BURHAN

Statistique : exercices, entraînement

Tronc Commun · Sciences · Mathématiques

Version en ligne : burhan.ma/tc/sciences/mathematiques/statistique/exercices/entrainement

Chapitre	Palier	Exercices
Statistique	Entraînement	7

Exercice 1

Les âges d'un club de randonnée

Voici les âges, en années, des 30 adhérents d'un club de randonnée :

34 ; 12 ; 47 ; 20 ; 25 ; 58 ; 31 ; 15 ; 28 ; 40 ; 19 ; 36 ; 24 ; 52 ; 11 ; 29 ; 33 ; 44 ; 17 ; 26 ; 38 ; 55 ; 21 ; 18 ; 35 ; 50 ; 27 ; 39 ; 45 ; 23.

On les regroupe dans les classes $[10;20[$, $[20;30[$, $[30;40[$ et $[40;60[$.

1.

Dresser le tableau des effectifs par classe. Dans quelle classe range-t-on l'adhérent de 20 ans ? celui de 40 ans ?

INDICE

Outil : la définition des classes. Le crochet ouvert à droite exclut la borne de droite. **Contrôle :** le total doit faire 30.

CORRIGÉ

Bornes. D'après la définition des classes, le crochet ouvert à droite exclut la borne de droite : la classe $[10;20[$ contient les âges de 10 inclus à 20 exclu. L'adhérent de 20 ans est donc dans $[20;30[$, et celui de 40 ans dans $[40;60[$.

Tableau. On compte les âges de chaque classe :

Âge (ans)	$[10;20[$	$[20;30[$	$[30;40[$	$[40;60[$
Effectif	6	9	7	8

Contrôle. $6+9+7+8=30$, l'effectif total.

Conclusion. Les effectifs des classes sont 6, 9, 7 et 8. L'adhérent de 20 ans est rangé dans $[20;30[$, celui de 40 ans dans $[40;60[$.

2.

Calculer l'amplitude de chaque classe. En prenant 10 ans pour amplitude de référence, calculer la hauteur de chaque rectangle de l'histogramme.

INDICE

Outil : l'aire de chaque rectangle est proportionnelle à l'effectif. **Attention :** la dernière classe contient deux amplitudes de référence.

CORRIGÉ

Amplitudes. Les trois premières classes ont pour amplitude 10 ans. La classe $[40;60[$ a pour amplitude $60-40=20$ ans.

Hauteurs. L'aire de chaque rectangle est proportionnelle à l'effectif. Pour les trois premières classes, la base est l'amplitude de référence : la hauteur est l'effectif, soit 6, 9 et 7. La dernière classe vaut deux amplitudes de référence, donc sa hauteur est :

$$\frac{8}{2}=4.$$

Conclusion. Les amplitudes sont 10, 10, 10 et 20 ans ; les hauteurs des rectangles sont 6, 9, 7 et 4.

3.

Décrire l'histogramme et préciser ce que représente un rectangle de base 10 ans et de hauteur 1.

INDICE

Début : chaque rectangle est posé sur sa classe, sans espace entre deux classes voisines.

