Plaquette 5 – Problèmes de synthèse : sport, santé et environnement

Statistiques descriptives — Seconde

Corrections détaillées

Boîte à outils

Tous les indicateurs en une page.

$$\overline{x} = \frac{\sum n_i x_i}{N} \qquad V = \overline{x^2} - (\overline{x})^2 \qquad \sigma = \sqrt{V}$$

Question posée	Indicateur à choisir
« valeur typique », série symétrique	moyenne
« valeur typique », valeurs extrêmes présentes	médiane
« régularité d'un procédé »	écart-type
« écart entre les gros et les petits »	quartiles, écart interquartile
« comparer deux unités différentes »	coefficient de variation

Choisir moyenne ou médiane.

— $\overline{x} \approx$ Méd : série à peu près symétrique, les deux se valent.
— $\overline{x} >$ Méd : queue vers les grandes valeurs — la médiane décrit mieux le cas courant.
— $\overline{x} <$ Méd : queue vers les petites valeurs.

Deux réflexes de rédaction.

— Toujours donner l'**unité** et préciser l'**arrondi**.
— Conclure par une phrase de contexte, pas par un nombre nu.

Moyennes pondérées et moyennes de sous-groupes.

$$\overline{x} = \frac{n_1\overline{x_1} + n_2\overline{x_2} + \cdots}{n_1 + n_2 + \cdots}$$

Ce n'est jamais la moyenne des moyennes, sauf si tous les effectifs sont égaux.

Contrôles systématiques.

— $\min \le Q_1 \le$ Méd $\le Q_3 \le \max$;
— $\overline{x}$ est comprise entre le minimum et le maximum ;
— la somme des fréquences vaut 1 ;
— $V \ge 0$.

Correction de l'exercice 1 – Temps de course d'un club

On compare moyenne et médiane. L'effectif, 11, est impair : la médiane est la 6^{e} valeur, 45 min. La somme des temps vaut 516, d'où

$$\overline{x} = \frac{516}{11} \approx 46{,}9 \text{ min}$$

La moyenne dépasse la médiane à cause du temps de 71 min, très isolé. Avec $Q_1 = 41$ (rang 3) et $Q_3 = 50$ (rang 9), l'écart interquartile vaut 9 et le seuil haut de la règle des 1,5 écart interquartile est $50 + 13{,}5 = 63{,}5$: le temps de 71 min est aberrant.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

L'indicateur qui décrit le mieux le club est donc la **médiane** : « la moitié du club court le 10 km en moins de 45 minutes ». Sans le dernier coureur, la moyenne chuterait de 2,4 min, la médiane de 1 min seulement.

Réponse. La médiane, 45 min : la moyenne (≈ 46,91) est déformée par le temps aberrant de 71 min.

Correction de l'exercice 2 – Deux sauteurs en longueur

Les deux sauteurs totalisent 37,00 m sur six essais : même moyenne, $\frac{37}{6} \approx 6{,}17$ m. On compare donc leur **dispersion**.

$$\text{étendue}_A = 6{,}25 - 6{,}10 = 0{,}15 \text{ m} \qquad \text{étendue}_B = 6{,}80 - 5{,}50 = 1{,}30 \text{ m}$$

Les écarts-types le confirment : environ 0,055 m pour A contre 0,435 m pour B, près de huit fois plus. A est très régulier, B irrégulier.

Dans une finale au **meilleur essai**, seul le plus grand saut compte : on choisit B, capable de 6,80 m quand A ne dépasse pas 6,25 m. Dans un concours à la **moyenne**, la régularité de A garantit le résultat : on choisit A. À moyenne égale, la dispersion décide selon la règle du concours.

Réponse. Même moyenne (≈ 6,17 m) : B pour une finale au meilleur essai, A pour un concours à la moyenne.

