## 1. Rappels de 3e : statistiques

**Ex. 1** ★☆☆ **Automatismes : les sorties à vélo d'Inès**

(a) Série rangée : 18 ; 18 ; 18 ; 22 ; 25 ; 27 ; 29 ; 32 ; 35 ; 41 (la valeur 18 apparaît trois fois, on l'écrit trois fois).

(b) Étendue $= 41 - 18 = 23$ km : les distances s'étalent sur 23 km.

(c) $N = 10$ est pair : $Me$ est la demi-somme des valeurs de rangs 5 et 6. $Me = \frac{25+27}{2} = 26$ km.

(d) $\overline{x} = \frac{18+32+25+18+41+27+22+35+18+29}{10} = \frac{265}{10} = 26{,}5$ km. Aucune sortie ne mesure 26,5 km : une moyenne n'est pas forcément une valeur de la série.

(e) $\frac{10}{4} = 2{,}5$, arrondi à 3 : $Q_1$ est la valeur de rang 3, $Q_1 = 18$ km. $\frac{3\times10}{4} = 7{,}5$, arrondi à 8 : $Q_3$ est la valeur de rang 8, $Q_3 = 32$ km.

(f) Écart interquartile $= Q_3 - Q_1 = 32 - 18 = 14$ km : la moitié centrale des sorties s'étale sur 14 km.

**Ex. 2** ★☆☆ **Automatismes : les livres lus pendant l'été**

(a) $N = 3 + 9 + 6 + 4 + 2 + 1 = 25$ : 25 élèves ont répondu.

(b) $f = \frac{4}{25} = 0{,}16 = 16\,\%$ : 16 % des élèves ont lu exactement 3 livres (on divise par l'effectif total 25).

(c) Le plus grand effectif est 9 : la valeur la plus fréquente est 1 livre.

(d) Effectifs cumulés : 3 ; 12 ; 18 ; 22 ; 24 ; 25.

(e) $N = 25$ est impair : $Me$ est la valeur de rang $\frac{25+1}{2} = 13$. L'effectif cumulé passe de 12 à 18 pour la valeur 2 : les rangs 13 à 18 valent 2, donc $Me = 2$ livres. Elle est différente de la valeur la plus fréquente, 1 livre.

(f) $\overline{x} = \frac{0\times3+1\times9+2\times6+3\times4+4\times2+5\times1}{25} = \frac{46}{25} = 1{,}84$ livre par élève (et non 2,5, la moyenne des valeurs 0 à 5 sans les effectifs).

**Ex. 3** ★★☆ **Les erreurs de Lucas au basket**

(a) Lucas oublie les effectifs et divise par le nombre de valeurs distinctes. $N = 5 + 8 + 2 + 3 + 2 = 20$ matchs. $\overline{x} = \frac{2\times5+4\times8+5\times2+7\times3+9\times2}{20} = \frac{91}{20} = 4{,}55$. Maëlle marque en moyenne 4,55 paniers par match.

**Ex. 3 (suite)**

(b) Non : l'étendue n'est pas le nombre de valeurs, c'est $\max - \min = 9 - 2 = 7$. Les nombres de paniers s'étalent sur 7 paniers.

(c) Effectifs cumulés : 5 ; 13 ; 15 ; 18 ; 20. $N = 20$ est pair : $Me$ est la demi-somme des rangs 10 et 11. Les rangs 6 à 13 valent 4, donc $Me = \frac{4+4}{2} = 4$. La moitié des matchs comptent au plus 4 paniers.

(d) Non. $\frac{20}{4} = 5$, entier : $Q_1$ est la valeur de rang 5. Les rangs 1 à 5 valent 2, donc $Q_1 = 2$ paniers (et non $\frac{2+4}{2} = 3$, qui n'est même pas une valeur de la série).

**Ex. 4** ★★☆ **Le tri des déchets de la famille Haddad**

(a) Il faut une somme de $6 \times 9 = 54$ kg. Or $9 + 6 + 11 + 8 + 7 = 41$, donc $x = 54 - 41 = 13$. La sixième semaine doit peser 13 kg.

(b) Série rangée : 6 ; 7 ; 8 ; 9 ; 11 ; 13. $N = 6$ est pair : $Me = \frac{8+9}{2} = 8{,}5$ kg (demi-somme des rangs 3 et 4). Elle est différente de la moyenne, 9 kg.

(c) Moyenne : $\frac{54+30}{7} = \frac{84}{7} = 12$ kg. Série rangée : 6 ; 7 ; 8 ; 9 ; 11 ; 13 ; 30, $N = 7$ : $Me$ est la valeur de rang 4, $Me = 9$ kg. La moyenne gagne 3 kg, la médiane seulement 0,5 kg.

(d) Faux. Rangée, la série devient 6 ; 7 ; 8 ; 9 ; 11 ; 13 ; 60 : la valeur de rang 4 est toujours 9, donc $Me = 9$ kg. Seule la moyenne change : $\frac{54+60}{7} = \frac{114}{7} \approx 16{,}29$ kg.

**Ex. 5** ★★★ **Les retards du TER de Sacha**

(a) On range d'abord la série, en gardant les valeurs répétées : 0 ; 0 ; 2 ; 2 ; 2 ; 3 ; 4 ; 5 ; 7 ; 9 ; 15 ; 31. $N = 12$ est pair. La médiane est donc la demi-somme des valeurs de rangs 6 et 7 : $Me = \frac{3+4}{2} = 3{,}5$ min. Au moins la moitié des jours, Sacha a eu au plus 3,5 min de retard. Au moins la moitié des jours, il a eu au moins 3,5 min de retard.

(b) $\frac{12}{4} = 3$ est un entier, donc $Q_1$ est la valeur de rang 3 : $Q_1 = 2$ min. $\frac{3\times12}{4} = 9$, donc $Q_3$ est la valeur de rang 9 : $Q_3 = 7$ min. L'écart interquartile vaut $7 - 2 = 5$ min : les retards centraux s'étalent sur 5 min.

**Ex. 5** **(suite)**

(c) Étendue $= 31 - 0 = 31$ min (et non 12, qui est le nombre de jours). Avec 16 au lieu de 31, la série rangée devient 0 ; 0 ; 2 ; 2 ; 2 ; 3 ; 4 ; 5 ; 7 ; 9 ; 15 ; 16. Les valeurs de rangs 3 et 9 restent 2 et 7, donc l'écart interquartile reste 5 min. L'étendue, elle, devient $16 - 0 = 16$ min. C'est donc l'étendue qui change : elle dépend de la valeur extrême, alors que l'écart interquartile ne la prend pas en compte.

(d) $\overline{x} = \frac{0+0+2+2+2+3+4+5+7+9+15+31}{12} = \frac{80}{12} \approx 6{,}67$ min. Cette moyenne ne fait pas partie de la série. Elle est presque le double de la médiane (3,5 min) parce que les deux gros retards (15 et 31 min) la tirent vers le haut. La médiane, elle, ne dépend que des rangs. Le retard moyen de Sacha est d'environ 6,67 min.

(e) Avec $N = 13$, la médiane est la valeur de rang $\frac{13+1}{2} = 7$. Six des 12 premières valeurs sont inférieures ou égales à 3 : 0 ; 0 ; 2 ; 2 ; 2 ; 3. Si $x \leqslant 3$, les sept plus petites valeurs sont ces six-là et $x$. La plus grande des sept est 3, donc $Me = 3$ min. Si $x \geqslant 4$, les rangs 1 à 6 restent occupés par 0 ; 0 ; 2 ; 2 ; 2 ; 3, et la valeur de rang 7 est 4 (que $x$ vaille 4 ou plus).
Donc $Me = 4$ min. Conclusion : la médiane vaut 3 min si $x \leqslant 3$ et 4 min si $x \geqslant 4$, même si le retard est énorme.

**Ex. 6** ★★★ **Fruits et légumes au lycée de Saint-Nazaire**

(a) Les effectifs cumulés croissants sont 4 ; 15 ; 20 ; 26 ; 29 ; 30, donc $N = 30$. $N$ est pair : la médiane est la demi-somme des valeurs de rangs 15 et 16. Les rangs 5 à 15 ont la valeur 1 et les rangs 16 à 20 ont la valeur 2.
Donc $Me = \frac{1+2}{2} = 1{,}5$ portion. Zoé a tort : 1 est la valeur la plus fréquente, mais ce n'est pas la médiane.

(b) $\overline{x} = \frac{0\times4+1\times11+2\times5+3\times6+4\times3+5\times1}{30} = \frac{56}{30} \approx 1{,}87$ portion. On divise par l'effectif total, 30. Aucun élève n'a pu manger exactement 1,87 portion, car les réponses sont des entiers : une moyenne n'est pas forcément une valeur de la série. En moyenne, un élève a mangé environ 1,87 portion.

(c) Au moins 3 portions : $6 + 3 + 1 = 10$ élèves sur 30. $f = \frac{10}{30} = \frac{1}{3} \approx 0{,}3333 \approx 33{,}33\,\%$. C'est exactement un tiers, et non moins d'un tiers : l'affirmation est fausse.

**Ex. 6** **(suite)**

(d) Noé fait la moyenne des deux moyennes, comme si les deux classes avaient le même effectif. Il faut additionner les sommes : 56 portions dans cette classe (question 2) et $25 \times 2{,}4 = 60$ portions dans l'autre. $\overline{x} = \frac{56+60}{30+25} = \frac{116}{55} \approx 2{,}11$ portions. En moyenne, un élève des deux classes a mangé environ 2,11 portions.

