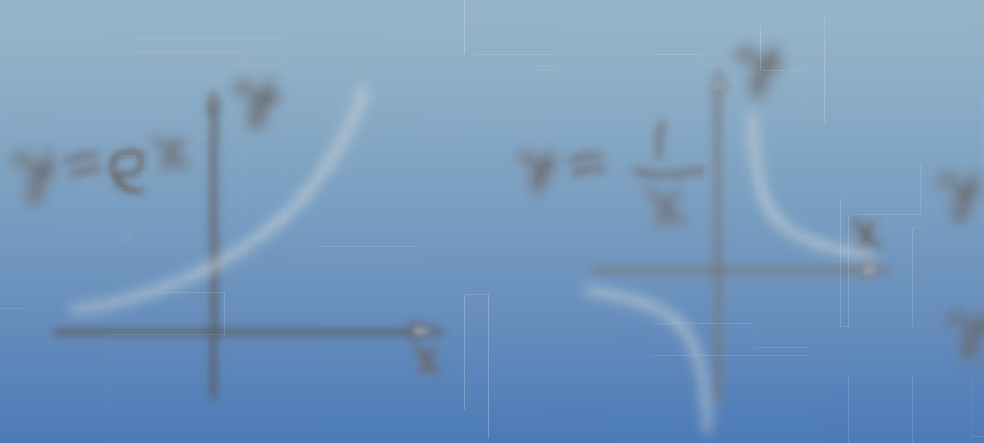


CHAPITRE 6

MODÉLISATION et RÉGRESSION

Appliquer la méthode de la droite de régression pour modéliser des données expérimentales à pas variable ou constant.

Les composantes particulières de l'élément de compétence visées par le présent chapitre sont:

- la reconnaissance du lien entre deux variables grâce à l'examen de données observées;
- la modélisation de données expérimentales à pas constant ou à pas variable;
- l'utilisation de papier semi-logarithmique ou logarithmique pour la détermination du lien entre des données expérimentales;
- la transformations de données visant à traduire en un lien affine une relation puissance, une relation exponentielle ou une relation logarithmique.

OBJECTIFS

6.1 Utiliser la droite de régression pour construire un modèle affine représentant des données numériques.

6.2 Utiliser un papier semi-log ou log-log pour détecter le type de lien entre les valeurs expérimentales de deux variables.

6.3 Utiliser la méthode des moindres carrés et les logarithmes pour modéliser des données expérimentales.

6.1 MODÉLISATION AFFINE

En modélisation affine, on utilise l'équation d'une droite comme modèle du lien entre les variables. Pour construire le modèle affine, il faut connaître les caractéristiques de la droite que sont la pente et l'ordonnée à l'origine. Elles sont parfois données dans la description verbale de la situation à modéliser, mais on peut aussi avoir à les dégager de valeurs correspondantes des variables. Dans le cas particulier où l'ordonnée à l'origine est nulle, la droite passe par l'origine du système d'axes et le modèle en est un de variation directement proportionnelle.

Modélisation et résolution de problèmes

REMARQUE

La réponse à un problème est la conclusion d'un travail. Elle doit être rédigée dans un français correct. Les phrases doivent être claires et comporter un sujet, un verbe et un complément.

À la fin de l'ouvrage, on donne des « éléments de réponse », mais pas nécessairement une réponse au sens de la conclusion d'un travail. Ces éléments de réponse ont pour but d'aider l'étudiante et l'étudiant à vérifier les résultats de ses manipulations et de ses calculs.

PROCÉDURE

Résolution d'un problème par modélisation affine

1. Identifier les variables du problème et les représenter par des symboles appropriés accompagnés des unités de mesure des variables.
2. Définir en compréhension la relation entre les variables en justifiant le choix du modèle.
3. Utiliser le modèle pour résoudre le problème.
   - 3.1 Reformuler la question (ou les questions) en utilisant les variables du problème.
   - 3.2 Effectuer les calculs et manipulations algébriques permettant de répondre à la question.
   - 3.3 Rédiger la réponse à la question posée.

EXEMPLE 6.1.1

Après avoir suivi un cours sur la modélisation affine, le propriétaire d'immeubles locatifs estime que le lien entre les coûts de chauffage mensuels et la température extérieure est de nature affine. Il a noté qu'en octobre, la température moyenne a été de 13°C et que les coûts de chauffage pour l'ensemble de ses logements a été de 1 340 $. Par ailleurs, en novembre, pour une température moyenne de 8°C, les coûts se sont élevés à 2 530 $.

a) En supposant que le phénomène est effectivement modélisable par un lien affine, construire un tel modèle, indiquer la signification des paramètres selon le contexte et représenter graphiquement le lien affine.

b) Déterminer le zéro de la fonction et interpréter le résultat selon le contexte.

c) Selon les données du service de météorologie local, la température moyenne durant le mois de janvier, au cours années précédentes, a été de –18°C. En utilisant le modèle, estimer les coûts de chauffage pour le mois de janvier suivant.

■ ***Solution***

a) **Identification des variables**

Soit C, les coûts mensuels de chauffage, et T, la température moyenne durant le mois.

Définition du lien entre les variables

Si l'hypothèse du propriétaire est exacte, la relation entre les coûts et la température est de la forme $C = aT + b$.

La pente de la droite représentant graphiquement le modèle affine cherché est :

$$a = \frac{1\,340 - 2\,530}{13 - 8} = -238 \text{ \$/°C}.$$

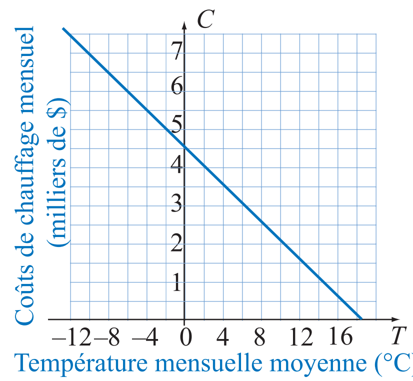
Le modèle est donc de la forme $C = -238T + b$. En substituant les coordonnées d'une des correspondances aux variables, on obtient la valeur de b :

$$2\,530 = -238 \times 8 + b, \text{ donc } b = 4\,434 \text{ \$}.$$

La fonction est donc $C(T) = -238T + 4\,434$ \$.

Le fait que la pente est de –238 \$/°C signifie que les coûts mensuels de chauffage diminuent de 238 \$ pour chaque augmentation de 1 °C de la température mensuelle moyenne. L'ordonnée à l'origine, 4 434 \$, est la valeurs des coûts mensuels de chauffage lorsque la température mensuelle moyenne est de 0 °C.

Utilisation du modèle



b) Le zéro de la fonction est la température mensuelle moyenne pour laquelle les coûts de chauffage mensuels sont nul :

$$-238T + 4\,434 = 0, \text{ donc } T = 18{,}63 \text{ °C}.$$

c) Si la température moyenne au mois de janvier est de –18 °C, les coûts de chauffage mensuels estimés à l'aide du modèle sont :

$$C(-18) = -238 \times -18 + 4\,434 = 8\,718 \text{ \$}.$$

Données à pas constant

Même si on peut décrire par un modèle affine un phénomène étudié lors d'une expérience de laboratoire, d'un sondage ou d'une recherche, il faut s'attendre à ce qu'il y ait une différence entre les valeurs observées et les valeurs fournies par le modèle. En effet, aucun modèle n'est une description exacte d'un phénomène expérimental. Une règle de correspondance obtenue à l'aide de données expérimentales est un **modèle empirique** dont la fiabilité dépend de la précision des données expérimentales. Lorsqu'on étudie la relation entre les variables d'un phénomène pour lequel on dispose de données empiriques, la représentation graphique se révèle un moyen efficace pour décider si le phénomène est descriptible par un modèle affine. En pratique, on a plusieurs couples formant un nuage de points et on cherche à déterminer la droite qui décrit le plus fidèlement possible le phénomène. Le cas le plus simple est celui où les données sont à pas constant, c'est-à-dire que les valeurs de la variable indépendante sont à intervalles réguliers.

Critère algébrique

Le taux de variation constant est une caractéristique du modèle affine. Dès que l'on peut déterminer que le taux de variation d'une grandeur donnée est constant, on conclut que le phénomène est modélisable par une fonction affine. Voici comment on procède dans le cas de données expérimentales à pas constant.

On suppose qu'on dispose de mesures expérimentales pour des valeurs à intervalles réguliers de la variable indépendante, et que le phénomène est descriptible par un modèle affine $f(x) = ax + b$. Pour chacun des couples, on a :

$$\begin{aligned} f(x+p) &= a(x+p)+b \\ &= ax+ap+b \\ &= ax+b+ap \\ &= f(x)+ap \end{aligned}$$

$$f(x+p)-f(x) = ap.$$

$$\frac{\Delta f}{p} = \frac{f(x+p)-f(x)}{p} = a.$$

Par conséquent, lorsque les valeurs de la variable indépendante sont à pas constant et que la différence entre les images consécutives de cette variable est proportionnelle au pas, on en conclut que le phénomène est descriptible par un modèle affine.

En bref, l'existence d'un lien affine dans le cas de données à pas constant est confirmée si

$$\Delta f = f(x+p)-f(x) = ap.$$

RégAffinIndus01

REMARQUE

Le critère algébrique est un moyen rapide de détecter un lien affine entre deux variables, mais il ne constitue pas le meilleur critère pour établir que le lien est bel et bien affine.

Dans une expérience de laboratoire portant sur des variables dont on sait que le lien est affine, le critère algébrique constitue une façon rapide de vérifier que les mesures sont prises correctement.

PROCÉDURE

Modélisation affine, données à pas constant

1. Représenter graphiquement les données. Si le nuage de points semble former une droite, on pose, sous toute réserve, l'hypothèse que le phénomène est descriptible par un modèle affine.
2. Calculer les différences $f(x+p) - f(p)$. Si elles sont relativement constantes, le paramètre a du modèle affine est la moyenne des différences divisée par le pas.
3. Calculer les ordonnées $b_i = f(x_i) - ax_i$ et déterminer le paramètre b en prenant la moyenne des valeurs de ces ordonnées.
4. Utiliser le modèle pour analyser le phénomène et répondre aux questions.

REMARQUE

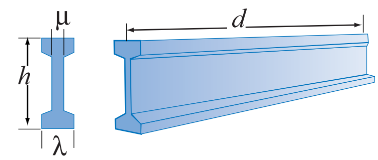
Il n'y a pas de recette universelle pour déterminer les chiffres à retenir pour les paramètres a et b d'un modèle. Quelques pistes de réflexion cependant. Il faut que la lecture du modèle soit simple et que les calculs ne soient pas alourdis inutilement.

Les valeurs calculées pour la variable dépendante ne doivent pas avoir plus de chiffres significatifs que les mesures de celle-ci, De plus, si les mesure de la variabledépendante ne comportent pas de décimales, les valeurs calculées ne devraient pas en avoir.

Dans le choix du nombre de chiffres à retenir, il faut tenir compte que le paramètre a est multiplié par la valeur de la variable indépendante et que le paramètre b est simplement additionné à ce produit.

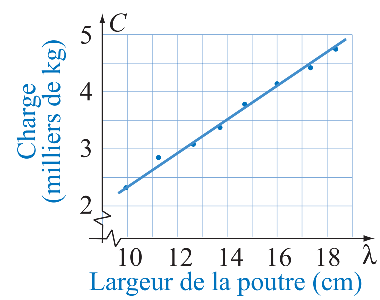
EXEMPLE 6.1.2

On veut fabriquer des poutres en I avec un nouveau matériau. La largeur μ de la bande centrale est le tiers de la largeur λ de la poutre et on veut déterminer la charge que peuvent supporter sans déformation des poutres de même longueur et de même épaisseur mais de différentes largeurs. Les mesures obtenues sont rassemblées dans le tableau ci-contre, où la charge C est en kilogrammes (kg) et la largeur λ d'une poutre est en centimètres (cm). Construire un modèle mathématique qui décrit le lien entre les variables.



