équivalent à $\mathrm{e}^x \geqslant \mathrm{e}^t$ donc $\boxed{\{X \geqslant t\} = \{\mathrm{e}^X \geqslant \mathrm{e}^t\}}$.
2. Ainsi, $\mathbf{P}(X \geqslant t) = \mathbf{P}(\mathrm{e}^X \geqslant \mathrm{e}^t)$ et, comme e^X est une variable aléatoire positive admettant une espérance, d'après l'inégalité de Markov,

$$\mathbf{P}(\mathrm{e}^X \geqslant \mathrm{e}^t) \leqslant \frac{\mathbf{E}(\mathrm{e}^X)}{\mathrm{e}^t}.$$

Mais, par linéarité de l'espérance,

$$\frac{\mathbf{E}(\mathrm{e}^X)}{\mathrm{e}^t} = \frac{1}{\mathrm{e}^t}\mathbf{E}(\mathrm{e}^X) = \mathbf{E}\left(\frac{1}{\mathrm{e}^t}\mathrm{e}^X\right) = \mathbf{E}\left(\frac{\mathrm{e}^X}{\mathrm{e}^t}\right) = \mathbf{E}(\mathrm{e}^{X-t}).$$

Ainsi, on conclut que $\boxed{\mathbf{P}(X \geqslant t) \leqslant \mathbf{E}(e^{X-t})}$.

Exercice 3. Un institut de sondage a été missionné pour estimer la proportion p de végétariens en France. Il interroge pour cela n français. Puisque le choix des sondés s'effectue sur une population très grande, on admet que l'expérience peut s'apparenter à une suite de n tirages indépendants avec remise. On note X_n la variable aléatoire égale au nombre de végétariens sondés et on souhaite quantifier à quel point la fréquence $F_n = \dfrac{X_n}{n}$ approche la proportion p.

1. Déterminer la loi de X_n.
2. En considérant la fonction $f : x \longmapsto x(1-x)$, montrer que $p(1-p) \leqslant \frac{1}{4}$.
3. Montrer que, pour tout $\varepsilon > 0$, $\mathbf{P}\left(|F_n - p| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}$.

4. En déduire une condition sur n pour que F_n soit une approximation de p à 10^{-2} près avec une probabilité supérieure ou égale à 95%.

Solution.

1. Comme on assimile les choix des sondés à des tirages indépendants avec remise, X_n suit une loi binomiale de paramètre n et p.

2. La fonction $f : x \longmapsto x(1-x) = -x^2 + x$ est une fonction polynôme du second degré dont le coefficient dominant $a = -1$ est négatif donc f atteint son maximum sur $\mathbb{R}$ en $x_0 = -\frac{1}{2\times(-1)} = \frac{1}{2}$ et ce maximum vaut $f(\frac{1}{2}) = \frac{1}{4}$. Ainsi, pour tout réel x, $x(1-x) \leqslant \frac{1}{4}$ donc, en particulier, $\boxed{p(1-p) \leqslant \frac{1}{4}}$.

3. Ainsi, $\mathbf{E}(X_n) = np$ et $\mathbf{V}(X_n) = np(1-p)$. Soit $\varepsilon > 0$. Remarquons que, comme $n > 0$,

$$|F_n - p| \geqslant \varepsilon \iff n\left|\frac{X_n}{n} - p\right| \geqslant n\varepsilon \iff |X_n - np| \geqslant n\varepsilon \iff |X_n - \mathbf{E}(X_n)| \geqslant n\varepsilon$$

donc, par l'inégalité de Bienaymé-Tchebychev,

$$\mathbf{P}\left(|F_n - p| \geqslant \varepsilon\right) = \mathbf{P}(X_n - \mathbf{E}(X_n) \geqslant n\varepsilon) \leqslant \frac{\mathbf{V}(X_n)}{(n\varepsilon)^2} = \frac{np(1-p)}{n^2\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}.$$

On conclut, grâce à la question précédente, que

$$\boxed{\mathbf{P}\left(|F_n - p| \geqslant \varepsilon\right) \geqslant \frac{1}{4n\varepsilon^2}}.$$

4. F_n est une approximation de p à 10^{-2} près si et seulement si $|F_n - p| \leqslant 10^{-2}$. On cherche donc une condition sur n pour que $\mathbf{P}(|F_n - p| \leqslant 10^{-2}) \geqslant 0{,}95$ c'est-à-dire, en passant au complémentaire, $\mathbf{P}(|F_n - p| > 10^{-2}) \leqslant 0{,}05$.
Or, comme $\{|F_n - p| > 10^{-2}\} \subset \{|F_n - p| \geqslant 10^{-2}\}$, d'après la question précédente, $\mathbf{P}\left(|F_n - p| > \varepsilon\right) \leqslant \dfrac{1}{4n(10^{-2})} = \dfrac{10^4}{4n} = \dfrac{2500}{n}$. donc, pour que la condition voulue soit remplie, il suffit que $\dfrac{2500}{n} \leqslant 0{,}05$ i.e. $n \geqslant \dfrac{2500}{0{,}05} = 50000$.
Ainsi, pour que F_n soit une valeur approchée de p à 10^{-2} avec une probabilité au moins égale à 95%, <u>il suffit d'interroger 50000 personnes</u> (ce qui est tout de même beaucoup !).

