

Plaquette 2 – Ajustement affine graphique

Statistiques à deux variables — Première générale

Corrections détaillées

Boîte à outils

Ajuster un nuage par une droite, à la main.

- **Coefficient directeur** de la droite passant par $A(x_A; y_A)$ et $B(x_B; y_B)$:

$$m = \frac{y_B - y_A}{x_B - x_A}$$

- **Équation** de cette droite, sous la forme point-pente puis réduite :

$$y = m(x - x_A) + y_A$$

- **Trois ajustements possibles à la main** : la droite joignant le **premier et le dernier** point du nuage ; la droite de **Mayer** (G_1G_2) ; une droite tracée **à la règle** au plus près du nuage, dont on lit ensuite deux points.
- **Contrôle indispensable**. Un bon ajustement passe par le **point moyen** $G(\bar{x}; \bar{y})$, ou très près de lui. On remplace x par $\bar{x}$ dans l'équation et on compare le résultat à $\bar{y}$.
- **Utiliser l'ajustement**. Estimer y connaissant x : on **remplace** x dans l'équation. Estimer x connaissant y : on **résout** l'équation.
- **Résidus**. L'écart entre la valeur observée et la valeur estimée, $y_i - (mx_i + p)$, mesure la qualité de l'ajustement point par point. De petits résidus, de signes alternés, signent un bon ajustement.
- **Interprétation**. Le coefficient directeur m s'exprime dans l'unité de y **par unité de** x . L'ordonnée à l'origine p est la valeur estimée pour $x = 0$: elle n'a de sens que si $x = 0$ est réaliste.
- **Changement de repère**. Quand les abscisses sont des années, on pose $x = 1$ pour la première année : les calculs deviennent plus simples, et il faut penser à **retraduire** le résultat en année.

Correction de l'exercice 1 – Droite joignant deux points du nuage

- a) On applique la formule du coefficient directeur :

$$m = \frac{y_F - y_A}{x_F - x_A} = \frac{31 - 12}{6 - 1} = \frac{19}{5} = 3,8$$

- b) On écrit l'équation sous forme point-pente avec le point A, puis on développe :

$$y = 3,8(x - 1) + 12 = 3,8x - 3,8 + 12 = 3,8x + 8,2$$

- c) On remplace x par 3,5 :

$$y = 3,8 \times 3,5 + 8,2 = 13,3 + 8,2 = 21,5$$

On obtient 21,5, alors que $\bar{y} = 21$: la droite **ne passe pas exactement** par G , mais elle en passe très près (à 0,5 millier d'euros).

Ce qu'il faut en penser. La droite « premier-dernier point » est rapide à construire, mais elle n'utilise que deux points sur six : rien ne garantit qu'elle passe par G . Les ajustements de Mayer et des moindres carrés, eux, y passent **par construction** — c'est leur avantage.

Réponse. a) $m = 3,8$ — b) $y = 3,8x + 8,2$ — c) non exactement : elle donne 21,5 au lieu de

21.

Correction de l'exercice 2 - Lire une droite d'ajustement tracée

a) La droite coupe l'axe des ordonnées en $(0; 3)$, et elle passe par $(4; 11)$ — deux points à coordonnées entières, faciles à lire sur la grille.

b) Le coefficient directeur :

$$m = \frac{11 - 3}{4 - 0} = \frac{8}{4} = 2$$

L'ordonnée à l'origine se lit directement : $p = 3$. L'équation de D est donc

$$y = 2x + 3$$

c) On remplace x par 7 :

$$y = 2 \times 7 + 3 = 17$$

L'estimation pour $x = 7$ est donc 17.

Le contrôle par le point moyen. Les six points donnent $\bar{x} = \frac{21}{6} = 3,5$ et

$\bar{y} = \frac{5+6+9+10+13+14}{6} = \frac{57}{6} = 9,5$. Or $2 \times 3,5 + 3 = 10$: la droite passe à 0,5 de G , ce qui est très correct pour un tracé à la règle ✓

Le conseil pratique. On choisit **toujours** deux points de la droite à coordonnées entières, et si possible éloignés l'un de l'autre : une erreur de lecture d'un demi-carreau se divise alors par la distance entre les deux points.