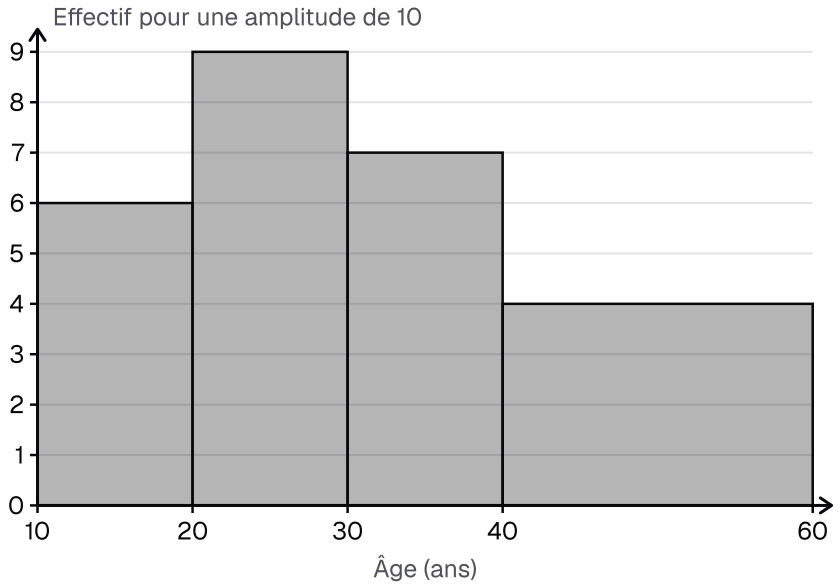
CORRIGÉ

Tracé. Sur l'axe des âges, on trace quatre rectangles accolés : de 10 à 20 de hauteur 6, de 20 à 30 de hauteur 9, de 30 à 40 de hauteur 7, de 40 à 60 de hauteur 4.

Unité d'aire. Un rectangle de base 10 ans, l'amplitude de référence, et de hauteur 1 représente un adhérent.

Contrôle sur le dernier rectangle. Son aire vaut $20 \times 4 = 80$; il représente donc $\dfrac{80}{10} = 8$ adhérents, l'effectif de sa classe.

Conclusion. L'histogramme compte quatre rectangles accolés, de hauteurs 6, 9, 7 et 4 ; un rectangle de base 10 ans et de hauteur 1 représente un adhérent.



L'histogramme des âges des adhérents : le rectangle de 40 à 60 ans est le plus large, et sa hauteur n'est que 4.

4.

Quelle est la classe modale ? La classe $[40;60[$ a pourtant un effectif plus grand que $[30;40[$: expliquer.

INDICE

Outil : la définition de la classe modale, qui regarde la hauteur des rectangles.

CORRIGÉ

Classe modale. D'après la définition de la classe modale, c'est $[20;30[$, dont le rectangle est le plus haut (9).

Explication. La classe $[40;60[$ compte 8 adhérents, mais répartis sur 20 ans : pour 10 ans, cela ne fait que 4 adhérents. Les adhérents y sont moins concentrés que dans $[30;40[$, qui en compte 7 sur 10 ans.

Conclusion. La classe modale est $[20;30[$. Un effectif plus grand, réparti sur une classe plus large, peut donner un rectangle plus bas.

Exercice 2

La consommation d'eau de cinquante foyers

Une régie de distribution d'eau relève la consommation journalière de 50 foyers, en litres :

Consommation (L)	$[100;150[$	$[150;200[$	$[200;250[$	$[250;300[$	$[300;400[$
Nombre de foyers	6	14	17	9	4

1.

Calculer les effectifs cumulés croissants, puis les fréquences cumulées croissantes.

INDICE

Outil : la définition des cumuls. Chaque fréquence cumulée est l'effectif cumulé divisé par 50.

CORRIGÉ

Effectifs cumulés croissants. D'après la définition des cumuls, on ajoute chaque effectif au cumul précédent :

- $[100;150[: 6$;
- $[150;200[: 6+14=20$;
- $[200;250[: 20+17=37$;
- $[250;300[: 37+9=46$;
- $[300;400[: 46+4=50$.

Fréquences cumulées croissantes. On divise chaque cumul par $N=50$: $0{,}12$; $0{,}4$; $0{,}74$; $0{,}92$ et 1.

Conclusion. Les effectifs cumulés croissants sont 6, 20, 37, 46 et 50 ; les fréquences cumulées croissantes valent $0{,}12$, $0{,}4$, $0{,}74$, $0{,}92$ et 1.

2.

Donner les points du polygone des effectifs cumulés croissants.

