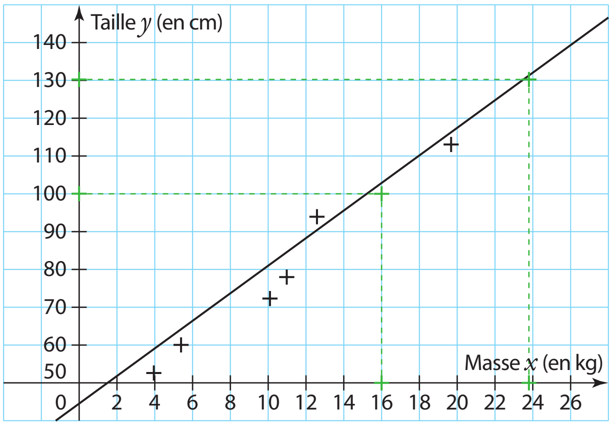
Cours

4 Interpolations et extrapolations

Lorsqu'on a pu réaliser un ajustement (affine ou exponentiel) satisfaisant pour le nuage de points d'une série statistique à deux variables x et y, alors on peut considérer, si cela a du sens dans la situation étudiée, que les quantités x et y ne sont pas complètement indépendantes l'une de l'autre et ont un lien qui peut être modélisé par une fonction f (affine ou exponentielle), qui correspond à l'ajustement trouvé, et telle que $y = f(x)$.
Cela permet d'estimer (graphiquement ou algébriquement) une valeur de la quantité y qui pourrait correspondre à une valeur théorique de la quantité x, non relevée initialement dans l'étude statistique. Et réciproquement on peut aussi estimer une valeur de x qui pourrait correspondre à une valeur donnée de y. Dans le cas où les valeurs choisies restent dans le « domaine d'étude » défini par l'échantillon, on dit que l'on réalise une **interpolation**. Dans le cas contraire, il s'agit alors d'une **extrapolation**.

Exemples

① Le nuage de points ci-contre représente la masse x (en kg) et la taille y (en cm) d'un enfant, relevés par son médecin traitant lors des 6 dernières visites médicales.
La forme du nuage a permis de réaliser un ajustement affine.
La droite d'ajustement obtenue permet de modéliser de façon approchée le lien entre les deux variables.
Ainsi on peut estimer graphiquement que si l'enfant avait été voir son médecin, entre la 5^{e} et la 6^{e} visite, au moment où il mesurait environ 100 cm, la pesée aurait probablement donné une valeur autour de 16 kg.

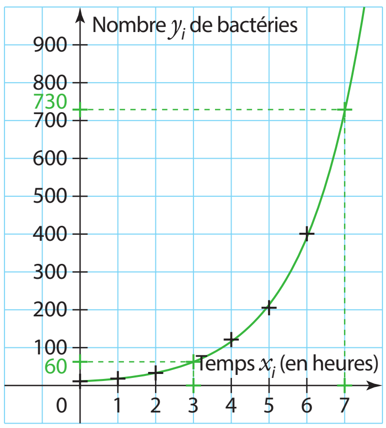


Les deux valeurs $y \approx 100$ et $x \approx 16$ font partie du domaine d'étude qui va de $y_{\text{min}} \approx 53$ à $y_{\text{max}} \approx 113$ d'une part, et de $x_{\text{min}} \approx 4$ à $x_{\text{max}} \approx 20$.
Alors l'estimation que nous venons de faire s'appelle une **interpolation**. Nous avons interpolé la valeur 16 kg à partir de la valeur 100 cm (on aurait tout aussi bien pu interpoler la valeur 100 cm à partir de la valeur 16 kg).
Si l'on souhaite faire une estimation en dehors du domaine d'étude (par exemple pour $y \approx 130$ on estime $x \approx 24$) on réalise une **extrapolation**.
Nous pouvons extrapoler (imaginer) la masse d'environ 24 kg de l'enfant lorsqu'il mesurera 130 cm.

② Dans un laboratoire, on étudie l'évolution d'une population de bactéries anaérobies.
Le nombre y_i de bactéries est relevé à différents instants x_i en heures.

Temps x_i (en heures)	0	1	2	4	5	6
Nombres y_i de bactéries	10	18	33	121	205	400

Après avoir représenté le nuage de points $(x_i\,;\,y_i)$ et constaté qu'il présentait une forme « ajustable » par une courbe « exponentielle », on a réalisé un ajustement exponentiel avec la méthode de la page précédente et on a tracé la courbe de la fonction obtenue. Cette courbe permet de modéliser de façon approchée le lien entre les deux variables. Ainsi, graphiquement, on peut estimer le nombre de bactéries qu'il y avait au bout de 3 heures : soit environ 60 (interpolation).
En supposant que l'évolution observée se prolongera d'heures en heures on peut estimer le nombre de bactéries qu'il y aura au bout de 7 heures : environ 730 (extrapolation).