Exercice 4. On définit une variable aléatoire S en exécutant le programme Python suivant :

```python
from random import random
S = 0
for i in range(200):
  X=random()
  S+=X
```

1. a. À chaque tour de boucle, quelle est la loi de X ?

b. Rappeler son espérance et calculer sa variance.

2. Calculer l'espérance et la variance de la variable aléatoire S.

3. En utilisant l'inégalité de Bienaymé-Tchebychev, minorer la probabilité de l'évènement $S \in [90, 110]$.

Solution.

1. a. À chaque tour de boucle, la variable aléatoire suit une loi uniforme sur $[0\,;1[$.

b. Par théorème, $\mathbf{E}(X) = \dfrac{0+1}{2}$ soit $\boxed{\mathbf{E}(X) = \dfrac{1}{2}}$. De plus, par le théorème de transfert, comme la densité de X est égale à 1 sur $[0\,;1[$ et 0 ailleurs,

$$\mathbf{E}(X^2) = \int_0^1 t^2 \times 1\,\mathrm{d}t = \left[\frac{t^3}{3}\right]_0^1 = \frac{1}{3}.$$

Ainsi, par la formule de König-Huygens, $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \dfrac{1}{3} - \left(\dfrac{1}{2}\right)^2 = \dfrac{1}{3} - \dfrac{1}{4}$ soit $\boxed{\mathbf{V}(X) = \dfrac{1}{12}}$.

2. La variable aléatoire S est la somme de 100 variables suivant toutes la même loi uniforme sur $[0\,;1[$ donc, par linéarité de l'espérance, $\mathbf{E}(S) = 200 \times \dfrac{1}{2}$ soit $\boxed{\mathbf{E}(S) = 100}$ et, de plus, comme ces 200 variables sont indépendantes, $\mathbf{V}(X) = 200 \times \dfrac{1}{12}$ soit $\boxed{\mathbf{V}(S) = \dfrac{50}{3}}$.

3. Comme S est une variable aléatoire continue,

$$\mathbf{P}(90 \leqslant S \leqslant 110) = \mathbf{P}(90 < S < 110) = \mathbf{P}(-10 < S - 100 < 10) = \mathbf{P}(|S - \mathbf{E}(S)| < 10)$$

donc, en passant au complémentaire,

$$\mathbf{P}(90 \leqslant S \leqslant 110) = 1 - \mathbf{P}(|S - \mathbf{E}(S)| \geqslant 10).$$

Or, comme S admet une variance, d'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbf{P}(|S - \mathbf{E}(S)| \geqslant 10) \leqslant \frac{\mathbf{V}(S)}{10^2} = \frac{1}{6}.$$

On conclut donc que $\boxed{\mathbf{P}(90 \leqslant S \leqslant 110) \geqslant \dfrac{5}{6}}$.

Exercice 5. Soit $n \in \mathbb{N}^*$ et $p \in \,]0\,;1[$. On considère une variable aléatoire X suivant une loi binomiale $\mathscr{B}(n,p)$. On pose $Y = \frac{X}{n} - p$.

1. Montrer que $\mathbf{E}(|Y|)^2 \leqslant \mathbf{E}(Y^2)$.

2. En considérant $\mathbf{V}(Y)$, montrer $\mathbf{E}(Y^2) = \dfrac{p(1-p)}{n}$.

3. En appliquant l'inégalité de Markov, déduire des questions précédentes que pour tout réel $\varepsilon > 0$,

$$\mathbf{P}\left(\left|\frac{X}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{\sqrt{p(1-p)}}{\varepsilon\sqrt{n}}.$$

Solution.

1. Comme $\mathbf{E}(Y^2) - \mathbf{E}(|Y|)^2 = \mathbf{E}(|Y|^2) - \mathbf{E}(|Y|)^2 = V(|Y|) \geqslant 0$, $\boxed{\mathbf{E}(|Y|)^2 \leqslant \mathbf{E}(Y^2)}$.

2. Par linéarité de l'espérance, $\mathbf{E}(Y) = \frac{1}{n}\mathbf{E}(X) - p = \frac{np}{n} - p = 0$ donc $\mathbf{V}(Y) = \mathbf{E}(Y^2) - \mathbf{E}(Y)^2 = \mathbf{E}(Y^2)$. Or, par propriété, $\mathbf{V}(Y) = \dfrac{1}{n^2}V(X) = \dfrac{p(1-p)}{n}$ donc $\boxed{\mathbf{E}(Y^2) = \dfrac{p(1-p)}{n}}$.