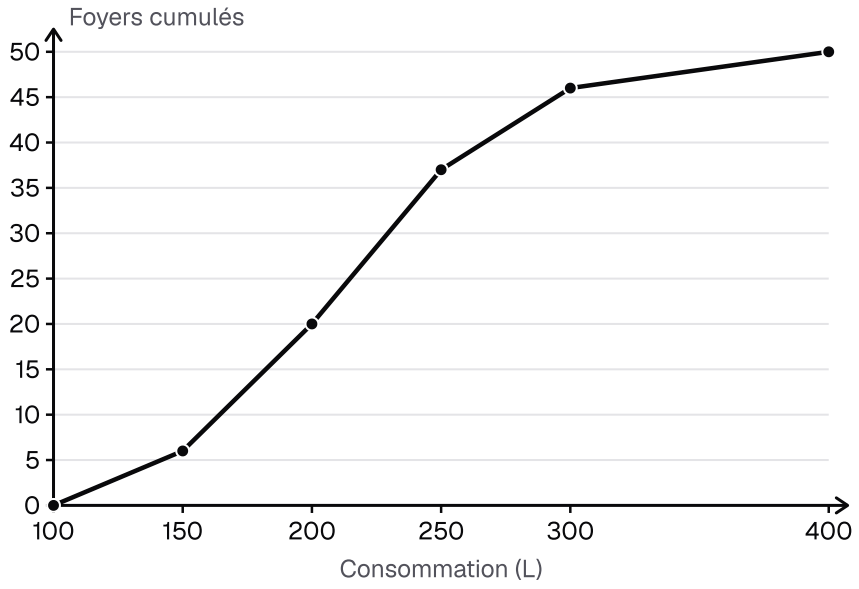
INDICE

Outil : la définition du polygone : chaque cumul est associé à la borne de droite de sa classe.

CORRIGÉ

D'après la définition du polygone des effectifs cumulés croissants, on part de la borne gauche de la première classe avec l'effectif 0, puis on associe à la borne droite de chaque classe son effectif cumulé.

Conclusion. On joint dans l'ordre les points $(100, 0)$, $(150, 6)$, $(200, 20)$, $(250, 37)$, $(300, 46)$ et $(400, 50)$.



Le polygone des foyers cumulés, tracé à partir du tableau de la régie : six sommets, de 100 à 400 litres.

3.

Estimer la consommation médiane par interpolation, arrondie au dixième de litre.

INDICE

Début : $\dfrac{N}{2} = 25$. Repère la classe $[a;b[$ dont les cumuls aux bornes vérifient $N_a < 25 \leq N_b$. **Ensuite :** calcule la part $\dfrac{25 - N_a}{N_b - N_a}$ de la classe à parcourir.

CORRIGÉ

Classe de la médiane. On a $\dfrac{N}{2} = 25$. Sur la classe $[200;250[$, les cumuls aux bornes vérifient $20 < 25 \leq 37$: la médiane est dans cette classe.

Part à parcourir. Elle vaut $\dfrac{25 - 20}{37 - 20} = \dfrac{5}{17}$ de la classe, dont l'amplitude est 50 L.

Interpolation. Sur ce segment du polygone :

$$\begin{aligned} M_e &= 200 + 50 \times \frac{5}{17} \\ &= 200 + \frac{250}{17} \\ &\approx 214{,}7 \text{ L}, \end{aligned}$$

arrondi au dixième. Contrôle : $200 < 214{,}7 \leq 250$.

Conclusion. La consommation médiane vaut environ $214{,}7$ litres par jour ; c'est une valeur approchée, obtenue par interpolation.

4.

Quelle proportion des foyers consomme au moins 250 litres par jour ?

INDICE

Outil : la fréquence cumulée croissante de la classe $[200;250[$ compte les foyers sous 250 L ; prends le complément à 1.

CORRIGÉ

Moins de 250 L. La fréquence cumulée croissante de la classe $[200;250[$ vaut $0{,}74$: c'est la proportion des foyers qui consomment moins de 250 L.