Correction de l'exercice 3 – Fréquences cardiaques

La somme des 20 valeurs vaut 1 420, donc $\overline{x} = 71$ bpm. La médiane est la demi-somme des 10^{e} et 11^{e} valeurs, 70 bpm ; Q_1 est la 5^{e} valeur, 64, et Q_3 la 15^{e}, 75. Résumé : 58 ; 64 ; 70 ; 75 ; 95.

L'écart-type, calculé à la calculatrice, vaut environ 8,64 bpm. L'intervalle $[\overline{x} - \sigma\,;\overline{x} + \sigma] \approx [62{,}4\,;79{,}6]$ contient 14 élèves sur 20, soit 70 %.

$$Q_3 + 1{,}5\,(Q_3 - Q_1) = 75 + 16{,}5 = 91{,}5$$

La valeur 95 dépasse ce seuil : elle est aberrante et mérite une nouvelle mesure. La classe a une fréquence cardiaque typique de 70 bpm, avec un cas isolé à vérifier.

Réponse. $\overline{x} = 71$ bpm, résumé 58 ; 64 ; 70 ; 75 ; 95, $\sigma \approx 8{,}64$; la valeur 95 est aberrante.

Correction de l'exercice 4 – Températures de deux villes

On calcule les moyennes : la somme des températures vaut 135 pour A et 165 pour B, donc

$$\overline{x_A} = \frac{135}{12} = 11{,}25 \text{ °C} \qquad \overline{x_B} = \frac{165}{12} = 13{,}75 \text{ °C}$$

L'amplitude annuelle (écart entre le mois le plus chaud et le plus froid) vaut $20 - 3 = 17$ °C pour A et $18 - 10 = 8$ °C pour B. Les écarts-types, environ 6,06 °C et 2,92 °C, confirment que A est deux fois plus dispersée.

B est donc à la fois plus douce en moyenne et beaucoup plus régulière : plus chaude en hiver (10 °C contre 3 °C), à peine moins chaude en été (18 °C contre 20 °C). C'est le profil d'un climat océanique, face au climat plus continental de A.

Réponse. A : 11,25 °C et amplitude 17 °C · B : 13,75 °C et amplitude 8 °C — B est plus douce et bien plus régulière.

Correction de l'exercice 5 – Consommation d'eau d'un immeuble

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

L'effectif total vaut 25. Avec les centres 25 ; 75 ; 125 ; 175 ; 225 :

$$\overline{x} = \frac{100 + 675 + 875 + 525 + 450}{25} = \frac{2\,625}{25} = 105 \text{ m}^3$$

Les effectifs cumulés valent 4 ; 13 ; 20 ; 23 ; 25. Le rang médian, 12,5, tombe dans $[50\,;100[$. En supposant les 9 logements de cette classe répartis régulièrement, on avance de $12{,}5 - 4 = 8{,}5$ logements sur 9 : Méd $\approx 50 + 50 \times \frac{8{,}5}{9} \approx 97{,}2$ m³.

La moyenne dépasse la médiane : quelques gros consommateurs la tirent vers le haut. À 3,80 € le m³, l'immeuble paie $2\,625 \times 3{,}80 = 9\,975$ €, soit 399 € par logement en moyenne.

Réponse. $\overline{x} = 105$ m³ et Méd $\approx 97{,}22$ m³ (facture totale 9975 € à 3,80 €/m³).

Correction de l'exercice 6 – Deux groupes de TP

On revient aux sommes de notes : $14 \times 12{,}5 = 175$ pour le premier groupe, $18 \times 10{,}5 = 189$ pour le second.

$$\overline{x} = \frac{175 + 189}{32} = \frac{364}{32} = 11{,}375$$

La moyenne est plus proche de 10,5, car le second groupe est plus nombreux ; la moyenne des moyennes, 11,5, serait fausse.

Pour une moyenne générale de 12, il faut une somme de $32 \times 12 = 384$, soit 20 points de plus. Répartis sur les 18 élèves du second groupe, cela fait $\frac{20}{18} \approx 1{,}11$ point chacun : sa moyenne devrait passer à environ 11,61. **Vérification** : $\frac{175 + 209}{32} = 12$.