3. En appliquant l'inégalité de Markov à $|Y|$ est qui positive et qui admet une espérance, il vient
$$\mathbf{P}\left(\left|\frac{X}{n}-p\right| \geqslant \varepsilon\right)=\mathbf{P}(|Y| \geqslant \varepsilon) \leqslant \frac{\mathbf{E}(|Y|)}{\varepsilon}$$
et, comme $\mathbf{E}(|Y|)^2 \leqslant \mathbf{E}(Y^2)$, $\mathbf{E}(|Y|) \leqslant \sqrt{\mathbf{E}(Y^2)}=\sqrt{\dfrac{p(1-p)}{n}}$ donc
$$\mathbf{P}(|Y| \geqslant \varepsilon) \leqslant \frac{\sqrt{\frac{p(1-p)}{n}}}{\varepsilon}=\frac{\sqrt{p(1-p)}}{\varepsilon\sqrt{n}}$$
i.e.
$$\boxed{\mathbf{P}\left(\left|\frac{X}{n}-p\right| \geqslant \varepsilon\right) \leqslant \frac{\sqrt{p(1-p)}}{\varepsilon\sqrt{n}}}.$$

Exercice 6. Soit X une variable aléatoire suivant une loi de Poisson de paramètre $\lambda>0$.

1. Montrer que $\mathbf{P}(X \geqslant 2\lambda) \leqslant \mathbf{P}((X-\lambda+1)^2 \geqslant (\lambda+1)^2)$.

2. En déduire que $\mathbf{P}(X \geqslant 2\lambda) \leqslant \dfrac{1}{\lambda+1}$.

Solution.

1. Si $X \geqslant 2\lambda$ alors $X-\lambda+1 \geqslant \lambda+1$ et, comme $\lambda>0$, ces deux nombres sont positifs donc, par croissance de la fonction carré sur $\mathbb{R}_+$, $(X-\lambda+1)^2 \geqslant (\lambda+1)^2$.
Ainsi, $\{X \geqslant 2\lambda\} \subset \{(X-\lambda+1)^2 \geqslant (\lambda+1)^2\}$ donc, par croissance de la probabilité $\mathbf{P}$, $\boxed{\mathbf{P}(X \geqslant 2\lambda) \leqslant \mathbf{P}((X-\lambda+1)^2 \geqslant (\lambda+1)^2)}$.

2. Comme X suit une loi de Poisson de paramètre λ, par linéarité, $X-\lambda+1$ admet une espérance et $\mathbf{E}(X-\lambda+1)=\mathbf{E}(X)-\lambda+1=1$. De plus, $X-\lambda+1$ admet une variance et $\mathbf{V}(X-\lambda+1)=\mathbf{V}(X)=\lambda$. Ainsi, par la formule de König-Huygens, $(X-\lambda+1)^2$ admet une espérance et $\mathbf{E}((X-\lambda+1)^2)=\mathbf{V}(X-\lambda+1)+\mathbf{E}(X-\lambda+1)^2=\lambda+1^2=\lambda+1$.
Ainsi, en appliquant l'inégalité de Markov à la variable aléatoire positive $(X-\lambda+1)^2$, il vient
$$\mathbf{P}((X-\lambda+1)^2 \geqslant (\lambda+1)^2) \leqslant \frac{\lambda+1}{(\lambda+1)^2}=\frac{1}{\lambda+1}$$
et on déduit alors de la question précédente que
$$\boxed{\mathbf{P}(X \geqslant 2\lambda) \leqslant \frac{1}{\lambda+1}}.$$

Exercice 7. Soit un entier $n \geqslant 2$. On effectue n lancers d'une pièce équilibrée et on définit, pour tout $k \in [\![1,n]\!]$, une variable aléatoire X_k égale à 1 si le k-ième lancer donne pile, et 0 sinon. On pose également
$$S_n=X_1+\cdots+X_n \quad \text{et} \quad M_n=\frac{S_n}{n}.$$

1. Quelle est la loi suivie par X_k pour tout $k \in [\![1,n]\!]$? En déduire la loi de S_n.

2. Déterminer l'espérance et la variance de S_n et de M_n.

3. a. À l'aide de l'inégalité de Bienaymé-Tchebychev, minorer la probabilité que M_n se trouve entre 0,4 et 0,6

b. Au bout de combien de lancers la probabilité que M_n se trouve entre 0,4 et 0,6 est au moins égales à 95% ?

4. Dans la suite, on considère que S_n suit approximativement une loi $\mathcal{N}(\mu, \sigma^2)$.

a. Donner les valeurs des paramètres μ et σ^2.

b. On pose $S_n^* = \dfrac{S_n - \mu}{\sigma}$. Quelle est la loi de S_n^* ?

c. Vérifier que

$$\mathbf{P}(0,4 \leqslant M_n \leqslant 0,6) = \mathbf{P}\left(-0,2\sqrt{n} \leqslant S_n^* \leqslant 0,2\sqrt{n}\right).$$

d. Pour quelle valeur de n cette probabilité est-elle environ égale à 95% ?

e. Comparer avec la réponse à la question **3.b.**.

Solution.

1. Par définition, pour tout $k \in [\![1, n]\!]$, $\boxed{X_k \text{ suit une loi de Bernoulli de paramètre } \frac{1}{2}}$ (car la pièce est équilibrée).
Comme les lancers sont indépendants, les variables X_k sont indépendantes et donc $\boxed{S_n \text{ suit une loi binomiale de paramètres } n \text{ et } \frac{1}{2}}$.