Charge selon la largeur

λ (cm)	C (kg)
10,0	2 475
11,2	2 790
12,4	3 095
13,6	3 405
14,8	3 705
16,0	4 015
17,2	4 320
18,4	4 625



■ *Solution*

Représentation graphique des données

La représentation graphique des données est un nuage de points qui semblent alignés. On pose donc l'hypothèse qu'il existe un lien affine entre les variables.

Calcul des différences et de la pente

Puisque les mesures ont été prises à des intervalles de 1,2 cm, les données sont à pas constant. En calculant la différence des images, on obtient un critère dont les valeurs sont inscrites dans le deuxième tableau.

Le fait que la différence des images est relativement constante confirme l'hypothèse d'un lien affine entre les variables. La pente est égale à la moyenne des différences divisée par la longueur du pas de la variable indépendante

$$a = \frac{307,14}{1,2} = 255,95.$$

Le modèle affine est donc de la forme

$$f(x) = 255,95x + b$$

ou, si on utilise les symboles des variables du problème,

$$C(\lambda_i) = 255,95\lambda_i + b.$$

Calcul des différences

λ (cm)	C (kg)	$C(\lambda_i + p) - C(\lambda_i)$
10,0	2 475	–
11,2	2 790	315
12,4	3 095	305
13,6	3 405	310
14,8	3 705	300
16,0	4 015	310
17,2	4 320	305
18,4	4 625	305
	Moyenne	307,14

Calcul de l'ordonnée

Il reste à calculer la valeur du paramètre b. Si on fait passer par chaque point du graphique une droite de pente 255,95, ces droites coupent l'axe vertical en des points distincts dont les ordonnées de ces points sont respectivement :

$$b_i = C(\lambda_i) - 255,95\lambda_i.$$

On calcule chacune de ces ordonnées en ajoutant simplement une colonne au tableau. Ce type de traitement est simple et rapide avec le logiciel Excel.

La moyenne des ordonnées calculées est 80,74. On acceptera donc comme modèle :

$$C(\lambda_i) = 255,95\lambda_i - 80,74 \text{ kg}.$$

En prenant connaissance des résultats, le responsable de la qualité s'étonne que l'ordonnée à l'origine soit négative dans ce contexte et il demande qu'une autre équipe refasse les tests.

Calcul de l'ordonnée

λ (cm)	C (kg)	ΔC_i	$C_i - 255,95\lambda_i$
10,0	2475	–	–84,50
11,2	2790	315	–76,64
12,4	3095	305	–78,78
13,6	3405	310	–75,92
14,8	3705	300	–83,06
16,0	4015	310	–80,20
17,2	4320	305	–82,34
18,4	4625	305	–84,48
Moyennes		307,14	–80,74

Données à pas variable

Dans ce qui suit, nous présentons trois méthodes applicables autant à des données à pas variable qu'à des données à pas constant, ainsi que des mesures utilisées pour juger de la précision d'un modèle mathématique.

Méthode graphique

RégAffinIndus02

La façon la plus simple et la plus rapide de construire un modèle affine à l'aide d'une règle transparente consiste à représenter les couples de données sur du papier quadrillé et à choisir parmi toutes les droites passant par deux des points obtenus celle qui semble le mieux décrire le phénomène. On utilise les coordonnées $(x_1; y_1)$ et $(x_2; y_2)$ des points choisis pour écrire l'équation de la droite retenue comme modèle, soit en appliquant la procédure de géométrie analytique :

$$\frac{y-y_1}{x-x_1}=\frac{y_2-y_1}{x_2-x_1}$$

soit en résolvant le système d'équations :

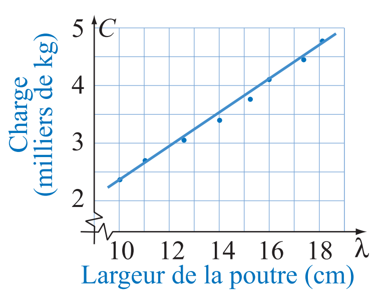
$$\begin{aligned} y_1 &= ax_1 + b \\ y_2 &= ax_2 + b. \end{aligned}$$

EXEMPLE 6.1.3

Une seconde équipe a refait les tests décrits dans le dernier exemple et elle a obtenu les valeurs inscrites dans le tableau présenté ci-contre. Construire un modèle mathématique décrivant le lien entre les variables.

Charge selon la largeur

λ (cm)	C (kg)
10,0	2 515
11,0	2 805
12,5	3 110
14,0	3 410
15,2	3 705
16,0	3 995
17,4	4 295
18,2	4 575



Solution

En représentant graphiquement les données du tableau, on constate que le nuage de points évoque une droite, mais les points ne sont pas parfaitement alignés, ce que peut expliquer des erreurs de mesure.

Par exemple, si la droite passant par les points (10; 2 515) et (16; 3 995) semble la plus apte à décrire la relation entre les variables, en appliquant la procédure de géométrie analytique, on a

$$\frac{C-2\ 515}{\lambda-10}=\frac{3\ 995-2\ 515}{16-10}.$$

En isolant C, on obtient le modèle

$$C(\lambda) = 246{,}67\lambda + 48.$$

L'équipe décide de mesurer la précision de ce modèle en calculant, pour chaque valeur de la variable indépendante, la différence entre la valeur observée (C_i) et la valeur donnée par le modèle mathématique (C_m). De telles différences sont appelées **résidus**.

La somme des carrés des résidus est une mesure de la précision du modèle mathématique. Pour le modèle de la seconde équipe, les résidus et leurs carrés sont inscrits dans le tableau présenté ci-contre.

Calcul des résidus

λ	C_i	C_m	R	R^2
10,0	2 515	2 515	0	0
11,0	2 805	2 761	44	1 936
12,5	3 110	3 131	–21	441
14,0	3 410	3 501	–91	8 464
15,2	3 705	3 797	–92	8 649
16,0	3 995	3 995	0	0
17,4	4 295	4 340	–45	2 025
18,2	4 575	4 537	38	1 444
				22 591

Méthode des données groupées

Cette méthode consiste à diviser les points en deux groupes contenant chacun la moitié, ou environ la moitié, des données. On calcule ensuite la valeur moyenne de la variable indépendante et de la variable dépendante dans chaque groupe. Les valeurs moyennes, représentées par $(\bar{x}_1; \bar{y}_1)$ et $(\bar{x}_2; \bar{y}_2)$, servent à écrire l'équation d'une droite à l'aide de la proportion

$$\frac{y-\bar{y}_1}{x-\bar{x}_1}=\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1}.$$

On peut également trouver les valeurs de a et b en solutionnant le système d'équations suivant :

$$\bar{y}_1 = a\bar{x}_1 + b$$
$$\bar{y}_2 = a\bar{x}_2 + b.$$

RégAffinIndus03

REMARQUE

Dans les exemples, les calculs sont effectués avec les chiffres retenus dans les tableaux. Si on effectue les calculs dans un tableur électronique, les valeurs calculées peuvent différer car le tableur conserve plus de chiffres, sans nécessairement les afficher, ce qui a une incidence sur le résultat des calculs.

EXEMPLE 6.1.4

Appliquer la méthode des données groupées aux résultats des tests de la seconde équipe, donnés dans le dernier exemple, pour construire un modèle affine décrivant la relation entre la largeur λ d'une poutre et la charge C qu'elle peut supporter.. Effectuer le calcul des résidus afin de mesurer la fiabilité du modèle obtenu.

Solution

En regroupant les données et en calculant les moyennes, on obtient

$$\bar{\lambda}_1=\frac{10{,}0+11{,}0+12{,}5+14{,}0}{4}=11{,}875;$$

$$\bar{C}_1=\frac{2\ 515+2\ 805+3\ 110+3\ 410}{4}=2\ 960;$$

$$\bar{\lambda}_2=\frac{15{,}2+16{,}0+17{,}4+18{,}2}{4}=16{,}70;$$

$$\bar{C}_2=\frac{3\ 705+3\ 995+4\ 295+4\ 575}{4}=4\ 142{,}5.$$

En appliquant la procédure de géométrie analytique, on a

$$\frac{C-2\ 960}{\lambda-11{,}875}=\frac{4\ 142{,}5-2\ 960}{16{,}70-11{,}875}.$$

En isolant C et en arrondissant le taux de variation à deux décimales et la constante à cinq chiffres significatifs, on obtient le modèle affine

$$C(\lambda) = 245{,}06\lambda + 50.$$

Le tableau présenté ci-contre donne la somme des carrés des résidus. On constate que ce modèle est plus fiable que celui que l'on obtient avec la méthode graphique puisque la somme des carrés des résidus est plus petite.

Charge selon la largeur

λ (cm)	C (kg)
10,0	2 515
11,0	2 805
12,5	3 110
14,0	3 410
15,2	3 705
16,0	3 995
17,4	4 295
18,2	4 575

REMARQUE

On arrondit la valeur du paramètre additif b au même nombre de décimales que dans les mseures de la charge et a, le paramètre multiplicatif à un chiffre significatif de plus.

Les largeurs des poutres sont des spécifications industrielles, on les traite comme des valeurs exactes.

Calcul des résidus

λ	C_i	C_m	R	R^2
10,0	2 515	2 501	14	196
11,0	2 805	2 746	59	3 481
12,5	3 110	3 113	–3	9
14,0	3 410	3 481	–71	5 041
15,2	3 705	3 775	–70	4 900
16,0	3 995	3 971	24	576
17,4	4 295	4 314	–19	361
18,2	4 575	4 510	65	4 225
				18 789

Méthode des moindres carrés

La méthode des moindres carrés consiste à calculer :

$\overline{x}$, la moyenne des valeurs de la variable indépendante x;

$\overline{y}$, la moyenne des valeurs de la variable dépendante y;

$\overline{x^2}$, la moyenne des carrés des valeurs de la variable indépendante x;

$\overline{xy}$, la moyenne des produits des valeurs des deux variables.

RégAffinIndus04

On obtient ensuite les paramètres a et b de la droite recherchée en solutionnant le système d'équations

$$\overline{y} = a\overline{x} + b$$
$$\overline{xy} = a\overline{x^2} + b\overline{x}.$$

EXEMPLE 6.1.5

En appliquant cette fois la méthode des moindres carrés aux résultats des tests de la seconde équipe, construire la méthode des données regroupées pour trouver un modèle affine décrivant la relation entre la largeur λ d'une poutre et la charge C qu'elle peut supporter. Effectuer le calcul des résidus afin de mesurer la fiabilité du modèle obtenu.

Solution

Charge selon la largeur

λ (cm)	C (kg)
10,0	2 515
11,0	2 805
12,5	3 110
14,0	3 410
15,2	3 705
16,0	3 995
17,4	4 295
18,2	4 575

Calcul des sommes

λ_i	C_i	λ_i^2	$\lambda_i C_i$
10,0	2 515	100,00	25 150
11,0	2 805	121,00	30 855
12,5	3 110	156,25	38 875
14,0	3 410	196,00	47 740
15,2	3 705	231,04	56 316
16,0	3 995	256,00	63 920
17,4	4 295	302,76	74 733
18,2	4 575	331,24	83 265
114,3	28 410	1 694,29	420 854

Le tableau ci-contre donne les sommes permettant le calcul des moyennes. En substituant dans :

$$\begin{cases} \overline{y} = a\overline{x} + b \\ \overline{xy} = a\overline{x^2} + b\overline{x} \end{cases}, \text{ on a } \begin{cases} \dfrac{28\ 410}{8} = a\dfrac{114,3}{8} + b \\ \dfrac{420\ 854}{8} = a\dfrac{1694,29}{8} + b\dfrac{114,3}{8} \end{cases}$$

d'où le système d'équations

$$28\ 410 = 114,3a + 8b$$
$$420\ 854 = 1\ 694,29a + 114,3b$$

En isolant b dans la première équation, on obtient :

$$b = \frac{28\ 410 - 114,3a}{8},$$

et, par substitution dans la deuxième équation, on a

$$420\ 854 = 1\ 694,29a + 114,3\left(\frac{28\ 410 - 114,3a}{8}\right)$$
$$3\ 366\ 832 = 13\ 554,32a + 3\ 247\ 263 - 13\ 064,49a$$
$$119\ 569 = 489,83a$$
$$a = 244,103...$$

En remplaçant a par sa valeur dans la première équation, on obtient

$$b = 63{,}6283...$$

En arrondissant a à quatre chiffres significatifs, on obtient $a = 244{,}1$ et en arrondissant b à deux décimales, on obtient $b = 63{,}63$.