Exercices résolus

Méthode 4 Effectuer des interpolations et des extrapolations

Énoncé

Pour l'achat d'une grosse quantité de ballons de handball, un fabricant propose un tarif dégressif selon la quantité d'articles commandés. Le tableau ci-contre présente un relevé des prix proposés.

Nombre de ballons x_i	100	500	1000	2000
Prix unitaire y_i (en €)	19,90	19	17,90	15,50

1. **Déterminer l'équation de la droite de régression de y en x et le coefficient de corrélation r.**
2. **D'après le coefficient r, l'ajustement affine de cette série statistique semble-t-il un bon choix ?**
3. **Déterminer le prix unitaire que devrait proposer le fabricant pour un achat de 1500 ballons.**
4. **Quelle quantité de ballons faudrait-il acheter pour obtenir un prix unitaire de 12 € ?**

Solution

1. et **2.** Avec la calculatrice on obtient : $y = -0{,}0023x + 20{,}16$. Et le coefficient de corrélation $r \approx -0{,}99978$. **1** Ce coefficient de corrélation est très proche de – 1, ce qui signifie que les points sont presque tous alignés et que l'ajustement affine de cette série statistique est donc particulièrement recommandé et très précis.

3. On peut effectuer une interpolation en utilisant l'équation de la droite d'ajustement. On obtient alors $y = -0{,}0023 \times 1\,500 + 20{,}16 = 16{,}71$.
Ainsi pour un achat de 1500 ballons, le prix unitaire que devrait proposer le fabricant pourrait être d'environ 16,70 €. **2**

4. Dans cette question, on sort du domaine d'étude (entre 15,50 € et 19,90 € d'après le tableau), on va donc réaliser une extrapolation, toujours à partir de l'équation de la droite d'ajustement. On a alors :

$$12 = -0{,}0023x + 20{,}16 \Leftrightarrow x = \frac{12 - 20{,}16}{-0{,}0023} \approx 3548\,.$$

Pour obtenir un prix unitaire de 12 €, il faudrait probablement acheter une quantité de ballons proche de 3550. **2**

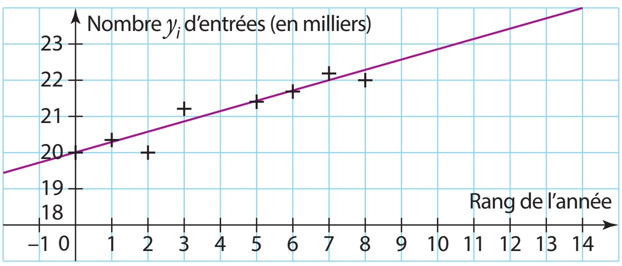
Conseils & Méthodes

1 Plus r est proche de -1, plus sa valeur absolue est proche de 1 : c'est l'indicateur d'une très bonne corrélation linéaire entre les deux variables étudiées.

2 Les résultats d'interpolation et d'extrapolation ne sont que des valeurs indicatives qui découlent de beaucoup d'approximations et d'ajustements successifs, donc ils doivent toujours être donnés avec beaucoup de précautions (utiliser le subjonctif et de termes comme « environ », « probablement »…)

À vous de jouer !

7 Le nuage de points $(x_i\,;y_i)$ suivant donne le nombre total d'entrées annuelle y_i (en milliers) d'un cinéma municipal selon l'année x_i depuis 2010 (année de rang 0).
Ce nuage a été ajusté par la droite d représentée en violet.
1. Il manque les données de 2014. Estimer graphiquement le nombre d'entrées que le cinéma a totalisé cette année-là.
2. Si la tendance observée se poursuit encore pendant plusieurs années, en quelle année le cinéma peut-il espérer atteindre les 24 000 spectateurs par an ?



8 Dans une école primaire, chaque année depuis 2014 (année de rang 0), on a relevé le nombre de livres empruntés à la bibliothèque de l'école.

Rang de l'année x_i	0	1	2	3	4	5
Nombre de livres y_i	201	187	191	162	163	134

1. Déterminer l'équation de la droite de régression de y en x et le coefficient de corrélation r.
2. D'après le coefficient r, l'ajustement affine de cette série statistique semble-t-il un bon choix ?
3. Peut-on estimer le nombre de livres qui seront empruntés à la bibliothèque en 2021 ?
4. Si la tendance se poursuit selon cet ajustement affine pendant les prochaines années, en quelle année pourrait-on avoir moins de 65 livres empruntés ?

