

## ACTIVITES 13. Variable aléatoire, loi des grands nombres

**[D]** Une **variable aléatoire**  $X$  sur l'univers  $\Omega = \{e_1; \dots; e_r\}$  est une fonction définie sur  $\Omega$  et à valeurs dans  $\mathbb{R}$ .

On note  $X(\Omega)$  l'ensemble des valeurs prises par  $X$ .

### exemple

On lance un dé, on s'intéresse au chiffre indiqué par la face supérieure du dé :  $\Omega = \{1; 2; 3; 4; 5; 6\}$ . Si on obtient un chiffre pair, alors on perd 3€, et si on obtient un chiffre impair, alors on gagne 5€. Le gain algébrique  $X$ , exprimé en €, est une variable aléatoire telle que :  $X(2) = X(4) = X(6) = -3$ ,  $X(1) = X(3) = X(5) = 5$ . L'ensemble des valeurs prises par  $X$  est  $X(\Omega) = \{-3; 5\}$ .

### **[D]** quelques définitions

- la **loi de probabilité** de  $X$  est donnée par le tableau suivant :

| valeurs $x_i$ prises par $X$ | $x_1$ | ... | $x_n$ |
|------------------------------|-------|-----|-------|
| $P(X = x_i)$                 | $p_1$ | ... | $p_n$ |

où  $p_1 + p_2 + \dots + p_n = 1$

- l'**espérance mathématique**  $E(X)$  de  $X$  est :

$$\blacktriangleright E(X) = \sum_{i=1}^n p_i x_i = \sum_{i=1}^n P(X = x_i) x_i$$

$E(X)$  est la moyenne théorique des valeurs prises par  $X$  lorsque l'expérience est répétée un très grand nombre de fois

- si  $X$  est une variable aléatoire prenant les valeurs  $x_1, \dots, x_n$  avec les probabilités  $p_1, \dots, p_n$ , et si  $a$  et  $b$  sont des constantes réelles, alors  $aX + b$  est la variable aléatoire qui prend les valeurs  $ax_1 + b, \dots, ax_n + b$  avec les mêmes probabilités  $p_1, \dots, p_n$  et on a :

$$\blacktriangleright E(aX + b) = aE(X) + b$$

**[i]** un jeu est **équitable**  $\stackrel{\text{déf}}{\Leftrightarrow} E(X) = 0$  (où  $X$  est le gain du joueur)

- la **variance** de  $X$  est le réel positif ou nul :

$$\blacktriangleright V(X) = \sum_{i=1}^n p_i (x_i - E(X))^2 = \sum_{i=1}^n P(X = x_i) (x_i - E(X))^2$$

- l'écart type de  $X$  est le réel positif ou nul :  $\blacktriangleright \sigma(X) = \sqrt{V(X)}$

### **[F]** quelques formules

- formule de **König - Huygens**

$$\blacktriangleright V(X) = \sum_{i=1}^n p_i x_i^2 - [E(X)]^2 = E(X^2) - [E(X)]^2$$

- pour  $a$  et  $b$  constantes réelles :

$$\blacktriangleright V(aX + b) = a^2 V(X), \text{ d'où } \sigma(aX + b) = |a| \sigma(X)$$

### **[D]** somme de deux variables

Soient  $X$  et  $Y$  deux variables aléatoires définies sur un même univers  $\Omega$ . Alors  $S = X + Y$  est la variable aléatoire définie sur  $\Omega$  par :

$$\forall \omega \in \Omega, S(\omega) = X(\omega) + Y(\omega)$$

**A01** On lance une pièce et un dé classique.

La face « pile » de la pièce porte le nombre 1, l'autre porte le nombre 2. On note  $X$  le nombre indiqué par la pièce,  $Y$  celui indiqué par le dé, et on pose :  $S = X + Y$ .

On admet que :  $E(X) = \frac{3}{2}$ ,  $V(X) = \frac{1}{4}$ ,  $E(Y) = \frac{7}{2}$  et  $V(Y) = \frac{35}{12}$ .

Calculer  $E(S)$  et  $V(S)$ , les comparer avec  $E(X) + E(Y)$  et  $V(X) + V(Y)$ .

rép :

| $s_k$        | 2              | 3              | 4              | 5              | 6              | 7              | 8              |
|--------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| $P(S = s_k)$ | $\frac{1}{12}$ | $\frac{2}{12}$ | $\frac{2}{12}$ | $\frac{2}{12}$ | $\frac{2}{12}$ | $\frac{2}{12}$ | $\frac{1}{12}$ |

$$E(S) = 5, V(S) = \frac{19}{6}$$

**A02** On lance une pièce de monnaie équilibrée,  $\Omega = \{\text{pile}; \text{face}\}$ .

$X$  est la variable aléatoire telle que  $X(\text{pile}) = 1$ ,  $X(\text{face}) = 0$ ,

$Y$  est la variable aléatoire telle que  $Y(\text{pile}) = 0$ ,  $Y(\text{face}) = 1$ .

- Donner la loi de probabilité de  $X$ , en déduire  $E(X)$  et  $V(X)$ .
- Donner la loi de probabilité de  $Y$ , en déduire  $E(Y)$  et  $V(Y)$ .
- On pose  $S = X + Y$ .
  - Vérifier que  $E(S) = 1$  et  $V(S) = 0$ .
  - A-t-on :  $E(S) = E(X) + E(Y)$ ,  $V(S) = V(X) + V(Y)$  ?

Soient  $X$  et  $Y$  deux variables aléatoires définies sur un même univers  $\Omega$ .

**[P]** Par **linéarité de l'espérance** on a :

►  $E(X + Y) = E(X) + E(Y)$ .