Réponse. a) $(0; 3)$ et $(4; 11)$ — b) $y = 2x + 3$ — c) environ 17.

Correction de l'exercice 3 - L'ajustement de Mayer

a) Pour les trois premières années :

$$\bar{x}_1 = \frac{1+2+3}{3} = 2 \quad \bar{y}_1 = \frac{12+15+19}{3} = \frac{46}{3} \approx 15,33$$

Pour les trois dernières :

$$\bar{x}_2 = \frac{4+5+6}{3} = 5 \quad \bar{y}_2 = \frac{22+27+31}{3} = \frac{80}{3} \approx 26,67$$

Donc $G_1(2; 15,33)$ et $G_2(5; 26,67)$.

b) Le coefficient directeur :

$$m = \frac{26,67 - 15,33}{5 - 2} = \frac{11,34}{3} = 3,78$$

L'équation s'écrit $y = 3,78(x - 2) + 15,33$, soit

$$y = 3,78x - 7,56 + 15,33 = 3,78x + 7,77$$

c) Pour $x = 3,5$:

$$y = 3,78 \times 3,5 + 7,77 = 13,23 + 7,77 = 21$$

On retrouve exactement $\bar{y} = 21$: la droite de Mayer passe bien par le point moyen ✓

L'avantage sur l'exercice 1. La droite « premier-dernier » donnait 21,5 en $x = 3,5$; celle de Mayer donne 21, la valeur exacte. Elle utilise les **six** points, et non deux : c'est un ajustement plus fiable, tout en restant calculable à la main.

Réponse. a) $G_1(2; 15,33)$ et $G_2(5; 26,67)$ — b) $y = 3,78x + 7,77$ — c) pour $x = 3,5$, on trouve $21 = \bar{y}$ ✓

Correction de l'exercice 4 - Utiliser une droite d'ajustement

a) On remplace x par 8 dans l'équation de la droite :

$$y = 4 \times 8 + 7 = 32 + 7 = 39$$

Le chiffre d'affaires estimé de l'année 8 est donc de 39 **milliers d'euros**, soit 39 000 €.

b) De même pour l'année 10 :

$$y = 4 \times 10 + 7 = 47$$

soit environ 47 000 €.

c) Le coefficient directeur s'interprète toujours « unité de y par unité de x ». Ici, y est en milliers d'euros et x en années :

$$m = 4 \text{ milliers d'euros par an}$$

D'après ce modèle, le chiffre d'affaires augmente donc d'environ 4 000 € **chaque année**.

La différence entre les deux estimations. L'année 8 est proche des données observées (années 1 à 6), l'année 10 l'est beaucoup moins : la seconde estimation est donc plus fragile. C'est la distinction entre interpolation et extrapolation, étudiée dans la plaquette 4.

Réponse. a) environ 39 000 € — b) environ 47 000 € — c) le chiffre d'affaires augmente d'environ 4 000 € par an.

Correction de l'exercice 5 - Calculer des résidus

a) On remplace x par chaque valeur dans l'équation de la droite :

$$y = 3,8x + 8,2$$

- pour $x = 2$: $7,6 + 8,2 = 15,8$;
- pour $x = 4$: $15,2 + 8,2 = 23,4$;
- pour $x = 5$: $19 + 8,2 = 27,2$.

b) Le résidu est la différence entre l'observation et l'estimation :

$$\text{résidu} = y_{\text{observé}} - y_{\text{estimé}}$$

- pour $x = 2$: $15 - 15,8 = -0,8$;
- pour $x = 4$: $22 - 23,4 = -1,4$;
- pour $x = 5$: $27 - 27,2 = -0,2$.

c) Oui, l'ajustement est de bonne qualité. Les trois résidus sont **petits** en valeur absolue (au plus 1,4) devant des chiffres d'affaires de 15 à 27 milliers d'euros, soit une erreur relative inférieure à 7 %. En revanche, ils sont tous **négatifs** ici : la droite surestime systématiquement ces trois années, signe qu'elle passe un peu au-dessus du nuage en son milieu — ce qui est cohérent avec l'exercice 1, où elle donnait 21,5 au lieu de 21 en $x = 3,5$.