↪ Exercices 30 à 31 p. 238

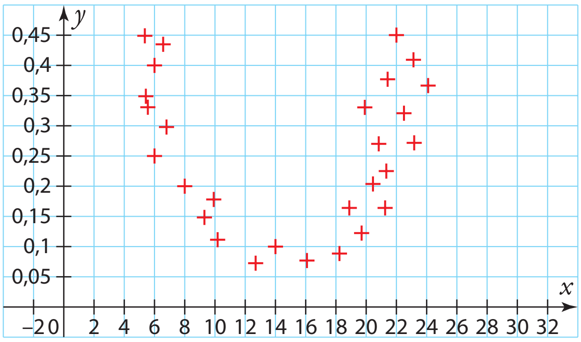
Exercices d'application

Nuage de points, ajustement

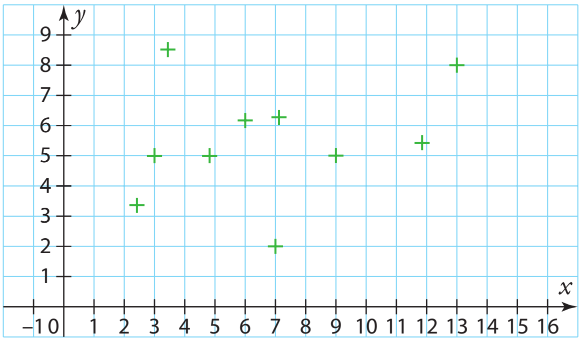
Méthode 1 p. 225

18 Parmi les nuages ci-dessous, lesquelles semblent « ajustables » par une fonction reconnaissable. Citer le nom de la forme du nuage et le nom de la fonction correspondante.

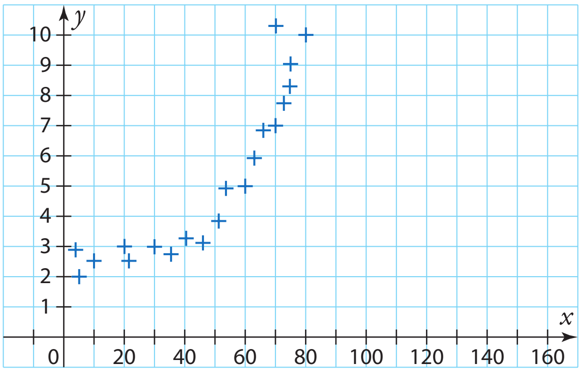
①



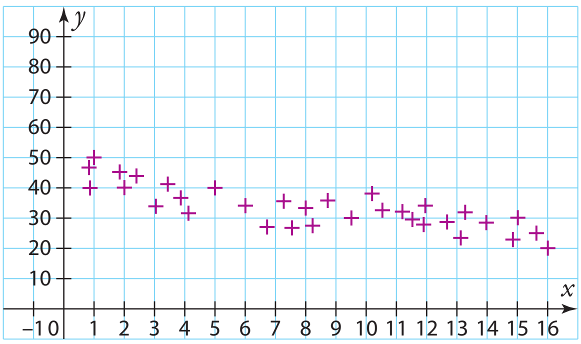
②



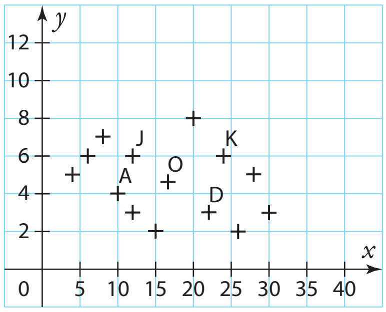
③



④



19 **1.** Dans le nuage de points suivant, parmi les points A, J, O, D, K, quel point semble être le point moyen du nuage ?



2. Peut-on trouver une corrélation entre les deux variables x et y de la série statistique double associée à ce nuage de points ?

20 Les températures, en degrés Celsius, de deux villes A et B ont été relevées à 7h00 tous les matins pendant une semaine.
Les résultats figurent dans le tableau ci-dessous.

Ville A : températures x_i (en °C)	7	2	– 1	– 3	0	3	6
Ville B : températures y_i (en °C)	4,5	0	– 3	– 4	– 1	1,5	3

1. Dans un repère orthonormé, représenter le nuage de points $M_i(x_i ; y_i)$.
2. Déterminer les coordonnées du point moyen G du nuage. Le placer dans le repère.
3. Ce nuage de points présente-t-il une forme caractéristique d'une fonction de référence ?
Si oui, tracer « au jugé » la courbe de la fonction qui passe au plus près des points du nuage et qui passe par G.

21 Le tableau ci-dessous représente une série statistique double (x , y).

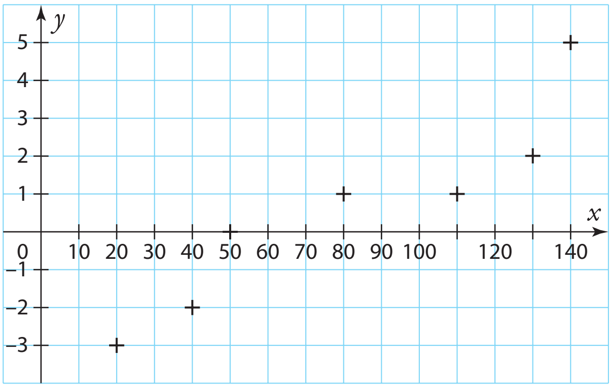
x_i	5	10	20	25	40	55	70
y_i	900	600	450	400	300	275	250

1. Dans un repère orthogonal, représenter le nuage de points $M(x_i ; y_i)$.
2. Déterminer les coordonnées du point moyen G du nuage. Le placer dans le repère.
3. Ce nuage de points présente-t-il une forme caractéristique d'une fonction de référence ? Si oui, tracer « au jugé » la courbe de la fonction qui passe au plus près des points du nuage et qui passe par G.