**[D]**  $X$  et  $Y$  sont **des variables indépendantes** lorsque, pour toute valeur  $x_i$  de  $X$  et toute valeur  $y_j$  de  $Y$ , les événements  $X = x_i$  et  $Y = y_j$  sont indépendants :  $P((X = x_i) \cap (Y = y_j)) = P(X = x_i) \times P(Y = y_j)$

**[F]** Si  $X$  et  $Y$  sont deux variables aléatoires **indépendantes** alors :

►  $V(X + Y) = V(X) + V(Y)$  (si  $X$  et  $Y$  sont **indépendantes**)

**[D]** Des variables aléatoires  $X_1, \dots, X_n$  sont **indépendantes** lorsque, pour toute suite de valeurs  $x_1, \dots, x_n$  prises respectivement par  $X_1, \dots, X_n$  :

$$P((X_1 = x_1) \cap \dots \cap (X_n = x_n)) = P(X_1 = x_1) \times \dots \times P(X_n = x_n)$$

**[P]** **lien avec la loi binomiale**

- si  $X_1, \dots, X_n$  sont  $n$  variables **indépendantes** suivant toutes la loi de Bernoulli  $\mathcal{B}(p)$ , alors  $S_n = X_1 + \dots + X_n$  suit la loi binomiale  $\mathcal{B}(n, p)$
- toute variable aléatoire suivant  $\mathcal{B}(n, p)$  est la somme de  $n$  variables aléatoires **indépendantes** suivant toutes la loi de Bernoulli  $\mathcal{B}(p)$
- ► si  $X$  suit  $\mathcal{B}(n, p)$  alors  $E(X) = np$  et  $V(X) = np(1 - p)$

**A03** Une urne contient : 3 jetons marqués 0, et 2 jetons marqués 1.

On tire au hasard un premier jeton de l'urne, sans le remettre, puis on tire un second jeton.

On note  $X$  la variable aléatoire « nombre indiqué par le premier jeton » et  $Y$  la variable aléatoire « nombre indiqué par le deuxième jeton ».

1. À l'aide d'un arbre pondéré, déterminer la loi de probabilité de  $X$  et celle de  $Y$ , en déduire :  $E(X)$ ,  $V(X)$ ,  $E(Y)$  et  $V(Y)$ .
2. Montrer que les variables  $X$  et  $Y$  **ne sont pas** indépendantes.
3. On pose :  $S = X + Y$ , justifier que  $S$  prend les valeurs 0, 1, 2 avec les probabilités  $\frac{3}{10}$ ,  $\frac{3}{5}$  et  $\frac{1}{10}$ , en déduire que :  $E(S) = \frac{4}{5}$  et  $V(S) = \frac{9}{25}$ .  
A-t-on :  $E(S) = E(X) + E(Y)$ ,  $V(S) = V(X) + V(Y)$  ?

**A04** On lance quatre fois de suite un dé et on note  $X$  le nombre de fois que le chiffre 6 est sorti sur l'ensemble de ces quatre lancers.

Déterminer  $E(X)$ ,  $V(X)$  et  $\sigma(X)$ .

rép : 2/3 et 5/9

**[D]** **quelques définitions**

- si  $x \geq 0$  alors  $|x| = x$ , si  $x < 0$  alors  $|x| = -x$
- pour tout réel  $x$ ,  $|x| \geq 0$
- $(O; \vec{i})$  est un axe normé lorsque  $\|\vec{i}\| = 1$
- si  $M$  et  $N$  sont deux points d'un axe normé :  $MN = |x_M - x_N|$   
où  $x_M$  et  $x_N$  désignent respectivement l'abscisse de  $M$  et celle de  $N$  ; la valeur absolue d'une différence s'interprète donc comme une distance entre deux points situés sur un même axe normé
- soient  $a$  et  $b$  deux réels tels que  $a \leq b$ , alors la **longueur** de l'intervalle  $[a, b]$  est :  $b - a$ , son **rayon** est  $r = \frac{|b-a|}{2}$ , son **centre** est  $c = \frac{a+b}{2}$
- on utilise souvent les équivalences, mémorisables avec un schéma :  
►  $a \leq x \leq b \Leftrightarrow |x - c| \leq r$     ►  $x \leq a$  ou  $x \geq b \Leftrightarrow |x - c| \geq r$   
►  $a < x < b \Leftrightarrow |x - c| < r$     ►  $x < a$  ou  $x > b \Leftrightarrow |x - c| > r$

**A05** Traduire à l'aide d'une valeur absolue :

1.  $3 < X < 7$
2.  $2,3 \leq X \leq 7,4$
3.  $X \leq 12,5$  ou  $X \geq 17,5$

rép :  $|X - 5| < 2$      $|X - 4,85| \leq 2,55$      $|X - 15| \geq 2,5$

**A06** Écrire sous la forme d'un encadrement de  $X$  :

1.  $|X - 3| < 0,2$
2.  $|X - 5,2| < 1$
3.  $|X - 3,2| \leq 2,1$

rép :  $2,8 < X < 3,2$      $4,2 < X < 6,2$      $1,1 \leq X \leq 5,3$

**A07** Écrire à l'aide d'une valeur absolue les événements  $\bar{E}$ ,  $\bar{F}$  et  $\bar{G}$ .

$E : |X - 5| < 1$      $F : |X - 7| \leq 0,5$      $G : |X - 4| > 0,8$

**[F]** ►  $A \subset B \Rightarrow P(A) \leq P(B)$

**A08** Les questions suivantes sont indépendantes.

1. On sait :  $P(|X - 3| \geq 1) \leq 0,2$  ; que dire de :  $P(|X - 3| > 1,5)$  ?

2. On sait :  $P(|X - 5| \geq 2) \leq 0,1$  ; que dire de :  $P(X \geq 7)$  ?

3. On sait :  $P(|X - 10| \geq 3) \leq 0,8$  ; que dire de :  $P(X < 7)$  ?

rép :  $P(|X - 3| > 1,5) \leq 0,2$  ;  $P(X \geq 7) \leq 0,1$  ;  $P(X < 7) \leq 0,8$

### **[P] Inégalité de Bienaymé-Tchebychev**

Soit  $X$  une variable aléatoire d'espérance  $E(X)$  et de variance  $V(X)$ .