(e) Pour une moyenne de 2 portions, il faut une somme de $2 \times 30 = 60$ portions. La somme actuelle est 56, il manque donc $60 - 56 = 4$ portions. Chaque élève qui passe de 1 à 3 portions en ajoute 2 : il en faut $\frac{4}{2} = 2$. Nouveaux effectifs : 4 ; 9 ; 5 ; 8 ; 3 ; 1, effectifs cumulés : 4 ; 13 ; 18 ; 26 ; 29 ; 30. Les rangs 15 et 16 valent 2, donc $Me = 2$ portions. Il faudrait que 2 élèves passent de 1 à 3 portions : la moyenne et la médiane vaudraient alors toutes les deux 2 portions.

**Ex. 7** ★★★ **Le javelot de Clémence et Yanis**

(a) Série rangée : 35 ; 36 ; 38 ; 39 ; 41 ; 41 ; 44 ; 47, avec $N = 8$. Minimum $= 35$ m et maximum $= 47$ m. $N$ est pair : $Me = \frac{39+41}{2} = 40$ m (rangs 4 et 5). $\frac{8}{4} = 2$ : $Q_1$ est la valeur de rang 2, $Q_1 = 36$ m. $\frac{3\times8}{4} = 6$ : $Q_3$ est la valeur de rang 6, $Q_3 = 41$ m.

(b) Clémence : $(40; 41 - 36) = (40; 5)$. Yanis : $(42; 45 - 37) = (42; 8)$. La médiane de Yanis est plus grande : au moins la moitié de ses lancers atteignent 42 m. Son écart interquartile est plus grand : 8 m contre 5 m. Yanis lance donc plus loin en général, mais Clémence est plus régulière, car ses lancers centraux sont plus resserrés.

(c) Vrai : $Q_1$ est la valeur d'un rang ($\frac{N}{4}$ arrondi au-dessus), c'est donc toujours une valeur de la série. Yanis a bien réussi un lancer de 37 m.

(d) Faux. $N = 8$ est pair : $Me$ est la demi-somme des valeurs de rangs 4 et 5, qui peuvent être différentes. Contre-exemple : 30 ; 37 ; 40 ; 41 ; 43 ; 45 ; 48 ; 50. Minimum 30, maximum 50. Rang $\frac{8}{4} = 2$ : $Q_1 = 37$. Rangs 4 et 5 : $Me = \frac{41+43}{2} = 42$. Rang $\frac{3\times8}{4} = 6$ : $Q_3 = 45$. Cette série donne exactement le diagramme de Yanis sans aucun lancer de 42 m.

## 2. *Moyenne et linéarité de la moyenne*

## Ex. 8 ★☆☆ Automatismes : calculer une moyenne

(a) $\overline{x} = \frac{7+11+9+13+10}{5} = \frac{50}{5} = 10$.

(b) Somme $= N \times \overline{x} = 8 \times 6{,}5 = 52$.

(c) $\overline{x} = \frac{2\times12+3\times9+1\times15}{2+3+1} = \frac{24+27+15}{6} = \frac{66}{6} = 11$ : Kenza a 11 de moyenne.

(d) $\overline{x} = 0{,}25 \times 2 + 0{,}5 \times 3 + 0{,}25 \times 4 = 0{,}5 + 1{,}5 + 1 = 3$.

(e) $\overline{x} = \frac{9+10+10+11+60}{5} = \frac{100}{5} = 20$. Une seule valeur (60) dépasse la moyenne : la valeur extrême 60 tire la moyenne vers le haut.

(f) Somme : $6 \times 12 = 72$. Nouvelle moyenne : $\frac{72+12}{7} = \frac{84}{7} = 12$. Elle est inchangée, car la valeur ajoutée est égale à la moyenne.

(g) Somme : $4 \times 15 = 60$. Nouvelle moyenne : $\frac{60+25}{5} = \frac{85}{5} = 17$. Elle augmente, car $25 > 15$.

## Ex. 9 ★☆☆ Automatismes : linéarité de la moyenne

(a) $\overline{y} = \overline{x} + 2 = 11{,}2 + 2 = 13{,}2$ : la nouvelle moyenne est 13,2 sur 20.

(b) $\overline{y} = 2 \times 6{,}3 = 12{,}6$ : la moyenne est 12,6 sur 20.

(c) $\overline{y} = 1{,}8 \times 15 + 32 = 27 + 32 = 59$ : la moyenne est de 59 °F.

(d) $\overline{y} = 3 \times 4{,}2 - 5 = 12{,}6 - 5 = 7{,}6$.

(e) On résout $1{,}8\overline{x} + 32 = 77$ : $\overline{x} = \frac{77-32}{1{,}8} = \frac{45}{1{,}8} = 25$. La moyenne est de 25 °C.

(f) $y = 60x$, donc $\overline{y} = 60 \times 2{,}5 = 150$ : la durée moyenne est de 150 min.

## Ex. 10 ★★☆ Le club de course à pied

(a) Juniors : $15 \times 42 = 630$ km. Adultes : $25 \times 58 = 1\,450$ km.

(b) $\overline{x} = \frac{630+1\,450}{15+25} = \frac{2\,080}{40} = 52$. Les adhérents ont couru en moyenne 52 km ce mois-ci.

(c) Non. Lina fait la moyenne des deux moyennes (50 km) comme si les deux groupes avaient le même effectif. Or il y a plus d'adultes, et ils courent davantage, donc la vraie moyenne (52 km) est plus proche de 58 km. Elle aurait dû passer par les sommes.

(d) Si les deux groupes ont le même effectif $N$ : $\frac{N\times42+N\times58}{2N} = \frac{42+58}{2} = 50$. La méthode de Lina ne marche que si les juniors et les adultes sont aussi nombreux.

## Ex. 11 ★★☆ La note qui manque à Malo

(a) Somme : $11 + 14 + 8 + 13 = 46$. Moyenne : $\frac{46}{4} = 11{,}5$. Malo a pour l'instant 11,5 sur 20 de moyenne.

(b) Somme visée : $5 \times 12 = 60$. Note nécessaire : $60 - 46 = 14$. Malo doit obtenir 14 sur 20.

(c) Somme visée : $5 \times 14 = 70$. Note nécessaire : $70 - 46 = 24$. Or une note ne dépasse pas 20 : Malo ne peut pas atteindre 14 de moyenne.

(d) Le total des coefficients vaut $4 + 2 = 6$, donc la somme pondérée visée est $6 \times 12 = 72$. On cherche $x$ tel que $46 + 2x = 72$, d'où $x = \frac{72-46}{2} = 13$. Il suffit à Malo d'obtenir 13 sur 20.

## Ex. 12 ★★★ La station météo du lycée

(a) Somme : $13 + 17 + 15 + 19 + 16 + 16 = 96$, donc $\overline{x} = \frac{96}{6} = 16$. La température moyenne est de 16 °C.

(b) Les valeurs en °F sont les $1{,}8x_i + 32$. Leur somme vaut $1{,}8 \times (x_1 + \ldots + x_6) + 6 \times 32 = 1{,}8 \times 6\overline{x} + 6 \times 32$. On divise par 6 : $\overline{y} = 1{,}8\,\overline{x} + 32 = 1{,}8 \times 16 + 32 = 28{,}8 + 32 = 60{,}8$. La température moyenne est de 60,8 °F.

(c) À Denver, $1{,}8\,\overline{x} + 32 = 57{,}2$, donc $\overline{x} = \frac{57{,}2-32}{1{,}8} = \frac{25{,}2}{1{,}8} = 14$ °C. Comme $16 > 14$, c'est le lycée de Montpellier qui a eu le plus chaud, de 2 °C en moyenne (soit $60{,}8 - 57{,}2 = 3{,}6$ °F).

(d) Somme : $96 + (-40) = 56$, donc $\overline{x} = \frac{56}{7} = 8$ °C. Série rangée : $-40$ ; 13 ; 15 ; 16 ; 16 ; 17 ; 19. Comme $N = 7$, la médiane est la 4e valeur : $Me = 16$ °C. Une seule valeur aberrante fait chuter la moyenne de 16 °C à 8 °C, mais la médiane reste à 16 °C. La moyenne est sensible aux valeurs extrêmes, la médiane ne l'est pas.

(e) Somme : $96 + 23 = 119$, donc $\overline{x} = \frac{119}{7} = 17$ °C. Elle augmente, car $23 > 16$. En °F : $1{,}8 \times 17 + 32 = 30{,}6 + 32 = 62{,}6$. La température moyenne des sept jours est de 17 °C, soit 62,6 °F.

## Ex. 13 ★★★ Les trajets du matin

(a) Nora fait la moyenne des trois moyennes, comme si chaque groupe pesait autant. Or les effectifs sont différents (24, 16 et 10). Il faut donc pondérer chaque moyenne par son effectif : la méthode de Nora est fausse.

(b) Sommes : $24 \times 18 = 432$ min, $16 \times 33 = 528$ min et $10 \times 9 = 90$ min. Total : $432 + 528 + 90 = 1\,050$ min. Moyenne : $\frac{1\,050}{50} = 21$. Le temps de trajet moyen des 50 élèves est de 21 min, et non de 20 min.

**Ex. 13 (suite)**

(c) Avec $k$ élèves de plus, la somme vaut $1\,050 + 9k$ et l'effectif vaut $50 + k$. On veut $\frac{1\,050+9k}{50+k} = 19$, soit $1\,050 + 9k = 950 + 19k$.
Donc $100 = 10k$ et $k = 10$. Vérification : $\frac{1\,050+90}{60} = \frac{1\,140}{60} = 19$. Il faut ajouter 10 élèves venant à vélo.

(d) Chaque temps $x$ de la ligne B devient $y = x - 6$, donc la moyenne de la ligne B devient $33 - 6 = 27$ min. Nouvelle somme : $432 + 16 \times 27 + 90 = 432 + 432 + 90 = 954$ min. Moyenne : $\frac{954}{50} = 19{,}08$. Le temps moyen des 50 élèves passe à 19,08 min, soit un peu plus de 19 min.