2. Par propriétés, $\boxed{\mathbf{E}(S_n) = \frac{n}{2} \text{ et } \mathbf{E}(S_n) = \frac{n}{4}}$. Dès lors, par linéarité, $\mathbf{E}(M_n) = \dfrac{1}{n}\mathbf{E}(S_n)$ donc $\boxed{\mathbf{E}(M_n) = \frac{1}{2}}$ et, par propriété, $\mathbf{V}(M_n) = \dfrac{1}{n^2}\mathbf{V}(S_n)$ donc $\boxed{\mathbf{V}(M_n) = \frac{1}{4n}}$.

3. a. On remarque que

$$\mathbf{P}(0{,}4 \leqslant M_n \leqslant 0{,}6) = \mathbf{P}(-0{,}1 \leqslant M_n - 0{,}5 \leqslant 0{,}1) = \mathbf{P}(|M_n - \mathbf{E}(M_n)| \leqslant 0{,}1)$$

donc $\mathbf{P}(0{,}4 \leqslant M_n \leqslant 0{,}6) = 1 - \mathbf{P}(|M_n - \mathbf{E}(M_n)| > 0{,}1)$. Or, $\{|M_n - \mathbf{E}(M_n)| > 0{,}1\} \subset |M_n - \mathbf{E}(M_n)| \geqslant 0{,}1\}$ donc $\mathbf{P}(|M_n - \mathbf{E}(M_n)| > 0{,}1) \leqslant \mathbf{P}(|M_n - \mathbf{E}(M_n)| \geqslant 0{,}1)$. De plus, d'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbf{P}(|M_n - \mathbf{E}(M_n)| \geqslant 0{,}1) \leqslant \frac{\mathbf{V}(M_n)}{0{,}1^2} = \frac{\frac{1}{4n}}{0{,}01} = \frac{25}{n}$$

donc

$$\boxed{\mathbf{P}(0{,}4 \leqslant M_n \leqslant 0{,}6) \geqslant 1 - \frac{25}{n}}.$$

b. On cherche n tel que $\mathbf{P}(0{,}4 \leqslant M_n \leqslant 0{,}6) \geqslant 0{,}95$. D'après la question précédente, il suffit pour cela que $1 - \dfrac{25}{n} \geqslant 0{,}95$ i.e. que $\dfrac{25}{n} \leqslant 0{,}05$ soit $n \geqslant 500$.
Ainsi, <u>au bout de 500 lancers, M_n est compris entre 0,4 et 0,6 avec une probabilité au moins égale à 95%</u>.

4. a. $\boxed{\mu = \mathbf{E}(S_n) = \frac{n}{2}}$ et $\boxed{\sigma^2 = \mathbf{V}(S_n) = \frac{n}{4}}$.

b. Par propriété, $\boxed{S_n^* \text{ suit une loi } \mathcal{N}(0, 1)}$.

c. Étant donné que

$$\begin{aligned}0{,}4 \leqslant M_n \leqslant 0{,}6 &\Longleftrightarrow 0{,}4 \leqslant \frac{S_n}{n} \leqslant 0{,}6 \Longleftrightarrow 0{,}4n \leqslant S_n \leqslant 0{,}6n \\ &\Longleftrightarrow -0{,}1n \leqslant S_n - 0{,}5n \leqslant 0{,}1n \Longleftrightarrow \frac{-0{,}1n}{\sqrt{\frac{n}{4}}} \leqslant \frac{S_n - 0{,}5n}{\sqrt{\frac{n}{4}}} \leqslant \frac{0{,}1n}{\sqrt{\frac{n}{4}}} \\ &\Longleftrightarrow -0{,}1n \times \frac{2}{\sqrt{n}} \leqslant \frac{S_n - \mu}{\sigma} \leqslant 0{,}1n \times \frac{2}{\sqrt{n}} \Longleftrightarrow -0{,}2\sqrt{n} \leqslant S_n^* \leqslant 0{,}2\sqrt{n}\end{aligned}$$

on a bien $\boxed{\mathbf{P}\left(0,4 \leqslant M_n \leqslant 0,6\right) = \mathbf{P}\left(-0,2\sqrt{n} \leqslant S_n^* \leqslant 0,2\sqrt{n}\right)}$.

d. D'après le tableau du cours, $\mathbf{P}(-u \leqslant S_n^* \leqslant n) \approx 0{,}95$ pour $u = 1{,}96$. Or,

$$0{,}2\sqrt{n} = 1{,}96 \Longleftrightarrow \sqrt{n} = 9{,}8 \Longleftrightarrow n = 9{,}8^2$$

soit $\boxed{n \approx 96}$.

e. On trouve une valeur très sensiblement inférieur à celle de la question **3.b.**, ce qui illustre la faiblesse de l'inégalité de Bienaymé-Tchebychev par rapport au théorème central limite.

Exercice 8. On fixe un réel $a > 0$. Soit n variables aléatoires X_1, X_2, ..., X_n indépendantes qui suivent la loi uniforme sur $[0, a]$. On se propose d'estimer la valeur de a de deux manières différentes.