Le modèle affine est alors :

$$C(\lambda) = 244{,}10\lambda + 64$$

Le tableau présenté ci-contre donne la somme des carrés des résidus. On constate que ce dernier modèle est le plus fiable des trois puisque la somme des carrés des résidus est la plus petite.

Calcul des résidus

λ	C_i	C_m	R	R^2
10,0	2 515	2 505	10	100
11,0	2 805	2 749	56	3 136
12,5	3 110	3 115	–5	25
14,0	3 410	3 481	–71	5 041
15,2	3 705	3 774	–69	4 761
16,0	3 995	3 969	25	625
17,4	4 295	4 311	–16	256
18,2	4 575	4 506	68	4 624
				18 568

Calcul des paramètres d'une droite de régression

Les valeurs moyennes des variables sont définies par

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n},\quad \overline{y} = \frac{\sum_{i=1}^{n} y_i}{n},\quad \overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n} \text{ et } \overline{xy} = \frac{\sum_{i=1}^{n} x_i y_i}{n}.$$

Si on remplace les variables par ces expressions dans les équations

$$\overline{y} = a\overline{x} + b \text{ et } \overline{xy} = a\overline{x^2} + b\overline{x}$$

et qu'on isole les paramètres a et b, on obtient

$$a = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2} \text{ et } b = \frac{\sum y_i - a\sum x_i}{n}.$$

RégAffinIndus05

REMARQUE

Nous utiliserons désormais, sans les démontrer, les expressions permettant de calculer directement les valeurs des paramètres a et b de la droite de régression.

On peut utiliser directement ces expressions en substituant aux variables les sommes du tableau pour obtenir les paramètres de la droite de régression. Cela évite d'avoir à résoudre le système d'équations à chaque fois. Ainsi, pour l'exemple 6.1.5, dont le tableau est reproduit ci-contre, on a :

$$a = \frac{n\sum \lambda_i C_i - \left(\sum \lambda_i\right)\left(\sum C_i\right)}{n\sum \lambda_i^2 - \left(\sum \lambda_i\right)^2} = \frac{8\times 420\ 854 - 114{,}3\times 28\ 410}{8\times 1\ 694{,}29 - (114{,}3)^2} = 244{,}103...$$

$$b = \frac{\sum C_i - a\sum \lambda_i}{n} = \frac{28\ 410 - 244{,}103...\times 114{,}3}{8} = 63{,}6275...$$

En arrondissant les valeurs de a et b, on obtient le modèle affine

$$C(\lambda) = 244{,}10\lambda + 64$$

Calcul des sommes

λ_i	C_i	λ_i^2	$\lambda_i C_i$
10,0	2 515	100,00	25 150
11,0	2 805	121,00	30 855
12,5	3 110	156,25	38 875
14,0	3 410	196,00	47 740
15,2	3 705	231,04	56 316
16,0	3 995	256,00	63 920
17,4	4 295	302,76	74 733
18,2	4 575	331,24	83 265
114,3	28 410	1 694,29	420 854

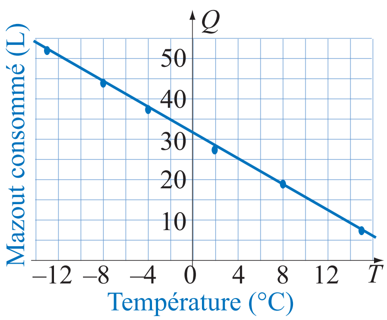
PROCÉDURE

Calcul des paramètres d'une droite de régression

1. Représenter graphiquement les données afin de s'assurer que le modèle affine est approprié.
2. Pour simplifier le traitement et la gestion des données, construire un tableau en réservant une colonne à chacune des grandeurs n, x, y, xy et x^2. La dernière ligne du tableau contient les sommations utilisées dans les formules de a et de b.

RégAffinIndus06

Coût de chauffage	
T_i	Q_i
−13	52,0
−8	44,0
−4	36,8
2	28,0
8	18,0
15	6,8



Calcul des sommes			
T	Q	TQ	T^2
−13	52,0	−676,0	169
−8	44,0	−352,0	64
−4	36,8	−147,2	16
2	28,0	56,0	4
8	18,0	144,0	64
15	6,8	102,0	225
0	185,6	−873,2	542

EXEMPLE 6.1.6

L'entrepreneur en construction pour lequel vous travaillez a décidé d'évaluer les coûts de chauffage des maisons qu'il construit afin de se servir de ce renseignement dans sa publicité. Il a noté, pour des périodes de 24 heures, la consommation moyenne de mazout en fonction de la température extérieure moyenne durant ces 24 heures. Les données qu'il a obtenues sont inscrites dans le tableau présenté ci-contre.

Trouver, par la méthode des moindres carrés, le modèle affine décrivant la relation entre la température et la quantité de mazout consommée.

Solution

Identification des variables

La quantité de mazout consommé Q (L) dépend de la température extérieure T (°C). La représentation graphique des données est un nuage de points (présenté ci-contre) qui évoque une droite, même si les points ne sont pas parfaitement alignés.

Définition du lien entre les variables

Pour déterminer la valeur des paramètres de la droite, il faut calculer les produits des valeurs correspondantes et le carré des valeurs de la variable indépendante, puis faire la somme des données et de ces résultats. On peut présenter tous les calculs dans un même tableau, dont la dernière ligne est réservée aux sommes des valeurs inscrites dans les colonnes. En utilisant les formules des paramètres, on obtient :

$$a = \frac{n\sum T_i Q_i - \left(\sum T_i\right)\left(\sum Q_i\right)}{n\sum T_i^2 - \left(\sum T_i\right)^2} = \frac{6\times(-873,2) - 0\times 185,6}{6\times 542 - (0)^2} = -1,611...$$

$$b = \frac{\sum Q_i - a\sum T_i}{n} = \frac{185,6 - (-1,611...)\times 0}{6} = 30,93...$$

Le modèle est donc $Q(T) = -1,611T + 30,93$.

Mesures de la précision du modèle

RégAffinIndus07

Le modèle mathématique construit est-il fiable? Il existe des mesures qui permettent de répondre partiellement à cette question. Ce sont la somme des carrés des résidus, le coefficient de corrélation et le coefficient de détermination.

Calcul des résidus

On effectue le calcul de la somme des carrés des résidus. On peut effectuer le calcul de la somme des résidus à l'aide du tableau utilisé pour déterminer les paramètres du modèle affine. Ainsi, pour le dernier exemple on obtient le tableau complémentaire suivant.

Consommation de mazout						
Valeurs observées				Valeurs du modèle	Résidus	Carrés des résidus
T_i	Q_i	T_iQ_i	T_i^2	$Q(T)$	R	R^2
–13	52,0	–676,0	169	51,873	0,123	0,015 069
–8	44,0	–352,0	64	43,822	0,178	0,031 722
–4	36,8	–147,2	16	37,378	–0,578	0,333 638
2	28,0	56,0	4	27,711	0,289	0,083 409
8	18,0	144,0	64	18,045	–0,045	0,002 005
15	6,8	102,0	225	6,767	0,033	0,001 070
0	185,6	–873,2	542			0,466 913

Coefficient de corrélation

Le coefficient de corrélation est une mesure de l'intensité du lien de linéarité entre deux variables. Il indique le degré de regroupement des points dans le voisinage de la droite. Il est défini par

$$r = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{\sqrt{n\sum {x_i}^2 - \left(\sum x_i\right)^2}\ \sqrt{n\sum {y_i}^2 - \left(\sum y_i\right)^2}}.$$

Ainsi, pour le dernier exemple, on a

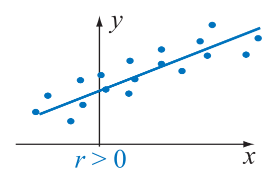
$$r = \frac{n\sum T_i Q_i - \left(\sum T_i\right)\left(\sum Q_i\right)}{\sqrt{n\sum {T_i}^2 - \left(\sum T_i\right)^2}\ \sqrt{n\sum {Q_i}^2 - \left(\sum Q_i\right)^2}}.$$

Le tableau du dernier exemple donne quatre des sommes apparaissant dans la formule de r. Il manque seulement $\sum Q_i^2$. On peut donc facilement calculer le coefficient de corrélation :

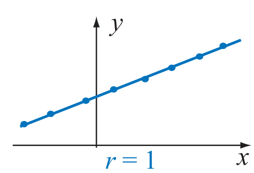
$$r = \frac{6\times(-873,2) - 0\times 185,6}{\sqrt{6\times 542 - (0)^2}\ \sqrt{6\times 7\ 148,48 - (185,6)^2}} = -0,999\ 8.$$

Le coefficient de corrélation linéaire r est un nombre compris entre –1 et 1 ($-1 \le r \le 1$). Lorsque $r = 0$ (corrélation nulle), le modèle affine n'est pas du tout approprié au phénomène. Lorsque r est proche de 1 ou de –1, le regroupement des points dans le voisinage de la droite est important. Si la valeur de r est positive, les variables varient dans un même sens, c'est-à-dire que la valeur de la variable dépendante augmente lorsque la valeur de la variable indépendante augmente. Si la valeur de r est négative, les valeurs des variables varient en sens inverse, c'est-à-dire que la valeur de la variable dépendante diminue lorsque la valeur de la variable indépendante augmente. L'exemple 6.1.6 illustre de dernier cas : la quantité de mazout consommée diminue lorsque la température augmente. De plus, le coefficient r est –0,999 8, ce qui est très proche de –1. La corrélation est donc très forte.

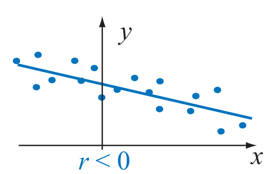
Le coefficient de détermination est le carré du coefficient de corrélation. Il est une mesure de la pertinence d'utiliser un modèle affine en faisant abstraction du fait que la corrélation peut être positive ou négative. C'est une mesure de l'adéquation entre le modèle et les données observées.



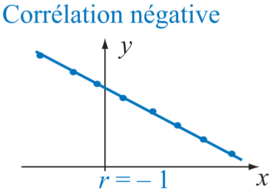
$r > 0$
Corrélation positive



$r = 1$
Corrélation positive parfaite



$r < 0$
Corrélation négative



$r = -1$
Corrélation négative parfaite

Droite de tendance

La droite de régression permet de construire un modèle simple, utilisé pour analyser des phénomènes ou décrire une tendance. On l'appelle alors **droite de tendance**. On distingue deux cas dans l'analyse de tendance, selon que les valeurs estimées sont à l'intérieur ou à l'extérieur de l'ensemble des données observées.

Interpolation

Lorsque les prévisions portent sur des valeurs à l'intérieur de l'intervalle des données, le processus est appelé **interpolation**. Généralement, les estimations par interpolation sont plutôt fiables.