**Ex. 14 ★★★ Vrai ou faux dans le groupe de Ben**

(a) Somme : $7 + 9 + 10 + 11 + 12 + 13 + 14 + 12 = 88$, donc $\overline{x} = \frac{88}{8} = 11$. Avec $y = 1{,}5x$, on obtient $\overline{y} = 1{,}5 \times 11 = 16{,}5$. Vrai : sur 30, la moyenne du groupe est de 16,5.

(b) La note $x$ devient $y = 1{,}5(x + 2) = 1{,}5x + 3$, donc $\overline{y} = 1{,}5 \times 11 + 3 = 19{,}5$ au lieu de 16,5. Les 2 points sont eux aussi multipliés par 1,5, donc la moyenne gagne 3 points. Faux.

(c) Si la note du nouvel élève est égale à la moyenne 11 : $\frac{88+11}{9} = \frac{99}{9} = 11$. La moyenne reste à 11. Faux : une note égale à la moyenne ne la change pas.

(d) $\frac{88+20}{9} = \frac{108}{9} = 12$, soit $12 - 11 = 1$ point de plus. Vrai : une seule note très élevée fait gagner 1 point à la moyenne de tout le groupe.

(e) On retire 7 et 14 : $\frac{88-7-14}{6} = \frac{67}{6} \approx 11{,}17$, ce qui est différent de 11. Faux : pour que la moyenne ne change pas, il faudrait que les deux notes retirées aient pour somme $2 \times 11 = 22$. Or $7 + 14 = 21$.

## 3. *Médiane, quartiles et diagramme en boîte*

**Ex. 15 ★☆☆ Automatismes : médiane et quartiles d'une série brute**

(a) Série rangée : 6 ; 9 ; 9 ; 11 ; 14 ; 17 ; 22 ; 25 ; 30. $N = 9$ est impair : rang $\frac{9+1}{2} = 5$, donc $Me = 14$ SMS.

(b) Série rangée : 37 ; 38 ; 39 ; 40 ; 40 ; 41 ; 41 ; 42. $N = 8$ est pair : demi-somme des rangs 4 et 5, donc $Me = \frac{40+40}{2} = 40$.

(c) $\frac{10}{4} = 2{,}5$, arrondi à 3 : $Q_1$ est la valeur de rang 3.
$\frac{3\times10}{4} = 7{,}5$, arrondi à 8 : $Q_3$ est la valeur de rang 8.

**Ex. 15 (suite)**

(d) Série rangée : 135 ; 138 ; 140 ; 147 ; 149 ; 152 ; 155 ; 158 ; 162 ; 165 ; 171. $\frac{11}{4} = 2{,}75$, arrondi à 3 : $Q_1 = 140$ g.
$\frac{3\times11}{4} = 8{,}25$, arrondi à 9 : $Q_3 = 162$ g. Écart interquartile : $162 - 140 = 22$ g.

(e) Étendue : $30 - 5 = 25$. Écart interquartile : $Q_3 - Q_1 = 19 - 11 = 8$.

(f) $N = 15$ est impair : $Me$ est la valeur de rang $\frac{15+1}{2} = 8$.
$\frac{15}{4} = 3{,}75$, arrondi à 4 : $Q_1$ est la valeur de rang 4.
$\frac{3\times15}{4} = 11{,}25$, arrondi à 12 : $Q_3$ est la valeur de rang 12.

**Ex. 16 ★☆☆ Automatismes : quartiles avec les effectifs cumulés**

(a) Effectifs cumulés : 4 ; $4 + 7 = 11$ ; $11 + 9 = 20$ ; $20 + 6 = 26$ ; $26 + 3 = 29$ ; $29 + 1 = 30$.

(b) Le dernier effectif cumulé donne $N = 30$ élèves.

(c) $N = 30$ est pair : demi-somme des rangs 15 et 16. Les rangs 12 à 20 ont la valeur 2, donc $Me = \frac{2+2}{2} = 2$ livres.

(d) $\frac{30}{4} = 7{,}5$, arrondi à l'entier supérieur : rang 8 (pas 7,5, ni la moyenne des rangs 7 et 8). Les rangs 5 à 11 ont la valeur 1, donc $Q_1 = 1$ livre.

(e) $\frac{3\times30}{4} = 22{,}5$, arrondi à 23. Les rangs 21 à 26 ont la valeur 3, donc $Q_3 = 3$ livres.

(f) $Q_3 - Q_1 = 3 - 1 = 2$ livres. Les élèves de rangs 8 à 23, soit au moins la moitié de la classe, ont lu entre 1 et 3 livres : le nombre de livres lus par les élèves du milieu de la série varie de 2 livres au plus.

(g) 26 élèves ont lu au plus 3 livres : $\frac{26}{30} \approx 0{,}867$, soit environ 86,7 % des élèves.

**Ex. 17 ★★☆ Attente au guichet : l'erreur d'Inès**

(a) Série rangée : 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 ; 10 ; 11 ; 13 ; 18 ; 25. $N = 12$ est pair : demi-somme des rangs 6 et 7, donc $Me = \frac{8+9}{2} = 8{,}5$. La moitié des voyageurs a attendu 8,5 min ou moins.

(b) Non. $Q_1$ n'est pas la médiane de la première moitié : $\frac{12}{4} = 3$, donc $Q_1$ est la valeur de rang 3, soit $Q_1 = 5$ min. D'ailleurs 5,5 n'est pas une valeur de la série, alors qu'un quartile en est toujours une.

(c) $\frac{3\times12}{4} = 9$ : $Q_3$ est la valeur de rang 9, soit $Q_3 = 11$ min. Écart interquartile : $11 - 5 = 6$ min. Environ 50 % des voyageurs, ceux du milieu de la série, ont attendu entre 5 et 11 min.

**Ex. 17** **(suite)**

(d) Avant : $\overline{x} = \frac{119}{12} \approx 9{,}92$ min et $Me = 8{,}5$ min. Après : $N = 13$, la somme vaut $119 + 50 = 169$ et $\overline{x} = \frac{169}{13} = 13$ min. La médiane est la valeur de rang $\frac{13+1}{2} = 7$, soit $Me = 9$ min. La médiane n'avance que d'un cran (8,5 à 9 min), alors que la moyenne gagne environ 3 min : la moyenne est sensible à cette attente extrême, la médiane très peu.

**Ex. 18** ★★☆ **Déchets ménagers : comparer deux boîtes**

(a) Dans l'ordre minimum, $Q_1$, $Me$, $Q_3$, maximum : 6 ; 10 ; 13 ; 22 ; 26.

(b) Quartier A : étendue $31 - 4 = 27$ kg ; écart interquartile $15 - 9 = 6$ kg. Quartier B : étendue $26 - 6 = 20$ kg ; écart interquartile $22 - 10 = 12$ kg.

(c) Non. L'étendue ne dépend que des deux valeurs extrêmes : celle de A est grande à cause du maximum, 31 kg. L'écart interquartile, qui porte sur les foyers du milieu de la série, vaut 12 kg pour B contre 6 kg pour A. Ce sont donc les quantités de déchets du quartier B qui sont les plus dispersées.

(d) Vrai. $\frac{3\times 40}{4} = 30$ : $Q_3 = 15$ kg est la valeur de rang 30. Les foyers de rangs 1 à 30 jettent donc tous au plus 15 kg : au moins 30 foyers jettent 15 kg ou moins.

(e) Non. Comme $18 > 15 = Q_3$, la nouvelle valeur reste placée après le rang 30 : les valeurs de rangs 1 à 30 ne changent pas. $Q_1$ (rang 10), $Me$ (rangs 20 et 21) et $Q_3$ (rang 30) restent 9 kg, 12 kg et 15 kg. Seul le maximum peut changer.

**Ex. 19** ★★★ **Trajets en bus : rangs, quartiles et pourcentages**

(a) Effectifs cumulés croissants : 4 ; 7 ; 16 ; 24 ; 28 ; 30, donc $N = 30$. Minimum : 0 ; maximum : 10. Médiane : demi-somme des rangs 15 et 16 ; les rangs 8 à 16 valent 4, donc $Me = \frac{4+4}{2} = 4$. $\frac{30}{4} = 7{,}5$, arrondi à 8 : $Q_1 = 4$. $\frac{3\times 30}{4} = 22{,}5$, arrondi à 23 : les rangs 17 à 24 valent 6, donc $Q_3 = 6$. Diagramme en boîte : 0 ; 4 ; 4 ; 6 ; 10 (trajets).

(b) Le rang 7 vaut 2 et le rang 8 vaut 4 : Nolan obtient $\frac{2+4}{2} = 3$. Première erreur : le rang 7,5 s'arrondit à l'entier supérieur, 8, on ne fait pas de moyenne de deux rangs. Seconde erreur : $Q_1$ est toujours une valeur de la série, or aucun élève n'a fait 3 trajets.
Donc $Q_1 = 4$ trajets.

**Ex. 19** **(suite)**

(c) Faux. Les élèves ayant fait au plus 4 trajets sont 16 (effectif cumulé), soit $\frac{16}{30} \approx 53{,}3\,\%$. La définition garantit seulement qu'au moins 25 % des valeurs sont inférieures ou égales à $Q_1$. Ici, plus de la moitié des élèves ont fait au plus 4 trajets.

(d) $Q_3 - Q_1 = 6 - 4 = 2$ trajets. Les élèves ayant fait 4 ou 6 trajets sont $9 + 8 = 17$, soit $\frac{17}{30} \approx 56{,}7\,\%$. Lou a tort : l'intervalle $[Q_1\,;Q_3]$ contient environ la moitié des valeurs, au moins 50 %, mais pas forcément exactement la moitié. Ici, 17 élèves sur 30 ont fait entre 4 et 6 trajets.