Au moins 250 L. On prend le complément à 1 : $1 - 0{,}74 = 0{,}26$.

Contrôle. Les deux dernières classes comptent $9 + 4 = 13$ foyers, et $\dfrac{13}{50} = 0{,}26$.

Conclusion. Une proportion $0{,}26$ des foyers, soit 13 foyers sur 50, consomme au moins 250 litres par jour.

Exercice 3

La durée de vie d'ampoules

Un fabricant teste 40 ampoules et relève leur durée de vie, en heures :

Durée x_i (h)	980	990	1000	1010	1020
Effectif n_i	3	7	12	10	8

On pose $y_i = \dfrac{x_i - 1000}{10}$.

1.

Calculer les valeurs y_i.

INDICE

Début : pour $x_1 = 980$, $y_1 = \dfrac{980 - 1000}{10}$.

CORRIGÉ

On applique $y_i = \dfrac{x_i - 1000}{10}$ à chaque valeur. Par exemple :

$$y_1 = \frac{980 - 1000}{10} = -2.$$

Conclusion. On obtient $y_1 = -2$, $y_2 = -1$, $y_3 = 0$, $y_4 = 1$ et $y_5 = 2$.

2.

Calculer la moyenne $\overline{y}$ de la série des y_i, avec les mêmes effectifs.

INDICE

Outil : la définition de la moyenne, avec $N = 40$.

CORRIGÉ

Somme des produits. Chaque y_i garde l'effectif de x_i :

$$\begin{aligned} &3 \times (-2) + 7 \times (-1) \\ &\quad + 12 \times 0 + 10 \times 1 \\ &\quad + 8 \times 2 \\ &= -6 - 7 + 0 + 10 + 16 \\ &= 13. \end{aligned}$$

Moyenne. D'après la définition de la moyenne, avec $N = 40$:

$$\overline{y} = \frac{13}{40} = 0{,}325.$$

Conclusion. La moyenne de la série des y_i vaut $0{,}325$.

3.

Exprimer x_i en fonction de y_i. En déduire la durée de vie moyenne $\overline{x}$.

INDICE

Outil : la propriété du changement affine, appliquée à la série des y_i.

CORRIGÉ

Expression de x_i. De $y_i = \dfrac{x_i - 1000}{10}$, on tire $10y_i = x_i - 1000$, donc $x_i = 10y_i + 1000$.

Moyenne. D'après la propriété du changement affine, avec $a = 10$ et $b = 1000$, on a $\overline{x} = 10\,\overline{y} + 1000$, soit $\overline{x} = 10 \times 0{,}325 + 1000 = 1003{,}25$ h.

Conclusion. Pour chaque valeur, $x_i = 10y_i + 1000$, et la durée de vie moyenne des ampoules vaut $1003{,}25$ heures.

4.

Contrôler ce résultat par la définition de la moyenne.

INDICE

Outil : calcule la somme des produits $n_i x_i$, puis divise par 40.

CORRIGÉ

Somme des produits. Les produits $n_i x_i$ valent $3 \times 980 = 2940$, $7 \times 990 = 6930$, $12 \times 1000 = 12000$, $10 \times 1010 = 10100$ et $8 \times 1020 = 8160$. Leur somme vaut :

$$\begin{aligned} &2940 + 6930 + 12000 \\ &\quad + 10100 + 8160 \\ &= 40130. \end{aligned}$$

Moyenne. D'après la définition de la moyenne, $\overline{x} = \dfrac{40130}{40} = 1003{,}25$ h.

Conclusion. On retrouve la même durée de vie moyenne, $1003{,}25$ heures ; le changement de variable y conduisait avec des calculs plus légers.

Exercice 4

Des temps de réaction

En biologie, on mesure le temps de réaction de 20 élèves à un signal lumineux, en centièmes de seconde :

Temps x_i	18	20	22	24	26
Effectif n_i	2	5	6	4	3

1.