Extrapolation

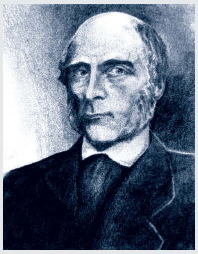
Si les prévisions portent sur des valeurs à l'extérieur de l'ensemble des données, le processus est appelé **extrapolation**. Il est à noter que la fiabilité est plus grande lorsqu'on fait des prédictions pour des valeurs proches de l'ensemble des données observées. Une prédiction portant sur une valeur éloignée de cet intervalle donne une estimation qui, sans être à rejeter, doit être utilisée avec circonspection. Dans les deux cas, il ne faut pas s'attendre à ce que le soit modèle plus précis que les données qu'il décrit.

Un peu *d'histoire*

FRANCIS GALTON

1822-1911

Influencés par les travaux de Charles Darwin (1809-1882), les statisticiens anglais de la fin du XIX^{e} siècle utilisèrent les statistiques dans des contextes plus proches de la biologie que de la sociologie, comme le faisaient les statisticiens du continent européen. Francis Galton, cousin de Darwin, s'intéressa à des questions statistiques liées à la génétique, l'hérédité et le comportement humain. Alors qu'Adolphe Quételet (1796-1874) avait réalisé des travaux sur des données biométriques de l'homme, comme le poids, la taille et le périmètre thoracique, et avait montré que ces données se répartissaient selon une courbe normale, Galton mena des recherches sur la **variabilité des caractères**, les différences entre individus et les moyens pour conserver et favoriser les meilleurs d'entre eux. Sa contribution majeure est la notion de **corrélation** et la mesure de celle-ci, le **coefficient de corrélation**.



Lors d'études sur l'hérédité, réalisées en 1877, Galton se rendit compte que des parents de petite taille avaient des enfants plus petits que la moyenne, mais plus grands que leurs parents. De même, des parents plus grands que la moyenne avaient des enfants plus grands que la moyenne, mais plus petits que leurs parents. Ce phénomène indique qu'il y a **corrélation** entre la taille des parents et celle des enfants, mais qu'il y a également une **régression** par rapport à la moyenne, d'où l'appellation **droite de régression**. La régression vers la moyenne est en inversement proportionnelle à la corrélation. Dans ses travaux sur l'eugénisme, Galton étudia la dispersion des résultats et élabora les notions de médiane et de quartile. À l'époque, les travaux de Galton étaient perçus comme une contribution importante dans la lutte de la science contre l'obscurantisme religieux. Ils furent malheureusement utilisés comme justification pour les exactions commises dans l'Allemagne nazie.

À partir de 1865, Galton se consacra à la statistique dans le but de quantifier les caractéristiques physiques, psychiques et comportementales de l'être humain, ainsi que leur évolution.

Darwin avait énoncé ses lois de l'évolution dans aucune considération du calcul des probabilités, mais ses théories ont assuré le triomphe d'une description probabiliste du monde. Galton fit le lien entre la théorie de la sélection naturelle et la recherche mathématique, consacrant une large partie de son activité à la défense de la théorie de l'évolution et cherchant à montrer qu'elle fournit des prévisions susceptibles d'être vérifiées.

6.2 Exercices

1. Un technicien en réparation d'appareils de chauffage applique un taux de 30 $ par demi-heure de travail, mais il charge un supplément de 20 $ pour le temps de déplacement.
   a) Construire un modèle mathématique décrivant le coût de la main-d'oeuvre pour les réparations effectuées par ce technicien.
   b) Déterminer le coût de la main-d'oeuvre pour une réparation qui a nécessité une demi-heure de travail.

2. Une personne désirant établir la correspondance entre le kilogramme et la livre se pèse à l'aide d'une balance graduée selon les deux échelles de mesure. Sur l'échelle graduée en kilogrammes, elle évalue sa masse à 70 kg et, sur l'échelle graduée en livres, elle fait une lecture de 154 lbs.
   a) À l'aide de ces données, établir la correspondance entre les deux unités de mesure.
   b) Esquisser le graphique de cette correspondance.
   c) Quel est l'équivalent en livres de 80 kg? de 100 kg?
   d) Si une personne a maigri de 8 lbs au cours du dernier mois, combien a-t-elle perdu de kilogrammes?

3. Vous contactez deux entrepreneurs paysagistes pour faire installer de la pelouse sur votre terrain. L'un d'eux charge 1,80 $ le mètre carré et des frais fixes de 120 $ et l'autre entrepreneur charge 2,10 $ le mètre carré, mais aucun frais fixes.
   a) Quelles sont la variable indépendante et la variable dépendante dans ce contexte ?
   b) Déterminer dans chacun des cas la fonction permettant d'évaluer le coût. Représenter graphiquement ces fonctions.
   c) Si la partie de votre terrain que vous désirez recouvrir de pelouse a une superficie de 300 m^2, lequel des deux entrepreneurs charge le moins cher?
   d) Quelle doit être la superficie du terrain à recouvrir pour qu'il soit plus avantageux de choisir l'autre entrepreneur?

4. L'entreprise qui vous emploie doit remplacer temporairement un appareil nécessitant des réparations dont la durée peut varier de deux à trois mois. Deux compagnies de location ont présenté une soumission. La première charge 10 $ par jour de location, tous services inclus; la deuxième exige 6 $ par jour et des frais fixes de 180 $. L'appareil est muni d'un dispositif qui détermine le nombre de jours d'utilisation en tenant compte seulement des jours ouvrables. Vous devez préparer une étude comparative des deux offres pour le conseil d'administration, qui doit choisir un fournisseur.
   a) Construire, pour chaque cas, un modèle mathématique décrivant le coût de la location en fonction de sa durée. Représenter graphiquement les deux modèles dans un même système d'axes.
   b) Quel est le coût d'une location de 30 jours? de 90 jours?
   c) Après analyse des modèles, quel fournisseur recommanderiez-vous au conseil d'administration ?

5. Deux cyclistes quittent simultanément deux endroits distants de 300 km et se dirigent l'un vers l'autre. André part du point A et roule à 22 km/h, alors que Bertrand part du point B et roule à 26 km/h.
   a) Exprimer, en fonction du temps, la distance entre le point A et chacun des cyclistes.
   b) Représenter graphiquement les deux fonctions dans un même système d'axes.
   c) Que représente l'abscisse du point de rencontre des deux droites ? Que représente l'ordonnée du point de rencontre des deux droites?
   d) Déterminer combien de temps les deux cyclistes mettent à se rencontrer.
   e) Déterminer la distance parcourue par chacun des cyclistes au moment de leur rencontre.

6. On a soumis des poutres d'un même matériau, de même longueur et de même épaisseur, mais de différentes largeurs à des essais pour déterminer la charge qu'elles peuvent supporter sans se déformer. Pour chacune des largeurs testées, on a noté la charge maximale avant déformation. Les données sont rassemblées dans le tableau suivant.

x (cm)	C (kg)
5	148
7	208
9	266
11	326
13	385
15	444
17	503

a) Construire un modèle mathématique de la correspondance entre les deux variables.
b) À l'aide du modèle, déterminer la charge que peut supporter une poutre dont la largeur est de 8 cm.

7. On réalise l'expérience suivante sur les échanges de chaleur. On plonge 25,0 g d'un alliage dans un bécher contenant 90,0 g d'eau à 25,82 ºC. La température finale T_f (lorsque les températures ont atteint leur point d'équilibre) est fonction de la température T_a de l'alliage au moment où on le plonge dans l'eau. Les températures en Celsius, mesurées au cours de divers essais sont données dans le tableau suivant.

T_a	T_f
120	27,7
110	27,4
100	27,2
90	26,9
80	26,7
70	26,4

a) Construire un modèle mathématique qui décrit le phénomène étudié.
b) Utiliser le modèle pour calculer la température finale de l'alliage dans le cas où sa température initiale est de 150°C.
c) Déterminer la température initiale de l'alliage si la température finale est de : 30°C, de 26°C.

8. Un entrepreneur en construction a décidé d'évaluer les coûts de chauffage des maisons qu'il bâtit afin de se servir de ce renseignement dans sa publicité. Il a noté la consommation moyenne de mazout durant des périodes de 24 heures en fonction de la température extérieure moyenne durant ces 24 heures. Les données qu'il a obtenues sont présentées dans le tableau suivant.

T (°F)	Q (L)
–11	48,0
–7	41,0
–1	32,0
2	27,0
6	20,0
12	11,0

a) Représenter graphiquement les données.
b) Construire un modèle affine décrivant la relation entre la température et la quantité de mazout consommée.
c) Évaluer la quantité de mazout consommée en une journée lorsque la température extérieure moyenne est de 9 °F.
d) Dans le cas où la température moyenne en janvier est de –12 °F, estimer la consommation de mazout durant ce mois.
e) Évaluer la quantité de mazout consommée en une journée lorsque la température extérieure moyenne est de –20 °F.

9. Vous travaillez pour une entreprise qui effectue l'entretien d'espaces à bureaux. Il est très important pour l'entreprise d'estimer le mieux possible le temps nécessaire à l'entretien d'un édifice avant de faire une soumission. Elle a donc noté la superficie des édifices dont elle fait déjà l'entretien, de même que le temps requis. Les données sont présentées dans le tableau suivant.

Superficie (m^2)	Nombre d'heures par semaine
87 000	320
81 000	400
69 000	260
64 000	388
60 000	325
51 000	284
44 000	227
39 000	180
28 000	125

a) Représenter graphiquement les données.
b) À l'aide des données, établir un modèle affine décrivant la relation entre le temps consacré à l'entretien et la superficie d'un édifice.
c) La compagnie doit soumissionner pour l'entretien d'un édifice de 56 000 m^2. Estimer le temps requis pour effectuer le travail à l'aide du modèle construit en b).
d) Calculer le coefficient de corrélation. Qu'est-ce qu'il indique ?

10. Vous devez finaliser une étude pour voir s'il y a un lien entre le nombre de logements mis en chantier et le taux hypothécaire annuel. L'étude porte plus précisément sur le mois de juin et les données du tableau suivant ont été recueillies.

Année	Taux hypothécaire	Nombre de logements
1981	18,55 %	9 000
1982	19,75 %	3 500
1983	13,00 %	10 100
1984	14,50 %	7 800
1985	11,75 %	8 800

a) Représenter graphiquement les données.
b) À l'aide de ces données, établir un modèle décrivant la relation entre le taux hypothécaire et le nombre de mises en chantier.
c) Calculer le coefficient de corrélation. Que vous indique ce coefficient?

11. Une association d'automobilistes a demandé à ses membres de noter la distance qu'ils ont parcourue et le coût d'utilisation de leur véhicule au cours de la dernière année, en incluant les coûts de l'immatriculation, des assurances, de l'essence et de l'entretien. L'association a dressé le tableau suivant à l'aide des informations reçues pour la voiture la plus populaire auprès de ses membres.

Distance (km)	Coût ($)
5 000	3 950
10 000	4 860
15 000	5 740
20 000	6 600
25 000	7 520
30 000	8 460

a) Construire un modèle mathématique décrivant la correspondance entre les deux variables.
b) Donner une mesure de la précision du modèle en calculant les résidus.
c) Prévoir, à l'aide du modèle, le coût d'utilisation de la voiture en question dans le cas où la distance parcourue en une année est de 45 000 km.

12. Une entreprise veut mettre sur le marché un nouveau modèle d'armoire avec serrure destinées à entreposer les médicaments hors de la portée des enfants. Elle a effectué une étude de marché afin de fixer le prix de ce produit. Les résultats de l'étude sur le prix et le volume estimé des ventes annuelles, sont présentés dans le tableau suivant.

Prix ($)	Volume des ventes
35	540
40	492
45	458
50	406
55	336
60	294

a) Déterminer la règle de correspondance entre le prix de l'article et le nombre de clients potentiels.
b) Estimer la précision du modèle à l'aide des résidus et du coefficient de corrélation.