Alors, pour tout réel  $\delta > 0$  :

$$\blacktriangleright P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

ou  $P(|X - \mu| \geq \delta) \leq \frac{V}{\delta^2}$  en notant  $\mu$  l'espérance et  $V$  la variance de  $X$ .

**A09**  $X$  variable aléatoire d'espérance  $\mu = 12$  et de variance  $V = 9$ .

Que peut-on dire de la probabilité de l'événement  $|X - 12| \geq 6$  ?

rép :  $P(|X - 12| \geq 6) \leq \frac{1}{4}$

**A10** Soit  $X$  une variable aléatoire d'espérance  $\mu = 10$  et de variance

$V(X) = 3$ . Montrer que :  $P(|X - 10| < 6) \geq \frac{11}{12}$ .

**A11**  $X$  est le chiffre d'affaire en € d'un glacier sur le créneau  $14h - 17h$ .

Une étude montre que  $E(X) = 160$  et  $\sigma(X) = 20$ .

Que peut-on dire de la probabilité que demain le chiffre d'affaire entre  $14h$  et  $17h$  soit strictement compris entre 120€ et 200€ ? rép :  $\geq 0,75$

### **[D] quelques définitions**

Soit  $X$  une variable aléatoire.

Un **échantillon** de taille  $n$  de la loi de  $X$  est une famille  $(X_1, \dots, X_n)$  de variables aléatoires indépendantes qui ont toutes la même loi que  $X$ .

La **somme**  $S_n$  et la **moyenne**  $M_n$  de cet échantillon sont respectivement définies par :

$$S_n = X_1 + \dots + X_n \text{ et } M_n = \frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$$

**[F]** Si  $(X_1, X_2, \dots, X_n)$  est un échantillon de la loi de  $X$ , alors :

$$\blacktriangleright E(S_n) = nE(X) \quad \blacktriangleright V(S_n) = nV(X) \quad \blacktriangleright \sigma(S_n) = \sqrt{n} \sigma(X)$$

$$\blacktriangleright E(M_n) = E(X) \quad \blacktriangleright V(M_n) = \frac{1}{n} V(X) \quad \blacktriangleright \sigma(M_n) = \frac{1}{\sqrt{n}} \sigma(X)$$

**A12** Justifier les formules précédentes.

**A13C** Lorsqu'une commande de téléviseur est passée par un client, on considère que le temps de livraison du téléviseur est modélisé par une variable aléatoire  $T$  égale à la somme de deux variables aléatoires  $T_1$  et  $T_2$  : la variable aléatoire  $T_1$  modélise la durée en jours pour l'acheminement du téléviseur depuis un entrepôt de stockage vers une plateforme de distribution et la variable aléatoire  $T_2$  modélise la durée en jours pour l'acheminement du téléviseur depuis cette plateforme jusqu'au domicile du client.

On admet que  $T_1$  et  $T_2$  sont deux variables indépendantes et on précise que :  $E(T_1) = 4$ ,  $V(T_1) = 2$ ,  $E(T_2) = 3$  et  $V(T_2) = 1$ .

1. Déterminer  $E(T)$  et  $V(T)$ .

2. Un client passe une commande de téléviseur sur internet.

Que peut-on dire de la probabilité que sa durée d'attente, exprimée en jours, appartienne à  $[4; 10]$  ?

rép : 1.  $E(T) = 7$ ,  $V(T) = 3$       2.  $P(4 \leq T \leq 10) \geq \frac{2}{3}$

**A14C** [d'après bac 2024] Un client arrive à une station d'essence et se dirige vers une pompe. Il constate que deux voitures sont devant lui, la première accédant à la pompe au moment de son arrivée. On désigne par  $T_1, T_2, T_3$  les variables aléatoires qui modélisent les temps passés en minute par chacun des trois clients, dans leur ordre d'arrivée, pour alimenter son véhicule entre l'instant où la pompe est disponible pour lui et celui où il la libère. On suppose que  $T_1, T_2, T_3$  sont des variables aléatoires indépendantes de même espérance 6 et de même variance 1. On note  $S$  la variable aléatoire correspondant au temps d'attente total passé à la station du troisième client entre son arrivée à la station et son départ de la pompe après avoir alimenté son véhicule.

1. Exprimer  $S$  en fonction de  $T_1$ ,  $T_2$  et  $T_3$ .
2. a. Déterminer l'espérance de  $S$  et interpréter ce résultat dans le contexte de l'exercice.  
b. Quelle est la variance du temps d'attente total  $S$  de ce troisième client ?
3. Montrer que la probabilité que le troisième client passe un temps strictement compris entre 14 et 22 minutes à la station est supérieure à 0,81.

