Terminale maths complémentaires / Statistiques à deux variables

benprof particulier.fr

42 exercices en 6 parties, du plus simple au plus difficile : ★☆☆ facile · ★★☆ moyen · ★★★ difficile. Le corrigé détaillé est sur **benprofparticulier.fr**.

1. *Rappels : moyenne, variance, écart type*

À retenir : $V(x)$ = moyenne des carrés des ÉCARTS $= \frac{1}{n}\sum x_i^2 - \overline{x}^2$, en unité², division par n ; $\sigma_x = \sqrt{V(x)}$. Si $y = ax + b$: $\overline{y} = a\overline{x} + b$, $\sigma_y = |a|\sigma_x$. Ex. : $y = -2x + 5$ et $\sigma_x = 3$ donnent $\sigma_y = 6$.

Ex. 1 ★☆☆ **Automatismes : moyenne, variance, écart type** — Réponds sans calculatrice.

a) Calcule la moyenne de la série 12 ; 15 ; 9 ; 14.

b) Calcule la variance de la série 3 ; 5 ; 7 ; 9 par la méthode des écarts.

c) Une série de 10 valeurs a pour moyenne 7. La somme des carrés des valeurs vaut 520. Calcule sa variance.

d) Une série de longueurs a pour variance 2,25 cm². Donne son écart type, avec l'unité.

e) Une série a pour moyenne 20 et pour écart type 3. On ajoute 5 à chaque valeur. Donne la nouvelle moyenne et le nouvel écart type.

f) Une série a pour moyenne $\overline{x} = 4$ et pour écart type $\sigma_x = 1{,}5$. On pose $y_i = -0{,}5x_i + 10$. Donne $\overline{y}$ et σ_y.

Ex. 2 ★☆☆ **Automatismes : vrai ou faux ?** — Pour chaque affirmation, dis si elle est vraie ou fausse, puis justifie.

a) « La variance de la série 2 ; 4 ; 6 est la moyenne de 4, 16 et 36. »

b) « Si les données sont des masses en kg, leur variance s'exprime en kg. »

c) « Pour toute série, la moyenne des carrés est supérieure ou égale au carré de la moyenne. »

d) « Si on multiplie toutes les valeurs par 3, la variance est multipliée par 3. »

e) « Deux séries qui ont la même moyenne ont le même écart type. »

f) « Si une série a un écart type nul, toutes ses valeurs sont égales. »

Ex. 3 ★☆☆ **Une série avec effectifs** — Dans un club de lecture de Rodez, 20 élèves ont noté combien de livres ils ont lus en un mois.

Livres lus x_i	Effectif n_i
0	1
1	4
2	6
3	5
4	4

a) Calcule le nombre moyen de livres lus $\overline{x}$.

b) Amine trouve $\overline{x} = 2$ en faisant la moyenne de 0, 1, 2, 3 et 4. Explique son erreur.

c) Calcule $\sum n_i x_i^2$.

d) Déduis-en $V(x)$ avec la formule de König-Huygens.

e) Donne l'écart type σ_x, arrondi au centième.

Ex. 4 ★★☆ **Trois erreurs d'élèves** — Ben fait peser cinq pommes, en grammes : 152 ; 148 ; 160 ; 145 ; 155. La somme des carrés des masses vaut 115 658.

a) Calcule la moyenne, puis la variance par la méthode des écarts.

b) Lina divise la somme des carrés des écarts par 4. Quelle valeur trouve-t-elle ? Explique son erreur.

c) Nolan applique König-Huygens en tapant « (somme des carrés − carré de la moyenne) ÷5 ». Quel résultat obtient-il ? Corrige son calcul.

d) Sami arrondit la variance à l'unité, puis en prend la racine carrée. Quel écart type annonce-t-il ? Donne l'écart type correct, arrondi au centième.

Ex. 5 ★★☆ **Températures : changer d'unité pour comparer** — Températures maximales relevées à Annecy cinq jours de suite, en °C : 12 ; 15 ; 14 ; 11 ; 13. À Clermont-Ferrand, les mêmes jours, une station annonce une moyenne de 57,2 °F et un écart type de 4,5 °F. On rappelle que $F = 1{,}8C + 32$. Arrondis au centième.

a) Calcule la moyenne et l'écart type des températures d'Annecy, en °C.

b) Déduis-en la moyenne et l'écart type des températures d'Annecy, en °F.

c) Pour convertir l'écart type de Clermont en °C, Inès calcule $\frac{4{,}5-32}{1{,}8}$. Corrige-la, puis donne la moyenne et l'écart type de Clermont en °C.

Ex. 5 (suite)

(d) Compare les températures des deux villes sur ces cinq jours.

Ex. 6 ★★★ **La moyenne, valeur la plus proche de toute la série** — Léna note le temps d'attente de son tram à Grenoble, en minutes, sur cinq matins : 4 ; 7 ; 5 ; 9 ; 10. Pour tout réel t, on pose $f(t) = \frac{1}{5}\left[(4-t)^2+(7-t)^2+(5-t)^2+(9-t)^2+(10-t)^2\right]$. Donne les arrondis au centième.

(a) Développe et réduis $f(t)$ sous la forme $\alpha t^2+\beta t+\gamma$.

(b) Détermine la valeur de t pour laquelle $f(t)$ est minimal, puis ce minimum. Que représentent ces deux nombres pour la série ?

(c) Calcule $f(0)-f(7)$ et interprète $f(0)$. Quelle formule du cours retrouves-tu ?

(d) Léna affirme que l'écart type des temps vaut $\sqrt{f(0)}$. A-t-elle raison ? Justifie, puis donne σ_x avec son unité.

(e) Pour une série quelconque $x_1,\dots,x_n$, montre que $f(t)=\frac{1}{n}\sum(x_i-t)^2$ est minimale en $t=\overline{x}$. Déduis-en que $\frac{1}{n}\sum x_i^2 \geqslant \overline{x}^2$.

Ex. 7 ★★★ **Portions de fromage : deux lignes de conditionnement** — Une fromagerie de Savoie conditionne des portions de 250 g. Sur la ligne A, Maëlle pèse 20 portions (masse m en grammes). Donne les arrondis au centième.

Masse m (g)	Effectif
246	2
248	5
250	8
252	4
254	1

(a) On pose $x = m - 250$. Calcule $\overline{x}$, puis déduis-en $\overline{m}$ sans repasser par les masses.