13. On a réalisé une étude de marché avant de commercialiser de petites remises de jardin, conçues pour être fixées à un mur de la maison ou du garage. L'étude de marché a été effectuée afin de fixer le prix du produit. Les résultats de l'étude sur le prix et le volume estimé des ventes annuelles sont présentés dans le tableau suivant.

Prix ($)	Volume des ventes
250	1 200
300	1 050
350	975
400	850
450	775
500	650

a) Déterminer la règle de correspondance entre le prix de l'article et le nombre de clients potentiels.
b) Estimer la précision du modèle à l'aide des résidus et du coefficient de corrélation.

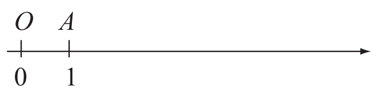
6.3 Échelles graphiques

Dans la présente section, nous appliquons la régression affine à l'élaboration d'une procédure d'analyse permettant de décider quel est le modèle le plus approprié et de définir ce dernier. Cette méthode repose sur la représentation graphique à échelle logarithmique. Nous allons donc d'abord, présenter les caractéristiques d'une telle échelle en la comparant à une échelle linéaire.

Échelle linéaire

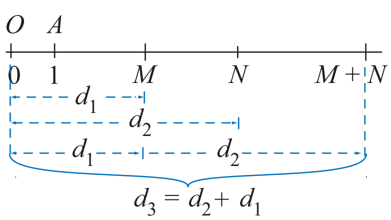
RégLoga01

Une échelle est dite **linéaire** si son pas est constant, c'est-à-dire si chaque nombre est situé à une distance de l'origine qui est proportionnelle à sa valeur. La droite représentée ci-contre comporte un point origine O et un point A qui détermine la valeur unitaire ou la longueur du pas de l'échelle.

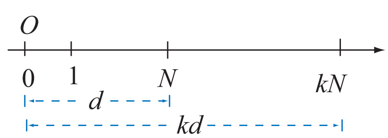


Pas unitaire d'une échelle linéaire

Si la droite est graduée selon la longueur unitaire $\overline{OA}$ et si M et N sont deux nombres positifs *situés* à des distances respectives d_1 et d_2 de l'origine, en respectant la proportionnalité, alors leur somme est un nombre $V = M + N$ représenté par un point situé à une distance $d_1 + d_2$ de l'origine. De plus, si un nombre $N > 0$ est situé à une distance d de l'origine, pour tout nombre $k > 0$, le nombre kN est situé à une distance kd de l'origine.

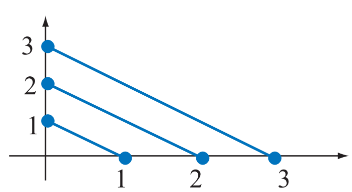


Somme sur une échelle linéaire



Produit sur une échelle linéaire

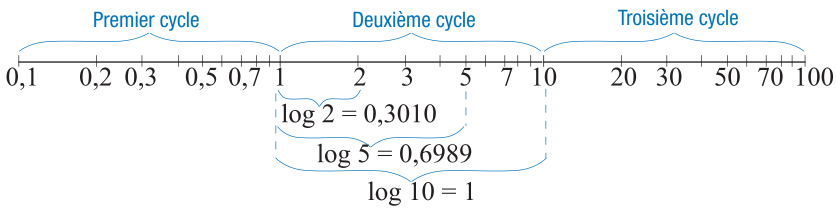
Dans un système d'axes gradués selon des échelles linéaires de pas différents, les segments de droite joignant deux points des axes de même valeur déterminent des triangles semblables, car la distance à l'origine de n'importe quel nombre est proportionnelle à sa valeur.



Proportionnalité de deux échelles linéaires

Échelle logarithmique

Nous avons déjà souligné que la droite est la représentation graphique la plus facile à reconnaître. Pour déceler un lien non affine entre deux variables, il est d'usage d'utiliser du papier quadrillé dont au moins l'une des deux échelles est graduée à l'aide du logarithme de base 10. Sur une échelle logarithmique, l'origine correspond au nombre 1, car (0; 0) = (0; log 1). La position des autres nombres est déterminée de telle sorte que leur distance à l'origine soit proportionnelle au logarithme du nombre. Ainsi, puisque le logarithme de base 10 de 5 est 0,698 9 et que le logarithme de 10 est 1, la distance de 1 à 5 correspond à 69,89 % de la distance de 1 à 10. Puisque le logarithme de 100 est 2, la distance de 1 à 100 est égale à deux fois la distance de 1 à 10; la distance entre 0,1 et 1 est égale à la distance entre 1 et 10 puisque le logarithme de 0,1 est égal à –1. Chacun des intervalles représentant une unité logarithmique est appelé **cycle**. Ainsi, l'intervalle de 0,1 à 1 est un cycle, tout comme les intervalles de 1 à 10 et de 10 à 100.



Graduations sur une échelle logarithmique

Du papier quadrillé suivant une échelle linéaire et une échelle logarithmique est appelé **papier semi-logarithmique** et un papier quadrillé suivant deux échelles logarithmiques est appelé **papier logarithmique**. Sur ces deux types de papiers, il n'y a pas de nombres indiquant les graduations; l'échelle commence à n'importe quel nombre suivant les besoins du problème. Dans les premiers exercices, nous indiquons les graduations pour permettre au lecteur de se familiariser avec ce genre de représentations graphiques. La caractéristique la plus intéressante est le fait que le graphique d'une fonction exponentielle sur du papier semi-logarithmique est une droite, comme l'illustre la représentation de la fonction $f(x) = 2^x$ sur un papier semi-logarithmique à deux cycles.

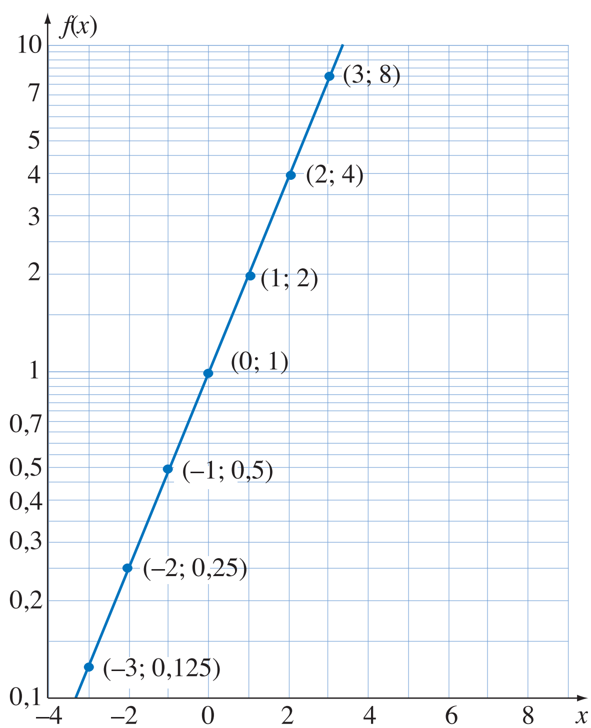
RégLoga02

REMARQUE

Dans la démarche d'analyse menant à la modélisation d'une liste de données, nous représentons ces données dans un système de référence bilinéaire, semi-logarithmique ou bi-logarithmique.

Dans le graphique, le point désigné par (2; 4) correspond en réalité au point (2; log 4) puisque sa distance à l'axe des x est proportionnelle au logarithme de la valeur de la variable dépendante.

x	$f(x)$
–3	0,125
–2	0,25
–1	0,5
0	1
1	2
2	4
3	8



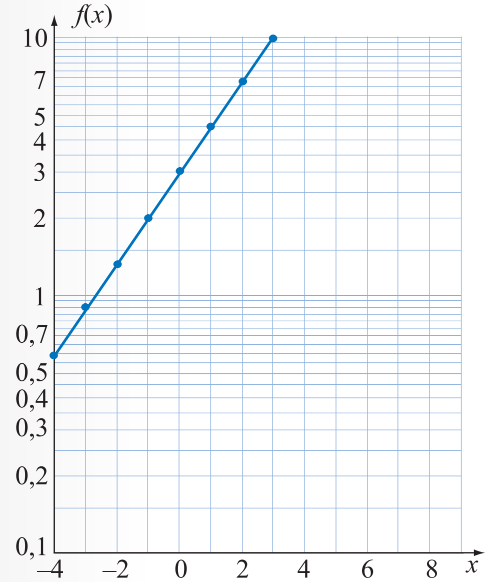
EXEMPLE 6.3.1

Représenter $f(x) = 3 \times 1{,}5^x$ sur du papier semi-logarithmique.

Solution

On calcule d'abord quelques correspondances que l'on inscrit dans un tableau comme le suivant.

x	$f(x)$
–4	0,59
–3	0,89
–2	1,33
–1	2
0	3
1	4,5
2	6,75
3	10,1
4	15,2
5	22,8



Même si la base de la fonction exponentielle est différente de 10, la représentation graphique est une droite. La raison en est fort simple. Soit, par exemple, une fonction exponentielle de la forme

$$y = ab^x.$$

On a

$\log y = \log(ab^x)$. Logarithme de chaque membre de l'équation.

$\log y = \log a + \log b^x$. Propriété du logarithme d'un produit.

$\log y = \log a + x \log b$. Propriété du logarithme d'une puissance.

$\log y = x \log b + \log a$. Commutativité de l'addition.

$Y = Ax + B$, où $Y = \log y$, $A = \log b$ et $B = \log a$.

Puisque $\log b$ et $\log a$ sont des constantes, il existe une relation affine entre x et $\log y$. C'est pourquoi la représentation graphique sur du papier semi-logarithmique donne une droite.

REMARQUE

Le texte ci-contre est particulièrement important. Lorsqu'on dit, par exemple, que :

$$\log [\alpha] = 21{,}7 \times 10^{-3} T + 36{,}8 \times 10^{-4}$$

définit une relation affine, il faut comprendre que la relation dont on parle est entre T et $\log [\alpha]$ alors que $\log [\alpha]$ est une fonction de T :

$$\log [\alpha] = f(T).$$

En prenant plutôt le logarithme naturel de chaque membre de l'équation $y = ab^x$, on obtient la relation

$$\ln y = x \ln b + \ln a.$$

Ainsi, qu'on effectue les calculs dans **l'une ou l'autre des bases, on obtient le même modèle.**

RégLoga03

Échelle logarithmique et modélisation

Les caractéristiques des échelles logarithmiques indiquent comment utiliser du papier semi-logarithmique pour reconnaître un lien exponentiel entre des variables et déterminer la règle de correspondance décrivant ce lien. L'exemple suivant illustre la marche à suivre.

EXEMPLE 6.3.2

Au cours d'une expérience de laboratoire, on a obtenu les grandeurs physiques inscrites dans le tableau présenté ci-contre.