Droite de régression, coefficient de corrélation linéaire

Méthode 2 p. 227

22 On considère le nuage de points suivant.



1. Un ajustement affine de ce nuage est-il pertinent ?
2. Déterminer, à l'aide de la calculatrice, l'équation réduite de la droite des moindres carrés et le coefficient de corrélation linéaire.

23 On considère la série statistique à deux variables (x, y).

Valeurs x_i	–11	–3	2	0	–5
Valeurs y_i	1100	1000	900	950	1050

1. Dans un repère orthogonal, représenter le nuage de points $(x_i ; y_i)$ et vérifier qu'il peut être ajusté par une droite.
2. Déterminer, à l'aide de la calculatrice, l'équation réduite de la droite d des moindres carrés et le coefficient de corrélation linéaire.
3. Représenter la droite d dans le repère.

24 Sous des conditions de température et de volume constants, on étudie la pression et la quantité de matière d'un gaz. Les résultats sont présentés dans ce tableau. Chimie

Nombre de moles x_i	0	10	20	30
Pression y_i (en kPa)	0	46	98	145

1. Dans un repère orthonormé, représenter le nuage de points $(x_i ; y_i)$ et vérifier qu'il peut être ajusté par une droite.
2. Déterminer, à l'aide de la calculatrice, l'équation réduite de la droite de régression d de y en x et le coefficient de corrélation linéaire.
3. Représenter la droite d dans le repère.

25 Les dépenses en communication d'une entreprise chaque année, depuis 2014, sont données dans ce tableau.

Années	2014	2015	2016	2017	2018	2019
Rang x_i	0	1	2	3	4	5
Dépenses y_i (en milliers d'euros)	11,5	11,2	10,7	10	9,9	9,5

1. Déterminer les coordonnées du point moyen du nuage.
2. Déterminer, l'équation de la droite d des moindres carrés et le coefficient de corrélation linéaire.
3. La droite d passe-t-elle par le point moyen du nuage ?

Ajustement exponentiel

Méthode 3 p. 229

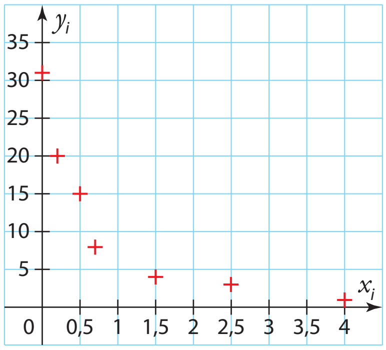
26 On injecte un médicament dans le sang d'une patiente en lui faisant une piqûre en intraveineuse.

La concentration y_i de ce médicament dans le sang, en microgrammes par millilitres est relevé à différents instants x_i en heures.

Temps x_i (en heures)	0	1	2	4	6	10
Concentration y_i (en µg · mL)	90	65	45	23	12	4

1. Représenter le nuage de points $(x_i ; y_i)$ sur la calculatrice ou un tableur et vérifier que sa forme peut être ajustée par une courbe de fonction exponentielle décroissante.
2. On pose $y_i' = \ln(y_i)$ pour tout entier i de 1 à 6.
Calculer les valeurs y_i'.
3. Représenter sur la calculatrice le nuage de points $(x_i ; y_i')$.
5. Déterminer l'équation de la droite de régression de y' en x.
6. En déduire une expression de y en fonction de x.

27 On considère le nuage de points $M_i(x_i ; y_i)$ suivant dont la forme suggère un ajustement exponentiel.



1. On pose, pour tout entier i de 1 à 7, $y_i' = \ln(y_i)$.
Recopier et compléter le tableau suivant.

x_i	0	0,2	0,5	0,7	1,5	2,5	4
y_i							
y_i'							

2. Déterminer l'équation de la droite de régression de y' en x et le coefficient de corrélation linéaire de la série statistique (x, y').
3. En déduire une expression de y en fonction de x.

Exercices d'application

28 On considère la série statistique à deux variables ci-contre.

Valeurs x_i	5	9	12	14
Valeurs y_i	0,02	0,07	0,32	1,20

1. On pose $y_i' = \ln(y_i)$ pour tout entier i de 1 à 4. Calculer les valeurs y_i'.
2. Représenter le nuage de points $(x_i\,;y_i')$ dans un repère et vérifier que sa forme peut être ajustée par une droite.
3. Déterminer l'équation de la droite de régression de y' en x.
4. En déduire une expression de y en fonction de x.