Réponse. $\overline{x} = \frac{364}{32} = 11{,}375$; il faudrait environ +1,11 point au second groupe.

Correction de l'exercice 7 – Notes et coefficients au bac

La moyenne d'un candidat est une moyenne pondérée par les coefficients, dont la somme vaut $16 + 10 + 6 + 8 = 40$:

$$\overline{x} = \frac{16 \times 13 + 10 \times 11 + 6 \times 9 + 8 \times 15}{40} = \frac{208 + 110 + 54 + 120}{40} = \frac{492}{40} = 12{,}3$$

Pour obtenir 13 de moyenne, il faut une somme pondérée de $40 \times 13 = 520$, soit 28 points de plus. Chaque point en maths en rapporte 16, donc il faut gagner $\frac{28}{16} = 1{,}75$ point : la note de maths doit être 14,75.

Vérification : $\frac{236 + 110 + 54 + 120}{40} = \frac{520}{40} = 13$. En histoire (coefficient 6), il faudrait gagner 4,67 points : on progresse plus efficacement dans une matière à fort coefficient.

Réponse. $\overline{x} = 12{,}3$; il lui faudrait 14,75 en maths.

Correction de l'exercice 8 – Pollution de l'air

La somme des 15 relevés vaut 390, donc $\overline{x} = 26$ µg/m³. L'effectif est impair : la médiane est la 8e valeur, 22 µg/m³. La moyenne dépasse la médiane, la série est étirée vers les fortes valeurs.

$$\frac{26 - 20}{20} = 30\,\%$$

La moyenne dépasse la recommandation de 30 %, et 8 jours sur 15 sont au-dessus de 20. Un seul jour (62) dépasse le seuil d'alerte.

Avec $Q_1 = 16$ et $Q_3 = 32$, le seuil haut de la règle des 1,5 écart interquartile vaut

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

$32 + 24 = 56$: la valeur 62 est statistiquement aberrante, mais c'est une vraie mesure, qu'on ne doit pas écarter. Même sans elle, la moyenne resterait à $\frac{328}{14} \approx 23{,}4$ µg/m³ : le dépassement de la recommandation ne tient pas au seul pic.

Réponse. $\overline{x} = 26$ et Méd = 22 µg/m³ ; un jour au-dessus du seuil, et le dépassement de la recommandation subsiste sans lui ($\approx 23{,}43$).

Correction de l'exercice 9 – Durée de vie d'ampoules

On compare deux critères. Pour la **durée**, F_2 l'emporte : 1 500 h en moyenne contre 1 200 h, soit 25 % de plus. Pour la **régularité**, F_1 l'emporte : son écart-type est quatre fois plus petit, et rapporté à la moyenne

$$\frac{100}{1\,200} \approx 8{,}3\,\% \qquad \text{contre} \qquad \frac{400}{1\,500} \approx 26{,}7\,\%$$

L'intervalle $[\overline{x} - 2\sigma\,;\overline{x} + 2\sigma]$, qui contient en général la très grande majorité des durées, vaut [1 000 ; 1 400] h pour F_1 et [700 ; 2 300] h pour F_2.

L'hôpital, qui planifie sa maintenance sur la durée la plus courte probable, choisira F_1 (1 000 h contre 700 h). Le particulier, qui veut la plus longue durée en moyenne, choisira F_2. Le bon choix dépend du critère de décision.

Réponse. F_1 pour l'hôpital ($CV \approx 8{,}3\,\%$, plancher 1000 h) et F_2 pour le particulier (+25 % de durée moyenne).