**Ex. 20** ★★★ **Bagages en soute : la médiane résiste**

(a) Série rangée : 9 ; 12 ; 15 ; 16 ; 18 ; 19 ; 21 ; 23 ; 23, $N = 9$. $Me$ est la valeur de rang $\frac{9+1}{2} = 5$ : $Me = 18$ kg. $\frac{9}{4} = 2{,}25$, arrondi à 3 : $Q_1 = 15$ kg. $\frac{3\times 9}{4} = 6{,}75$, arrondi à 7 : $Q_3 = 21$ kg.

(b) Étendue : $23 - 9 = 14$ kg ; écart interquartile : $21 - 15 = 6$ kg. Avec 4 kg au lieu de 9 kg, ce bagage resterait le plus léger : seul le rang 1 changerait. $Q_1$ (rang 3) vaudrait toujours 15 kg et $Q_3$ (rang 7) toujours 21 kg, donc l'écart interquartile resterait 6 kg. Seule l'étendue changerait : $23 - 4 = 19$ kg. L'écart interquartile n'est pas sensible à une valeur extrême, l'étendue si.

(c) Somme des neuf masses : 156 kg. Pour une moyenne de 19 kg sur 10 bagages, il faut une somme de $10 \times 19 = 190$ kg, donc un bagage de $190 - 156 = 34$ kg. Série : 9 ; 12 ; 15 ; 16 ; 18 ; 19 ; 21 ; 23 ; 23 ; 34. $N = 10$ : $Me = \frac{18+19}{2} = 18{,}5$ kg, pas 19 kg. La moyenne passe d'environ $\frac{156}{9} \approx 17{,}3$ kg à 19 kg, alors que la médiane ne gagne que 0,5 kg : un seul bagage lourd tire la moyenne vers le haut, la médiane bouge à peine.

(d) Avec 10 valeurs, la médiane est la demi-somme des valeurs de rangs 5 et 6. Si $x \leqslant 16$ : ces rangs valent 16 et 18, donc $Me = \frac{16+18}{2} = 17$. Si $16 < x \leqslant 18$ : ils valent $x$ et 18, donc $Me = \frac{x+18}{2}$, entre 17 et 18. Si $18 < x < 19$ : ils valent 18 et $x$, donc $Me$ est entre 18 et 18,5. Si $x \geqslant 19$ : ils valent 18 et 19, donc $Me = 18{,}5$. Quelle que soit la masse du dixième bagage, la médiane vaut au moins 17 kg et au plus 18,5 kg.

**Ex. 21** ★★★ **Natation : ce que disent deux boîtes**

(a) Dauphins : $Me = 34$ s, écart interquartile $38 - 31 = 7$ s. Orques : $Me = 35$ s, écart interquartile $36 - 33 = 3$ s. En médiane, les Dauphins sont un peu plus rapides (34 s contre 35 s). Mais les temps des Orques sont bien plus resserrés : ce club est plus homogène.

**Ex. 21** (suite)

(b) Faux. $N = 36$ est pair : $Me = 35$ est la demi-somme des valeurs de rangs 18 et 19. Comme la valeur de rang 19 est supérieure ou égale à celle de rang 18, elle vaut au moins 35. Les nageurs de rangs 19 à 36, soit $36 - 19 + 1 = 18$ nageurs, mettent donc au moins 35 s. Au plus 18 Orques sur 36, donc au plus la moitié, nagent en moins de 34 s.

(c) Dauphins : $N = 40$, $Me = 34$ est la demi-somme des valeurs de rangs 20 et 21. La valeur de rang 20 vaut donc au plus 34 : les nageurs de rangs 1 à 20 mettent 34 s ou moins, soit au moins 20 Dauphins. Orques : d'après la question précédente, les 18 nageurs de rangs 19 à 36 mettent au moins 35 s, donc au plus 18 Orques nagent en 34 s ou moins. Comme $20 > 18$, il y a plus de Dauphins que d'Orques à nager en 34 s ou moins.

(d) Orques : $\frac{36}{4} = 9$, donc $Q_1 = 33$ s est la valeur de rang 9. Les nageurs de rangs 9 à 36, soit $36 - 9 + 1 = 28$ nageurs, mettent au moins 33 s, plus que Timéo : $\frac{28}{36} \approx 77{,}8\,\%$, donc au moins 75 %. Dauphins : $\frac{40}{4} = 10$, donc les nageurs de rangs 1 à 10 mettent au plus $Q_1 = 31$ s et sont plus rapides que Timéo. Seuls les rangs 21 à 40, soit 20 nageurs (50 %), mettent à coup sûr au moins 34 s. Les rangs 11 à 20 sont entre 31 s et 34 s, donc on ne peut pas savoir s'ils sont plus lents que Timéo. On ne peut pas l'affirmer pour les Dauphins.

## 4. *Écart type et comparaison de deux séries*

**Ex. 22** ★☆☆ **Automatismes : variance et écart type**

(a) Toutes les valeurs sont égales à la moyenne 7, donc tous les écarts sont nuls : $\sigma = 0$.

(b) $\overline{x} = \frac{3+6+8+11}{4} = 7$. Écarts : $-4$ ; $-1$ ; $1$ ; $4$. Leur moyenne vaut $\frac{-4-1+1+4}{4} = 0$. Elle vaut toujours 0, car les écarts négatifs compensent les positifs. C'est pour ça qu'on élève les écarts au carré.

(c) $\sigma = \sqrt{6{,}25} = 2{,}5$ kg. La variance est en kg², l'écart type en kg.

(d) Faux : $\sigma$ est la racine carrée d'une variance, donc $\sigma \geqslant 0$.

(e) Ajouter 3 cm ne change pas les écarts à la moyenne, donc l'écart type reste égal à 4 cm.

(f) Chaque écart est multiplié par 2, donc $\sigma = 2 \times 1{,}20 = 2{,}40$ €.

**Ex. 23** ★☆☆ **Écart type à partir d'un tableau d'effectifs**

(a) $\overline{x} = \frac{2\times0+3\times1+4\times2+1\times4}{10} = \frac{15}{10} = 1{,}5$ livre.

(b) Écarts : $-1{,}5$ ; $-0{,}5$ ; $0{,}5$ ; $2{,}5$. Carrés : $2{,}25$ ; $0{,}25$ ; $0{,}25$ ; $6{,}25$. Carrés × effectifs : $2 \times 2{,}25 = 4{,}5$ ; $3 \times 0{,}25 = 0{,}75$ ; $4 \times 0{,}25 = 1$ ; $1 \times 6{,}25 = 6{,}25$. Somme : 12,5.

(c) $V = \frac{12{,}5}{10} = 1{,}25$.

(d) $\sigma = \sqrt{1{,}25} \approx 1{,}12$ livre. Cet écart type mesure la dispersion des emprunts autour de la moyenne de 1,5 livre.

(e) $V$ s'exprime en livres², une unité qui n'a pas de sens concret. $\sigma$ s'exprime en livres, comme les données : on peut donc l'interpréter directement.

**Ex. 24** ★★☆ **Bus ou vélo : qui est le plus régulier ?**

(a) Inès : $\frac{14+26+18+22+15+25}{6} = \frac{120}{6} = 20$ min. Hugo : $\frac{19+21+20+18+22+20}{6} = \frac{120}{6} = 20$ min.

(b) Écarts à 20 : $-1$ ; $1$ ; $0$ ; $-2$ ; $2$ ; $0$. Carrés : $1$ ; $1$ ; $0$ ; $4$ ; $4$ ; $0$. $V = \frac{10}{6} \approx 1{,}67$, donc $\sigma = \sqrt{\frac{10}{6}} \approx 1{,}29$ min.

(c) Les deux ont le même temps moyen (20 min). L'écart type de Hugo (1,29 min) est bien plus petit que celui d'Inès (4,65 min). Les trajets à vélo de Hugo sont donc beaucoup plus réguliers que les trajets en bus d'Inès.

(d) Inès : entre $20 - 4{,}65 = 15{,}35$ et $20 + 4{,}65 = 24{,}65$ min, seuls 18 et 22 conviennent, soit 2 trajets sur 6. Hugo : entre $20 - 1{,}29 = 18{,}71$ et $20 + 1{,}29 = 21{,}29$ min, on trouve 19 ; 21 ; 20 ; 20, soit 4 trajets sur 6. Oui : les temps de Hugo sont plus resserrés autour de la moyenne, donc ses trajets à vélo sont plus réguliers.

**Ex. 25** ★★☆ **Températures : programme, conversion et comparaison**

(a) $m = \frac{50}{5} = 10$. Écarts : $-2$ ; $2$ ; $0$ ; $4$ ; $-4$, donc $s = 4 + 4 + 0 + 16 + 16 = 40$. La fonction renvoie $\frac{40}{5} = 8$ (Python affiche `8.0`). Lina a tort : c'est la variance, en °C², pas l'écart type. Elle a importé `sqrt` mais a oublié de prendre la racine carrée.

(b) Dernière ligne : `return sqrt(s / n)`. On obtient $\sigma = \sqrt{8} \approx 2{,}83$ °C.

(c) Moyenne : $1{,}8 \times 10 + 32 = 50$ °F. Le $+32$ ne change pas l'écart type, donc $\sigma = 1{,}8 \times \sqrt{8} \approx 5{,}09$ °F.

**Ex. 25 (suite)**

(d) On ne mélange pas les couples : on compare (moyenne ; $\sigma$) de Colmar avec (moyenne ; $\sigma$) de Brest. Colmar : (10 °C ; 2,83 < *spanstyle* = "*white* − *space* : *nowrap*" > °C) < /*span* > ; Brest : (11,5 °C ; 1,4 < *spanstyle* = "*white* − *space* : *nowrap*" > °C) < /*span* >. À 7 h, il fait en moyenne 1,5 °C de plus à Brest, et les températures y sont plus régulières (écart type environ deux fois plus petit).