Interpolation, extrapolation

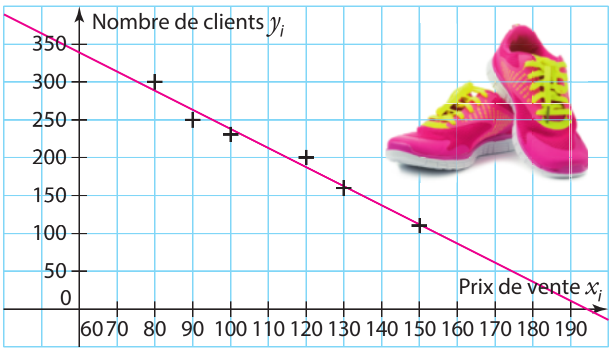
Méthode 4 p. 231

29 Dans une entreprise on étudie conjointement l'évolution du salaire moyen et du chiffre d'affaires. Les données obtenues pendant plusieurs années fournissent une série statistique à deux variable x (montant du salaire moyen en euros) et y (montant du chiffre d'affaires en milliers d'euros) dont le nuage de points $M_i(x_i\,;y_i)$ a pu être ajusté à l'aide de la droite d'équation $y = 5{,}3x - 5\,900$.

Les affirmations suivantes sont-elles vraies ou fausses ?

	V	F
1. En 2003, si le salaire moyen dans l'entreprise était égal à 1 500 €, alors on peut estimer que le chiffre d'affaires devait être d'environ 2,05 millions d'euros.	☐	☐
2. En 2015, l'entreprise a réalisé 10 millions d'euros de chiffre d'affaires. On peut donc estimer que le salaire moyen dans l'entreprise devait être environ égal à 1 115 euros.	☐	☐

30 À l'occasion de la sortie d'un nouveau modèle de baskets, une grande enseigne de distribution réalise une enquête. Le nuage de points $(x_i\,;y_i)$ ci-contre présente les résultats : en ordonnées, le nombre y_i de personnes prêtes à acheter les baskets selon le prix x_i proposé par l'enseigne en abscisses. On considère que la droite d (représentée en rose sur le graphique) est un bon ajustement du nuage.



1. Donner, par lecture graphique, une estimation du nombre de personnes interrogées qui seraient prêtes à acheter les nouvelles baskets :
a) au prix de 110 €. **b)** au prix de 70 €.
2. Estimer le prix à partir duquel aucune des personnes interrogées ne serait prête à acheter les baskets.

31 Dans le tableau suivant est donné le revenu y_i d'une entreprise qui fabrique du matériel de haute technologie chaque année de rang x_i depuis 2009 (année de rang 0).

Rang de l'année x_i	0	1	2	3	4	5
Revenu y_i (en millions d'euros)	10,7	13,3	16,7	20,9	26,1	32,7

Le nuage de points associé à cette série statistique présente une forme « exponentielle croissante ».
Un ajustement exponentiel a été réalisé et a permis d'obtenir la relation $y = 10{,}67e^{0{,}224x}$.
1. Donner une estimation du revenu de l'entreprise en 2016.
2. En quelle année le revenu pourrait-t-il dépasser 100 millions d'euros ?

32 Dans le service de soins d'un hôpital, on a lancé en 2012, un plan de luttes contre les maladies nosocomiales.
On a relevé chaque année le pourcentage de malades ayant contracté une telle maladie.

Année	2012	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	0	1	2	3	4	5	6
Pourcentage de malades y_i	6,9	6,4	6,2	5,1	4,8	3,9	4

1. À la calculatrice on obtient l'équation de la droite d d'ajustement du nuage $(x_i\,;y_i)$ par la méthode des moindres carrés :

$$y = -0{,}54x + 6{,}95$$

ainsi que le coefficient de corrélation linéaire r de la série statistique $(x\,,y)$: $r \approx -0{,}978$.
Peut-on en déduire que la droite d est un bon modèle d'ajustement du nuage de points ?
2. On considère que la tendance observée pendant ces sept années se poursuit encore quelques années plus tard.
a) Estimer le pourcentage de malades ayant contracté une maladie nosocomiale dans cet hôpital en 2019.
b) En quelle année le pourcentage de malades ayant contracté une maladie nosocomiale dans cet hôpital pourrait-il être inférieur à 1,5 % ?

33 L'évolution annuelle, entre 2009 et 2014, du prix x_i (en euros) d'un paquet de cigarettes et des ventes y_i (en milliards d'unités) de la marque la plus vendue est donnée dans le tableau suivant.

Prix x_i	5,35	5,65	6	6,3	6,7	7
Ventes y_i	55	54,8	54,1	51,5	47,5	45

1. Déterminer l'équation de la droite des moindres carrés du nuage de points $(x_i\,;y_i)$.
2. En considérant que cette équation est un bon modèle de corrélation des variables x_i et y_i, déterminer les ventes de cigarettes si le prix du paquet atteignait 10 €.