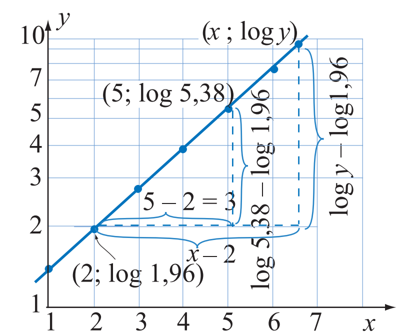
a) Vérifier l'hypothèse d'un lien exponentiel entre les variables x et y.

b) Déterminer la règle de correspondance décrivant le lien entre les deux variables.

x	y
1	1,40
2	1,96
3	2,74
4	3,84
5	5,38
6	7,53

Solution

a) On représente les données dans un système d'axes semi-logarithmique. Puisque les valeurs de la variable dépendante s'échelonnent de 1,40 à 7,53, on a besoin d'un seul cycle. Le graphique sur du papier semi-logarithmique dont l'échelle logarithmique est verticale est une droite, ce qui indique l'existence d'un lien exponentiel entre les variables.



b) Pour trouver la description algébrique du lien entre les variables, on applique la méthode des moindres carrés en utilisant le logarithme des valeurs de la variable dépendante. On peut effectuer les calculs dans la base naturelle ou la base 10. Si on choisit la base e, on obtient le tableau présenté ci-contre. Pour déterminer les paramètres du modèle par la méthode des moindres carrés, on prend le logarithme des valeurs de la variable dépendante. En effectuant les calculs en base e, on obtient :

$$A = \frac{n\sum x_i \ln y_i - \left(\sum x_i\right)\left(\sum \ln y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$= \frac{6\times 30{,}614\ 9 - 21\times 7{,}064\ 4}{6\times 91 - 21\times 21} = 0{,}336\ 5.$$

$$B = \frac{\sum \ln y_i - A\left(\sum x_i\right)}{n}$$

$$= \frac{7{,}064\ 4 - 0{,}336\ 5\times 21}{6} = -0{,}000\ 48.$$

x	y	$\ln y$	$x \ln y$	x^2
1	1,40	0,336 5	0,336 5	1
2	1,96	0,672 9	1,345 9	4
3	2,74	1,008 0	3,023 9	9
4	3,84	1,345 5	5,381 9	16
5	5,38	1,682 7	8,413 4	25
6	7,53	2,018 9	12,113 4	36
21	22,85	7,064 4	30,614 9	91

La valeur de B est négligeable, compte tenu de la précision des données de départ. Le lien entre les variables est donc :

$$\ln y = 0{,}3365x.$$

On trouve le lien exponentiel en exprimant cette équation sous forme exponentielle

$$y = e^{0{,}3365x} = 1{,}4^x.$$

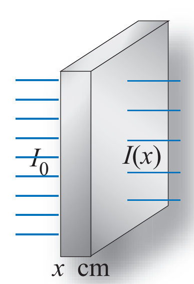
EXEMPLE 6.3.3

RégLoga04

On désire analyser la capacité d'absorption des rayons X d'un matériau. Pour ce faire, on bombarde des plaques de différentes épaisseurs et on mesure l'intensité $I(x)$ des radiations à la sortie des plaques. En prenant que $I_0 = 1$ unité comme intensité des radiations à l'entrée, on a obtenu les mesures inscrites dans le tableau pour différentes épaisseurs x, en centimètres.

a) Quel type de correspondance relie les variables.

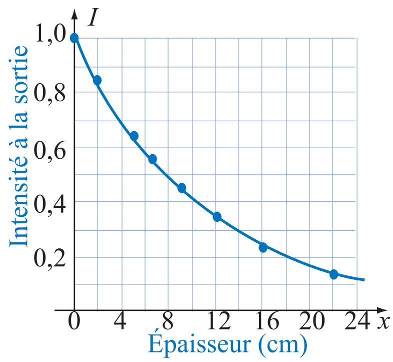
b) Déterminer la règle de correspondance.

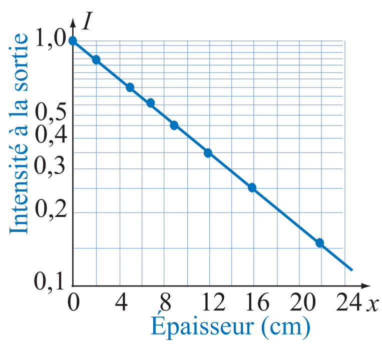


x	$I(x)$
0,00	1
2,00	0,84
5,00	0,65
6,50	0,57
9,00	0,46
12,00	0,35
16,00	0,25
22,00	0,15

■ *Solution*

a) On représente les données sur du papier bilinéaire, ce qui donne la représentation ci-contre. Le graphique représenté ci-contre étant une courbe, on conclut qu'il ne s'agit pas d'une correspondance affine. La représentation graphique sur du papier semi-logarithmique dont l'échelle logarithmique est verticale est une droite, ce qui confirme l'existence d'un lien exponentiel entre les variables.





b) Pour trouver la description algébrique du lien exponentiel entre les variables, on calcule le logarithme des valeurs de la variable dépendante, ce qui donne le tableau présenté ci-contre. On obtient par régression

x	I	$\ln I$	$x \ln I$	x^2
0,00	1,00	0,0000	0,0000	0,00
2,00	0,84	–0,1744	–0,3487	4,00
5,00	0,65	–0,4308	–2,1539	25,00
6,50	0,57	–0,5621	–3,6538	42,25
9,00	0,46	–0,7765	–6,9888	81,00
12,00	0,35	–1,0498	–12,5979	144,00
16,00	0,25	–1,3863	–22,1807	256,00
22,00	0,15	–1,8971	–41,7366	484,00
72,50	4,27	–6,2770	–89,6604	1 036,25

$$A = \frac{n\sum x_i \ln I_i - \left(\sum x_i\right)\left(\sum \ln I_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$= \frac{8\times(-89{,}660\ 4)-(72{,}50\times -6{,}277\ 0)}{8\times 1036{,}25 - 72{,}50\times 72{,}50} \approx -0{,}086\ 427\ 919...$$

$$B = \frac{\sum \ln I_i - A\left(\sum x_i\right)}{n}$$

$$= \frac{-6{,}2770-(-0{,}086\ 427\times 72{,}50)}{8} \approx -0{,}001\ 377\ 198.$$

La valeur de B est négligeable compte tenu de la précision des données. Elle est théoriquement nulle, car la valeur initiale de I est 1 et une plaque d'épaisseur nulle n'absorbe pas de rayons X. Donc,

$$\ln I = -0{,}086\ 428x.$$

Ce qui s'écrit sous forme exponentielle,

$$I = e^{-0{,}086\ 428x} = 0{,}917^x.$$

Le modèle est $I(x) = 0{,}917^x$. Il est à noter que la plupart des calculatrices sont munies de fonctions permettant de calculer directement les paramètres A et B.

REMARQUE

Dans les exemples, les calculs sont effectués avec les chiffres retenus dans les tableaux. Si on effectue les calculs dans un tableur électronique, les valeurs calculées peuvent différer car le tableur conserve plus de chiffres, sans nécessairement les afficher, ce qui a une incidence sur le résultat des calculs.

RégLoga05

Fonction puissance

La représentation graphique sur du papier logarithmique permet également de reconnaître une fonction puissance ou une fonction logarithmique. Une fonction puissance est de la forme $y = ax^b$. En prenant le logarithme de chaque membre de l'équation, on obtient :

$\log y = \log(ax^b)$

$\log y = \log a + \log x^b$. Propriété du logarithme d'un produit.

$\log y = \log a + b \log x$. Propriété du logarithme d'une puissance.

$\log y = b \log x + \log a$. Commutativité de l'addition.

En posant $Y = \log y$, $A = b$, $X = \log x$ et $B = \log a$, on a $Y = AX + B$.

Il existe donc entre $\log x$ et $\log y$, une correspondance affine que la représentation des données sur du papier logarithmique met en évidence.

Fonction logarithmique

En ce qui concerne la fonction logarithmique, l'équation $y = a \log x + b$, indique clairement qu'il existe une relation affine entre y et $\log x$, représentée symboliquement par $y = AX + B$ où $A = a$, $X = \log x$ et $B = b$. Cette relation est mise en évidence par sa représentation sur du papier semi-logarithmique, la variable indépendante étant portée sur l'échelle logarithmique. Si le nuage de points évoque une droite, le modèle est logarithmique.

EXEMPLE 6.3.4

On a obtenu en laboratoire les données présentées dans le tableau ci-contre.

x	$I(x)$
3,00	0,94
4,00	0,53
5,00	0,34
6,00	0,24
7,00	0,17
8,00	0,13
9,00	0,10

a) Quel type de correspondance relie les deux variables.

b) Déterminer la règle de correspondance.

Solution

a) La représentation graphique sur du papier bilinéaire ou du papier semi-logarithmique est une courbe.

La représentation sur du papier log-log laisse penser qu'une fonction puissance pourrait décrire le lien entre les deux variables. On établit donc une relation affine entre $\ln x$ et $\ln I$. Les résultats des calculs préliminaires sont rassemblés dans le tableau suivant.

	x	I	$\ln x$	$\ln I$	$\ln x \ln I$	$(\ln x)^2$
	3,00	0,94	1,098 6	–0,061 9	–0,068 0	1,206 9
	4,00	0,53	1,386 3	–0,634 9	–0,880 1	1,921 8
	5,00	0,34	1,609 4	–1,078 8	–1,736 3	2,590 3
	6,00	0,24	1,791 8	–1,427 1	–2,557 0	3,210 4
	7,00	0,17	1,945 9	–1,772 0	–3,448 1	3,786 6
	8,00	0,13	2,079 4	–2,040 2	–4,242 5	4,324 1
	9,00	0,10	2,197 2	–2,302 6	–5,059 3	4,827 8
Σ	42,00	2,45	12,108 6	–9,317 5	–17,991 3	21,867 9

$$A = \frac{n\sum \ln x_i \ln I_i - \left(\sum \ln x_i\right)\left(\sum \ln I_i\right)}{n\sum (\ln x_i)^2 - \left(\sum \ln x_i\right)^2}$$

$$= \frac{7\times(-17{,}991\,3) - (12{,}108\,6 \times -9{,}317\,5)}{7\times 21{,}867\,9 - (12{,}108\,6)^2} \approx -2{,}031\,439\,381$$

$$B = \frac{\sum \ln I_i - A\left(\sum \ln x_i\right)}{n}$$

$$= \frac{-9{,}317\,4 - (-2{,}031\,439\,381 \times 12{,}108\,7)}{7} \approx 2{,}182\,912\,413$$

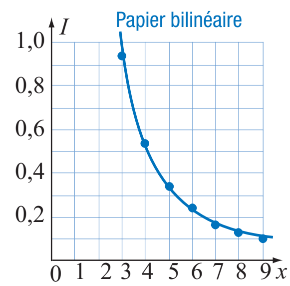
b) La relation affine est $\ln I = -2{,}032\,2 \ln x + 2{,}184\,3$; donc,

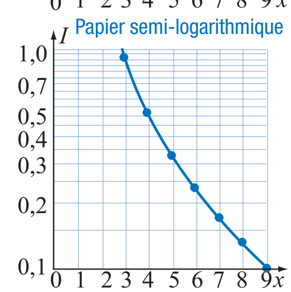
$$\ln I = \ln x^{-2{,}032\,2} + 2{,}184\,3$$

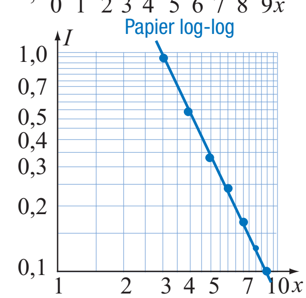
$$I = e^{2{,}184\,3} x^{-2{,}032\,2} = 8{,}8844 x^{-2{,}032\,2}.$$

Compte tenu de la précision des données, le modèle est :

$$I = \frac{8{,}87}{x^{2{,}03}}.$$





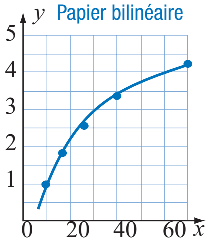


Les valeurs des paramètres A et B varient selon le nombre de chiffres significatifs retenus pour les calculs, particulièrement si y interviennent des logarithmes. Idéalement, on devrait retenir tous les chiffres et arrondir seulement à la fin des calculs en tenant compte de la précision des mesures à modéliser.

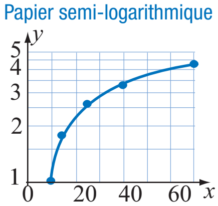
RégLoga06

x	y
10	1,00
15	1,70
25	2,59
40	3,41
70	4,38

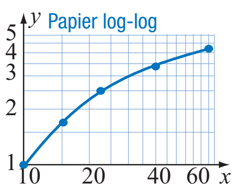
Papier bilinéaire



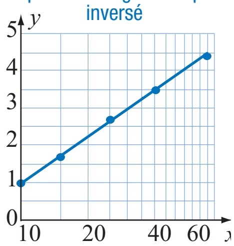
Papier semi-logarithmique



Papier log-log



Papier semi-logarithmique inversé



EXEMPLE 6.3.5

On a obtenu les données du tableau présenté ci-contre en laboratoire.

a) Quel type de correspondance relie les deux variables.

b) Déterminer la règle de correspondance.