**Ex. 26 ★★★ Retards d'un TER : une valeur extrême**

(a) $\overline{x} = \frac{1+5+3+4+6+5}{6} = \frac{24}{6} = 4$ min. Écarts : $-3$ ; $1$ ; $-1$ ; $0$ ; $2$ ; $1$. Carrés : $9$ ; $1$ ; $1$ ; $0$ ; $4$ ; $1$, de somme $16$. $V = \frac{16}{6} \approx 2{,}67$ min² (unité au carré, sans interprétation directe), puis $\sigma = \sqrt{\frac{16}{6}} \approx 1{,}63$ min. Il ne faut pas oublier la racine carrée.

(b) $\overline{x} = \frac{24+32}{7} = \frac{56}{7} = 8$ min. Le retard de 32 min est très loin de la moyenne : son écart vaut 24 et son carré 576, sur une somme totale de 688. Cette seule valeur extrême gonfle la variance, donc l'écart type, qui passe de 1,63 min à 9,91 min.

(c) Série rangée : $1$ ; $3$ ; $4$ ; $5$ ; $5$ ; $6$ ; $32$, avec $N = 7$. $Me$ est la $4^{e}$ valeur, donc $Me = 5$ min. $\frac{7}{4} = 1{,}75$ : on prend le rang 2, donc $Q_1 = 3$ min. $\frac{3\times 7}{4} = 5{,}25$ : on prend le rang 6, donc $Q_3 = 6$ min. Écart interquartile : $6 - 3 = 3$ min. La panne ne change presque pas ces indicateurs.

(d) Léo compare une moyenne (celle de A) avec une médiane (celle de B) : ces indicateurs ne se comparent pas entre eux. De plus, il ne regarde pas la dispersion. Il faut comparer des couples cohérents, (moyenne ; écart type) ou (médiane ; écart interquartile), pris de la même façon pour les deux lignes.

(e) La ligne A a une valeur extrême (la panne), qui fausse la moyenne et l'écart type : avec (moyenne ; $\sigma$), A (8 min ; 9,91 < *spanstyle* = "*white* − *space* : *nowrap*" > min) < /*span* > paraîtrait pire que B (6,5 min ; 2,4 < *spanstyle* = "*white* − *space* : *nowrap*" > min) < /*span* >. Le couple (médiane ; écart interquartile) est donc le mieux adapté. A : (5 min; 3 min) ; B : (6 min; 4 min). Conclusion : en dehors de la panne exceptionnelle, le TER de la ligne A a des retards plus courts (5 min au lieu de 6 min) et plus réguliers que celui de la ligne B.

**Ex. 27 ★★★ Notes converties : ce que devient l'écart type**

(a) $\overline{x} = \frac{3+5+6+8+8}{5} = \frac{30}{5} = 6$. Écarts : $-3$ ; $-1$ ; $0$ ; $2$ ; $2$. Carrés : $9$ ; $1$ ; $0$ ; $4$ ; $4$, de somme $18$. $V = \frac{18}{5} = 3{,}6$, donc $\sigma_x = \sqrt{3{,}6} \approx 1{,}90$ point.

(b) Faux. Par linéarité, $\overline{z} = \overline{x} + 1$, donc $z_i - \overline{z} = (x_i + 1) - (\overline{x} + 1) = x_i - \overline{x}$. Les écarts sont inchangés, donc la variance et l'écart type aussi : $\sigma_z = \sigma_x \approx 1{,}90$ point. Le bonus décale toutes les notes sans les rapprocher ni les écarter.

(c) $\overline{y} = 2 \times 6 + 1 = 13$. Série rangée : $3$ ; $5$ ; $6$ ; $8$ ; $8$, donc $Me_x = 6$. L'ordre est conservé ($a = 2 > 0$), donc $Me_y = 2 \times 6 + 1 = 13$. Enfin $\sigma_y = 2\,\sigma_x = 2\sqrt{3{,}6} \approx 3{,}79$ points : le $+1$ ne change pas $\sigma$, mais le $\times 2$ le double. Sur 20, la moyenne est 13 et l'écart type environ 3,79 points.

(d) $\overline{y} = 2\,\overline{x} + 1$, donc $\overline{x} = \frac{14{,}6-1}{2} = 6{,}8$. $\sigma_y = 2\,\sigma_x$ (le $+1$ ne compte pas), donc $\sigma_x = \frac{3{,}2}{2} = 1{,}6$. Sur 10, cette classe a une moyenne de 6,8 et un écart type de 1,6 point.

**Ex. 28 ★★★ Handball : un programme pour départager deux joueuses**

(a) Maëlle : $\overline{x} = \frac{2\times 1+3\times 2+4\times 4+5\times 2+6\times 1}{10} = \frac{40}{10} = 4$ buts. Salomé : $\overline{x} = \frac{1\times 2+2\times 2+4\times 2+6\times 2+7\times 2}{10} = \frac{40}{10} = 4$ buts. Les deux moyennes sont égales : la moyenne ne suffit pas pour les départager.

(b) À la fin, `N` est la somme des effectifs, soit le nombre de matchs (10). `s` est la somme des carrés des écarts à la moyenne, chacun multiplié par son effectif. `len(V)` vaut 5 : c'est le nombre de valeurs différentes, pas le nombre de matchs. Diviser par 5 donnerait une variance fausse.

(c) Dernière ligne : `return sqrt(s / N)` (sans la racine, on obtiendrait la variance). Maëlle : écarts $-2$ ; $-1$ ; $0$ ; $1$ ; $2$. Carrés × effectifs : $4 \times 1 + 1 \times 2 + 0 \times 4 + 1 \times 2 + 4 \times 1 = 12$. $V = \frac{12}{10} = 1{,}2$, donc $\sigma = \sqrt{1{,}2} \approx 1{,}10$ but. La fonction renvoie environ 1,10.

(d) Les moyennes sont égales : on utilise le couple (moyenne ; écart type). Maëlle : $(4; 1{,}10)$ ; Salomé : $(4; 2{,}28)$. Avec la même moyenne, la série de Maëlle a un écart type plus de deux fois plus petit. Conclusion : les deux joueuses marquent en moyenne 4 buts par match, mais Maëlle est la plus régulière ; c'est elle que l'entraîneur devrait choisir.

**Ex. 28 (suite)**

(e) La valeur 4 est égale à la moyenne. La moyenne reste $\frac{40+4}{11} = 4$ buts. L'écart de cette valeur est nul, donc la somme des carrés reste 12, mais on divise maintenant par 11 : $\sigma = \sqrt{\frac{12}{11}} \approx 1{,}04$ but. L'écart type diminue légèrement : Maëlle devient encore un peu plus régulière.

## 5. *Séries en classes : moyenne, effectifs cumulés, classe médiane*

**Ex. 29** ★☆☆ **Automatismes : amplitude, centre, histogramme**

(a) Amplitude : $20 - 15 = 5$ min. Centre : $\frac{15+20}{2} = 17{,}5$ min.

(b) $20 = 140 - 120$, c'est l'amplitude. Le centre vaut $\frac{120+140}{2} = 130$.

(c) Les bornes sont à $\frac{4}{2} = 2$ du centre : $26 - 2 = 24$ et $26 + 2 = 28$. C'est la classe $[24\,;28[$.

(d) Les amplitudes sont égales, donc les hauteurs sont proportionnelles aux effectifs : 1 cm correspond à $\frac{12}{3} = 4$ valeurs. Effectif : $4 \times 4{,}5 = 18$.

(e) Ce sont les valeurs des classes $[0\,;10[$ et $[10\,;20[$ : $4 + 9 = 13$ valeurs, c'est l'effectif cumulé de $[10\,;20[$.

(f) Non : on sait seulement que le maximum est dans $[20\,;30[$, donc qu'il est supérieur ou égal à 20 et strictement inférieur à 30. Les valeurs ne sont pas connues une par une.

**Ex. 30** ★☆☆ **Automatismes : effectifs et fréquences cumulés**

(a) $9$ ; $9 + 13 = 22$ ; $22 + 15 = 37$ ; $37 + 2 = 39$ ; $39 + 1 = 40$.

(b) C'est l'effectif cumulé de $[10\,;20[$ : 22 élèves.

(c) $40 - 37 = 3$ élèves
(37 élèves ont un trajet strictement inférieur à 30 min).

(d) $\frac{37}{40} = 0{,}925$, soit 92,5 % des élèves mettent moins de 30 min.

(e) Premier point $(0\,;0)$, puis les bornes supérieures : $(10\,;22{,}5)$, $(20\,;55)$, $(30\,;92{,}5)$, $(40\,;97{,}5)$, $(50\,;100)$.

(f) $\frac{N}{2} = \frac{40}{2} = 20$. Le premier effectif cumulé qui atteint 20 est 22 : la classe médiane est $[10\,;20[$. Ce n'est pas la classe de plus grand effectif, qui est $[20\,;30[$ (15 élèves).

**Ex. 31** ★★☆ **Retards de trains : moyenne estimée ou exacte ?**

(a) Tom a pris l'amplitude 4 de chaque classe au lieu de son centre : son calcul donne $\frac{120}{30} = 4$, ce qui n'a pas de sens. Avec les centres 2 ; 6 ; 10 ; 14 (répartition uniforme) : $\overline{x} \approx \frac{12\times2+9\times6+6\times10+3\times14}{30} = \frac{24+54+60+42}{30} = \frac{180}{30} = 6$. Le retard moyen est d'environ 6 min.