Compléter le tableau par les lignes $n_i x_i$ et $n_i x_i^2$, puis calculer le temps moyen.

INDICE

Début : $2 \times 18 = 36$ et $2 \times 18^2 = 648$.

CORRIGÉ

Tableau. On complète colonne par colonne ; la dernière colonne donne les totaux :

Temps x_i	18	20	22	24	26	**Total**
$n_i x_i$	36	100	132	96	78	442
$n_i x_i^2$	648	2000	2904	2304	2028	9884

Moyenne. D'après la définition de la moyenne, $\overline{x} = \dfrac{442}{20} = 22{,}1$ centièmes de seconde.

Conclusion. Le temps de réaction moyen vaut $22{,}1$ centièmes de seconde.

2.

Calculer la variance par la formule, puis l'écart type arrondi au centième.

INDICE

Outil : la formule de la variance : moyenne des carrés moins carré de la moyenne.

CORRIGÉ

Variance. Par la formule de la variance, moyenne des carrés moins carré de la moyenne :

$$\begin{aligned} V &= \frac{9884}{20} - 22{,}1^2 \\ &= 494{,}2 - 488{,}41 \\ &= 5{,}79. \end{aligned}$$

Écart type. $\sigma = \sqrt{5{,}79} \approx 2{,}41$ centièmes de seconde, arrondi au centième.

Conclusion. La variance vaut $5{,}79$ et l'écart type environ $2{,}41$ centièmes de seconde.

3.

Combien d'élèves ont un temps compris entre $\overline{x} - \sigma$ et $\overline{x} + \sigma$? Quelle fraction de la classe cela représente-t-il ?

INDICE

Début : calcule les bornes $22{,}1 - 2{,}41$ et $22{,}1 + 2{,}41$, puis repère les valeurs du tableau situées entre elles.

CORRIGÉ

Bornes. On calcule $22{,}1 - 2{,}41$ et $22{,}1 + 2{,}41$: les bornes valent environ $19{,}69$ et $24{,}51$.

Valeurs comprises. Les temps 20, 22 et 24 sont entre ces bornes ; 18 et 26 sont en dehors. Cela fait $5 + 6 + 4 = 15$ élèves.

Conclusion. 15 élèves ont un temps compris entre $\overline{x} - \sigma$ et $\overline{x} + \sigma$; cela représente $\dfrac{15}{20} = 0{,}75$ de la classe, soit trois élèves sur quatre.

Exercice 5

Le tarif des colis

Un transporteur pèse cinq colis : 2 kg, 4 kg, 5 kg, 7 kg et 7 kg. Il facture chaque colis $y = 2x + 3$ dirhams, où x est la masse en kilogrammes.

On note $\overline{x}$, V_x et σ_x la moyenne, la variance et l'écart type des masses ; $\overline{y}$, V_y et σ_y ceux des tarifs.

1.

Calculer la masse moyenne $\overline{x}$, puis la variance V_x des masses par la définition.

INDICE

Outil : la définition de la moyenne, puis celle de la variance.

CORRIGÉ

Moyenne. D'après la définition de la moyenne, $\overline{x} = \dfrac{2 + 4 + 5 + 7 + 7}{5} = \dfrac{25}{5} = 5$ kg.

Variance. Les écarts à la moyenne sont -3, -1, 0, 2 et 2 ; leurs carrés valent 9, 1, 0, 4 et 4. D'après la définition de la variance, $V_x = \dfrac{9 + 1 + 0 + 4 + 4}{5} = \dfrac{18}{5} = 3{,}6$.

Conclusion. La masse moyenne vaut $\overline{x} = 5$ kg et la variance des masses $V_x = 3{,}6$.

2.

Calculer les cinq tarifs y_i, puis le tarif moyen $\overline{y}$ de deux façons.

INDICE

Outil : la définition de la moyenne, puis la propriété du changement affine avec $a = 2$ et $b = 3$.