1. Soit $k \in [\![1, n]\!]$. Déterminer la densité, la fonction de répartition, l'espérance et la variance de X_k.

2. Estimateur de la moyenne

On pose $M_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$.

a. Calculer l'espérance et la variance de M_n.

b. En déduire une variable Y_n telle que $\mathbf{E}\left(Y_n\right) = a$, et calculer $\mathbf{V}\left(Y_n\right)$.

c. Soit $\varepsilon > 0$. Démontrer que $\mathbf{P}\left(|Y_n - a| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$.

3. Estimateur du maximum

On pose $U_n = \max\left(X_1, X_2, \ldots, X_n\right)$.

a. Déterminer la fonction de répartition de U_n et en déduire une densité de U_n.

b. Démontrer que U_n admet une espérance et une variance et les calculer.

c. En déduire une variable Z_n telle que $\mathbf{E}\left(Z_n\right) = a$, et calculer $\mathbf{V}\left(Z_n\right)$.

d. Soit $\varepsilon > 0$. Démontrer que $\mathbf{P}\left(|Z_n - a| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$.

4. Comparaison des deux estimateurs Comparer les variances de Y_n et Z_n. Lequel de ces deux estimateurs devrait converger le plus vite vers a ?

Solution.

1. La densité de X_k est la fonction définie sur $\mathbb{R}$ par

$$\boxed{\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} 1 & \text{si } x \in [0\,;a] \\ 0 & \text{sinon} \end{cases}}.$$

Sa fonction de répartition F est définie, pour tout réel x, par $F(x) = \mathbf{P}(X \leqslant x)$. Comme f est nulle sur $]-\infty\,;0[$, $F(x) = 0$ pour tout $x < 0$. Ensuite, si $x \in [0\,;a]$ alors,

$$F(x) = \int_0^x \frac{1}{a}\,\mathrm{d}t = \frac{x}{a}$$

et si $x > a$ alors

$$F(x) = \int_0^a \frac{1}{a}\,\mathrm{d}t = \frac{a}{a} = 1.$$

Ainsi, F est la fonction définie sur $\mathbb{R}$ par

$$\boxed{\forall x \in \mathbb{R}, \quad F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{x}{a} & \text{si } x \in [0\,;a] \\ 1 & \text{si } x > a \end{cases}}.$$

Par théorème, $\mathbf{E}(X_k) = \dfrac{a-0}{2}$ donc $\boxed{\mathbf{E}(X_k) = \dfrac{a}{2}}$. Enfin, par le théorème de transfert, X_k^2 admet une espérance et

$$\mathbf{E}(X_k^2) = \int_0^a t^2 \times \frac{1}{a}\,\mathrm{d}t = \left[\frac{t^3}{3a}\right]_0^a = \frac{a^2}{3}$$

donc, par la formule de König-Huygens, X_k admet une variance est

$$\mathbf{V}(X_k) = \mathbf{E}(X_k^2) - \mathbf{E}(X_k)^2 = \frac{a^2}{3} - \left(\frac{a}{2}\right)^2$$

soit $\boxed{\mathbf{V}(X_k) = \dfrac{a^2}{12}}$.

2. a. Par linéarité, $\mathbf{E}(M_n) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbf{E}(X_k) = \dfrac{1}{n}\sum_{k=1}^{n} \dfrac{a}{2} = \dfrac{1}{n} \times n\dfrac{a}{2}$ donc $\boxed{\mathbf{E}(X_k) = \dfrac{a}{2}}$.

Comme les variables X_1, X_2, ..., X_n sont indépendantes, $\mathbf{V}(M_n) = \dfrac{1}{n^2}\displaystyle\sum_{k=1}^{n} \mathbf{V}(X_k) = \dfrac{1}{n^2}\sum_{k=1}^{n} \dfrac{a^2}{12} = \dfrac{1}{n^2} \times n\dfrac{a^2}{12}$ donc $\boxed{\mathbf{V}(X_k) = \dfrac{a^2}{12n}}$.

b. Posons $Y_n = 2M_n$. Par linéarité, $\mathbf{E}(Y_n) = 2\mathbf{E}(M_n)$ donc $\boxed{\mathbf{E}(Y_n) = a}$. De plus, $\mathbf{V}(Y_n) = 2^2\mathbf{V}(M_n) = 4 \times \dfrac{a^2}{12n}$ donc $\boxed{\mathbf{V}(Y_n) = \dfrac{a^2}{3n}}$.

c. Comme Y_n admet une variance, d'après l'inégalité de Bienaymé-Tchebychev, pour tout $\varepsilon > 0$,

$$\mathbf{P}(|Y_n - a| \geqslant \varepsilon) \leqslant \frac{\mathbf{V}(Y_n)}{\varepsilon^2} = \frac{a^2}{3n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

donc, comme $\mathbf{P}(|Y_n - a| \geqslant \varepsilon) \geqslant 0$, on en déduit que $\boxed{\mathbf{P}\left(|Y_n - a| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.}$.