Rép : 2.  $E(S) = 18 \text{ min}$ ,  $V(S) = 3 \text{ min}^2$  (ou pas d'unité)

**P** Si  $M_n$  est la moyenne d'un échantillon de taille  $n$  de la loi de  $X$ , alors pour tout  $\delta > 0$  :

- **inégalité de concentration** :  $P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$  (🔴 $n$ )
- **loi faible des grands nombres** :  $\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0$

**A15C** Démontrer les deux points précédents.

**A16C** On lance  $n$  fois une pièce de 2 € et l'on désigne par  $F_n$  la variable aléatoire qui associe à cette expérience la fréquence d'apparition de « Pile », autrement dit la proportion de lancers ayant donné « Pile ». Déterminer un entier naturel  $n$  tel que la probabilité de l'événement  $0,48 < F_n < 0,52$  est supérieure ou égale à 0,95.

rép :  $P(|F_n - 0,5| \geq 0,02) \leq \frac{625}{n}$ , il suffit que  $1 - \frac{625}{n} \geq 0,95$  ;  $n \geq 12\,500$

**A17C** [d'après bac 2025] On note  $X$  la variable aléatoire donnant le nombre de points marqués par Victor lors d'un match de basket-ball. On admet que  $E(X) = 22$  et  $V(X) = 65$ .

Victor joue  $n$  matches, où  $n$  est un nombre entier strictement positif. On note  $X_1, X_2, \dots, X_n$  les variables aléatoires donnant le nombre de points marqués au cours des 1<sup>er</sup>, 2<sup>e</sup>, ...,  $n$ -ième matches, on admet que  $X_1, X_2, \dots, X_n$  sont indépendantes et suivent la même loi que celle de  $X$ , on pose :

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

1. Dans cette question, on prend  $n = 50$ .  
a. Que représente la variable aléatoire  $M_{50}$  ?  
b. Démontrer que  $P(|M_{50} - 22| \geq 3) \leq \frac{13}{90}$ .  
c. En déduire que  $P(19 < M_{50} < 25) > 0,85$ .
2. Indiquer si l'affirmation suivante est vraie ou fausse : « il n'existe aucun entier naturel  $n$  tel que  $P(|M_n - 22| \geq 3) < 0,01$  ».

rép : 1.c.  $P(|M_{50} - 22| < 3) \geq \frac{77}{90} > 0,85$  2. Fausse (2 méthodes)

**A18** [d'après bac 2025, extrait]

La probabilité qu'une connexion soit **instable** est égale à 0,145.

1. Dans le but de détecter les dysfonctionnements de serveurs, on étudie un échantillon de 50 connexions au réseau, ces connexions étant choisies au hasard. On suppose que le nombre de connexions est suffisamment important pour que ce choix puisse être assimilé à un tirage avec remise. On désigne par  $X$  la variable aléatoire égale au nombre de connexions instables au réseau de l'entreprise, dans cet échantillon de 50 connexions.  
a. On admet que la variable aléatoire  $X$  suit une loi binomiale. Préciser ses paramètres.  
b. Donner la probabilité qu'au plus huit connexions soient instables. On donnera la valeur arrondie au millièm.
2. Dans cette question, on constitue désormais un échantillon de  $n$  connexions, toujours dans les mêmes conditions, où  $n$  désigne un entier naturel strictement positif. On note  $X_n$  la variable aléatoire égale aux nombres de connexions instables et on admet que  $X_n$  suit une loi binomiale de paramètres  $n$  et 0,145.  
a. Donner l'expression en fonction de  $n$  de la probabilité  $p_n$  qu'au moins une connexion de cet échantillon soit instable.  
b. Déterminer la plus petite valeur de l'entier naturel  $n$  telle que la probabilité  $p_n$  est supérieure ou égale à 0,99.

rép :  $1 - 0,855^n$  puis  $n \geq \frac{\ln(0,01)}{\ln(0,855)}$ , c'est à partir de  $n = 30$

3. On s'intéresse à la variable aléatoire  $F_n$  égale à la fréquence de connexions instables dans un échantillon de  $n$  connexions, où  $n$  désigne un entier naturel strictement positif. On a donc  $F_n = \frac{X_n}{n}$ , où  $X_n$  est la variable aléatoire définie à la question 2.

- Calculer l'espérance  $E(F_n)$  et justifier que  $(F_n) = \frac{0,123975}{n}$ .
- Vérifier que :  $P(|F_n - 0,145| \geq 0,1) \leq \frac{12,5}{n}$ .
- Un responsable de l'entreprise étudie un échantillon de 1 000 connexions et constate que, pour cet échantillon,  $F_{1000} = 0,3$ . Il soupçonne un dysfonctionnement des serveurs. A-t-il raison ?

Rép on est dans **b.**  $n = 1000$  ( $\leq 1,25\%$ ), oui

**A13C** Lorsqu'une commande de téléviseur est passée par un client, on considère que le temps de livraison du téléviseur est modélisé par une variable aléatoire  $T$  égale à la somme de deux variables aléatoires  $T_1$  et  $T_2$  : la variable aléatoire  $T_1$  modélise le nombre de jours pour l'acheminement du téléviseur depuis un entrepôt de stockage vers une plateforme de distribution et la variable aléatoire  $T_2$  modélise le nombre de jours pour l'acheminement du téléviseur depuis cette plateforme jusqu'au domicile du client.

On admet que  $T_1$  et  $T_2$  sont deux variables indépendantes et on précise que :  $E(T_1) = 4$ ,  $V(T_1) = 2$ ,  $E(T_2) = 3$  et  $V(T_2) = 1$ .

1. Déterminer  $E(T)$  et  $V(T)$ .

$$E(T) = E(T_1 + T_2) = E(T_1) + E(T_2) \text{ (linéarité)}$$

Or,  $E(T_1) = 4$  et  $E(T_2) = 3$ , donc  $E(T) = 4 + 3 = 7$ .

$$\boxed{E(T) = 7}$$

$$V(T) = V(T_1 + T_2) = V(T_1) + V(T_2) \text{ (variables indépendantes)}$$

Or,  $V(T_1) = 2$  et  $V(T_2) = 1$ , donc  $V(T) = 2 + 1 = 3$ .

$$\boxed{V(T) = 3}$$

2. Un client passe une commande de téléviseur sur internet.

Justifier que la probabilité que sa durée d'attente, exprimée en jours, appartiennent à  $[4; 10]$  est supérieure ou égale à  $\frac{2}{3}$ .