(b) Calcule $V(x)$ par König-Huygens. Pourquoi taper $\overline{x}^2$ sans parenthèses à la calculatrice donne-t-il un résultat faux ?

(c) Justifie que $\sigma_m = \sigma_x$, puis calcule σ_m avec son unité.

(d) Une élève affirme : « avec $d = 250 - m$, la masse manquante, on a $\sigma_d = -\sigma_m$ ». Corrige-la, puis donne $\overline{d}$ et σ_d.

(e) Sur la ligne B : $\overline{m} = 250{,}1$ g et $\sigma_m = 3{,}4$ g. Une ligne est conforme si $\overline{m} \in [249;251]$ et $\sigma_m \leqslant 2{,}5$ g. Compare les deux lignes et conclus.

2. Nuage de points et point moyen

À retenir : Nuage : je décris sa forme, son sens et les points isolés (un nuage allongé ne prouve pas un lien affine). $G(\overline{x};\overline{y})$: je calcule les deux moyennes séparément. Ex. : $(1;4)$, $(2;6)$, $(3;11)$ donnent $G(2;7)$ et non $(2;6)$.

Ex. 8 ★★★ **Automatismes : le point moyen d'une série** — À Sète, un glacier relève sur cinq jours la température maximale x_i (en °C) et le nombre y_i de cornets vendus.

Jour	Température x_i (°C)	Cornets y_i
Lundi	14	32
Mardi	18	45
Mercredi	21	51
Jeudi	25	68
Vendredi	27	74

(a) Calcule la température moyenne $\overline{x}$.

(b) Calcule le nombre moyen de cornets vendus $\overline{y}$.

(c) Donne les coordonnées du point moyen G de la série.

(d) Calcule la médiane des x_i et celle des y_i. Le point des médianes est-il le point G ? Justifie.

(e) Le point G est-il un point du nuage ? Justifie.

(f) Le glacier s'est trompé pour vendredi : il a vendu 84 cornets, et non 74. Quelle coordonnée de G change ? Donne le nouveau point moyen.

Ex. 9 ★★★ **Automatismes : vrai ou faux sur le nuage** — Pour chaque affirmation, réponds par vrai ou faux et justifie.

(a) « Quand un nuage de points est allongé, il existe forcément un lien affine entre x et y. »

(b) L'axe des ordonnées d'un graphique commence à 480. Le premier point est en $y = 492$, le dernier, tout en haut, en $y = 510$. « y a plus que doublé. »

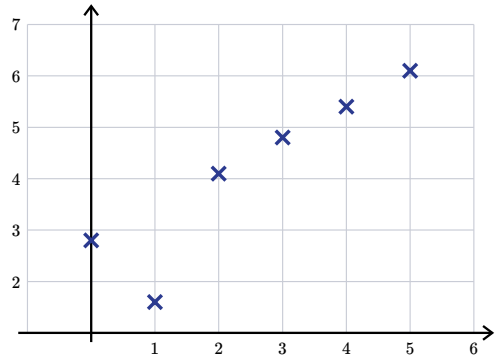
(c) « Le point moyen de la série $(2;10)$, $(4;14)$, $(6;30)$ est $(4;14)$, le point du milieu. »

(d) « Si on ajoute à une série un point confondu avec son point moyen G, le point moyen ne change pas. »

(e) « Si on ajoute 3 à chaque valeur y_i sans changer les x_i, le point moyen monte de 3. »

(f) x désigne le rang de l'année, avec $x = 1$ pour 2020. « L'année 2026 correspond à $x = 6$. »

Ex. 10 ★★☆ **Autopartage à Grenoble : lire un nuage** — Un service d'autopartage de Grenoble suit son nombre d'abonnés y (en milliers), x désignant le rang de l'année, avec $x = 0$ pour 2019. Arrondis au centième.



(a) Lis sur le nuage le nombre d'abonnés en 2022.

(b) Décris le nuage (forme, sens, point isolé) en comparant les variations d'une année sur l'autre. Un ajustement affine te paraît-il adapté ? Justifie.

(c) Calcule les coordonnées du point moyen G.

(d) Élise affirme : « Le point de 2024 est tout en haut, celui de 2020 tout en bas : les abonnés ont été multipliés par plus de 8. » A-t-elle raison ? Justifie.

Ex. 11 ★★☆ **Course à pied à Annecy : corrige l'élève** — Dans un club d'Annecy, on relève pour cinq coureurs leur entraînement x_i (en heures par semaine) et leur temps y_i sur 10 km (en minutes).

Coureur	Temps y_i (min)	Entraînement x_i (h)
A	58	3
B	52	5
C	50	6
D	46	8
E	43	9

Nolan écrit : « Je place le temps sur 10 km en abscisse, car c'est la première colonne. G est le point du milieu du tableau, celui du coureur C : $G(50\,;6)$. »

(a) Quelle variable faut-il placer en abscisse ? Justifie, puis corrige la première erreur de Nolan.

(b) Explique l'erreur de Nolan sur G, puis calcule les coordonnées exactes de G.

(c) Le nuage est décroissant. Traduis cette tendance dans le contexte.

(d) Le coureur B dit : « Sur le graphique, mon point est au-dessus de G, donc je cours mieux que la moyenne du club. » A-t-il raison ? Justifie.

Ex. 12 ★★☆ **Engrais près de Chartres : sous-groupes et point ajouté** — Une coopérative près de Chartres teste sur six parcelles une dose d'engrais x_i (en kg/ha) et relève le rendement y_i en blé (en quintaux par hectare). Arrondis au centième.

Dose x_i (kg par ha)	Rendement y_i (q par ha)
40	52
60	58
80	64
100	69
120	73
140	77

(a) Calcule les coordonnées du point moyen G des six parcelles.

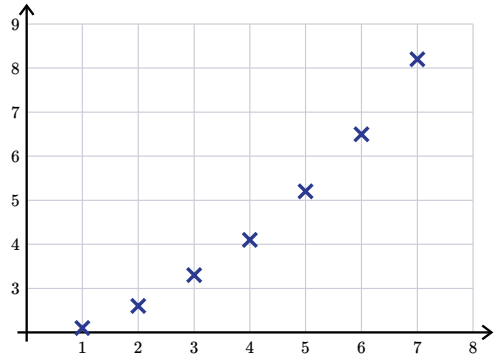
(b) Calcule le point moyen G_1 des trois parcelles les moins fertilisées, puis le point moyen G_2 des trois autres.