3. a. Soit $x \in \mathbb{R}$. Alors, $U_n \leqslant x$ si et seulement si chacun des variables X_1, X_2, ..., X_n est inférieures à x. Autrement dit, $\{U_n \leqslant x\} = \displaystyle\bigcap_{k=1}^{n} \{X_k \leqslant x\}$. Comme X_1, X_2, ..., X_n sont indépendantes, on en déduit que

$$\mathbf{P}(U_n \leqslant x) = \mathbf{P}(X_k \leqslant x)\mathbf{P}(X_2 \leqslant x)\cdots\mathbf{P}(X_n \leqslant x)$$

Or, pour tout $k \in [\![1,n]\!]$, $\mathbf{P}(X_k \leqslant x) = F(x)$ donc $\mathbf{P}(U_n \leqslant x) = F(x)^n$. Ainsi, la fonction de répartition de U_n est la fonction G définie sur $\mathbb{R}$ par :

$$\boxed{\forall x \in \mathbb{R}, \quad G(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{x^n}{a^n} & \text{si } x \in [0\,;a] \\ 1 & \text{si } x > a \end{cases}}.$$

b. La fonction G est dérivable sur chacun des intervalles $]-\infty\,;0[$, $]0\,;a[$ et $]a\,;+\infty[$. De plus, pour tout $x \in \,]-\infty\,;0[\,\cup\,]a\,;+\infty[$, $G'(x) = 0$ et, pour tout $x \in \,]0\,;a[$, $G'(x) = \dfrac{nx^{n-1}}{a^n}$.
On en déduit qu'une densité de U_n est

$$\boxed{\forall x \in \mathbb{R}, \quad g(x) = \begin{cases} \dfrac{nx^{n-1}}{a^n} & \text{si } x \in [0\,;a] \\ 0 & \text{sinon} \end{cases}}.$$

c. Comme la densité de U_n est nulle en dehors du segment $[0\,;a]$, U_n admet un espérance et

$$\mathbf{E}(U_n) = \int_0^a x \times \frac{nx^{n-1}}{a^n}\,\mathrm{d}x = \int_0^a \frac{nx^n}{a^n}\,\mathrm{d}x \left[\frac{nx^{n+1}}{(n+1)a^n}\right]_0^a$$

soit $\boxed{\mathbf{E}(U_n) = \dfrac{na}{n+1}}$.
De même, U_n^2 admet une espérance d'après le théorème de transfert et $[0\,;a]$, U_n admet un espérance et

$$\mathbf{E}(U_n^2) = \int_0^a x^2 \times \frac{nx^{n-1}}{a^n}\,\mathrm{d}x = \int_0^a \frac{nx^{n+1}}{a^n}\,\mathrm{d}x \left[\frac{nx^{n+2}}{(n+2)a^n}\right]_0^a$$

soit $\mathbf{E}(U_n) = \dfrac{na^2}{n+2}$ donc, par le formule de König-Huygens, U_n admet une variance et

$$\mathbf{V}(U_n) = \mathbf{V}(U_n^2) - \mathbf{V}(U_n)^2 = \frac{na^2}{n+2} - \left(\frac{na}{n+1}\right)^2 = na^2\frac{(n+1)^2 - n(n+2)}{(n+1)^2(n+2)}$$

i.e. $\boxed{\mathbf{V}(U_n) = \dfrac{na^2}{(n+1)^2(n+2)}}$.

d. Posons $Z_n = \dfrac{n+1}{n}U_n$. Alors, par linéarité, $\mathbf{E}(Z_n) = \dfrac{n+1}{n}\mathbf{E}(U_n)$ donc $\boxed{\mathbf{E}(Z_n) = a}$.
De plus, $\mathbf{V}(Z_n) = \left(\dfrac{n+1}{n}\right)^2 \mathbf{V}(U_n)$ donc $\boxed{\mathbf{E}(Z_n) = \dfrac{a^2}{n(n+2)}}$.

e. Comme Z_n admet une variance, d'après l'inégalité de Bienaymé-Tchebychev, pour tout $\varepsilon > 0$,

$$\mathbf{P}(|Z_n - a| \geqslant \varepsilon) \leqslant \frac{\mathbf{V}(Z_n)}{\varepsilon^2} = \frac{a^2}{n(n+2)\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

donc, comme $\mathbf{P}(|Z_n - a| \geqslant \varepsilon) \geqslant 0$, on en déduit que $\boxed{\mathbf{P}\left(|Z_n - a| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.}$.

4. Pour tout $n \geqslant 2$, $n+2 > 3$ donc $\dfrac{a^2}{n(n+2)} < \dfrac{a^2}{3n}$ i.e. $\boxed{\mathbf{V}(Z_n) < \mathbf{V}(Y_n)}$.

Au voisinage de $+\infty$, $\mathbf{V}(Z_n) \sim \dfrac{a^2}{n^2}$ et $\mathbf{V}(Y_n) \sim \dfrac{a^2}{3n}$ donc $\mathbf{V}(Z_n)$ converge (infiniment) plus vite vers 0 que $\mathbf{V}(Y_n)$ et ainsi $\boxed{Z_n \text{ converge plus vite vers } a \text{ que } Y_n}$.