Solution

a) La représentation graphique sur du papier bilinéaire, sur papier semi-logarithmique ou log-log est une courbe. Cependant, en portant les valeurs de la variable indépendante sur l'axe logarithmique d'un papier semi-logarithmique, on obtient une droite. Le modèle logarithmique est donc approprié.

b) Les résultats des calculs préliminaires sont rassemblées dans le tableau suivant.

	x	y	$\log x$	$y \log x$	$(\log x)^2$
	10	1,00	1,000	1,000	1,000
	15	1,70	1,176	1,999	1,383
	25	2,59	1,398	3,621	1,954
	40	3,41	1,602	5,463	2,567
	70	4,38	1,845	8,081	3,404
Σ	160	13,08	7,021	20,165	10,308

$$A = \frac{n\sum(\log x_i)\, y_i - \left(\sum \log x_i\right)\left(\sum y_i\right)}{n\sum(\log x_i)^2 - \left(\sum \log x_i\right)^2}$$

$$= \frac{5\times(20,165) - (7,021\times 13,08)}{5\times 10,308 - (7,021)^2} \approx 4,003\,599\,994$$

$$B = \frac{\sum y_i - A\left(\sum \log x_i\right)}{n}$$

$$= \frac{13,08 - (4,003\,599\,994\times 7,021)}{5} \approx -3,005\,855$$

Le modèle est donc : $y = 4,004 \log x - 3,01$.

Il n'est pas toujours nécessaire de représenter les données sur autant d'échelles différentes. Si on sait quel type de relation lie les variables est connue, on choisit tout de suite une échelle appropriée. On peut cependant utiliser du papier linéaire tout en sachant que la relation est exponentielle, logarithmique ou de puissance. Il faut alors effectuer des calculs sur les données expérimentales pour déterminer les valeurs à porter dans le graphique. La démarche élaborée dans le présent chapitre est très utile pour analyser des données expérimentales.

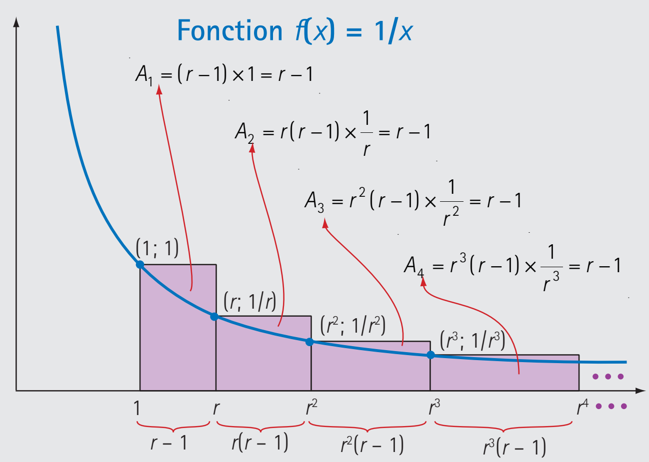
Un peu d'histoire

AVÈNEMENT DES LOGARITHMES

En développant les logarithmes, Napier a fait beaucoup plus que développer des méthodes facilitant les calculs. Il a permis la découverte d'une nouvelle fonction et ouvert la voie à de nouvelles techniques de modélisation.

Grégoire de Sant-Vincent 1584-1667

En 1647, Grégoire de Saint-Vincent (1584-1687) travaille sur la quadrature de l'hyperbole[1] et met en évidence une nouvelle fonction (NH St-Vincent). En 1661, Christiaan Huygens (1629-1695) remarque que cette fonction se trouve être une fonction logarithme particulière : le logarithme naturel. La notion de fonction et la correspondance entre les fonctions exponentielles et les fonctions logarithmes apparaissent après le travail de Leibniz sur la notion de fonction, en 1697.



La somme des bases des rectangles est une progression géométrique, et la somme des aires est une progression arithmétique, puisque les rectangles ont la même aire.

1. La quadrature d'une surface est la recherche d'un carré ayant même aire que la surface en question. La quadrature de l'hyperbole porte sur l'aire entre la courbe et l'axe horizontal.

Le mathématicien anglais d'origine galloise Edmund Gunter (1581-1626) développe l'échelle logarithmique (NH ÉchelleLog) et l'utilise dans une première version de règle à calcul. Jumelée à la régression linéaire, l'échelle logarithmique constitue un outil de modélisation important du lien entre deux variables (NH Log-Modélisation).

Henrietta Leavitt 1868-1921

C'est en utilisant une représentation de mesures à l'aide d'une échelle logarithmique qu'Henrietta Leavitt (NH Leavitt) a pu déterminer la relation entre la luminosité et la période des céphéïdes. La relation période-luminosité découverte par Leavitt est à la base d'une méthode d'évaluation des distances des amas stellaires et des galaxies dans l'Univers. C'est la méthode utilisée par Edwin Hubble pour conclure à l'expansion de l'univers.

C'est également en utilisant une échelle logarithmique que Vilfredo Pareto a établi la relation entre le revenu familial et le nombre de familles ayant ce revenu (NH Pareto) . En fait, on retrouve les exponentielles et les logarithmes dans plusieurs domaines, la croissance des capitaux en gestion, la croissance des populations en biologie et la radioactivité en physique. On les retrouve également en chimie, notamment dans les équations d'Arrhenius (NH Arrhenius).

La découverte de la radioactivité et la description de la quantité de matière radioactive en fonction du temps par une fonction exponentielle a mis à la disposition des savants un moyen de datation très efficace. La fonction inverse, qui est une fonction logarithmique, permet de calculer le temps écoulé depuis le début de la désintégration de la matière radioactive. Cette découverte a facilité les démarches de datation en archéologie et en sciences de la nature (NH Log-Datation01-02).

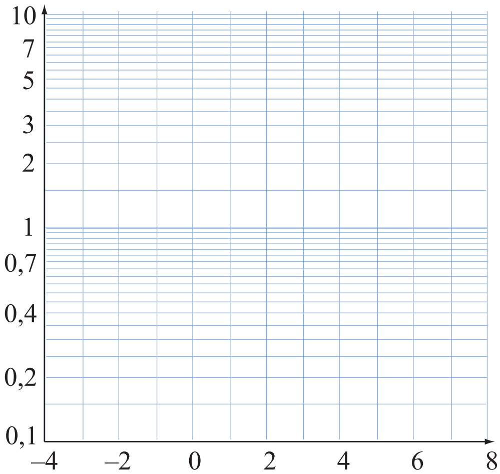
Vilfredo Pareto 1848-1923

Notes et vidéos historiques disponibles gratuitement à :

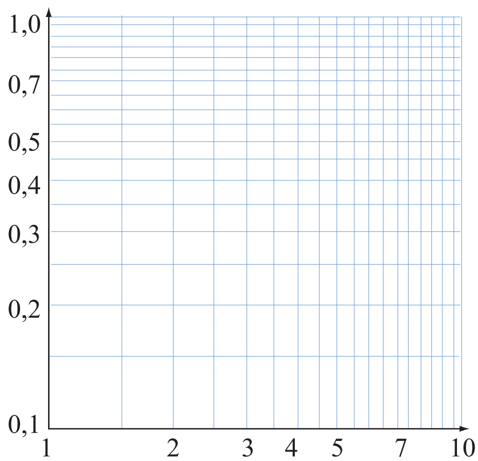
http://www.lozedion.com/complements-dinfo/

6.4 Exercices

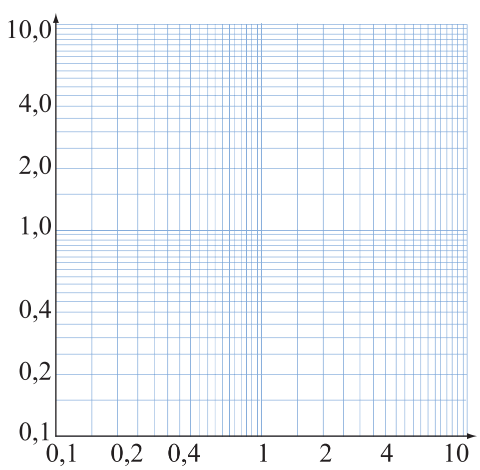
1. Représenter $f(x) = 1{,}8^x$ sur du papier semi-logarithmique.



2. Représenter $f(x) = 1/x$ sur du papier log-log.



3. Représenter $f(x) = 3x^2$ sur du papier log-log.



4. On a mesuré la vitesse angulaire N (en tours par minute) d'une roue d'entraînement à différents instants t (en minutes) après avoir coupé le courant.

Temps (min)	Vitesse angulaire, N (r/min)
0,5	22,75
1,0	12,80
1,5	7,24
2,0	4,10
2,5	2,32
3,0	1,31

a) Quel type de correspondance relie les deux variables ?
b) Déterminer la règle de correspondance entre les deux variables.

5. La pression atmosphérique (p, en kilopascals) dépend de l'altitude (h, en kilomètres) au-dessus du niveau de la mer. On a pris les mesures regroupées dans le tableau suivant.

Altitude (km)	Pression (kPa)
0,5	95,15
1,0	89,36
1,5	83,93
2,0	78,82
2,5	74,02
3,0	69,52
3,5	65,29
4,0	61,32

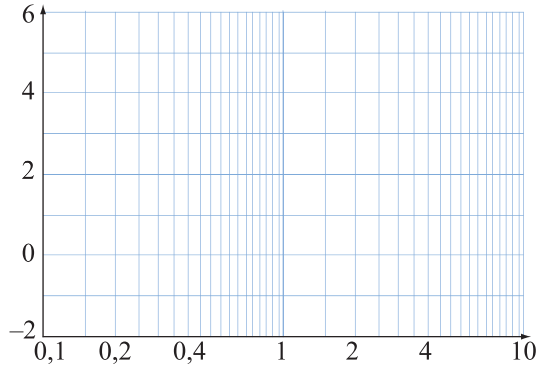
a) Quel type de correspondance relie les deux variables ?
b) Déterminer la règle de correspondance entre les deux variables.

6. On a obtenu les données suivantes en laboratoire.

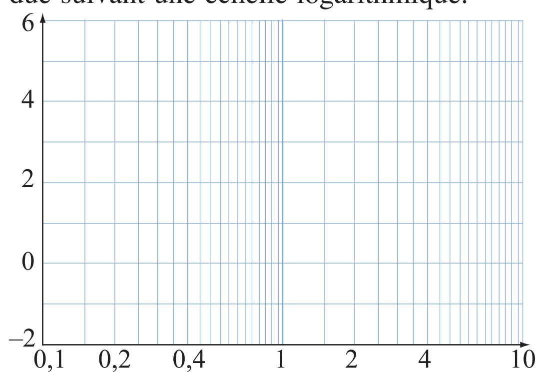
E	I
2,00	37,73
10,00	97,73
20,00	148,95
30,00	190,59
40,00	227,02
50,00	260,01
60,00	290,49

a) Quel type de correspondance relie les deux variables ?
b) Déterminer la règle de correspondance entre les deux variables.

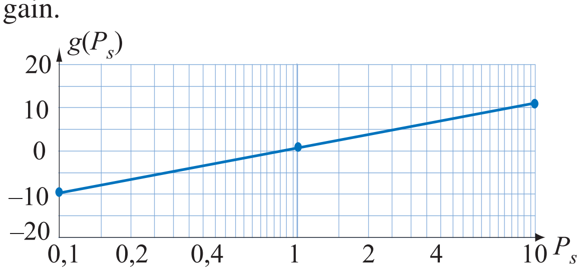
7. Représenter la fonction $f(x) = \ln x$ dans le système suivant dont l'axe horizontal est gradué suivant une échelle logarithmique.