(c) On ajoute une septième parcelle : 160 kg/ha, 74 q/ha. Calcule le nouveau point moyen G' des sept parcelles.

(d) Compare les variations de rendement d'une dose à la suivante pour les sept parcelles. Un ajustement affine te paraît-il adapté ? Justifie.

Ex. 13 ★★★ **Application nantaise : un nuage qui se courbe** — Une start-up nantaise suit le nombre y_i d'utilisateurs actifs de son application (en milliers). x_i désigne le rang du trimestre, avec $x = 1$ pour le premier trimestre 2024. Arrondis au centième.

Rang x_i	Utilisateurs y_i (milliers)
1	2,1
2	2,6
3	3,3
4	4,1
5	5,2
6	6,5
7	8,2



(a) Décris le nuage (forme, sens, point isolé). Calcule les écarts successifs $y_{i+1} - y_i$, puis dis si un ajustement affine te paraît adapté. Justifie.

Ex. 13 (suite)

(b) Calcule les quotients successifs $\frac{y_{i+1}}{y_i}$, arrondis au centième. Que constates-tu ? Quel type de modèle cela suggère-t-il ?

(c) Léa affirme que le point moyen est le point du 4e trimestre, puisqu'il est au milieu du tableau. Calcule G, puis corrige son affirmation.

(d) On ajoute le trimestre de rang $x = 8$: de quel trimestre s'agit-il ? Combien d'utilisateurs faut-il alors pour que l'ordonnée du point moyen passe à 5,2 ? Donne aussi sa nouvelle abscisse.

Ex. 14 ★★★ **Vélos électriques d'occasion à Rennes : sous-groupes et point ajouté** — Inès, revendeuse à Rennes, relève l'âge x_i (en années) et le prix de vente y_i (en euros) de huit vélos électriques d'occasion. Arrondis au centième.

Vélo	Âge x_i (années)	Prix y_i (€)
V1	1	1900
V2	1	1750
V3	2	1550
V4	3	1400
V5	4	1150
V6	5	1000
V7	6	850
V8	8	600

(a) En t'appuyant sur le tableau, décris l'allure du nuage (sens, forme, point isolé), puis calcule les coordonnées du point moyen G.

(b) Malo prend comme point moyen le milieu des points extrêmes $(1\,;1\,900)$ et $(8\,;600)$. Calcule ce point et explique pourquoi Malo se trompe.

(c) Calcule le point moyen G_1 des vélos V1 à V4, puis le point moyen G_2 des vélos V5 à V8. Démontre que G est le milieu de $[G_1G_2]$ et explique pourquoi c'est le cas ici.

(d) On ajoute un neuvième vélo, âgé de 10 ans. Inès affirme que seule l'abscisse de G va changer. Quel devrait être le prix de ce vélo pour qu'elle ait raison ? Ce prix est-il vraisemblable ? Justifie.

(e) En réalité, ce vélo se vend 450 €. Calcule le point moyen G' des neuf vélos.

3. *Droite des moindres carrés*

À retenir : $a = \frac{\text{cov}(x,y)}{V(x)}$ (pas σ_x), $b = \overline{y} - a\overline{x}$ avec a non arrondi ; contrôle $a\overline{x} + b = \overline{y}$. Résidu $e_i = y_i - (ax_i + b)$: somme nulle, somme des carrés minimale. Ex. : cov $= 6$, $V(x) = 3$, $G(1\,;5)$ donnent $y = 2x + 3$.

Ex. 15 ★★★ **Automatismes : a et b à partir des indicateurs** — Questions indépendantes, sans calculatrice. Dans chaque cas, il s'agit de la droite des moindres carrés (droite de régression de y en x).

(a) On donne $\overline{x} = 4$, $\overline{y} = 20$, $\text{cov}(x,y) = 6$ et $V(x) = 2{,}5$. Détermine l'équation de la droite.

(b) Une droite des moindres carrés a pour coefficient directeur $a = -3$, et le point moyen de la série est $G(5\,;14)$. Calcule b.

(c) On donne $\sigma_x = 2$, $\text{cov}(x,y) = 10$, $\overline{x} = 3$ et $\overline{y} = 40$. Calcule a puis b.

(d) On a $V(x) = 4$, $V(y) = 9$ et $\text{cov}(x,y) = 5$. Hugo écrit $a = \frac{9}{4}$. A-t-il raison ? Justifie.

(e) On a $a = \frac{7}{3}$, $\overline{x} = 2$ et $\overline{y} = 10$. Calcule b au centième, puis refais le calcul en arrondissant d'abord a à 2,33. Conclus.

(f) Le point moyen d'une série est $G(6\,;10)$. Vrai ou faux : $y = 1{,}5x + 2$ peut être la droite des moindres carrés de cette série. Justifie.

Ex. 16 ★★★ **Automatismes : résidus et estimations** — Questions indépendantes, sans calculatrice. Le résidu d'un point $(x_i\,;y_i)$ est $e_i = y_i - (ax_i + b)$.

(a) Pour la droite $y = 2x + 3$, calcule le résidu du point $(5\,;14)$.

(b) Pour la même droite, calcule le résidu du point $(4\,;9)$. Ce point est-il au-dessus ou en dessous de la droite ?

(c) Pour la droite des moindres carrés d'une série de 4 points, trois résidus valent 0,4 ; $-1{,}2$ et 0,5. Calcule le quatrième.

(d) Les résidus d'un ajustement valent 1 ; -2 ; 0,5 et 0,5. Calcule la somme de leurs carrés.

(e) Vrai ou faux : la droite des moindres carrés est la droite qui rend minimale la somme des résidus. Justifie.

(f) La droite de régression de y en x est $y = 0{,}8x + 12$. Estime y pour $x = 15$, puis la valeur de x pour laquelle le modèle prévoit $y = 20$.

Ex. 17 ★★★ **Publicité et chiffre d'affaires** — Élodie tient un glacier à Sète. Sur 4 mois, elle note son budget publicitaire x_i (en centaines d'euros) et son chiffre d'affaires mensuel y_i (en milliers d'euros). Calculs à la main ; arrondis au millième.

Budget x_i	Chiffre d'affaires y_i
2	31
3	35
6	40
9	46

(a) Calcule les coordonnées du point moyen G.