Correction de l'exercice 10 – Retards de train

L'effectif total vaut $12 + 6 + 5 + 4 + 2 + 1 = 30$. La somme des retards est $0 + 12 + 25 + 40 + 60 + 90 = 227$ min, d'où

$$\overline{x} = \frac{227}{30} \approx 7{,}6 \text{ min}$$

Les effectifs cumulés valent 12 ; 18 ; 23 ; 27 ; 29 ; 30. La médiane est la demi-somme des 15^{e} et 16^{e} valeurs, qui valent toutes deux 2 : Méd = 2 min. Les trains à l'heure représentent $\frac{12}{30} = 40\,\%$, et $\frac{27}{30} = 90\,\%$ ont au plus 10 min de retard.

La moyenne est presque quatre fois la médiane : le seul train de 90 min pèse à lui seul 40 % de la somme des retards. La compagnie communiquera sur les 90 %, les usagers retiendront ce train ; la répartition complète est l'information la plus honnête.

Réponse. $\overline{x} \approx 7{,}57$ min, Méd = 2 min, 40 % de trains à l'heure et 90 % sous 10 min.

Correction de l'exercice 11 – Poids de paquets de café

La somme des 10 masses vaut 2 500 g, donc $\overline{x} = 250$ g : la première condition (moyenne au moins égale à 250 g) est respectée, de justesse.

Les écarts à 250 g sont -2 ; -1 ; 0 ; 1 ; 2 ; -2 ; 0 ; 3 ; -3 ; 2. Le plus grand, en valeur absolue, vaut 3 g : tous les paquets sont dans la tolérance de ±5 g. Le lot est **conforme**.

$$\sigma = \sqrt{\frac{4 + 1 + 0 + 1 + 4 + 4 + 0 + 9 + 9 + 4}{10}} = \sqrt{3{,}6} \approx 1{,}90 \text{ g}$$

La tolérance représente plus de 2,5 écarts-types : le procédé est régulier. Mais la moyenne est exactement à la limite ; un léger dérèglement la ferait passer sous 250 g, d'où l'habitude de viser un peu plus, par exemple 251 g.

Réponse. Oui : $\overline{x} = 250$ g et l'écart maximal est de 3 g (< 5), avec $\sigma \approx 1{,}90$ g.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l’exercice 12 – Évolution d’une moyenne de classe

La somme des trois premières notes vaut $3 \times 11 = 33$. Pour une moyenne de 12 sur quatre devoirs, il faut une somme de $4 \times 12 = 48$:

$$\text{note nécessaire} = 48 - 33 = 15$$

Pour 13 de moyenne, il faut $4 \times 13 = 52$, donc une note de $52 - 33 = 19$. **Vérification** : $\dfrac{33+15}{4} = 12$ et $\dfrac{33+19}{4} = 13$.

Gagner un point de moyenne sur quatre devoirs demande 4 points de plus au dernier. Viser 14 exigerait $56 - 33 = 23$, impossible sur 20 : la meilleure moyenne atteignable est $\dfrac{33+20}{4} = 13{,}25$.

Réponse. 15 pour atteindre 12, 19 pour atteindre 13 ; 14 est hors d’atteinte (maximum 13,25).

Correction de l’exercice 13 – Rendements agricoles

On calcule d’abord les moyennes : $\dfrac{30{,}5}{5} = 6{,}1$ t/ha pour P_1 et $\dfrac{31{,}5}{5} = 6{,}3$ t/ha pour P_2. En moyenne, P_2 est légèrement meilleure.

Mais le prêt impose au moins 5,5 t/ha **chaque année** : c’est le rendement le plus faible qui compte.

$$\min(P_1) = 5{,}8 > 5{,}5 \qquad \min(P_2) = 4{,}0 < 5{,}5$$

P_1 respecte la condition les cinq années ; P_2 a échoué une année sur cinq. Les écarts-types (environ 0,2 t/ha contre 1,4 t/ha) traduisent cette différence de régularité. On conseille P_1 : on accepte environ 3 % de rendement moyen en moins pour supprimer le risque de défaut.