Il s'agit de calculer  $P(4 \leq T \leq 10)$ .

Avec les notations du cours :

$$c = \frac{4 + 10}{2} = \frac{14}{2} = 7 \text{ et } r = \frac{10 - 4}{2} = \frac{6}{2} = 3$$

L'événement  $4 \leq T \leq 10$  s'écrit donc aussi :  $|T - 7| \leq 3$

donc le contraire est  $|T - 7| > 3$ .

D'après l'inégalité de Bienaymé-Tchebychev, pour tout  $\delta > 0$  :

$$P(|T - E(T)| \geq \delta) \leq \frac{V(T)}{\delta^2}$$

Or  $E(T) = 7$  et  $V(T) = 3$  donc :

$$\forall \delta > 0, P(|T - 7| \geq \delta) \leq \frac{3}{\delta^2}$$

Pour  $\delta = 3$  on obtient :  $P(|T - 7| \geq 3) \leq \frac{3}{3^2}$

c'est-à-dire :  $P(|T - 7| \geq 3) \leq \frac{1}{3}$

donc :

$$\begin{aligned} 1 - P(|T - 7| < 3) &\leq \frac{1}{3} \Leftrightarrow -P(|T - 7| < 3) \leq \frac{1}{3} - 1 \\ &\Leftrightarrow -P(|T - 7| < 3) \leq -\frac{2}{3} \end{aligned}$$

En multipliant chaque membre par  $(-1) < 0$ , on obtient :

$$P(|T - 7| < 3) \geq \frac{2}{3}$$

autrement dit :  $\frac{2}{3} \leq P(|T - 7| < 3)$

Or :  $|T - 7| < 3 \subset |T - 7| \leq 3$  donc :

$$P(|T - 7| < 3) \leq P(|T - 7| \leq 3)$$

On a donc :  $\frac{2}{3} \leq P(|T - 7| < 3) \leq P(|T - 7| \leq 3)$

d'où :  $\frac{2}{3} \leq P(|T - 7| \leq 3)$ , autrement dit :

$$\boxed{P(T \in [4; 10]) \geq \frac{2}{3}}$$

**A14C** [d'après bac 2024, extrait] ... deux voitures sont devant lui, la première accédant à la pompe au moment de son arrivée.

$T_1, T_2, T_3$  les variables aléatoires qui modélisent les temps passés en minute par chacun des trois clients, dans leur ordre d'arrivée, pour alimenter son véhicule entre l'instant où la pompe est disponible pour lui et celui où il la libère ;  $T_1, T_2, T_3$  sont des variables aléatoires indépendantes de même espérance 6 et de même variance 1.

$S$  = variable aléatoire correspondant au temps d'attente total passé à la station du troisième client entre son arrivée à la station et son départ de la pompe après avoir alimenté son véhicule.

**1. Exprimer  $S$  en fonction de  $T_1, T_2$  et  $T_3$ .**

On a immédiatement :  $S = T_1 + T_2 + T_3$ .

**2. a. Déterminer l'espérance de  $S$  et interpréter ce résultat dans le contexte de l'exercice.**

$$E(S) = E(T_1 + T_2 + T_3) = E(T_1) + E(T_2) + E(T_3) \text{ (linéarité)}$$

Or,  $E(T_1) = E(T_2) = E(T_3) = 6$ , donc :

$$E(S) = 6 + 6 + 6 = 18 \quad \boxed{E(S) = 18}$$

La moyenne du temps avant de sortir de la pompe à essence d'un client qui arrive à une station et constate que deux voitures sont devant lui, la première accédant à la pompe au moment de son arrivée, est de **18 minutes**.

**b. Quelle est la variance du temps d'attente total  $S$  de ce troisième client ?**

$$V(S) = V(T_1 + T_2 + T_3) = V(T_1) + V(T_2) + V(T_3)$$

(les variables aléatoires  $T_1, T_2$  et  $T_3$  sont indépendantes)

or,  $V(T_1) = V(T_2) = V(T_3) = 1$ , donc :

$$V(S) = 1 + 1 + 1 = 3.$$

$$\boxed{V(S) = 3}$$

**3. Montrer que la probabilité que le troisième client passe un temps strictement compris entre 14 et 22 minutes à la station est supérieure à 0,81.**

Il s'agit de montrer que  $P(14 < S < 22) > 0,81$ .

Avec les notations habituelles :

$$c = \frac{14 + 22}{2} = \frac{36}{2} = 18 \text{ et } r = \frac{22 - 14}{2} = \frac{8}{2} = 4$$

L'événement  $14 < S < 22$  s'écrit donc aussi :  $|S - 18| < 4$ .

D'après l'inégalité de Bienaymé-Tchebychev, pour tout  $\delta > 0$  :

$$P(|S - E(S)| \geq \delta) \leq \frac{V(S)}{\delta^2}$$

Pour  $E(S) = 18, V(S) = 3$  et  $\delta = 4$  on obtient :

$$P(|S - 18| \geq 4) \leq \frac{3}{4^2} \Leftrightarrow P(|S - 18| \geq 4) \leq 0,1875$$

Le contraire de l'événement  $|S - 18| \geq 4$  est  $|S - 18| < 4$

donc l'inégalité précédente s'écrit aussi :

$$1 - P(|S - 18| < 4) \leq 0,1875 \Leftrightarrow -P(|S - 18| < 4) \leq 0,1875 - 1$$

$$\Leftrightarrow -P(|S - 18| < 4) \leq -0,8125$$

En multipliant chaque membre par  $(-1) < 0$  on obtient :

$$P(|S - 18| < 4) \geq 0,8125$$

Or :  $0,8125 > 0,81$  donc :  $P(|S - 18| < 4) \geq 0,8125 > 0,81$

d'où :  $P(|S - 18| < 4) > 0,81$  autrement dit :

$$\boxed{P(14 < S < 22) > 0,81}$$

$\boxed{P} \forall \delta > 0 :$

• **inégalité de concentration** :  $P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$