(b) Calcule $\text{cov}(x,y)$ et $V(x)$.

Ex. 17 (suite)

(c) Déduis-en l'équation de la droite des moindres carrés (b calculé avec a non arrondi). Contrôle qu'elle passe par G.

(d) Estime le chiffre d'affaires pour un budget de 700 €.

(e) Quel budget le modèle prévoit-il pour atteindre 50 milliers d'euros ? Cette estimation est-elle fiable ?

Ex. 18 ★★☆ **L'erreur de Chloé** — Chloé étudie le temps de révision x (en h) et le score y (sur 100) de lycéens à un test blanc. On a $\overline{x} = 6$, $\overline{y} = 50$, $\sigma_x = 2$ et $\mathrm{cov}(x, y) = 8$. Elle écrit : « $a = \frac{\mathrm{cov}(x,y)}{\sigma_x} = 4$ et $b = \overline{y} + a\overline{x} = 74$, donc $y = 4x + 74$. »

(a) Vérifie si la droite de Chloé passe par le point moyen G. Que peux-tu en conclure ?

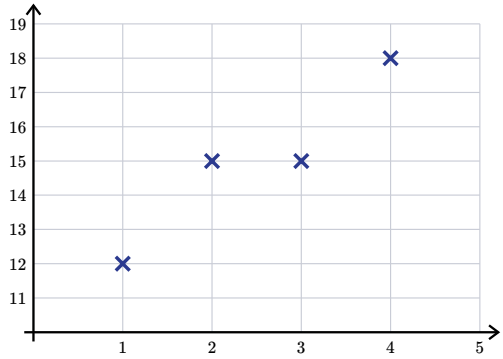
(b) Identifie les deux erreurs de Chloé.

(c) Détermine l'équation correcte de la droite des moindres carrés et vérifie qu'elle passe par G.

(d) Avec la droite correcte, estime le score d'un lycéen qui révise 8 h.

(e) Un ami de Chloé affirme : « Puisqu'on a trouvé une droite d'ajustement, les points du nuage sont alignés. » A-t-il raison ? Justifie.

Ex. 19 ★★☆ **Deux droites en concurrence** — Un club de sport d'Annecy relève sa consommation d'électricité y (en MWh, valeurs entières) pour les années de rang $x = 1$ à 4. Nadia propose la droite $D_1 : y = 1{,}8x + 10{,}5$ et Théo la droite $D_2 : y = 2x + 10$.



(a) Lis les valeurs de y sur le nuage, puis vérifie que D_1 et D_2 passent toutes deux par le point moyen G.

(b) Calcule les résidus de chaque droite, puis les sommes de leurs carrés S_1 et S_2.

(c) Théo : « La somme des $|e_i|$ est plus petite pour D_2, donc D_2 est la meilleure droite au sens des moindres carrés. » A-t-il raison ? Justifie.

(d) On fixe $a = 1{,}8$. Écris $S(b) = \sum(y_i - 1{,}8x_i - b)^2$ sous forme de trinôme en b, puis trouve la valeur de b qui le rend minimal.

Ex. 20 ★★★ **Démonstration guidée : pourquoi cette droite ?** — Dans un jardin partagé de Besançon, Inès relève sur 4 parcelles la dose d'engrais x_i (en kg par are) et la récolte y_i (en dizaines de kg par are). Calculatrice interdite. On note $S(a, b) = \sum(y_i - ax_i - b)^2$. En développant, on obtient $S(a, b) = 4b^2 - 2b\sum(y_i - ax_i) + \sum(y_i - ax_i)^2$.

Dose x_i	Récolte y_i
2	3
4	6
5	8
9	11

(a) On fixe a. Exprime $\sum(y_i - ax_i)$ en fonction de a, puis trouve la valeur de b qui rend minimal le trinôme en b. Qu'en déduis-tu pour la droite et le point moyen G ?

(b) On remplace b par cette valeur. À l'aide des écarts $x_i - \overline{x}$ et $y_i - \overline{y}$, écris S comme un trinôme en a. Déduis-en a et b, en valeurs exactes puis au millième.

(c) Yanis prend $a = \frac{\mathrm{cov}(x,y)}{\sigma_x}$. Calcule la valeur qu'il obtient (au centième), puis montre avec S qu'elle ne convient pas.

(d) Lina affirme : « La droite $y = x + 2$ convient aussi, car la somme de ses résidus est nulle. » A-t-elle raison ? Justifie.

(e) Ilan arrondit a au centième avant de calculer b. Quelle valeur de b trouve-t-il ? Explique son erreur.

Ex. 21 ★★★ **Vélos en libre-service : du modèle à la prévision** — Une ville relève le nombre d'abonnés y_i (en milliers) à son service de vélos en libre-service pour l'année $2018 + x_i$ (données fictives). Calculatrice autorisée ; arrondis au millième, sauf mention contraire.

Rang x_i	Abonnés y_i (milliers)
1	4,2
2	5,1
3	5,7
4	6,8
5	7,4
6	8,6

(a) Calcule les coordonnées du point moyen G. À la calculatrice, donne l'équation de la droite des moindres carrés de y en x, puis contrôle qu'elle passe par G.

(b) La calculatrice donne $r \approx 0{,}996$. Mehdi en déduit que les points sont alignés. A-t-il raison ? Justifie avec le résidu de l'année de rang 3.

(c) Calcule le résidu de l'année de rang 6. Sans calculer les autres, donne la somme des six résidus et justifie.

(d) Le service vise 12 000 abonnés. Kevin veut calculer la droite de x en y. Explique pourquoi on ne fait pas ainsi, puis donne l'année visée selon le modèle.

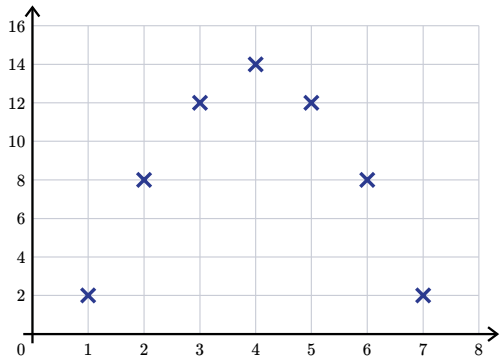
Ex. 21 **(suite)**

(e) Cette prévision est-elle fiable ? Donne deux raisons d'être prudent.