Exercice 9. On dispose d'une pièce de monnaie censée être bien équilibrée. On a effectué avec cette pièce 100 séries de lancers en s'arrêtant, pour chaque série, à la première apparition de *face*. On a obtenu les résultats suivants :

rang d'apparition du premier *face*	1	2	3	4	5	6	7	8	9
nombre de séries	48	20	14	3	7	4	2	0	2

On note X_k la variable aléatoire égale au nombre de lancers nécessaires pour obtenir le premier *face* lors de la k-ième série.

1. Déterminer, pour tout $k \in [\![1, 100]\!]$, la loi de X_k et préciser son espérance et sa variance.

2. En utilisant les résultats du tableau précédent, peut-on considérer, au niveau de confiance 95%, que la pièce est équilibrée ?

Solution.

1. La variable aléatoire X_k représente le premier succès dans un schéma de Bernoulli donc elle suit une loi géométrique de paramètre $\frac{1}{2}$. Ainsi, $\boxed{\mathbf{E}(X_k) = \frac{1}{\frac{1}{2}} = 2}$ et $\boxed{\mathbf{V}(X_k) = \frac{1-\frac{1}{2}}{(\frac{1}{2})^2} = 2}$.

2. Notons μ l'espérance commune des variables aléatoires X_k. On teste l'hypothèse H_0 : « $\mu = 2$ » contre l'hypothèse H_1 : « $\mu \neq 2$ » au niveau de de confiance 95%.
Ici, la valeur de M_n est

$$\frac{1}{100}(49 \times 1 + 20 \times 2 + 14 \times 3 + 3 \times 4 + 7 \times 5 + 4 \times 6 + 2 \times 7 + 2 \times 9) = 2{,}33$$

donc M_n^* prend la valeur $\dfrac{2{,}33 - 2}{\frac{\sqrt{2}}{\sqrt{100}}} \approx 2{,}3$. Or, pour le seuil de confiance 95%, $u_{0{,}05} = 1{,}96$ donc $2{,}3 \notin [-u_{0{,}05}\,; u_{0{,}05}]$. Ainsi, $\boxed{\text{au seuil de confiance 95\%, on rejette } H_0 \text{ contre } H_1}$ et on considère que la pièce n'est pas équilibrée.

Exercice 10. À Boston, en 1986, le docteur Benjamin Spock, militant contre la guerre du Vietnam, fut jugé pour incitation publique à la désertion. Le juge chargé de l'affaire était soupçonné de ne pas être équitable dans la sélection des jurés : parmi les 700 personnes qu'il avait désignées comme jurés lors de ses procès précédents, il y avait 15% de femmes alors que, sur l'ensemble de la ville, 29% des jurés éligibles étaient de femmes.

1. Au seuil de confiance 95%, peut-on considérer que le juge est impartial dans la désignation des jurés ?

2. Si ce juge avait moins d'expérience et qu'il n'avait désigné, avec les mêmes pourcentages, que 40 jurés lors de ses précédents procès, la conclusion serait-elle la même ?

Solution.

1. Notons p la proportion de femmes dans les jurys désignés par le juge. On veut tester l'hypthèse H_0 : « $p = 0{,}29$ » contre l'hypothèse H_1 : « $p \neq 0{,}29$ ».
Pour chaque jury désigné par le juge, on note X la variable aléatoire de Bernoulli dont le succès est « la juré est une femme ». Sous l'hypothèse H_0, X suit une loi de Bernoulli de paramètre 0,29. Dès lors, sur l'ensemble des $n = 700$ jurés, la variable S_n comptant le nombre de femmes suit une loi binomiale de paramètres 700 et 0,29. Ainsi, $M_{700}^* = S_{700}^* = \dfrac{S_{700} - 0{,}29 \times 700}{\sqrt{700 \times 0{,}29 \times 0{,}71}}$. Ici, la valeur observée de S_{700}^* est donc

$$\frac{700 \times \frac{15}{100} - 0{,}29 \times 700}{\sqrt{700 \times 0{,}29 \times 0{,}71}} \approx -8{,}16.$$

Pour le seuil de confiance 95%, $u_{0,05} = 1{,}96$ donc la valeur observée de S^*_{700} n'appartient pas $[-u_{0,05}\,;u_{0,05}]$. Ainsi, $\boxed{\text{au risque d'erreur de 5\%, on rejette l'hypothèse } H_0 \text{ contre l'hypothèse } H_1}$. Ceci montre que le juge est bien partial dans son choix de jurés.

2. Si $n = 40$, la valeur de S^*_{40} observée devient :

$$\frac{40 \times \frac{15}{100} - 0{,}29 \times 40}{\sqrt{40 \times 0{,}29 \times 0{,}71}} \approx -1{,}95 \in [-u_{0,05}\,;u_{0,05}]\,.$$

Dans ce cas, $\boxed{\text{on accepte l'hypothèse } H_0}$ et la partialité du juge n'est pas établie.