• **loi faible des grand nombres** :  $\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0$

**A15C Démontrer les deux points précédents.**

• **inégalité de concentration**

D'après l'inégalité de Bienaymé-Tchebychev, pour tout  $\delta > 0$  :

$$P(|M_n - E(M_n)| \geq \delta) \leq \frac{V(M_n)}{\delta^2}$$

Or, on a montré précédemment que :

$$E(M_n) = E(X) \text{ et } V(M_n) = \frac{1}{n} V(X)$$

donc l'inégalité précédente s'écrit aussi :

$$P(|M_n - E(X)| \geq \delta) \leq \frac{\frac{1}{n} V(X)}{\delta^2}$$



$$\Leftrightarrow P(|M_n - E(X)| \geq \delta) \leq \frac{1}{n} V(X) \times \frac{1}{\delta^2}$$

$$\Leftrightarrow P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$$

• loi faible des grand nombres :

$\delta^2 > 0$  donc :  $\lim_{n \rightarrow +\infty} (n\delta^2) = +\infty$ , puis par limite d'un quotient :

$$\lim_{n \rightarrow +\infty} \frac{V(X)}{n\delta^2} = 0$$

On a :

$$\begin{cases} \forall n \in \mathbb{N}^*, 0 \leq P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2} \\ \lim_{n \rightarrow +\infty} 0 = 0 = \lim_{n \rightarrow +\infty} \frac{V(X)}{n\delta^2} = 0 \end{cases}$$

donc d'après le théorème des gendarmes on en déduit que :

$$\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0$$

**A16C** On lance  $n$  fois une pièce de 2€ et l'on désigne par  $F_n$  la variable aléatoire qui, à ces  $n$  lancers, associe la fréquence de « Pile » obtenue.

Déterminer un entier naturel  $n$  tel que la probabilité de l'événement  $0,48 < F_n < 0,52$  est supérieure ou égale à 0,95.

On note  $X$  la variable aléatoire qui prend la valeur 0 lorsque l'on obtient « Face » en lançant une pièce et la valeur 1 si on obtient « Pile », la loi de probabilité de  $X$  est donnée par :

|              |               |               |
|--------------|---------------|---------------|
| $x_i$        | 0             | 1             |
| $P(X = x_i)$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

$$E(X) = \sum_{i=1}^2 P(X = x_i) x_i = \frac{1}{2}(0) + \frac{1}{2}(1) = \frac{1}{2} = 0,5$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} \times 0^2 + \frac{1}{2} \times 1^2 - \left(\frac{1}{2}\right)^2 = \frac{1}{4} = 0,25$$

Autre méthode :  $X$  suit la loi de Bernoulli de paramètre  $p = 0,5$  d'où :  $E(X) = p = 0,5$  et  $V(X) = p(1-p) = 0,5(1-0,5) = 0,25$

Considérons l'événement  $0,48 < F_n < 0,52$  ; avec les notations habituelles on a :

$$c = \frac{0,48 + 0,52}{2} = 0,5 \text{ et } r = \frac{0,52 - 0,48}{2} = \frac{0,04}{2} = 0,02$$

L'événement  $0,48 < F_n < 0,52$  s'écrit aussi  $|F_n - 0,5| < 0,02$  donc le contraire est :  $|F_n - 0,5| \geq 0,02$ .

D'après l'inégalité de concentration on a :

$$\forall \delta > 0, P(|F_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$$

Or  $E(X) = 0,5$ ,  $V(X) = 0,25$  donc :

$$\forall \delta > 0, P(|F_n - 0,5| \geq \delta) \leq \frac{0,25}{n\delta^2}$$

Pour  $\delta = 0,02$  on obtient :

$$P(|F_n - 0,5| \geq 0,02) \leq \frac{0,25}{n(0,02)^2}$$

Puis on obtient :

$$1 - P(|F_n - 0,5| < 0,02) \leq \frac{0,25}{0,0004n}$$

$$\Leftrightarrow -P(|F_n - 0,5| < 0,02) \leq \frac{0,25}{0,0004n} - 1$$

puis en multipliant par  $(-1) < 0$  :

$$P(|F_n - 0,5| < 0,02) \geq -\frac{0,25}{0,0004n} + 1$$

Pour que  $P(|F_n - 0,5| < 0,02) \geq 0,95$  il suffit que :

$$-\frac{0,25}{0,0004n} + 1 \geq 0,95 \Leftrightarrow -\frac{0,25}{0,0004n} \geq 0,95 - 1$$

$$\Leftrightarrow -\frac{0,25}{0,0004n} \geq -0,05 \Leftrightarrow \frac{0,25}{0,0004n} \leq 0,05$$

$$\Leftrightarrow 0,25 \leq 0,05 \times 0,0004 \times n \Leftrightarrow \frac{0,25}{0,05 \times 0,0004} \leq n \Leftrightarrow 12\,500 \leq n$$

Il suffit donc que  $n \geq 12\,500$ .

**A17C**  $X$  donne le nombre de points marqués par Victor lors d'un match de basket-ball,  $E(X) = 22$  et  $V(X) = 65$ . Victor joue  $n \in \mathbb{N}^*$  matchs,  $X_1, X_2, \dots, X_n$  variables aléatoires donnant le nombre de points marqués au cours des 1<sup>er</sup>, 2<sup>e</sup>, ...,  $n$ -ième matchs,  $X_1, X_2, \dots, X_n$  sont indépendantes et suivent la même loi que celle de  $X$ ,

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

1. Dans cette question, on prend  $n = 50$ .

a. Que représente la variable aléatoire  $M_{50}$  ?