Ce qu'on attend d'un bon ajustement. Des résidus petits **et** de signes alternés : si tous ont le même signe sur une partie du nuage, la droite est mal placée, ou bien le modèle affine n'est pas le bon.

Réponse. a) 15,8 ; 23,4 ; 27,2 — b) résidus $-0,8$; $-1,4$; $-0,2$ — c) oui, les écarts sont petits, mais tous négatifs : la droite passe un peu au-dessus.

Correction de l'exercice 6 - Choisir entre deux ajustements

a) Avec $D_1: y = 2x + 3$, les estimations sont 5, 7, 9 et 11. Les résidus (observé moins

estimé) :

$$5 - 5 = 0 \quad ; \quad 6 - 7 = -1 \quad ; \quad 9 - 9 = 0 \quad ; \quad 10 - 11 = -1$$

b) Avec $D_2: y = 1,5x + 4$, les estimations sont 5,5, 7, 8,5 et 10. Les résidus :

$$5 - 5,5 = -0,5 \quad ; \quad 6 - 7 = -1 \quad ; \quad 9 - 8,5 = 0,5 \quad ; \quad 10 - 10 = 0$$

c) On compare les sommes des **carrés** des résidus :

$$D_1: 0 + 1 + 0 + 1 = 2 \quad D_2: 0,25 + 1 + 0,25 + 0 = 1,5$$

La somme est plus petite pour D_2 : c'est donc D_2 qui ajuste **mieux** le nuage, au sens des moindres carrés.

Pourquoi élever au carré. Si l'on additionnait les résidus tels quels, les positifs compenseraient les négatifs : la somme vaudrait -2 pour D_1 et -1 pour D_2 , sans mesurer l'ampleur des écarts. Le carré supprime les signes et pénalise davantage les gros écarts — c'est exactement le critère que la calculatrice minimise pour obtenir la droite de régression.

Réponse. a) résidus $0 ; -1 ; 0 ; -1$ — b) résidus $-0,5 ; -1 ; 0,5 ; 0$ — c) somme des carrés 2 contre 1,5 : D_2 est meilleure.

Correction de l'exercice 7 - Tester si une droite passe par le point moyen

a) Avec $n = 4$:

$$\bar{x} = \frac{2 + 4 + 6 + 8}{4} = \frac{20}{4} = 5 \quad \bar{y} = \frac{7 + 12 + 16 + 21}{4} = \frac{56}{4} = 14$$

Le point moyen est $G(5 ; 14)$.

b) On remplace x par 5 dans la première équation :

$$y = 2,3 \times 5 + 2,4 = 11,5 + 2,4 = 13,9$$

On obtient 13,9 et non 14 : la droite **ne passe pas** par G , elle passe 0,1 en dessous.

c) Avec la seconde équation :

$$y = 2,3 \times 5 + 3 = 11,5 + 3 = 14,5$$

On obtient 14,5 : cette droite ne passe pas non plus par G , elle passe 0,5 au-dessus.

Comment corriger. Pour qu'une droite de coefficient directeur 2,3 passe exactement par $G(5 ; 14)$, son ordonnée à l'origine doit vérifier

$$14 = 2,3 \times 5 + p \Leftrightarrow p = 14 - 11,5 = 2,5$$

La bonne droite est donc $y = 2,3x + 2,5$. C'est la méthode générale : la pente vient du nuage, et G fixe l'ordonnée à l'origine.

Réponse. a) $G(5 ; 14)$ — b) non, elle donne 13,9 — c) non, elle donne 14,5 ; la bonne serait $y = 2,3x + 2,5$.