CORRIGÉ

Tarifs. Pour $x = 2$, le tarif vaut $2 \times 2 + 3 = 7$ dirhams. De même, les cinq tarifs sont 7, 11, 13, 17 et 17 dirhams.

Par la définition. La somme des tarifs vaut 65, donc $\overline{y} = \dfrac{65}{5} = 13$ dirhams.

Par le changement affine. Avec $a = 2$ et $b = 3$, $\overline{y} = 2\,\overline{x} + 3$, soit $\overline{y} = 2 \times 5 + 3 = 13$ dirhams.

Conclusion. Les deux calculs concordent : le tarif moyen vaut 13 dirhams.

3.

Comparer chaque écart $y_i - \overline{y}$ à l'écart $x_i - \overline{x}$. En déduire V_y à partir de V_x, puis σ_y à partir de σ_x.

INDICE

Début : $y_i - \overline{y} = (2x_i + 3) - (2\,\overline{x} + 3)$. **Ensuite :** élève au carré, puis applique la définition de la variance.

CORRIGÉ

Écarts. Pour chaque colis :

$$\begin{aligned} y_i - \overline{y} &= (2x_i + 3) \\ &\quad - (2\,\overline{x} + 3) \\ &= 2\,(x_i - \overline{x}). \end{aligned}$$

Chaque écart est doublé : les écarts des tarifs valent -6, -2, 0, 4 et 4.

Variance. Chaque carré d'écart est multiplié par $2^2 = 4$. D'après la définition de la variance, leur moyenne l'est aussi :

$$V_y = 4V_x = 4 \times 3{,}6 = 14{,}4.$$

Contrôle direct : $\dfrac{36 + 4 + 0 + 16 + 16}{5} = \dfrac{72}{5} = 14{,}4$.

Écart type. $\sigma_y = \sqrt{4V_x} = 2\,\sigma_x$, soit $\sigma_y = \sqrt{14{,}4} \approx 3{,}79$ dirhams, arrondi au centième.

Conclusion. Chaque écart $y_i - \overline{y}$ est le double de l'écart $x_i - \overline{x}$; donc $V_y = 4V_x$ et $\sigma_y = 2\,\sigma_x$. Ajouter 3 décale tous les tarifs sans changer leur dispersion ; multiplier par 2 double l'écart type.

Exercice 6

Un mois sans relevé

En géographie, on étudie les précipitations d'une station de montagne sur un semestre, en millimètres. Le relevé de mars a été perdu :

Mois	janvier	février	mars	avril	mai	juin
Précipitations (mm)	42	35	x	18	9	4

On sait que la moyenne mensuelle du semestre vaut 22 mm.

1.

Écrire une équation vérifiée par x, puis la résoudre.

INDICE

Outil : la définition de la moyenne : la somme des six valeurs vaut 6×22.

CORRIGÉ

D'après la définition de la moyenne, la somme des six valeurs, divisée par 6, vaut 22 :

$$\frac{42 + 35 + x + 18 + 9 + 4}{6} = 22.$$

En multipliant les deux membres par 6 et en regroupant les nombres connus, on obtient $108 + x = 132$. Ainsi $x = 132 - 108 = 24$.

Conclusion. L'équation $108 + x = 132$ a pour solution 24 : il est tombé 24 mm en mars.

2.

Déterminer la médiane et l'étendue des précipitations du semestre.

INDICE

Méthode : range d'abord les six valeurs ; $N = 6$ est pair.

CORRIGÉ

Série rangée. 4 ; 9 ; 18 ; 24 ; 35 ; 42.

Médiane. $N = 6$ est pair : d'après la définition de la médiane, M_e est la moyenne des valeurs de rangs 3 et 4, soit $M_e = \dfrac{18 + 24}{2} = 21$ mm.

Étendue. D'après la définition de l'étendue, elle vaut $42 - 4 = 38$ mm.

Conclusion. Sur le semestre, la médiane des précipitations vaut 21 mm et l'étendue 38 mm.