$M_{50}$  représente la moyenne par match du nombre de points marqués par Victor pour une série de 50 matchs.

b. Démontrons que  $P(|M_{50} - 22| \geq 3) \leq \frac{13}{90}$ .

D'après l'inégalité de concentration, pour tout  $\delta > 0$  :

$$P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$$

Pour  $n = 50$ ,  $E(X) = 22$ ,  $V(X) = 65$  et  $\delta = 3$  on obtient :

$$P(|M_{50} - 22| \geq 3) \leq \frac{65}{50 \times 3^2}$$

$$\text{Or, } \frac{65}{50 \times 3^2} = \frac{5 \times 13}{5 \times 10 \times 9} = \frac{13}{90}, \text{ donc :}$$

$$\boxed{P(|M_{50} - 22| \geq 3) \leq \frac{13}{90}}$$

c. En déduire que  $P(19 < M_{50} < 25) > 0,85$ .

Avec les notations habituelles :

$$c = \frac{19 + 25}{2} = 22 \text{ et } r = \frac{25 - 19}{2} = \frac{6}{2} = 3$$

Il s'agit donc de montrer que :  $P(|M_{50} - 22| < 3) > 0,85$ .

L'événement contraire de  $|M_{50} - 22| < 3$  est  $|M_{50} - 22| \geq 3$  donc :  $P(|M_{50} - 22| < 3) = 1 - P(|M_{50} - 22| \geq 3)$ .

Or, on a montré en b. que :  $P(|M_{50} - 22| \geq 3) \leq \frac{13}{90}$ ,

donc en multipliant par  $(-1) < 0$  :

$$-P(|M_{50} - 22| \geq 3) \geq -\frac{13}{90}$$

puis en ajoutant 1 à chaque membre :

$$1 - P(|M_{50} - 22| \geq 3) \geq 1 - \frac{13}{90} \Leftrightarrow P(|M_{50} - 22| < 3) \geq \frac{77}{90}$$

Or,  $\frac{77}{90} \approx 0,856$  et  $0,856 > 0,85$  donc :

$$P(|M_{50} - 22| < 3) \geq \frac{77}{90} > 0,85$$

d'où :  $P(|M_{50} - 22| < 3) > 0,85$  autrement dit :

$$\boxed{P(19 < M_{50} < 25) > 0,85}$$

2. Indiquer si l'affirmation est vraie ou fausse : « il n'existe aucun entier naturel  $n$  tel que  $P(|M_n - 22| \geq 3) < 0,01$  ».

La loi faible des grands nombres indique que :

$$\forall \delta > 0, \lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0$$

Pour  $E(X) = 22$  et  $\delta = 3$  on obtient :

$$\lim_{n \rightarrow +\infty} P(|M_n - 22| \geq 3) = 0$$

donc,  $\forall \varepsilon > 0$ , il existe  $n_0 \in \mathbb{N}$  tel que :  $\forall n \in \mathbb{N}$ , si  $n \geq n_0$

alors  $P(|M_n - 22| \geq 3) < \varepsilon$ , autrement dit :

$$P(|M_n - 22| \geq 3) < \varepsilon$$

En particulier pour  $\varepsilon = 0,01$  : il existe  $n_0 \in \mathbb{N}$  tel que,

pour tout  $n \in \mathbb{N}$ , si  $n \geq n_0$  alors  $P(|M_n - 22| \geq 3) < 0,01$ .

L'existence d'un tel  $n_0 \in \mathbb{N}$  montre que l'affirmation est **fausse**.

**A18C** La probabilité qu'une connexion soit instable est égale à 0,145.

1. ... échantillon de 50 connexions au réseau choisies au hasard, ce choix est assimilé à un tirage avec remise,  $X$  = nombre de connexions instables dans cet échantillon de 50 connexions.

a.  $X$  suit une loi binomiale, paramètres ?

$X$  suit la loi binomiale de paramètres  $n = 50$  et  $p = 0,145$

b. Donner la probabilité qu'au plus huit connexions soient instables. On donnera la valeur arrondie au millièmes.

Il s'agit de déterminer  $P(X \leq 8)$ .

En utilisant les probabilités cumulées croissantes sur la calculatrice on obtient :

$$P(X \leq 8) \approx 0,704$$



[illegible]

2. On constitue désormais un échantillon de  $n$  connexions,  $X_n$  donne les nombres de connexions instables, on admet que  $X_n$  suit une loi binomiale de paramètres  $n$  et 0,145.

On note  $Y$  la variable aléatoire qui prend la valeur 1 lorsqu'une connexion est instable et 0 sinon : elle suit une loi de Bernoulli de paramètre  $p = 0,145$ . On a donc (cours) :

Or,  $F_n$  est la fréquence des connexions instables dans un échantillon de taille  $n$ .

$$E(F_n) = E(Y), \text{ donc : } \boxed{E(F_n) = 0,145}$$

- b. Vérifier que :  $P(|F_n - 0,145| \geq 0,1) \leq \frac{12,5}{n}$ .**

$$P(|F_n - E(F_n)| \geq \delta) \leq \frac{V(F_n)}{\delta^2}$$

$$\begin{aligned}
P(|F_n - 0,145| \geq 0,1) &\leq \frac{0,123975}{\frac{n}{0,1^2}} \\
&\Leftrightarrow P(|F_n - 0,145| \geq 0,1) \leq \frac{0,123975}{\frac{n}{1}} \\
&\Leftrightarrow P(|F_n - 0,145| \geq 0,1) \leq \frac{0,123975}{n} \times \frac{100}{1} \\
&\Leftrightarrow P(|F_n - 0,145| \geq 0,1) \leq \frac{12,3975}{n} \\
\text{Or, } \frac{12,3975}{n} &\leq \frac{12,5}{n} \text{ donc :}
\end{aligned}$$

$$P(|F_n - 0,145| \geq 0,1) \leq \frac{12,5}{n}$$

- c. Un responsable étudie un échantillon de 1 000 connexions et constate que, pour cet échantillon,  $F_{1000} = 0,3$ .  
Il soupçonne un dysfonctionnement des serveurs.  
A-t-il raison ?