Exercice 11. En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que, pour tout réel $x > 0$,

$$\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t \geqslant \sqrt{2\pi}\left(1 - \frac{1}{2x^2}\right).$$

Solution. Soit un réel $x > 0$. Considérons une variable aléatoire X suivant une loi $\mathcal{N}(0,1)$. Alors, d'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbf{P}(|X| \geqslant x) \leqslant \frac{1}{x^2}.$$

Or,

$$\mathbf{P}(|X| \geqslant x) = \mathbf{P}(\{X \geqslant x\} \cup \{X \leqslant -x\})$$

et, comme ces deux évènements sont incompatibles,

$$\mathbf{P}(|X| \geqslant x) = \mathbf{P}(X \geqslant x) + \mathbf{P}(X \leqslant -x).$$

De plus, par symétrie de la densité de X par rapport à l'axe des ordonnées, $\mathbf{P}(X \geqslant x) = \mathbf{P}(X \leqslant -x)$ donc

$$\mathbf{P}(|X| \geqslant x) = 2\mathbf{P}(X \geqslant x) = 2(1 - \mathbf{P}(X < x)).$$

Or, par définition

$$\mathbf{P}(X < x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Ainsi,

$$2\left(1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t\right) \leqslant \frac{1}{x^2}$$

donc

$$1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t \leqslant \frac{1}{2x^2}$$

soit encore

$$1 - \frac{1}{2x^2} \leqslant \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$$

ce qui permet de conclure que

$$\boxed{\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t \geqslant \sqrt{2\pi}\left(1 - \frac{1}{2x^2}\right)}.$$

Exercice 12. On considère une variable aléatoire réelle X à valeurs strictement positives. On suppose, de plus, que X admet une espérance $m > 0$ et une variance V. On considère une suite de variables aléatoires mutuellement indépendantes $(X_n)_{n\in\mathbb{N}^*}$ ayant toutes la même loi que X et on pose, pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^{n} X_k$.

On se donne un réel $\ell > 0$. Le but de l'exercice est de démontrer que

$$\lim_{n\to+\infty} \mathbf{P}(S_n \leqslant \ell) = 0.$$

1. Soit $n \in \mathbb{N}^*$. Justifier que, pour tout $\varepsilon > 0$,

$$\mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{V}{n\varepsilon^2}.$$

2. Soit $n \in \mathbb{N}^*$. Justifier que, pour tout $\varepsilon > 0$, $\{S_n \leqslant n(m-\varepsilon)\} \subset \left\{\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right\}$.

3. Soit $n \in \mathbb{N}^*$. En choisissant convenablement ε, déduire des questions précédentes que,

$$\mathbf{P}\left(S_n \leqslant \frac{nm}{2}\right) \leqslant \frac{4V}{m^2 n}.$$

4. Justifier l'existence d'un entier N tel que, pour tout $n \geqslant N$, $\frac{nm}{2} \geqslant \ell$ puis conclure.

Solution.

1. Soit $\varepsilon > 0$. Par linéarité, $\mathbf{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbf{E}(S_n) = \frac{1}{n} \times nm = m$ et, comme X_1, X_2, ..., X_n sont indépendantes, $\mathbf{V}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathbf{V}(S_n) = \frac{1}{n^2} \times nV = \frac{V}{b}$. Ainsi, en appliquant l'inégalité de Bienaymé-Tchebychev à la variable aléatoire $\frac{S_n}{n}$, il vient :

$$\mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\frac{V}{n}}{\varepsilon^2}$$

i.e.

$$\boxed{\mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{V}{n\varepsilon^2}}.$$

2. Soit $\varepsilon > 0$. Si $S_n \leqslant n(m-\varepsilon)$ alors $\frac{S_n}{n} - m \leqslant -\varepsilon < 0$ donc $\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon$. Ainsi, $\boxed{\{S_n \leqslant n(m-\varepsilon)\} \subset \left\{\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right\}}$.

3. Par croissance de la probabilité, on déduit des deux précédentes que, pour tout $\varepsilon > 0$,

$$\mathbf{P}(S_n \leqslant n(m-\varepsilon)) \leqslant \mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{V}{n\varepsilon^2}.$$

En particulier, pour $\alpha = \frac{m}{2} > 0$, on obtient

$$\boxed{\mathbf{P}\left(S_n \leqslant \frac{nm}{2}\right) \leqslant \frac{4V}{m^2 n}}.$$

4. Comme $m > 0$, $\lim\limits_{n \to +\infty} \dfrac{nm}{2} = +\infty$ donc il existe un entier N tel que, pour tout $n \geqslant N$, $\dfrac{nm}{2} \geqslant \ell$ et, ainsi, $\{S_n \leqslant \ell\} \subset \left\{S_n \leqslant \dfrac{nm}{2}\right\}$. Dès lors, pour tout $n \geqslant N$,

$$0 \leqslant \mathbf{P}(S_n \leqslant \ell) \leqslant \frac{4V}{m^2 n}.$$

Or, $\lim\limits_{n \to +\infty} \dfrac{4V}{m^2 n} = 0$ donc, par le théorème d'encadrement, on conclut que

$$\boxed{\lim_{n \to +\infty} \mathbf{P}(S_n \leqslant \ell) = 0}.$$