Réponse. P_1 : moyenne un peu plus faible (6,1 contre 6,3) mais minimum $5{,}8 > 5{,}5$, alors que P_2 tombe à 4,0.

Correction de l’exercice 14 – Sondage et échantillons de tailles inégales

Le taux global est la proportion de favorables parmi tous les interrogés :

$$\frac{72+30}{150+50} = \frac{102}{200} = 51\,\%$$

Les taux par zone valent $\dfrac{72}{150} = 48\,\%$ en ville et $\dfrac{30}{50} = 60\,\%$ à la campagne ; leur moyenne simple vaut 54 %.

La différence vient du poids de chaque zone : la ville représente $\dfrac{150}{200} = 75\,\%$ de l’échantillon. Le taux global est une moyenne **pondérée** : $0{,}75 \times 48 + 0{,}25 \times 60 = 51$. Avec deux échantillons de même taille, on retrouverait 54 %. Si la population réelle est moitié urbaine, l’échantillon surreprésente la ville, et un redressement est nécessaire.

Réponse. 51 % au global, contre 54 % pour la moyenne des deux taux : la ville pèse 75 % de l’échantillon.

Correction de l’exercice 15 – Choisir entre trois logements

On lit dans chaque résumé la place du budget de 650 € par rapport aux quartiles.

— Quartier A : la médiane vaut exactement 650 € ; environ la moitié des logements est dans le budget.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

— Quartier B : $Q_1 = 520$ et Méd = 700 ; entre un quart et la moitié des logements est sous 650 €, mais la dispersion est énorme (écart interquartile 360 €).
— Quartier C : $Q_1 = 640$ et Méd = 660 ; un peu plus d'un quart des logements, très serrés autour de 650 €.

$$Q_3 - Q_1 = 100 \text{ (A)} \qquad 360 \text{ (B)} \qquad 40 \text{ (C)}$$

On conseille de chercher d'abord en A, où une annonce sur deux convient avec une dispersion modérée. B peut offrir une bonne affaire (minimum 450 €), au prix de nombreuses visites ; C laisse très peu de marge sous le budget.

Réponse. En A : sa médiane vaut exactement 650 €, donc la moitié du marché tient dans le budget, avec une dispersion modérée.

Correction de l'exercice 16 – Synthèse : bilan d'une saison sportive

a) La somme des 20 scores vaut 1 480, donc $\overline{x} = 74$ points. La médiane est la demi-somme des 10^{e} et 11^{e} valeurs, $\dfrac{72+74}{2} = 73$. Q_1 est la 5^{e} valeur, 65, et Q_3 la 15^{e}, 80. Résumé : 52 ; 65 ; 73 ; 80 ; 102.

b) La moyenne des carrés vaut $\dfrac{112\,296}{20} = 5\,614{,}8$, d'où

$$V = 5\,614{,}8 - 74^2 = 138{,}8 \qquad \sigma = \sqrt{138{,}8} \approx 11{,}78 \text{ points}$$

c) Les scores d'au moins 75 points sont 75 ; 77 ; 78 ; 80 ; 82 ; 85 ; 88 ; 92 ; 102 : 9 victoires sur 20, soit 45 %. C'est cohérent avec une médiane inférieure à 75. Les défaites les plus serrées sont à 74, 72 et 71 points.

d) Passer de 74 à 76 points de moyenne, c'est progresser de $\dfrac{2}{74} \approx 2{,}7\,\%$. Avec 2 points de plus à chaque match, seule la défaite à 74 points deviendrait une victoire (10 victoires) ; gagner aussi celles à 72 et 71 demanderait 3 et 4 points de plus.

Réponse. a) $\overline{x} = 74$, Méd = 73, résumé 52 ; 65 ; 73 ; 80 ; 102 · b) $\sigma \approx 11{,}78$ · c) 9 victoires sur 20 (45 %) · d) +2,7 %, soit 2 points par match.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).