$$|F_{1000} - 0,145| = |0,3 - 0,145| = 0,155 \geq 0,1$$

donc d'après la question b. :

$$\begin{aligned}
P(|F_{1000} - 0,145| \geq 0,1) &\leq \frac{12,5}{1000} \\
&\Leftrightarrow P(|F_{1000} - 0,145| \geq 0,1) \leq 0,0125
\end{aligned}$$

La probabilité étant particulièrement faible : **le responsable a raison** de soupçonner un dysfonctionnement des serveurs.

### Preuve Bienaymé-Tchebychev (cas variable discrète)

#### • Étape 1 : inégalité de Markov

Soit  $X$  une variable aléatoire positive ou nulle d'espérance  $E(X)$ .

Alors, pour tout réel  $\delta > 0$ , on a :

$$P(X \geq \delta) \leq \frac{E(X)}{\delta}$$

#### Preuve

On note  $x_1, \dots, x_n$  les valeurs prises par  $X$ , alors on a :

$$E(X) \stackrel{\text{def}}{=} \sum_{i=1}^n P(X = x_i)x_i = \sum_{x_i < \delta} P(X = x_i)x_i + \sum_{x_i \geq \delta} P(X = x_i)x_i$$

Résumons :

$$E(X) = \sum_{x_i < \delta} P(X = x_i)x_i + \sum_{x_i \geq \delta} P(X = x_i)x_i$$

Or, tous les  $x_i$  sont positifs ou nuls et une probabilité est un réel positif ou nul, donc pour tous les  $x_i$  on a  $P(X = x_i)x_i \geq 0$ , et comme une somme de nombres tous positifs ou nuls est positive ou nulle, on a :

$$\sum_{x_i < \delta} P(X = x_i)x_i \geq 0$$

puis en ajoutant à chaque membre le nombre :

$$\sum_{x_i \geq \delta} P(X = x_i)x_i$$

on obtient :

$$\sum_{x_i < \delta} P(X = x_i)x_i + \sum_{x_i \geq \delta} P(X = x_i)x_i \geq \sum_{x_i \geq \delta} P(X = x_i)x_i$$

c'est – à – dire :

$$E(X) \geq \sum_{x_i \geq \delta} P(X = x_i)x_i$$

Or, tous les  $x_i$  tels que  $x_i \geq \delta$  sont plus grands ou égaux à  $\delta$  :

$$x_i \geq \delta$$

donc pour ces  $x_i$  on a, en multipliant par le réel positif ou nul

$P(X = x_i)$  :

$$P(X = x_i)x_i \geq P(X = x_i)\delta$$

puis par somme membre à membre :

$$\sum_{x_i \geq \delta} P(X = x_i)x_i \geq \sum_{x_i \geq \delta} P(X = x_i)\delta$$

On a :

$$E(X) \geq \sum_{x_i \geq \delta} P(X = x_i)x_i \text{ et } \sum_{x_i \geq \delta} P(X = x_i)x_i \geq \sum_{x_i \geq \delta} P(X = x_i)\delta$$

donc :

$$E(X) \geq \sum_{x_i \geq \delta} P(X = x_i)\delta$$

Or,

$$\sum_{x_i \geq \delta} P(X = x_i)\delta = \delta \times \sum_{x_i \geq \delta} P(X = x_i) = \delta \times P(X \geq \delta)$$

Donc :

$$E(X) \geq \delta \times P(X \geq \delta)$$

puis en divisant par  $\delta > 0$  :

$$\frac{E(X)}{\delta} \geq P(X \geq \delta)$$

autrement dit :

$$P(X \geq \delta) \leq \frac{E(X)}{\delta}$$

#### • Étape 2 : on applique l'inégalité de Markov

Soit  $X$  une variable aléatoire et un réel  $\delta > 0$ , posons  $\mu = E(X)$ .

La variable aléatoire  $(X - \mu)^2$  est positive ou nulle et  $\delta^2 > 0$

donc d'après l'inégalité de Markov :

$$P((X - \mu)^2 \geq \delta^2) \leq \frac{E((X - \mu)^2)}{\delta^2} \quad (*)$$

Or, d'une part on a les équivalences :

$$(X - \mu)^2 \geq \delta^2 \Leftrightarrow \sqrt{(X - \mu)^2} \geq \sqrt{\delta^2} \Leftrightarrow |X - \mu| \geq |\delta| \\ \Leftrightarrow |X - \mu| \geq \delta \quad (\delta > 0)$$

et d'autre part on a les égalités :

$$E((X - \mu)^2) = E(X^2 - 2X\mu + \mu^2) = E(X^2) + E(-2\mu X) + \mu^2$$

$$\begin{aligned}
&= E(X^2) + (-2\mu)E(X) + \mu^2 = E(X^2) - 2\mu \times \mu + \mu^2 \\
&= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - (E(X))^2 \\
&= V(X) \text{ (formule de König)}
\end{aligned}$$

Donc l'inégalité (\*) s'écrit aussi :

$$P(|X - \mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

ou encore

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

#### Conclusion

**Pour toute variable aléatoire  $X$  et tout réel  $\delta > 0$ , on a :**

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

**(inégalité de Bienaymé-Tchebychev)**