4. *Coefficient de corrélation, corrélation et causalité*

À retenir : $r = \dfrac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y}$ est entre -1 et 1, a le signe de a et ne dépend pas des unités. $|r|$ proche de 1 : lien affine fort, pas forcément causal. $r \approx 0$: pas de lien affine, mais un lien courbé reste possible. Ex. : $r = -0{,}97$: nuage décroissant presque aligné.

Ex. 22 ★☆☆ **Automatismes : vrai ou faux sur r** — Pour chaque affirmation, dis si elle est vraie ou fausse et justifie. Le nuage donne la fréquentation y d'un marché de plein air (en centaines de visiteurs) selon le rang x du mois, de mars ($x = 1$) à septembre ($x = 7$).



(a) Une calculatrice affiche $r = 1{,}08$: c'est possible si les points sont très bien alignés.

(b) Si $r = 0$, la droite de régression est parallèle à l'axe des abscisses.

(c) $r = 0{,}95$ signifie que 95 % des points sont sur la droite de régression.

(d) Pour la série du marché, $\operatorname{cov}(x, y) = 0$, donc la fréquentation ne dépend pas du mois.

(e) Si l'on convertit x de kilomètres en mètres, r est multiplié par 1 000.

(f) Si $r = -0{,}98$, les points sont proches d'une droite décroissante.

(g) Chez des enfants de 6 à 11 ans, la pointure et la vitesse de lecture ont $r = 0{,}9$: avoir de grands pieds fait mieux lire.

Ex. 23 ★☆☆ **Automatismes : calculer r et l'exploiter** — Arrondis les valeurs de r au millième.

(a) On donne $\operatorname{cov}(x, y) = 6{,}3$, $\sigma_x = 2{,}5$ et $\sigma_y = 2{,}8$. Calcule r.

(b) On donne $\operatorname{cov}(x, y) = -14{,}4$, $V(x) = 16$ et $V(y) = 16{,}2$. Calcule r.

(c) Une calculatrice affiche la droite $y = -2{,}1x + 40$ et $r^2 = 0{,}8836$, le carré de r. Donne r.

Ex. 23 **(suite)**

(d) On donne $r = 0{,}8$, $\sigma_x = 2$ et $\sigma_y = 5$. Calcule le coefficient directeur a de la droite de régression.

(e) Avec $V(x) = 3$, $V(y) = 10$ et $\operatorname{cov}(x, y) = 5{,}4$, Léna arrondit $\sigma_x \approx 1{,}7$ et $\sigma_y \approx 3{,}2$, puis trouve $r \approx 0{,}993$. Corrige son calcul.

(f) Une série peut-elle avoir $\operatorname{cov}(x, y) = -3$ et une droite de régression de coefficient directeur $a = 0{,}5$? Justifie.

Ex. 24 ★★☆ **Entraînement et temps au 3 000 m** — Un club d'athlétisme de Rennes relève, pour quatre coureurs, le temps d'entraînement hebdomadaire x_i (en h) et le temps y_i réalisé au 3 000 m (en min). Le club juge un ajustement affine justifié si $|r| \geqslant 0{,}9$.

Coureur	x_i (h)	y_i (min)
1	2	15,4
2	4	14,6
3	6	13,9
4	8	12,9

(a) Détermine le point moyen G, puis calcule $\operatorname{cov}(x, y)$.

(b) Calcule $V(x)$ et $V(y)$, puis r arrondi au millième.

(c) Interprète le signe de r et indique, avec le critère du club, si un ajustement affine est justifié.

(d) Détermine la droite de régression de y en x, puis estime le temps d'un coureur qui s'entraînerait 20 h par semaine. Cette estimation est-elle fiable ? Justifie.

Ex. 25 ★★☆ **Chauffage et point aberrant** — À Grenoble, Sacha relève pendant six jours la température extérieure moyenne x_i (en °C) et la consommation de chauffage y_i (en kWh) de sa maison. On juge un ajustement affine justifié si $|r| \geqslant 0{,}9$. Utilise la calculatrice, arrondis r au millième et les coefficients au centième.

Jour	x_i (°C)	y_i (kWh)
1	0	62
2	2	54
3	4	51
4	6	20
5	8	37
6	10	29

(a) Donne r et la droite de régression de y en x pour les six jours.

(b) Avec le seuil de l'énoncé, l'ajustement affine est-il justifié ? Justifie.

(c) Le relevé du jour 4 est dû à une panne du compteur. Sans ce jour, donne la nouvelle valeur de r et la nouvelle droite.

Ex. 25 **(suite)**

(d) Estime la consommation pour une température de 5 °C avec chacune des deux droites, puis commente l'écart obtenu.

Ex. 26 ★★☆ **Piscines et cambriolages** — Pour cinq communes d'un même département, on relève la population (en milliers d'habitants), le nombre x_i de piscines publiques et le nombre y_i de cambriolages en un an. On parle de forte corrélation linéaire si $|r| \geqslant 0{,}9$.

Commune	Population (milliers)	x_i	y_i
A	8	1	40
B	15	2	72
C	22	2	112
D	40	4	196
E	65	6	330

(a) À la calculatrice, donne r arrondi au millième, puis interprète-le avec le seuil de l'énoncé.

(b) Hugo conclut : « Construire des piscines fait augmenter les cambriolages. » A-t-il raison ? Justifie.

(c) Calcule, pour chaque commune, le nombre de cambriolages pour 1 000 habitants (au centième). Que constates-tu ?

(d) Rédige une conclusion prudente sur le lien de cause à effet entre x et y.

Ex. 27 ★★★ **Le câble de la passerelle** — Sur une passerelle suspendue près d'Annecy, un technicien mesure la hauteur y_i (en m) du câble porteur au-dessus du tablier à la position x_i (en m). Cette position est comptée depuis le milieu de la passerelle, négative à gauche. On juge un ajustement affine pertinent si $|r| \geqslant 0{,}99$. Calculatrice autorisée.

x_i (m)	y_i (m)
-30	20,3
-20	9,8
-10	4,1
0	2,0
10	3,9
20	10,1
30	19,8

(a) À la calculatrice, donne r entre x et y au millième. Mehdi en conclut : « la hauteur du câble ne dépend pas de la position ». A-t-il raison ? Justifie.

(b) On pose $z = x^2$. Calcule les z_i, puis donne r entre z et y au dix-millième. Conclus avec le seuil de l'énoncé.