(b) Moyenne pondérée par les moyennes des classes : $\overline{x} = \frac{12\times1{,}5+9\times5{,}8+6\times9{,}5+3\times13}{30} = \frac{18+52{,}2+57+39}{30} = \frac{166{,}2}{30} = 5{,}54$. Le retard moyen de ces trains est de 5,54 min.

(c) La valeur exacte est 5,54 min, calculée avec les vraies moyennes des classes ; 6 min est une estimation, notée « $\approx$ ». Dans chaque classe, la moyenne est inférieure au centre ($1{,}5 < 2$, $5{,}8 < 6$, $9{,}5 < 10$, $13 < 14$) : les retards sont plutôt dans le bas des classes. Supposer une répartition uniforme surestime donc la moyenne.

**Ex. 32** ★★☆ **Déchets ménagers : la médiane à partir du polygone**

(a) Fréquences des classes : 16 % ; $44 - 16 = 28$ % ; $62 - 44 = 18$ % ; $94 - 62 = 32$ % ; $100 - 94 = 6$ %. Effectifs : $0{,}16 \times 50 = 8$ ; $0{,}28 \times 50 = 14$ ; $0{,}18 \times 50 = 9$ ; $0{,}32 \times 50 = 16$ ; $0{,}06 \times 50 = 3$ foyers.

(b) Non. En 6, la fréquence cumulée vaut 44 % : moins de la moitié des foyers jette moins de 6 kg, donc la médiane n'est pas dans $[3\,;6[$. La fréquence cumulée atteint 50 % pour la première fois en 9 (62 %). La classe médiane est donc $[6\,;9[$, même si 44 % est plus proche de 50 % que 62 %.

(c) $\frac{N}{2} = 25$. Avant la classe $[6\,;9[$, l'effectif cumulé est $C = 8 + 14 = 22$ ; il manque $25 - 22 = 3$ foyers sur les 9 de la classe, d'amplitude 3 kg. $Me \approx 6 + 3 \times \frac{25-22}{9} = 6 + 1 = 7$. La moitié des foyers jette environ 7 kg de déchets non recyclés par semaine, ou moins.

(d) Vrai. 16 % des foyers jettent moins de 3 kg et 94 % moins de 12 kg.
Donc $94 - 16 = 78$ % des foyers jettent entre 3 kg (inclus) et 12 kg (exclu), soit $14 + 9 + 16 = 39$ foyers sur 50. Comme $78\,\% > 75\,\%$, c'est bien plus des trois quarts.

**Ex. 33** ★★★ **Course de 10 km : l'histogramme ne dit pas tout**

(a) Amplitude : $45 - 40 = 5$ min pour chaque classe. Comme toutes les classes ont la même amplitude, les hauteurs sont proportionnelles aux effectifs : 2 cm correspondent à 10 coureurs, donc 1 cm à 5 coureurs. Effectifs : $2 \times 5 = 10$ ; $6{,}8 \times 5 = 34$ ; $6 \times 5 = 30$ ; $3{,}2 \times 5 = 16$ ; $2 \times 5 = 10$. Donc $N = 10 + 34 + 30 + 16 + 10 = 100$ coureurs.

(b) Effectifs cumulés croissants : 10 ; 44 ; 74 ; 90 ; 100. $\frac{N}{2} = 50$. Le cumul 44 n'atteint pas 50 ; le premier cumul qui atteint 50 est 74, celui de $[50\,;55[$. Léa a tort : la classe médiane est $[50\,;55[$, pas la classe de plus grand effectif.

(c) Avant la classe médiane, $C = 44$ ; dans la classe, $n = 30$. $Me \approx 50 + 5 \times \frac{50-44}{30} = 50 + 5 \times 0{,}2 = 51$. La moitié des coureurs a mis environ 51 min ou moins.

(d) Centres : 42,5 ; 47,5 ; 52,5 ; 57,5 ; 62,5. $\overline{x} \approx \frac{10\times42{,}5+34\times47{,}5+30\times52{,}5+16\times57{,}5+10\times62{,}5}{100} = \frac{425+1\,615+1\,575+920+625}{100} = \frac{5\,160}{100} = 51{,}6$. Le temps moyen est d'environ 51,6 min.

(e) Les deux affirmations sont fausses. Seuls 44 coureurs sur 100 ont mis moins de 50 min, soit 44 % et non la moitié. Le temps médian estimé est d'environ 51 min, mais ce n'est qu'une estimation (répartition uniforme supposée) : la médiane exacte est inconnue. Le dernier coureur est dans $[60\,;65[$ : son temps est strictement inférieur à 65 min, et sa valeur exacte est inconnue.

**Ex. 34** ★★★ **Consommation d'électricité : un effectif effacé**

(a) $N = 28 + x$ et la somme des consommations vaut $8 \times 160 + 245x + 14 \times 352 + 6 \times 430 = 8\,788 + 245x$. On résout $8\,788 + 245x = 293{,}2(28 + x)$, soit $8\,788 + 245x = 8\,209{,}6 + 293{,}2x$, donc $578{,}4 = 48{,}2x$ et $x = 12$. La classe $[200\,;300[$ compte 12 foyers ($N = 40$).

(b) Centres : 150 ; 250 ; 350 ; 450. $\overline{x} \approx \frac{8\times150+12\times250+14\times350+6\times450}{40} = \frac{1\,200+3\,000+4\,900+2\,700}{40} = \frac{11\,800}{40} = 295$. Malik obtient environ 295 kWh : il remplace la moyenne de chaque classe par son centre (répartition uniforme supposée), alors que les vraies moyennes (160, 245…) ne sont pas les centres. Seul 293,2 kWh est exact.

**Ex. 34** (suite)

(c) Fréquences : $\frac{8}{40} = 20\,\%$ ; $\frac{12}{40} = 30\,\%$ ; $\frac{14}{40} = 35\,\%$ ; $\frac{6}{40} = 15\,\%$. Fréquences cumulées : 20 % ; 50 % ; 85 % ; 100 %. 20 % des foyers consomment moins de 200 kWh : la fréquence cumulée se place sur la borne supérieure. Points : $(100\,;0)$, $(200\,;20\,\%)$, $(300\,;50\,\%)$, $(400\,;85\,\%)$, $(500\,;100\,\%)$.

(d) La fréquence cumulée en 400 est 85 % : 85 % des foyers consomment moins de 400 kWh. Moins de 300 kWh : 50 %, donc au moins 300 kWh : $100 - 50 = 50\,\%$ des foyers.

(e) La fréquence cumulée atteint 50 % en 300 : $Me \approx 300$ kWh. Contrôle : $\frac{N}{2} = 20$, la classe médiane est $[200\,;300[$ (premier cumul $\geqslant 20$) et $Me \approx 200 + 100 \times \frac{20-8}{12} = 300$. C'est une estimation : on ne connaît pas les consommations individuelles, donc la médiane exacte reste inconnue.

**Ex. 35** ★★★ **Sommeil à l'internat : internes contre externes**

(a) 27 est l'effectif cumulé de $[7\,;8[$ : 27 internes ont dormi strictement moins de 8 h. Effectifs des internes : 3 ; $13 - 3 = 10$ ; $27 - 13 = 14$ ; $37 - 27 = 10$ ; $40 - 37 = 3$.

(b) Centres : 5,5 ; 6,5 ; 7,5 ; 8,5 ; 9,5. Internes : $\overline{x} \approx \frac{3\times5{,}5+10\times6{,}5+14\times7{,}5+10\times8{,}5+3\times9{,}5}{40} = \frac{300}{40} = 7{,}5$ h. Externes : $\overline{x} \approx \frac{9\times5{,}5+21\times6{,}5+18\times7{,}5+9\times8{,}5+3\times9{,}5}{60} = \frac{426}{60} = 7{,}1$ h. Les internes dorment en moyenne environ 7,5 h, les externes environ 7,1 h.

(c) Internes : $\frac{40}{2} = 20$ ; cumuls 3 ; 13 ; 27 : la classe médiane est $[7\,;8[$ et $Me \approx 7 + 1 \times \frac{20-13}{14} = 7{,}5$ h. Externes : $\frac{60}{2} = 30$ ; cumuls 9 ; 30 : 30 atteint 30, la classe médiane est $[6\,;7[$ et $Me \approx 6 + 1 \times \frac{30-9}{21} = 7$ h. La moitié des internes dort environ 7,5 h ou moins, contre 7 h pour les externes.

(d) Internes : $40 - 27 = 13$ ont dormi au moins 8 h, soit $\frac{13}{40} = 32{,}5\,\%$. Externes : $9 + 3 = 12$, soit $\frac{12}{60} = 20\,\%$. Moyenne, médiane et part des élèves dormant au moins 8 h sont plus élevées chez les internes : ce sont eux qui dorment le plus.

(e) Non. Les groupes n'ont pas le même effectif (40 et 60). Le calcul $\frac{7{,}5+7{,}1}{2} = 7{,}3$ donne autant de poids aux 40 internes qu'aux 60 externes. On ajoute les sommes $N \times \overline{x}$ : $\overline{x} \approx \frac{40\times7{,}5+60\times7{,}1}{100} = \frac{300+426}{100} = 7{,}26$. La durée moyenne de sommeil des 100 élèves est d'environ 7,26 h. Ici, l'écart est faible parce que les deux moyennes sont proches, mais la méthode de Ben est fausse.

## 6. *Tableau croisé d'effectifs, à compléter à partir de pourcentages*

**Ex. 36** ★☆☆ **Pourcentage de quoi ?**

(a) Les 25 % portent sur tout le lycée : $0{,}25 \times 560 = 140$. Le lycée compte 140 internes.

(b) Par différence : $560 - 140 = 420$. Le lycée compte 420 externes.

(c) Les 35 % portent sur les internes : $0{,}35 \times 140 = 49$. Il y a 49 internes en seconde.

(d) Les 45 % portent sur les externes : $0{,}45 \times 420 = 189$. 189 externes viennent en bus.