Correction de l'exercice 8 - Construire une droite passant par G

a) La droite a pour équation $y = 3,5x + p$, et elle doit passer par $G(10 ; 46)$: les coordonnées de G vérifient donc l'équation.

$$46 = 3,5 \times 10 + p \Leftrightarrow 46 = 35 + p \Leftrightarrow p = 11$$

b) L'équation réduite est donc

$$y = 3,5x + 11$$

Contrôle : pour $x = 10$, $y = 35 + 11 = 46$ ✓ La droite passe bien par G .

c) Pour $x = 15$:

$$y = 3,5 \times 15 + 11 = 52,5 + 11 = 63,5$$

L'estimation est donc 63,5.

La logique de la construction. Deux informations suffisent à déterminer une droite : une **direction** (le coefficient directeur, donné par l'allure du nuage) et un **point** (ici G , imposé par la théorie). C'est exactement la démarche de l'ajustement affine.

Réponse. a) $p = 11$ — b) $y = 3,5x + 11$ — c) environ 63,5.

Correction de l'exercice 9 - Interpréter les deux coefficients

a) Le coefficient directeur s'exprime en unité de y **par** unité de x , donc ici en milliers d'abonnés par mois :

$$m = 1,8 \text{ millier d'abonné par mois}$$

Le service gagne donc environ 1 800 abonnés chaque mois, d'après le modèle.

b) L'ordonnée à l'origine est la valeur estimée pour $x = 0$, c'est-à-dire au moment du lancement :

$$p = 12$$

Le service comptait donc 12 000 abonnés au lancement — ce qui est plausible s'il s'agit du basculement d'un ancien service.

c) Un an et demi représente 18 mois, donc $x = 18$:

$$y = 1,8 \times 18 + 12 = 32,4 + 12 = 44,4$$

soit environ 44 400 **abonnés**.

Quand l'ordonnée à l'origine n'a pas de sens. Si x était une taille en centimètres et y une masse, p serait la masse d'un individu de taille nulle : un nombre sans signification, qui ne sert qu'à positionner la droite. On n'interprète p que lorsque $x = 0$ correspond à une situation réelle.

Réponse. a) environ 1 800 nouveaux abonnés par mois — b) 12 000 abonnés au lancement — c) environ 44 400 abonnés.

Correction de l'exercice 10 - Ajuster un nuage décroissant

a) Le coefficient directeur :

$$m = \frac{5 - 20}{14 - 2} = \frac{-15}{12} = -1,25$$

L'équation, écrite à partir du point (2 ; 20) :

$$y = -1,25(x - 2) + 20 = -1,25x + 2,5 + 20 = -1,25x + 22,5$$

b) Les deux moyennes, avec $n = 7$:

$$\bar{x} = \frac{2 + 4 + 6 + 8 + 10 + 12 + 14}{7} = \frac{56}{7} = 8 \quad \bar{y} = \frac{20 + 18 + 15 + 13 + 10 + 8 + 5}{7} = \frac{89}{7} \approx 12,71$$

Pour $x = 8$, la droite donne $y = -10 + 22,5 = 12,5$, à comparer à 12,71 : l'écart est de 0,21 m³, donc la droite passe **très près** de G . L'ajustement est acceptable.

c) Le coefficient directeur vaut $-1,25$ m³ par degré : chaque degré **gagné** à l'extérieur fait **baisser** la consommation d'environ 1,25 m³ par jour. Le signe négatif traduit la corrélation négative déjà observée.

L'erreur d'interprétation à éviter. Dire « la consommation baisse de 1,25 % par degré » : le coefficient directeur n'est pas un pourcentage, c'est une **quantité** exprimée dans l'unité de y . Ici, 1,25 m³ par degré.

Réponse. a) $y = -1,25x + 22,5$ — b) $G(8; 12,71)$ et la droite donne 12,5 : très proche — c) chaque degré supplémentaire fait baisser la consommation d'environ 1,25 m³/jour.