(c) Donne la droite de régression de y en z (a au dix-millième, b au millième), puis déduis-en une expression de y en fonction de x.

Ex. 27 **(suite)**

(d) Estime la hauteur du câble à 25 m du milieu, avec les valeurs non arrondies de la calculatrice. Est-ce une interpolation ou une extrapolation ?

(e) Nour reprend l'étude avec les hauteurs en cm et les positions en dm. Sans aucun calcul, donne les deux coefficients de corrélation qu'elle obtient. Justifie.

Ex. 28 ★★★ **Glaces et sauvetages en mer** — Dans une station balnéaire de Vendée, les secouristes relèvent pendant six semaines d'été la température maximale moyenne t_i (en °C), les ventes de glaces x_i (en centaines) d'un glacier de la plage et le nombre y_i d'interventions de sauvetage en mer. On parle de corrélation forte si $|r| \geqslant 0{,}9$. Calculatrice autorisée.

Semaine	t_i	x_i	y_i
1	22	15	5
2	25	26	7
3	28	33	8
4	31	44	10
5	27	30	7
6	23	20	5

(a) À la calculatrice, donne r entre x et y au millième et la droite de régression de y en x (coefficients au millième). Interprète r avec le seuil de l'énoncé.

(b) Inès affirme : « $r \approx 0{,}98$, donc 98 % des points sont sur la droite. » Calcule le résidu de la semaine 2, puis dis si elle a raison. Justifie.

(c) À la calculatrice, donne au millième le coefficient de corrélation entre t et x, puis celui entre t et y.

(d) Yanis remarque que r entre x et y est plus grand qu'entre t et y. Il conclut : « Ce sont les glaces, et non la chaleur, qui causent les sauvetages. » A-t-il raison ? Justifie.

(e) Un élu propose de limiter la vente de glaces pour réduire le nombre de sauvetages. En t'appuyant sur les questions précédentes, discute cette proposition.

5. *Interpoler et extrapoler avec un ajustement*

À retenir : Plage $[x_{\min} ; x_{\max}]$: dedans, c'est une interpolation (fiable) ; dehors, c'est une extrapolation « si la tendance se poursuit ». Pour un seuil, tu résous $ax + b > y_0$ (le sens change si $a < 0$), tu prends l'entier supérieur, puis tu donnes l'année : $x = 8$ avec $x = 1$ en 2019 donne 2026.

Ex. 29 ★☆☆ **Automatismes : interpolation ou extrapolation ?** — À Besançon, un journal papier relève chaque année son nombre d'abonnés y (en milliers), de 2018 ($x = 1$) à 2024 ($x = 7$). Il retient l'ajustement $y = -1{,}8x + 52$. Pour chaque question, précise s'il s'agit d'une interpolation ou d'une extrapolation.

(a) Estime le nombre d'abonnés pour $x = 3$.

(b) Calcule le rang x de l'année 2027, puis estime le nombre d'abonnés en 2027.

(c) Estime le nombre d'abonnés en 2017.

(d) Calcule le rang de l'année 2047, puis l'estimation donnée par l'ajustement. Commente.

Ex. 30 ★☆☆ **Automatismes : atteindre un seuil** — Dans chaque question, x est un entier positif.

(a) Résous $2{,}5x + 12 > 30$ et donne le plus petit entier x qui convient.

(b) Résous $-0{,}6x + 9 < 4$ et donne le plus petit entier x qui convient.

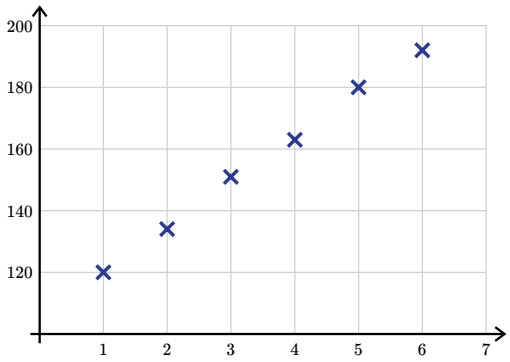
(c) On a $y = 1{,}35x + 40{,}2$, avec $x = 1$ en 2021. En quelle année y dépasse-t-il 50 pour la première fois ?

(d) Pour $y = -3{,}2x + 75$, Léo affirme que $y \leqslant 40$ dès $x = 10$. Vrai ou faux ? Justifie, puis donne le plus petit entier x qui convient.

(e) Noé résout $-0{,}4x + 20 > 14$ et trouve $x > 15$. Repère son erreur et corrige-la.

Ex. 31 ★★☆ **Une coopérative solaire à Aurillac** — Une coopérative d'Aurillac relève sa production annuelle d'électricité y (en MWh), de 2015 ($x = 1$) à 2020 ($x = 6$). Calculatrice autorisée.

Rang x_i	Production y_i (MWh)
1	120
2	134
3	151
4	163
5	180
6	192



(a) Donne à la calculatrice une équation de la droite de régression de y en x, avec les coefficients arrondis au centième.

(b) Estime la production en 2023, arrondie au centième de MWh. S'agit-il d'une interpolation ou d'une extrapolation ?

(c) En 2023, la production réellement mesurée a été de 228 MWh. Calcule l'écart entre la prévision et cette valeur, en % de la valeur réelle, arrondi au centième.

(d) Selon ce modèle, à partir de quelle année la production dépasse-t-elle 250 MWh ?

Ex. 32 ★★☆ **La fibre à Millau** — Dans un quartier de Millau, on relève le taux y de foyers raccordés à la fibre (en %), de 2020 ($x = 1$) à 2024 ($x = 5$).

Rang x_i	Taux y_i (%)
1	62
2	68
3	73
4	79
5	85

(a) Calcule les coordonnées du point moyen G du nuage.

(b) On admet que la droite des moindres carrés a pour coefficient directeur $a = 5{,}7$. Déduis-en son ordonnée à l'origine b.

(c) Estime le taux de raccordement en 2029 avec cette droite. Commente le résultat.

Ex. 32 **(suite)**

(d) À partir de quelle année ce modèle donne-t-il un taux supérieur à 100 % ? Que peux-tu en conclure ?

Ex. 33 ★★☆ **La chaîne vidéo de Maëlys** — Maëlys relève le nombre y d'abonnés de sa chaîne vidéo (en milliers) du mois 1 au mois 8. Avec $z = \ln y$, elle obtient l'ajustement $z = 0{,}18x + 2{,}1$, où x est le numéro du mois. Arrondis au centième.