3.

Juillet, sans aucune pluie, s'ajoute à la série avec 0 mm. Calculer la nouvelle moyenne, arrondie au dixième, et la nouvelle médiane.

INDICE

Début : la somme des valeurs ne change pas, mais N devient 7, impair.

CORRIGÉ

Moyenne. Juillet ajoute 0 mm : la somme reste 132 mm, mais l'effectif devient $N = 7$. D'après la définition de la moyenne, $\overline{x} = \dfrac{132}{7} \approx 18{,}9$ mm, arrondi au dixième.

Médiane. La série rangée devient 0 ; 4 ; 9 ; 18 ; 24 ; 35 ; 42. Comme $N = 7$ est impair, la médiane est la valeur de rang $\dfrac{7+1}{2} = 4$: $M_e = 18$ mm.

Conclusion. Avec juillet, la moyenne passe à environ $18{,}9$ mm et la médiane à 18 mm ; l'étendue passe à 42 mm.

Exercice 7

Deux variétés de blé

Un agronome compare deux variétés de blé, chacune semée sur six parcelles. Il relève les rendements, en quintaux par hectare :

- variété A : 30 ; 32 ; 35 ; 33 ; 31 ; 37 ;
- variété B : 25 ; 40 ; 28 ; 38 ; 36 ; 31.

1.

Calculer la moyenne et la médiane de chaque série.

INDICE

Outil : la définition de la moyenne, puis celle de la médiane ; chaque série compte 6 valeurs.

CORRIGÉ

Moyennes. Chaque série compte 6 valeurs, de somme 198. D'après la définition de la moyenne, $\overline{x}_A = \dfrac{198}{6} = 33$ et $\overline{x}_B = \dfrac{198}{6} = 33$ quintaux par hectare.

Médianes. $N = 6$ est pair : la médiane est la moyenne des valeurs de rangs 3 et 4 de la série rangée.

- Série A rangée : 30 ; 31 ; 32 ; 33 ; 35 ; 37, donc $M_e = \dfrac{32 + 33}{2} = 32{,}5$.
- Série B rangée : 25 ; 28 ; 31 ; 36 ; 38 ; 40, donc $M_e = \dfrac{31 + 36}{2} = 33{,}5$.

Conclusion. Les deux variétés ont la même moyenne, 33 quintaux par hectare ; la médiane vaut $32{,}5$ pour A et $33{,}5$ pour B.

2.

Calculer l'étendue et l'écart moyen de chaque série.

INDICE

Outil : les définitions de l'étendue et de l'écart moyen, avec les moyennes de la question 1.

CORRIGÉ

Étendues. D'après la définition de l'étendue : $37 - 30 = 7$ pour A et $40 - 25 = 15$ pour B.

Écarts moyens. D'après la définition de l'écart moyen, on fait la moyenne des distances à la moyenne 33 :

- pour A, les distances sont 3, 1, 2, 0, 2 et 4, de somme 12, donc $e_A = \dfrac{12}{6} = 2$;
- pour B, ce sont 8, 7, 5, 5, 3 et 2, de somme 30, donc $e_B = \dfrac{30}{6} = 5$.

Conclusion. La variété A a pour étendue 7 et pour écart moyen 2 ; la variété B, 15 et 5 (en quintaux par hectare).

3.

Quelle variété conseiller à un agriculteur qui veut un rendement prévisible d'une parcelle à l'autre ?

INDICE

Méthode : compare la position, puis la dispersion, puis conclus.

CORRIGÉ

Position. Les moyennes sont égales et les médianes proches : en moyenne, les deux variétés rendent autant.

Dispersion. L'étendue et l'écart moyen de A sont plus petits que ceux de B.

Conclusion. La variété A donne des rendements plus réguliers d'une parcelle à l'autre, pour un même rendement moyen : c'est elle qu'il faut conseiller à cet agriculteur.