Correction de l'exercice 11 - Estimer x connaissant y

a) Ici on connaît y et on cherche x : il faut **résoudre** l'équation, et non substituer.

$$4x + 7 = 51 \Leftrightarrow 4x = 44 \Leftrightarrow x = 11$$

Le chiffre d'affaires atteint donc 51 milliers d'euros à l'année 11, et le dépasse à partir de l'année 12.

b) De même :

$$4x + 7 = 60 \Leftrightarrow 4x = 53 \Leftrightarrow x = 13,25$$

La valeur 60 est atteinte « au cours » de l'année 13,25, qui n'existe pas comme rang entier.

c) Parce que le rang d'une année est un **entier**. À l'année 13, le modèle donne $4 \times 13 + 7 = 59$ milliers d'euros, donc moins de 60 ; à l'année 14, il donne 63 milliers d'euros. Le seuil est donc dépassé à partir de l'**année 14** : on arrondit vers le haut, car arrondir à 13 donnerait une année où le seuil n'est pas atteint.

Le réflexe de vérification. Après avoir arrondi, on **teste** les deux entiers voisins dans l'équation. C'est le seul moyen de ne pas se tromper d'une année, et c'est ce que le correcteur attend.

Réponse. a) $x = 11$, donc dépassé à partir de l'année 12 — b) $x = 13,25$, donc à partir de l'année 14 — c) le rang est un entier, et $x = 13$ donne $59 < 60$.

Correction de l'exercice 12 - Ajuster en écartant un point aberrant

a) Le point (4 ; 2) est aberrant : les cinq autres valeurs suivent une progression régulière (10, 13, 16, puis 22, 25), alors que celle du mois 4 s'effondre à 2. Une explication plausible : le musée a été fermé une grande partie du mois (travaux, grève, sinistre).

b) Sans ce point, il reste cinq couples. Les moyennes :

$$\bar{x} = \frac{1+2+3+5+6}{5} = \frac{17}{5} = 3,4 \quad \bar{y} = \frac{10+13+16+22+25}{5} = \frac{86}{5} = 17,2$$

Le point moyen est donc $G(3,4; 17,2)$.

c) Le coefficient directeur de la droite joignant (1 ; 10) et (6 ; 25) :

$$m = \frac{25-10}{6-1} = \frac{15}{5} = 3$$

Son équation : $y = 3(x - 1) + 10 = 3x + 7$.

Pour $x = 3,4$:

$$y = 3 \times 3,4 + 7 = 10,2 + 7 = 17,2$$

C'est exactement $\bar{y}$: la droite passe par le point moyen de la série **corrigée** ✓

Ce qu'aurait donné le point aberrant. Avec les six points, $\bar{y} = \frac{88}{6} \approx 14,67$: le point moyen descend de plus de 2,5 centaines de visiteurs, et tout ajustement construit sur lui sous-estimerait systématiquement la fréquentation. D'où l'importance de repérer les points aberrants **avant** d'ajuster.

Réponse. a) le point (4 ; 2), mois de fermeture probable — b) $G(3,4; 17,2)$ — c) $y = 3x + 7$, qui donne exactement 17,2 en $x = 3,4$ ✓

Correction de l'exercice 13 - Changer le repérage des années

a) Avec $n = 5$:

$$\bar{x} = \frac{0 + 1 + 2 + 3 + 4}{5} = \frac{10}{5} = 2 \quad \bar{y} = \frac{140 + 152 + 167 + 178 + 193}{5} = \frac{830}{5} = 166$$

Le point moyen est $G(2; 166)$.

b) Le coefficient directeur de la droite joignant $(0; 140)$ et $(4; 193)$:

$$m = \frac{193 - 140}{4 - 0} = \frac{53}{4} = 13,25$$

L'ordonnée à l'origine se lit directement, puisque le premier point a pour abscisse 0 : $p = 140$. D'où

$$y = 13,25x + 140$$

Contrôle avec G : $13,25 \times 2 + 140 = 26,5 + 140 = 166,5$, à comparer à 166 — l'écart est de 0,5, la droite passe très près de G ✓

c) L'année 2027 correspond à $x = 2027 - 2019 = 8$:

$$y = 13,25 \times 8 + 140 = 106 + 140 = 246$$

La production estimée pour 2027 est donc d'environ 246 unités.

L'intérêt du changement de repère. Poser $x = 0$ pour la première année donne directement l'ordonnée à l'origine, et évite de manipuler des nombres à quatre chiffres. En contrepartie, il faut **retraduire** : $x = 8$ signifie 2027, et non l'année 8.