(a) Exprime y en fonction de x sous la forme $y = k\,\mathrm{e}^{0{,}18x}$, avec k arrondi au centième.

(b) Estime le nombre d'abonnés au mois 6. Interpolation ou extrapolation ?

(c) Selon ce modèle, à partir de quel mois la chaîne dépasse-t-elle 100 000 abonnés ?

(d) Maëlys affirme : « Au mois 24, j'aurai 614 000 abonnés, c'est sûr. » A-t-elle raison ? Justifie.

Ex. 34 ★★★ **Enneigement près de Font-Romeu : jusqu'où croire le modèle ?** — Une station de ski près de Font-Romeu relève chaque hiver le nombre y de jours où l'enneigement permet de skier, de l'hiver 2017 ($x = 1$) à l'hiver 2023 ($x = 7$). Calculatrice autorisée.

Rang x_i	Jours y_i
1	118
2	112
3	109
4	101
5	97
6	93
7	86

(a) Donne à la calculatrice une équation de la droite de régression de y en x (coefficients au centième) et r au millième. Un ajustement affine est-il justifié ?

(b) Avec cette équation aux coefficients arrondis, estime le nombre de jours d'enneigement de l'hiver 2027. Est-ce une interpolation ou une extrapolation ?

(c) L'hiver 2025, on a observé 71 jours. Calcule l'estimation du modèle pour 2025, puis son écart relatif à la valeur observée, en %, au dixième.

(d) Selon ce modèle, à partir de quel hiver l'enneigement passerait-il sous 50 jours ? Justifie par une inéquation.

(e) Hugo affirme : « D'après le modèle, il n'y aura plus aucun jour d'enneigement dès l'hiver 2039. » Vrai ou faux ? Justifie, puis commente.

Ex. 35 ★★★ **Une culture de levures : affine ou exponentiel ?** — Inès, technicienne dans un laboratoire de Rennes, mesure la concentration y d'une culture de levures (en milliers de cellules par mL), x heures après le début de l'expérience. Calculatrice autorisée.

Temps x_i (h)	y_i
1	12
2	18
3	27
4	40
5	61

(a) Donne à la calculatrice la droite de régression de y en x (coefficients au dixième), puis estime la concentration à 3 h 30. Interpolation ou extrapolation ?

(b) On pose $z = \ln y$. Donne la droite de régression de z en x (coefficients au millième), puis écris $y = k\,\mathrm{e}^{ax}$ avec k arrondi à l'unité.

(c) À 8 h, Inès mesure 198 milliers de cellules par mL. Calcule la prévision de chaque modèle et son écart relatif à cette mesure, en %, au dixième.

(d) Avec le modèle exponentiel, détermine à partir de quelle heure entière la concentration dépasse 500 milliers de cellules par mL.

(e) Inès écrit : « Avec le modèle exponentiel, à 24 h, il y aura certainement plus de 100 millions de cellules par mL. » Vrai ou faux ? Justifie.

6. *Ajustement par changement de variable*

À retenir : Nuage courbé : on pose $z = \ln y$ (si $y > 0$), $z = \frac{1}{x}$ ou $\ln y$ en $\ln x$, on ajuste, puis on revient à y. Ex. : $z = 0{,}1x + 2$ donne $y = \mathrm{e}^2 \times (\mathrm{e}^{0{,}1})^x$, et non $\mathrm{e}^{0{,}1x} + \mathrm{e}^2$; le taux vaut $\mathrm{e}^a - 1$.

Ex. 36 ★☆☆ **Automatismes : passer par $z = \ln y$ et revenir à y** — Arrondis au centième sauf mention contraire.

(a) Inès a relevé $y = 3{,}5$; 0 ; 12. Elle veut poser $z = \ln y$ pour toute la série. Est-ce possible ? Justifie.

(b) Calcule $z = \ln y$ pour $y = 45$, au millième.

(c) La droite de régression de $z = \ln y$ en x est $z = 0{,}2x + 1{,}5$. Écris y sous la forme $y = k \times q^x$.

(d) Avec cette même droite, Léa écrit $y = \mathrm{e}^{0{,}2x} + \mathrm{e}^{1{,}5}$. A-t-elle raison ? Justifie avec $x = 2$.

(e) Pour $x = 8$, cette droite donne $z = 3{,}1$. Matéo conclut : « $y \approx 3{,}1$ ». Corrige sa réponse.

(f) Pour une autre série, on trouve $z = -0{,}05x + 3$. De quel pourcentage y évolue-t-il quand x augmente de 1 ?

Ex. 37 ★☆☆ **Automatismes : choisir le bon changement de variable** — Arrondis au centième sauf mention contraire.

(a) Le nuage de $(x\,;y)$ ressemble à une branche d'hyperbole qui se rapproche d'une droite horizontale. Quel changement de variable choisis-tu ?

(b) Pour un nuage de ce type, après ce changement de variable, on obtient la droite de régression $y = 24z + 1{,}5$. Estime y pour $x = 8$.

(c) On propose le modèle $y = k \times x^{p}$ avec $x > 0$ et $y > 0$. Quelles variables poses-tu pour te ramener à un ajustement affine ?

(d) Après ce changement de variables, la calculatrice donne la droite $Y = 1{,}5X + 0{,}7$, où Y dépend de y et X de x. Donne p et k.

(e) Avec le modèle puissance de la question précédente, estime y pour $x = 4$.

(f) Vrai ou faux ? « y est multiplié par 1,08 à chaque pas de x, donc on pose $z = \dfrac{1}{x}$. » Justifie.

(g) La droite de régression de $z = \ln y$ en x est $z = 0{,}3x + 2$. Pour quelles valeurs de x a-t-on $y \geqslant 100$?

Ex. 38 ★★☆ **Élimination d'un médicament** — On mesure la concentration y (en μg/L) d'un médicament dans le sang d'un patient, x heures après l'injection. Utilise la calculatrice.