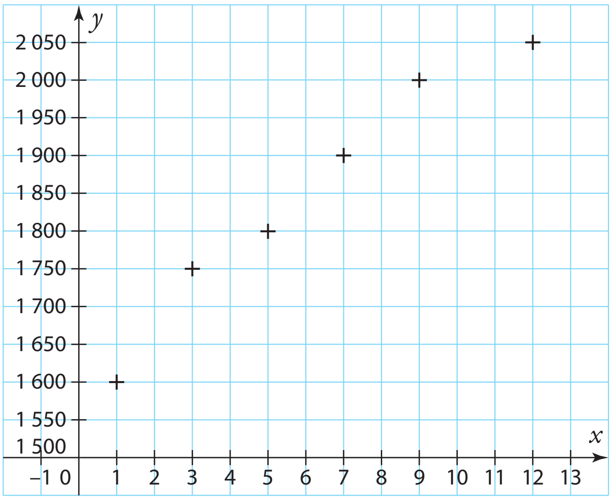
Ajustement affine d'un nuage de points

34 Le tableau ci-contre représente une série statistique double $(x ; y)$.

x_i	0	1	2	3	4
y_i	7	5	4	2	2

1. Dans un repère orthonormé, représenter le nuage de points $M(x_i ; y_i)$.
2. Déterminer les coordonnées du point moyen G du nuage. Le placer dans le repère.
3. Soit d la droite passant par G et de coefficient directeur $-\frac{1}{2}$.
a) Déterminer l'équation réduite de la droite d.
b) Représenter d dans le repère.
c) d est-elle un bon ajustement du nuage ?

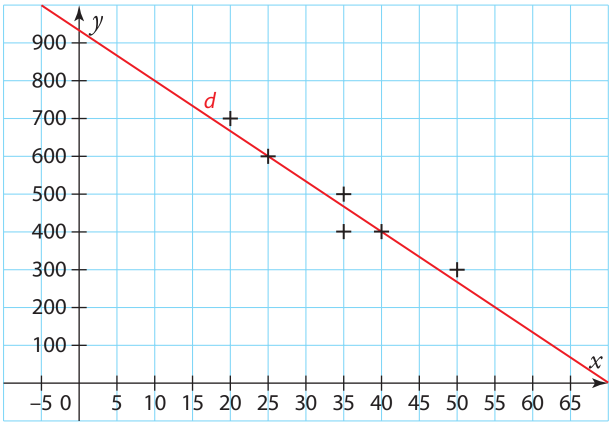
35 On considère la série statistique double $(x ; y)$ donnée par son nuage de points $(x_i ; y_i)$ suivant.



On admet qu'on peut réaliser un ajustement affine de ce nuage de points.

Les affirmations suivantes sont-elles vraies ou fausses ?	**V**	**F**
1. Le point moyen G du nuage a pour coordonnées $\left(\frac{37}{6} ; 1850\right)$	☐	☐
2. L'équation de la droite d'ajustement affine d obtenue par la méthode des moindres carrés et dont les coefficients ont été arrondis à l'unité près, est $y = 41x + 1598$	☐	☐
3. Le coefficient de corrélation linéaire r_1 de la série (x , y) est très proche de 1.	☐	☐
4. Si on retire du nuage de points le point (1 ;1 600) correspondant aux valeurs $x_1 = 1$ et $y_1 = 1\,600$ de la série statistique double (x , y) alors le coefficient de corrélation linéaire r_2 de la nouvelle série statistique obtenue est inférieur au coefficient de corrélation linéaire r_1.	☐	☐
5. La droite d ne passe pas par le point moyen G.	☐	☐

36 Le nuage de points suivant a été ajusté par la droite d d'équation $y = 13{,}5x + 937{,}5$.



1. Déterminer les coordonnées du point moyen.
2. La droite d semble passer par deux points du nuage. Est-ce vraiment le cas ?
3. La droite d choisie pour ajuster le nuage passe-t-elle par le point moyen G ?

37 Thème 9
Dans un hôpital, on a relevé la tension artérielle x_i (en mm de mercure) et l'âge y_i (en années) de 16 patients.
Les résultats sont présentés dans le tableau suivant.

Âge x_i (en années)	36	40	43	45	48	49	50	54
Tension y_i (en mm de mercure)	120	108	112	128	110	122	121	128
Âge x_i (en années)	57	57	58	59	61	65	66	67
Tension y_i (en mm de mercure)	140	130	145	131	136	140	140	134

1. Représenter le nuage de points $M_i(x_i ; y_i)$ (pour i entier de 1 à 16) associé à cette série statistique dans un repère orthonormé.
(On commencera à graduer l'axe des abscisses à 30 et l'axe des ordonnées à 100).
2. On décide de réaliser un ajustement affine par la méthode de Mayer.
On considère deux sous-nuages : celui des huit personnes les plus jeunes (points M_1 à M_8) et celui des huit personnes les plus âgées (points M_9 à M_{16}).
a) Déterminer le point moyen G_1 du premier sous-nuage $M_i(x_i ; y_i)$ pour i entier de 1 à 8.
b) Déterminer le point moyen G_2 du deuxième sous-nuage $M_i(x_i ; y_i)$ pour i entier de 9 à 16.
c) Placer les points G_1 et G_2 dans le repère et tracer la droite (G_1G_2).
3. Déterminer l'équation réduite de la droite (G_1G_2), dite « droite de Mayer ».