Réponse. a) $G(2; 166)$ — b) $y = 13,25x + 140$ — c) environ 246 unités en 2027 ($x = 8$).

Correction de l'exercice 14 - Deux séries, deux pentes

a) Le chiffre d'affaires « au départ » correspond à $x = 0$, c'est-à-dire à l'ordonnée à l'origine : 40 pour A, 25 pour B. C'est donc le **magasin A** qui partait le plus haut, avec 40 000 € contre 25 000 €.

b) La vitesse de progression est donnée par le coefficient directeur : 2,5 pour A, 4 pour B. C'est donc le **magasin B** qui progresse le plus vite, de 4 000 € par mois contre 2 500 € par mois.

c) On cherche le rang à partir duquel B dépasse A, donc on résout l'inéquation :

$$4x + 25 > 2,5x + 40 \Leftrightarrow 4x - 2,5x > 40 - 25 \Leftrightarrow 1,5x > 15 \Leftrightarrow x > 10$$

Le magasin B dépasse donc A à partir du **onzième mois**. Au mois 10, les deux sont à égalité : $2,5 \times 10 + 40 = 65$ et $4 \times 10 + 25 = 65$ ✓

Les deux coefficients, deux informations distinctes. L'ordonnée à l'origine dit où l'on part, le coefficient directeur dit à quelle vitesse on avance. Partir plus bas avec une meilleure pente finit toujours par payer — la seule question est : au bout de combien de temps ?

Réponse. a) le magasin A (40 000 € contre 25 000 €) — b) le magasin B (4 000 €/mois) — c) à partir du onzième mois (égalité au mois 10, à 65 000 €).

Correction de l'exercice 15 - Synthèse : ventes de vélos électriques

a) Avec $n = 6$:

$$\bar{x} = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3,5 \quad \bar{y} = \frac{34 + 41 + 52 + 58 + 69 + 76}{6} = \frac{330}{6} = 55$$

Le point moyen est $G(3,5; 55)$.

b) Pour les trois premiers trimestres :

$$\bar{x}_1 = 2 \quad \bar{y}_1 = \frac{34 + 41 + 52}{3} = \frac{127}{3} \approx 42,33$$

Pour les trois derniers :

$$\bar{x}_2 = 5 \quad \bar{y}_2 = \frac{58 + 69 + 76}{3} = \frac{203}{3} \approx 67,67$$

Le coefficient directeur de (G_1G_2) :

$$m = \frac{67,67 - 42,33}{5 - 2} = \frac{25,34}{3} \approx 8,45$$

L'équation : $y = 8,45(x - 2) + 42,33$, soit

$$y = 8,45x - 16,9 + 42,33 = 8,45x + 25,43$$

c) Pour $x = 3,5$:

$$y = 8,45 \times 3,5 + 25,43 = 29,58 + 25,43 = 55,01$$

On retrouve 55 aux arrondis près : la droite passe bien par G ✓

d) Pour le huitième trimestre, $x = 8$:

$$y = 8,45 \times 8 + 25,43 = 67,6 + 25,43 = 93,03 \approx 93 \text{ vélos}$$

e) On résout l'inéquation :

$$8,45x + 25,43 > 100 \Leftrightarrow 8,45x > 74,57 \Leftrightarrow x > 8,82$$

Le seuil serait donc franchi au **neuvième trimestre**. Contrôle :

$$8,45 \times 9 + 25,43 = 76,05 + 25,43 = 101,48 > 100 \text{ ✓ alors que le huitième donne 93.}$$

Le bilan de la plaquette. Un ajustement affine se construit en trois gestes — une pente, un point, une équation — et se contrôle toujours en vérifiant le passage par G . Il sert ensuite dans les deux sens : estimer y en substituant, estimer x en résolvant.

Réponse. a) $G(3,5; 55)$ — b) $G_1(2; 42,33)$, $G_2(5; 67,67)$ et $y = 8,45x + 25,43$ — c) elle donne 55 en $x = 3,5$ ✓ — d) environ 93 vélos — e) au neuvième trimestre.