Heure x_i	Concentration y_i
0	480
1	400
2	315
3	252
4	200
5	162

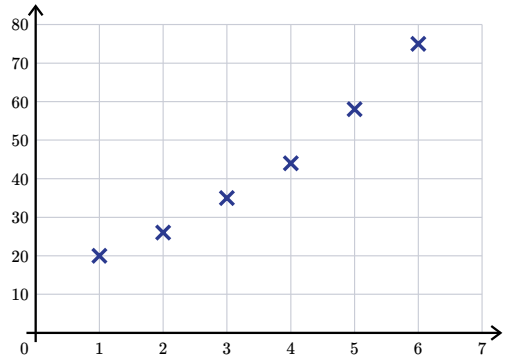
(a) Justifie que l'on peut poser $z = \ln y$. Donne la droite de régression de z en x (coefficients au millième) et r au centième.

(b) Déduis-en y en fonction de x sous la forme $y = k \times q^{x}$ (k à l'unité, q au millième).

(c) De quel pourcentage la concentration baisse-t-elle chaque heure, selon ce modèle ?

(d) Le médicament n'agit plus sous 50 μg/L. Selon ce modèle, à partir de combien d'heures est-ce le cas ? Ce résultat est-il fiable ?

Ex. 39 ★★☆ **Panneaux solaires : affine ou exponentiel ?** — Voici la puissance y (en kWc) des panneaux solaires installés dans une commune, selon le rang x de l'année. Pour l'ajustement affine, la calculatrice donne $y = 10{,}86x + 5$ avec $r \approx 0{,}984$.



Rang x_i	Puissance y_i
1	20
2	26
3	35
4	44
5	58
6	75

(a) Calcule $z = \ln y$ au millième, puis donne la droite de régression de z en x (coefficients au millième) et r au millième.

(b) Déduis-en $y = k \times q^{x}$ (k au dixième, q au millième) et le taux d'évolution annuel donné par ce modèle.

(c) L'année de rang 7, on mesure 96 kWc. Quel ajustement est le plus fiable ? Justifie par le calcul.

(d) Noé arrondit au dixième, écrit $z = 0{,}3x + 2{,}7$ et trouve 121,5 kWc pour le rang 7. Explique son erreur.

Ex. 40 ★★☆ **Recharge d'une voiture électrique** — Un garage de Rennes relève la durée y (en min) de recharge d'une voiture selon la puissance x (en kW) de la borne. Le nuage a la forme d'une hyperbole : on pose $z = \dfrac{1}{x}$. La calculatrice donne $r \approx 0{,}9996$ pour la série $(z\,;y)$.

Puissance x_i	Durée y_i
10	250
20	135
40	70
50	61

(a) Calcule les valeurs $z_i = \dfrac{1}{x_i}$, puis les coordonnées du point moyen G de la série $(z\,;y)$.

(b) Calcule $V(z)$ et $\operatorname{cov}(z, y)$. Déduis-en la droite de régression de y en z, sous la forme $y = \alpha z + \beta$ (α à l'unité, β au dixième).

Ex. 40 **(suite)**

(c) Exprime y en fonction de x, puis estime la durée de recharge sur une borne de 25 kW.

(d) Quelle puissance faut-il au minimum pour recharger en moins de 45 min ? Ce résultat est-il fiable ?

Ex. 41 ★★★ **Croissance des bars : un modèle puissance** — Maëlle, biologiste dans une ferme aquacole près de Sète, mesure la longueur x (en cm) et la masse y (en g) de cinq bars. Elle propose le modèle $y = k\,x^{p}$. Pour la régression, fais calculer à la calculatrice les listes $\ln x$ et $\ln y$ sans les arrondir.

Longueur x_i (cm)	Masse y_i (g)	u_i = ln x_i	v_i = ln y_i
20	98	…	…
25	185	…	…
30	330	…	…
35	510	…	…
40	775	…	…

(a) On suppose $y = k\,x^{p}$ avec $k > 0$ et $x > 0$. Montre que $\ln y$ est une fonction affine de $\ln x$, puis explique pourquoi $\ln x_i$ et $\ln y_i$ existent ici.

(b) Complète le tableau (u_i et v_i au millième). Donne ensuite la droite de régression de v en u (coefficients au millième) et r au dix-millième.

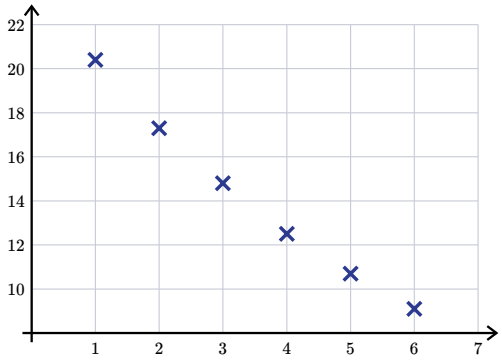
(c) Malo écrit : « $v = pu + b$, donc $y = x^{p} + \mathrm{e}^{b}$ ». Explique son erreur, puis donne k au dix-millième et l'expression de y en fonction de x.

(d) Estime au gramme près la masse d'un bar de 50 cm. Ninon a arrondi p et $\ln k$ au dixième et trouve environ 1 535 g : d'où vient l'écart ?

(e) D'après ce modèle, à partir de quelle longueur (au mm près) un bar dépasse-t-il 1 kg ? Résous l'inéquation avec ln, puis commente la fiabilité.

Ex. 42 ★★★ **Cote d'une citadine d'occasion** — Yanis, vendeur dans un garage de Clermont-Ferrand, relève la cote y (en milliers d'euros) d'un modèle de citadine selon son âge x (en années). Calculatrice autorisée ; pour les régressions, garde les valeurs non arrondies.

Âge x_i (ans)	Cote y_i (k€)	z_i = ln y_i
1	20,4	…
2	17,3	…
3	14,8	…
4	12,5	…
5	10,7	…
6	9,1	…



(a) Donne la droite de régression de y en x (coefficients au centième) et r au millième. Ce r suffit-il à valider l'ajustement affine ? Justifie avec le nuage.

(b) Complète le tableau (z_i au millième). Donne ensuite la droite de régression de z en x (coefficients au dix-millième) et r au millième.

(c) Écris y sous la forme $k \times q^{x}$ (k au centième, q au millième). Déduis-en le taux d'évolution annuel de la cote, en %.

(d) Une citadine de 8 ans est cotée 6,6 k€. Compare les deux ajustements avec cette valeur. Que prédit le modèle affine pour 10 ans ?

(e) Selon le modèle exponentiel, à partir de quel âge (en années entières) la cote passe-t-elle sous 5 000 € ? Résous l'inéquation avec ln.