Exercices d'entraînement

Droite des moindres carrées et corrélation linéaire

Méthode 5 p. 232

38 On a relevé le taux de chômage des jeunes (16 à 24 ans) en France, entre 2010 et 2015.

Années	2010	2011	2012	2013	2014	2015
Rang x_i de l'année	0	1	2	3	4	5
Taux y_i de chômage (en %)	7,8	8,1	8	8,7	9	9,2

1. Calculer la variance var(x) de la variable x (rang de l'année) et en déduire l'écart-type σ_x.
2. Calculer la variance var(y) de la variable y (taux de chômage) et en déduire l'écart-type σ_y.
3. Calculer la covariance σ_{xy} des variables x et y.
4. En déduire l'équation de la droite de régression de y en x, et la valeur du coefficient de corrélation linéaire de la série statistique $(x\,,y)$.

39 Thème 9

Le tableau ① suivant donne les températures maximales x prévues par un site de météorologie et les températures y maximales réelles observées, pour chaque jour de la semaine du lundi 9 au dimanche 15 mars. Le tableau ② donne les températures prévues x' par le site météo et les températures y' réelles observées le lundi 9 mars, toutes les quatre heures. Toutes les valeurs sont données en degré Celsius.

Tableau ①

Jours de la semaine	L	M	Me	J	V	S	D
Températures prévues x_i (en °C)	12	15	16	14	12	11	8
Températures réelles y_i (en °C)	13	14	14	10	13	7	7

Tableau ②

Heures	2h	6h	10h	14h	18h	22h
Températures prévues x_i' (en °C)	2	3	6	12	12	9
Températures réelles y_i' (en ° C)	2,5	3	7	13	12,5	8

1. Sur la calculatrice, représenter les nuages de points associés aux séries statistiques (x,y) et $(x'\,,y')$.
2. Pour chaque nuage, déterminer l'équation de la droite d'ajustement par la méthode des moindres carrés, et représenter-la.
3. Déterminer le coefficient de corrélation linéaire r_1 de la série statistique (x,y), puis celui, noté r_2 de la série statistique $(x'\,,y')$.
4. On considère l'affirmation : « Les prévisions météo à un jour sont meilleures que les prévisions à une semaine ». Vos résultats confirment-ils cette affirmation.
Pourquoi ?

40 Une série statistique à deux variables $(x\,,y)$ est telle que cov$(x\,,y) \approx 75$, var$(x) \approx 124$ et $var(y) \approx 236$. Est-il pertinent de chercher une corrélation linéaire entre les variables x et y ? Si oui, quel serait le coefficient directeur de la droite de régression de y en x ?

41 Une série statistique à deux variables (x,y) est telle que cov$(x\,,y) \approx -105$, var$(x) \approx 112$ et var$(y) \approx 109$. Est-il envisageable d'ajuster le nuage de points associé à cette série statistique par une droite ? Si oui, quel est le coefficient directeur de la droite obtenue par la méthode des moindres carrés ?

Ajustements, extrapolation et interpolation

42 En France, pour une femme, l'âge moyen du premier mariage était de
- 34 ans en 2012,
- 34,9 ans en 2014,
- 35,4 ans en 2016,
- 35,9 ans en 2018.

1. Dresser un tableau avec ces données. (Pour les années, on pourra au choix les utiliser telles quelles ou utiliser leur rang en prenant l'année 2012 comme rang 0).
2. Réaliser un ajustement affine de ces données.
(Les coefficients seront arrondis au dixième).
3. À l'aide de cet ajustement, donner une estimation, arrondie au dixième, de ce que pourrait être l'âge moyen du premier mariage pour une femme en France en 2022 ?
4. Si cette « tendance » devait se prolonger assez longtemps à l'avenir, en quelle année l'âge moyen du premier mariage pour une femme pourrait atteindre 40 ans ?

43 TICE

Dans le tableau suivant sont présentées les ventes annuelles de smartphones d'un fabricant depuis 2007. Il manque les données des années 2009 et 2014.

	A	B	C
1	Années	Rang x_i de l'année	Nombre y_i de smartphones
2	2007	0	1,5
3	2008	1	5
4	2009	2	
5	2010	3	20
6	2011	4	55
7	2012	5	85
8	2013	6	114
9	2014	7	
10	2015	8	256

1. Représenter sur le tableur le nuage de points $M_i(x_i\,;y_i)$ de cette série statistique et vérifier qu'il peut être ajusté par la courbe $\mathscr{C}$ d'une fonction exponentielle.
2. Démontrer que la courbe $\mathscr{C}$ qui ajuste « au mieux » le nuage a pour équation $y = 2{,}5e^{0{,}641x}$.
3. Si on considère que l'évolution des ventes a été la même depuis 2007, donner une estimation du nombre de smartphones vendus en 2009 et 2014.

44 Thème 9

Les tableaux suivants donnent l'évolution de la population allemande y et de la population turque z, en millions d'habitants, pendant les années 2000 de rang x allant de 0 à 9.