(e) Lina a multiplié par $N = 560$ alors que les 35 % portent sur les 140 internes. D'ailleurs, $196 > 140$ : il ne peut pas y avoir plus d'internes en seconde que d'internes.

(f) $\frac{49}{560} = 0{,}0875$. Les internes de seconde représentent 8,75 % des élèves du lycée.

**Ex. 37** ★☆☆ **Un tableau en pourcentages du total**

(a) Par la ligne : $50 - 13 - 25 = 12$. Par la colonne : $30 - 18 = 12$. Les deux calculs concordent : 12 % des spectateurs ont vu une comédie au tarif réduit.

(b) $0{,}22 \times 400 = 88$. 88 spectateurs ont vu un film d'action au plein tarif.

(c) $0{,}35 \times 400 = 140$. 140 spectateurs ont vu un film d'animation.

(d) Les pourcentages portent tous sur le même total, on peut donc faire le quotient : $\frac{25}{50} = 0{,}5$. La moitié des spectateurs au tarif réduit ont vu un film d'animation.

(e) Faux. Les 10 % portent sur tous les spectateurs. Parmi ceux qui ont vu un film d'animation : $\frac{10}{35} \approx 0{,}286$, soit environ 28,6 % ont payé le plein tarif.

(f) $0{,}18 \times 230 = 41{,}4$ : ce n'est pas un nombre entier de spectateurs. L'effectif réel est 41 ($\frac{41}{230} \approx 17{,}8\,\%$) ou 42 ($\frac{42}{230} \approx 18{,}3\,\%$), et le pourcentage affiché est donc arrondi.

**Ex. 38** ★★☆ **Pourcentages emboîtés dans un réseau de bus**

(a) On commence par la colonne Total, car les 75 % et les 55 % portent sur les abonnés et sur les non-abonnés. Abonnés : $0{,}4 \times 1\,200 = 480$ ; non-abonnés : $1\,200 - 480 = 720$. Ensuite, les abonnés satisfaits : $0{,}75 \times 480 = 360$, et les non-abonnés satisfaits : $0{,}55 \times 720 = 396$. Par différences : $480 - 360 = 120$ et $720 - 396 = 324$. Totaux des colonnes : $360 + 396 = 756$ satisfaits et $120 + 324 = 444$ non satisfaits.

(b) Par la colonne Total : $480 + 720 = 1\,200$. Par la ligne Total : $756 + 444 = 1\,200$. Les deux totaux concordent.

(c) $\frac{756}{1\,200} = 0{,}63$ : 63 % des usagers sont satisfaits, et non $\frac{75+55}{2} = 65\,\%$. Le journal a fait la moyenne de pourcentages qui portent sur des groupes d'effectifs différents (480 et 720). Les non-abonnés, moins satisfaits, sont plus nombreux, donc ils pèsent davantage.

(d) $\frac{360}{756} \approx 0{,}476$. Parmi les usagers satisfaits, environ 47,6 % sont abonnés.

**Ex. 39** ★★☆ **Du fichier Python au tableau croisé**

(a) La fonction compte les rangs où l'on lit à la fois « moyen » et « oui » : ce sont les rangs 2, 6 et 9. L'appel renvoie 3 : 3 salariés ont un niveau moyen et passent plus de 3 h par jour sur un écran.

(b) Plus de 3 h : 2, 3, 1, soit 6 au total. 3 h au plus : 1, 1, 2, soit 4 au total. Totaux des colonnes : $2 + 1 = 3$, $3 + 1 = 4$ et $1 + 2 = 3$. Contrôle : $3 + 4 + 3 = 10$ et $6 + 4 = 10$. Le niveau d'activité est une variable qualitative ordinale : ses modalités ont un ordre naturel, et on le respecte.

(c) Avec `or`, la fonction compte les salariés de niveau faible OU à plus de 3 h d'écran. On note $A$ « niveau faible » et $B$ « plus de 3 h » : $n(A) = 3$ et $n(B) = 6$. En faisant $3 + 6 = 9$, Inès compte deux fois les 2 salariés qui sont dans les deux cas. $n(A \cup B) = 3 + 6 - 2 = 7$ : l'appel renvoie 7.

(d) $\frac{1}{6} \approx 0{,}167$. Parmi les salariés qui passent plus de 3 h par jour sur un écran, environ 16,7 % ont un niveau d'activité élevé.

**Ex. 40** ★★★ **Une inconnue dans le tableau de la médiathèque**

(a) Les 40 % portent sur les $x$ adultes : $0{,}4x$ adultes empruntent des BD. Il y a $1\,500 - x$ jeunes, et les 70 % portent sur eux : $0{,}7(1\,500 - x)$ jeunes empruntent des BD.

(b) $0{,}4x + 0{,}7(1\,500 - x) = 717$, soit $0{,}4x + 1\,050 - 0{,}7x = 717$, donc $-0{,}3x = -333$ et $x = \frac{333}{0{,}3} = 1\,110$. La médiathèque compte 1 110 adultes inscrits.

(c) Jeunes : $1\,500 - 1\,110 = 390$. BD : adultes $0{,}4 \times 1\,110 = 444$, jeunes $0{,}7 \times 390 = 273$ (contrôle : $444 + 273 = 717$). Pas de BD : $1\,110 - 444 = 666$ adultes, $390 - 273 = 117$ jeunes, soit $666 + 117 = 783$ au total. Contrôle : $717 + 783 = 1\,500$ et $1\,110 + 390 = 1\,500$.

(d) Parmi les 717 emprunteurs de BD, la part des jeunes est $\frac{273}{717} \approx 0{,}381$, soit environ 38,1 %. Nathan a tort : les 70 % portent sur les jeunes, pas sur les emprunteurs. La plupart des emprunteurs de BD sont des adultes (444 sur 717).

**Ex. 41** ★★★ **Satisfaction au camping : une échelle ordonnée**

(a) Familles : $0{,}6 \times 250 = 150$ ; autres : $250 - 150 = 100$. Familles : satisfaites $0{,}5 \times 150 = 75$, très satisfaites $0{,}2 \times 150 = 30$, peu satisfaites $150 - 75 - 30 = 45$. Autres : satisfaits $0{,}4 \times 100 = 40$, très satisfaits $0{,}36 \times 100 = 36$, peu satisfaits $100 - 40 - 36 = 24$. Totaux : $45 + 24 = 69$, $75 + 40 = 115$, $30 + 36 = 66$. Contrôle : $69 + 115 + 66 = 250$ et $150 + 100 = 250$.

(b) Les 33 % porteraient sur les 150 familles : $0{,}33 \times 150 = 49{,}5$. Or un nombre de séjours est forcément entier. Cette donnée est donc une erreur de lecture ou de saisie.

(c) On note $F$ « séjour en famille » et $T$ « très satisfait ». La case « famille et très satisfait » (30) appartient aux deux groupes : on ne la compte qu'une fois. $n(F \cup T) = 150 + 66 - 30 = 186$, et $\frac{186}{250} = 0{,}744$. Ainsi, 74,4 % des séjours sont en famille ou très satisfaits.

(d) « Au moins satisfait » signifie satisfait ou très satisfait. Familles : $\frac{75+30}{150} = \frac{105}{150} = 0{,}7$, soit 70 %. Autres : $\frac{40+36}{100} = 0{,}76$, soit 76 %. Les autres séjours sont un peu plus souvent au moins satisfaits que les séjours en famille.

**Ex. 42** ★★★ **Trajets domicile-travail : du pourcentage du total à l'effectif**

(a) Voiture seule en périphérie : $60 - 15 - 12 = 33$ % ; total ville : $100 - 60 = 40$ % ; covoiturage en ville : $40 - 12 - 22 = 6$ %. Totaux : covoiturage $6 + 15 = 21$ %, vélo ou bus $22 + 12 = 34$ %. Contrôle : $45 + 21 + 34 = 100$ et $40 + 60 = 100$ ; la colonne voiture : $12 + 33 = 45$.

(b) La case « périphérie et vélo ou bus » vaut 12 % du total : $0{,}12 \times N = 60$, donc $N = \frac{60}{0{,}12} = 500$. Covoitureurs : $0{,}21 \times 500 = 105$. L'entreprise a interrogé 500 salariés, dont 105 covoitureurs.

(c) Faux. Le journal a comparé les cases 33 % et 12 % ( $\frac{33}{12} = 2{,}75$). Ce sont des pourcentages du total, pas des groupes « ville » et « périphérie ». Parmi les habitants de la périphérie : $\frac{33}{60} = 0{,}55$, soit 55 % ; parmi ceux de la ville : $\frac{12}{40} = 0{,}3$, soit 30 %. Or $\frac{0{,}55}{0{,}3} \approx 1{,}83$ : la voiture seule est environ 1,8 fois plus fréquente en périphérie, pas presque trois fois.

(d) Parmi les covoitureurs (21 % du total), ceux de la périphérie représentent 15 % du total : $\frac{15}{21} \approx 0{,}714$ (ou $\frac{75}{105}$). Environ 71,4 % des covoitureurs habitent en périphérie.

## 7. *Fréquences marginales et conditionnelles*

**Ex. 43** ★☆☆ **Automatismes : spectateurs d'un festival**

(a) Sans « parmi », on divise par $N = 320$ : $\frac{100}{320} = 0{,}3125$, soit 31,25 %, environ 31,3 % des spectateurs.

(b) Case « réduit et train » divisée par $N$ : $\frac{54}{320} = 0{,}16875$, soit environ 16,9 % de tous les spectateurs.

(c) « Parmi les tarifs réduits » donne le dénominateur 120 : $\frac{54}{120} = 0{,}45$, soit 45 % des spectateurs au tarif réduit.