Année	2000	2001	2002	2003	2004
Rang de l'année x_i	0	1	2	3	4
Population allemande y_i	82,2	82,3	82,4	82,5	82,5
Population turque z_i	66,9	64,6	65,6	66,4	67,2
Année	2005	2006	2007	2008	2009
Rang de l'année x_i	5	6	7	8	9
Population allemande y_i	82,5	82,4	82,3	82 ,2	82
Population turque z_i	68	68,9	69,7	70,6	71,5

1. Représenter dans le même repère (éventuellement sur un tableur ou la calculatrice) le nuage de points $M_i(x_i\,;y_i)$ associé à la série statistique à deux variables $(x\,,y)$ et le nuage de points $P_i(x_i\,;z_i)$ associé à la série statistique à deux variables $(x\,,z)$.

2. On réalise un ajustement affine de chacun de ces nuages.

a) Déterminer l'équation de la droite de régression de y en x.

b) Déterminer l'équation de la droite de régression de z en x.

3. Si on suppose que l'évolution observée pour chacune de ces séries se poursuit encore les années suivantes, en quelle année la population de la Turquie pourrait dépasser la population de l'Allemagne ?

Autres ajustements

Méthode 6 p. 233

45 On a relevé, dans le tableau suivant, la masse y_i (en kg) d'un nourrisson selon son âge x_i (en mois).

Age x_i (en mois)	0	1	2	3	4	5	7	9	12
Masse y_i (en kg)	3,01	5,04	6,22	7,06	7,71	8,25	9,09	9,61	10,12

1. Représenter le nuage de points $(x_i\,;y_i)$ sur un tableur ou une calculatrice et vérifier qu'un ajustement affine de ce nuage n'est pas pertinent.

2. On pose, pour tout entier i compris entre 1 et 9, $y_i' = e^{y}$; calculer les valeurs y_i'.

3. Représenter le nuage de points $(x_i\,;y_i')$.
Vérifier qu'il peut être ajusté par une droite.

4. Déterminer la droite de régression de y' en x.
En déduire une expression de y en fonction de x.

5. En admettant que l'évolution du poids du nourrisson va continuer selon la même tendance pendant encore quelques mois, donner une estimation du poids du bébé lorsqu'il aura atteint l'âge de 18 mois.

46 La consommation de carburant et la vitesse d'un véhicule sont bien évidemment liés par une relation de causalité.
Le tableau suivant présente pour un type de poids-lourd, quelques valeurs y_i de consommation de carburant en litres pour 100 km selon certaines valeurs x_i de la vitesse en km·h^{-1}.
L'objectif de cet exercice est de déterminer le lien de corrélation entre ces deux variables.

Vitesse x_i (en km·h^{-1})	30	40	50	60	70	80	90	100	110
Consommation y_i (en L pour 100 km)	43	37	34	31,8	31	32	34,5	38	42

1. Représenter le nuage de points $M_i(x_i\,;y_i)$ (pour i entier de 1 à 9) dans un repère orthonormé.

2. Vérifier que ce nuage de points peut être ajusté par une parabole ayant un sommet proche du point $M_5(70\,;31)$.

3. Soit f la fonction définie sur $]0\,;120[$ par :

$$f(x) = a(x-70)^2 + 31$$

où a est un réel strictement positif.

a) Expliquer pourquoi la courbe représentative $\mathscr{C}_f$ de la fonction f pourrait être un bon modèle d'ajustement au nuage.

b) En admettant que la courbe $\mathscr{C}_f$ passe par le point $M_5(60\,;31{,}8)$, démontrer que le réel a est égal à $\frac{1}{125}$.

4. Si on prend la courbe $\mathscr{C}_f$ pour modèle d'ajustement du nuage, quelle consommation peut-on prévoir sur ce type de poids-lourd lorsqu'on roule à 20 km·h^{-1} ?

47 Thème 9

Une société vend des vélos électriques.
Ses bénéfices annuels (en milliers d'euros) ont évolué comme indiqué dans le tableau suivant.

Année	2012	2013	2014	2015	2016	2017
Rang de l'année x_i	0	1	2	3	4	5
Bénéfice y_i (en k€)	64	75	100	113	125	127

1. a) Construire le nuage de points $(x_i\,;y_i)$ associé à la série statistique dans un repère orthonormé.
On prendra comme unités :
- 2 cm pour une unité sur l'axe des abscisses,
- 1 cm pour 10 unités sur l'axe des ordonnées.

b) Donner les coordonnées du point moyen G du nuage et le placer dans le repère.

2. En observant le nuage de point, on envisage de l'ajuster par une courbe $\mathscr{C}_f$ représentative de la fonction f définie sur $[0\,;6]$ par :

$$f(x) = -2x^2 + 23x + 63.$$

a) Étudier les variations de la fonction f sur l'intervalle $[0\,;6]$.

b) Tracer la courbe $\mathscr{C}_f$ dans le repère de la question **1**.

c) Quelle prévision peut-on faire pour le bénéfice en 2005 avec ce modèle d'ajustement ?

D'après bac