(d) Le dénominateur devient 100 : $\frac{54}{100} = 0{,}54$, soit 54 %. Le numérateur est le même (54), mais la population de référence a changé (100 au lieu de 120).

**Ex. 44** ★☆☆ **Automatismes : musique ou sport ?**

(a) Instrument sans club : $50 - 16 = 34$. Total sans instrument : $200 - 50 = 150$. Sans instrument, en club : $84 - 16 = 68$. Total sans club : $200 - 84 = 116$. Sans instrument ni club : $150 - 68 = 82$ (vérification : $116 - 34 = 82$).

**Ex. 44 (suite)**

(b) Lina compte deux fois les 16 élèves qui font les deux. Le OU est inclusif : $50 + 84 - 16 = 118$ élèves.

(c) $\frac{118}{200} = 0{,}59$, soit 59 % des élèves.

(d) « Ni… ni » est le contraire du OU : $200 - 118 = 82$ élèves (c'est bien la case « pas d'instrument, pas de club »). Fréquence : $\frac{82}{200} = 0{,}41$, soit 41 % des élèves.

(e) NON : $200 - 50 = 150$ élèves, donc $\frac{150}{200} = 0{,}75$, soit 75 % des élèves.

**Ex. 45 ★★☆ Le casque à vélo et un titre de presse**

(a) On divise par le nombre de blessés : $\frac{28}{45} \approx 0{,}622$, soit environ 62,2 %. Le chiffre est exact : plus de 60 % des blessés portaient un casque.

(b) Avec casque : $\frac{28}{480} \approx 0{,}058$, soit environ 5,8 %. Sans casque : $\frac{17}{120} \approx 0{,}142$, soit environ 14,2 %.

(c) Les deux groupes n'ont pas la même taille : 480 cyclistes avec casque contre 120 sans. Il faut comparer des fréquences conditionnelles « parmi les cyclistes avec ou sans casque », pas des effectifs. Le titre inverse aussi le conditionnement : il donne la part des casqués parmi les blessés. Conclusion : sans casque, la fréquence des blessés est environ 2,4 fois plus grande (14,2 % contre 5,8 %). L'enquête ne montre pas que le casque est inutile.

(d) $\frac{45}{600} = 0{,}075$, soit 7,5 %. Ce pourcentage est compris entre 5,8 % et 14,2 % : c'est une moyenne des deux fréquences, pondérée par 480 et 120. Il est plus proche de 5,8 % parce que le groupe avec casque est le plus nombreux. Conclusion : 7,5 % des cyclistes suivis ont été blessés.

**Ex. 46 ★★☆ Tri et compost : un filtre à corriger**

(a) La condition avec `or` est vraie pour les foyers 1, 3, 4, 5, 6 et 8 : la fonction renvoie 6. Elle compte les foyers qui trient OU compostent. Vérification : 5 foyers trient, 4 compostent, 3 font les deux, et $5 + 4 - 3 = 6$.

(b) Ligne 6 : `if tri[i] == "oui" and compost[i] == "oui":`. Seuls les foyers 3, 4 et 8 vérifient les deux conditions : la fonction renvoie 3.

(c) `if tri[i] == "oui" and not compost[i] == "oui":` (ou `compost[i] == "non"`). Ici, le `not` porte seulement sur le compost. Les foyers 1 et 6 sont comptés : la fonction renvoie 2. Vérification : 5 foyers trient, dont 3 compostent, et $5 - 3 = 2$.

**Ex. 46 (suite)**

(d) Parmi les 5 foyers qui trient : $\frac{3}{5} = 0{,}6$, soit 60 % compostent. Parmi les 4 foyers qui compostent : $\frac{3}{4} = 0{,}75$, soit 75 % trient. Conclusion : le numérateur est le même, mais pas le dénominateur, donc les deux fréquences sont différentes.

**Ex. 47 ★★★ Le club de natation et la compétition**

(a) Filles : $0{,}4 \times 250 = 100$, donc garçons : $250 - 100 = 150$. Filles en compétition : $0{,}3 \times 100 = 30$, filles en loisir : $100 - 30 = 70$. Garçons en compétition : $0{,}2 \times 150 = 30$, garçons en loisir : $150 - 30 = 120$. Total compétition : $30 + 30 = 60$. Total loisir : $70 + 120 = 190$. Contrôle : $60 + 190 = 250$.

(b) Faux. Les effectifs sont égaux (30 et 30), mais les deux groupes n'ont pas la même taille. Parmi les filles : $\frac{30}{100} = 0{,}3$, soit 30 %. Parmi les garçons : $\frac{30}{150} = 0{,}2$, soit 20 %. La compétition attire donc une part plus grande des filles que des garçons.

(c) $f(\text{compétition}) = \frac{60}{250} = 0{,}24$. Note $p$ la part de filles, avec $0 \leqslant p \leqslant 1$. La fréquence de la compétition vaut $p \times 0{,}3 + (1-p) \times 0{,}2 = 0{,}2 + 0{,}1p$. Ici, $0{,}4 \times 0{,}3 + 0{,}6 \times 0{,}2 = 0{,}12 + 0{,}12 = 0{,}24$. Comme $0 \leqslant 0{,}1p \leqslant 0{,}1$, cette fréquence est toujours comprise entre 0,2 et 0,3. Conclusion : dans ce club, 24 % des licenciés font de la compétition, une valeur bien comprise entre 20 % et 30 %.

(d) Le 30 % est la part des compétitrices parmi les filles, pas la part des filles parmi les compétiteurs. Parmi les compétiteurs : $\frac{30}{60} = 0{,}5$.
Donc 50 % des compétiteurs sont des filles. Fille et compétition : $\frac{30}{250} = 0{,}12$.
Donc 12 % des licenciés sont des filles qui font de la compétition.

(e) Avec $x$ nouvelles nageuses, la fréquence devient $\frac{30+x}{100+x}$. On veut $30 + x \geqslant 0{,}4(100 + x)$, soit $0{,}6x \geqslant 10$, donc $x \geqslant \frac{10}{0{,}6} \approx 16{,}7$. Contrôle : $\frac{46}{116} \approx 0{,}397$, c'est trop peu, et $\frac{47}{117} \approx 0{,}402$ convient. Le club doit recruter au moins 17 nageuses en compétition.

**Ex. 48 ★★★ Sport et sommeil : attention aux conclusions**

(a) On sait que $n(A \cup B) = n(A) + n(B) - n(A \cap B)$, donc $n(A \cap B) = 150 + 210 - 270 = 90$. Il y a 90 salariés qui font du sport chaque semaine et qui dorment au moins 7 h par nuit.

**Ex. 48** **(suite)**

(b) « Ni $A$ ni $B$ », c'est NON ($A$ OU $B$) : $380 - 270 = 110$ salariés. Fréquence : $\frac{110}{380} \approx 0{,}289$.
Donc environ 28,9 % des salariés ne font pas de sport chaque semaine et dorment moins de 7 h par nuit.

(c) Bons dormeurs sportifs : 90. Bons dormeurs non sportifs : $210 - 90 = 120$ (Karim a raison sur les effectifs). Non-sportifs : $380 - 150 = 230$. Parmi les sportifs : $\frac{90}{150} = 0{,}6$, soit 60 % de bons dormeurs. Parmi les non-sportifs : $\frac{120}{230} \approx 0{,}522$, soit 52,2 %. Karim compare les effectifs de deux groupes de tailles différentes (150 et 230). La part de bons dormeurs est plus grande chez les sportifs : ses chiffres ne montrent pas que le sport n'aide pas à dormir
(ils ne prouvent pas non plus que le sport en est la cause).

(d) Faux. Parmi les bons dormeurs, la part des sportifs vaut $\frac{90}{210} \approx 0{,}429$, soit 42,9 % : c'est moins de la moitié. L'article a inversé le « parmi » : c'est parmi les sportifs qu'une majorité (60 %) dort au moins 7 h. Phrase corrigée : « La plupart des salariés sportifs dorment au moins 7 h par nuit. »

**Ex. 49** ★★★ **Vélo ou bus selon la météo**

(a) Les jours de pluie sont les jours 1, 4 et 7, avec les modes B, B, V.
Donc `n` vaut 3, `c` vaut 1, et la fonction renvoie $\frac{1}{3} \approx 0{,}333$ (Python affiche 0.3333333333333333). C'est la fréquence des trajets à vélo parmi les jours de pluie : Chloé prend le vélo environ 33,3 % des jours de pluie.

(b) Il y a 7 jours secs, et Chloé prend le vélo 5 de ces jours. La fonction renvoie $\frac{5}{7} \approx 0{,}714$. Quand il fait sec, elle prend le vélo environ 71,4 % des jours, contre 33,3 % quand il pleut. Sur ces 10 jours, elle prend nettement plus souvent le vélo par temps sec.

(c) On échange les deux tests : `if mode[k] == "V":` sur la ligne 5 et `if meteo[k] == m:` sur la ligne 7. Chloé roule à vélo les jours 2, 3, 6, 7, 8 et 10, soit 6 jours, dont 1 jour de pluie. La fonction renvoie $\frac{1}{6} \approx 0{,}167$ : seulement 16,7 % des jours à vélo sont des jours de pluie. Le numérateur ne change pas, mais le dénominateur si.

**Ex. 49** **(suite)**

(d) La condition de Léo signifie « pas sec OU pas à vélo », c'est-à-dire NON (sec ET vélo). Il y a 5 jours secs à vélo, donc elle compte $10 - 5 = 5$ jours. La bonne condition est `not (meteo[k] == "S" or mode[k] == "V")`.
Vérification : $n(\text{sec} \cup \text{vélo}) = 7 + 6 - 5 = 8$, donc $10 - 8 = 2$. Il y a 2 jours ni secs ni à vélo : ce sont les jours 1 et 4, pluie et bus.