8. Représenter la fonction $f(x) = 2 \log 2x$ dans le système suivant dont l'axe horizontal est gradué suivant une échelle logarithmique.



9. Le gain d'un composant électrique est illustré par le graphique suivant. Décrire algébriquement ce gain.



10. On a obtenu les données suivantes en laboratoire.

c_x	E
20	11,49
50	14,24
90	16,00
130	17,10
150	17,53
210	18,54

Déterminer, à l'aide d'une droite de régression le modèle qui décrit le mieux la relation entre les deux variables.

11. Un réservoir cylindrique de 12 m de hauteur et de 2,5 m de diamètre contient un liquide. On laisse s'écouler le liquide par une valve située à la base du cylindre. On note la vitesse d'écoulement du liquide et la hauteur de celui-ci dans le réservoir. Les données sont rassemblées dans le tableau suivant.

h (m)	v (m^3/min)
1	0,50
2	0,71
3	0,87
4	1,00
5	1,12
6	1,22
7	1,32

Déterminer à l'aide d'une droite de régression le modèle qui décrit le mieux la relation entre les deux variables et en déduire par extrapolation la vitesse d'écoulement du liquide quand le niveau est de 10 m.

12. On a mesuré le courant dans un circuit comprenant une source de tension constante et une résistance variable en faisant varier la résistance. On a obtenu les valeurs inscrites dans le tableau suivant :

Résistance (Ω)	Courant (A)
1,2	6,1
1,7	4,3
2,2	3,3
2,7	2,7
3,2	2,3
3,7	2,0

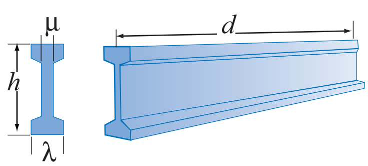
a) Quelle est la variable indépendante dans ce problème ?
b) Représenter graphiquement les données obtenues expérimentalement. Quel modèle mathématique la représentation graphique suggère-t-elle d'employer pour décrire la relation entre les deux variables ?
c) Décrire mathématiquement la relation en procédant par régression.

13. On a mesuré le volume occupé par 32 g d'ammoniac à 0 °C en faisant varier la pression exercée sur e gaz. On a obtenu les valeurs regroupées dans le tableau suivant.

Pression (kPa)	Volume (L)
10	420,0
20	210,0
40	105,0
60	70,0
80	52,5
100	42,0

a) Quelle est la variable indépendante dans ce problème ?
b) Représenter graphiquement les données obtenues expérimentalement. Quel modèle mathématique la représentation graphique suggère-t-elle d'employer pour décrire la relation entre les deux variables ?
c) Décrire mathématiquement la correspondance en procédant par régression.

14. On a fabriqué des poutres en I avec un nouveau matériau. La largeur μ de la bande centrale est le tiers de la largeur total λ. On veut déterminer la charge que peuvent supporter des poutres de même longueur et de même largeur mais de différentes épaisseur h, sans déformation.

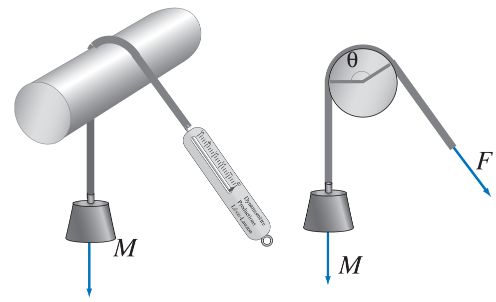


Le tableau suivant donne les mesures prises; la charge C est en kilogrammes (kg) et l'épaisseur h de la poutre en centimètres (cm).

Épaisseur (cm)	Charge (kg)
4	2 190
6	4 930
8	8 770
10	13 700
11	16 580
12	19 730
13	23 150

a) Montrer qu'il existe un lien entre la charge et le carré de l'épaisseur h d'une poutre.
b) Utiliser le modèle pour déterminer la charge que peut supporter une poutre de 9 cm d'épaisseur.

15. On a tenté d'établir la relation entre la force nécessaire pour équilibrer une masse M en utilisant une corde enroulée sur une poutre ronde, et l'angle d'enroulement θ de la corde.

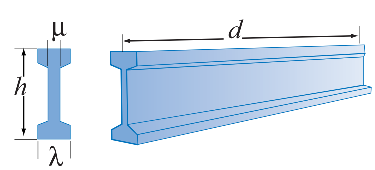


À l'aide d'un dynamomètre, on a mesuré la force minimale nécessaire pour équilibrer la masse M, qui exerce une force de 500 N selon l'angle d'enroulement de la corde sur la poutre. Les valeurs obtenues sont rassemblées dans le tableau suivant.

Angle (nombre de tours)	Force équilibrante
0,4	390
0,8	304
1,0	269
1,5	197
2,1	136
2,7	94
3,2	70
4,1	41
5,6	19

a) Montrer qu'il existe un lien exponentiel entre les deux variables.
b) Quelle force permettra d'équilibrer 800 N en faisant deux tours complets avec la corde?
c) Combien de tours faut-il faire pour que 50 N équilibrent 400 N ?

16. On a fabriqué des poutres en I ont été fabriquées avec un nouveau matériau. La largeur μ de la bande centrale est le tiers de la largeur totale λ. On veut déterminer la charge que peuvent supporter des poutres de même épaisseur et de même largeur, mais de différentes longueur d, sans déformation.



Le tableau suivant donne les mesures prises; la charge C est en kilogrammes (kg) et l'épaisseur h de la poutre en centimètres (cm).

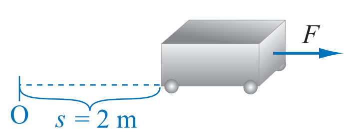
Longueur (m)	Charge (kg)
2	12 490
3	8 340
4	6 250
5	5 000
6	4 170
7	3 570
8	3 130
9	2 780
10	2 500

Montrer qu'il existe a un lien inversement proportionnel entre la charge et la longueur d d'une poutre.

Exercices de synthèse

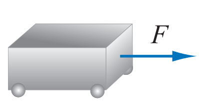
1. On a vidangé le réservoir illustré ci-contre et pour le remplir à nouveau l'opérateur ouvre la valve de la conduite principale. L'indicateur de débit donne une lecture de 15 L/min et l'opérateur referme la valve après 1 h 40 min.
   a) Calculer le volume de liquide dans le réservoir 60 minutes après l'ouverture de la valve.
   b) Construire un modèle décrivant le volume de liquide dans le réservoir en fonction du temps.
   c) Représenter graphiquement ce modèle.
   d) Le réservoir est-il plein après 1 h 40 min ?

2. Un mobile est à une distance de 2 m d'un point de référence fixe et s'en éloigne en ligne droite à une vitesse constante de 0,4 m/s.



   a) Déterminer la distance parcourue par le mobile en trois secondes.
   b) Déterminer et représenter graphiquement la fonction décrivant la position du mobile par rapport au point de référence au temps t.

3. Un mobile, initialement au repos, a une accélération constante de 0,2 m/s^2.



   a) Déterminer la variation de vitesse du mobile en trois secondes.
   b) Représenter graphiquement la fonction décrivant la vitesse au temps t.

4. Pour répondre à la demande en période de forte consommation d'eau potable, une municipalité a fait installer un réservoir près de l'usine de traitement des eaux. Ce réservoir devrait permettre d'accumuler des réserves durant les périodes de moindre consommation et de répondre à la demande durant les périodes de forte consommation.
   a) Le réservoir étant initialement vide (il vient d'être installé), on ouvre une vanne qui a un débit de 150 litres à la minute. Représenter graphiquement la fonction débit.
   b) Représenter graphiquement et évaluer le volume de liquide dans le réservoir au bout de 4 min.
   c) Représenter graphiquement et évaluer le volume de liquide dans le réservoir en fonction du temps t.
   d) Représenter graphiquement la fonction décrivant le volume de liquide au temps t.

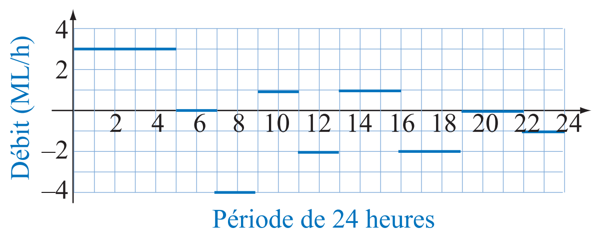
6. Un concierge a détecté une fuite dans la plomberie. Il a installé un récipient pour recueillir l'eau qui s'échappe par cette fuite. Après 40 minutes, il mesure la quantité d'eau dans le récipient et constate qu'il s'est accumulé 1,2 litre d'eau.
   a) Quel est le débit de cette fuite en litres par minute ?
   b) Quel est le modèle mathématique décrivant la quantité d'eau dans le récipient en fonction du temps ?

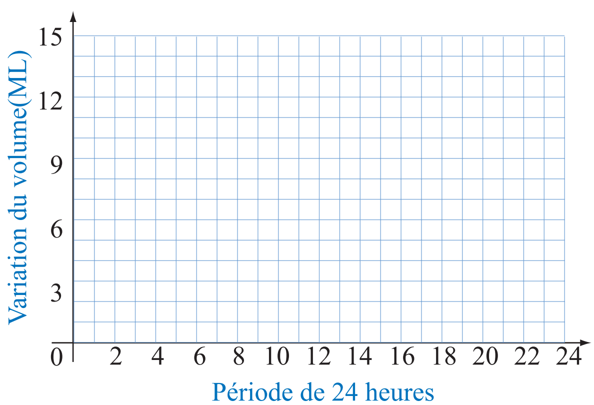
6. Une automobile se déplace à 60 km/h. Le conducteur freine et l'automobile s'arrête en 20 s. Quelle est l'accélération moyenne de l'automobile en m/s^2 durant cet intervalle de temps ?

7. On laisse tomber un caillou du sommet d'un édifice. Sachant que l'accélération due à l'attraction terrestre est de 9,8 m/s^2, trouver sa vitesse quatre secondes après le début de la chute.

8. Un réservoir d'une capacité de 250 kL est plein. On ouvre la valve d'une conduite de vidange qui a un débit de 15 L/min. Décrire le volume de liquide dans le réservoir en fonction du temps.

9. Pour répondre à la demande en période de forte consommation d'eau potable, une municipalité a fait installer un réservoir de 15 ML près de l'usine de traitement des eaux. Ce réservoir devrait permettre d'accumuler des réserves durant les périodes de moindre consommation et de répondre à la demande durant les périodes de forte consommation.





   a) Le graphique suivant donne le débit net pour une période de 24 heures. Représenter graphiquement la variation du volume de liquide durant cette période.
   b) Quel est le volume de liquide dans le réservoir à 4 h du matin ? à 12 h ? à 21 h ?
   c) Quel est le débit moyen durant l'intervalle [0; 6] ? durant l'intervalle [0; 12] ? durant l'intervalle [0; 24] ?
   d) Selon vous, le réservoir a-t-il une capacité suffisante pour répondre à la demande ?

10. Un mobile a une accélération uniforme de 0,5 m/s^2 pendant 40 secondes, après quoi son accélération est nulle.
   a) Représenter graphiquement la fonction décrivant l'accélération durant les 60 premières secondes.
   b) Sachant que sa vitesse initiale était nulle, déterminer la fonction décrivant la vitesse au temps t pendant les 60 premières secondes.
   c) Représenter graphiquement cette fonction.
   d) Quelle est la relation entre les graphiques de la fonction accélération et de la fonction vitesse ?
   e) Quelle est la vitesse atteinte à 5 s ? à 18 s ?

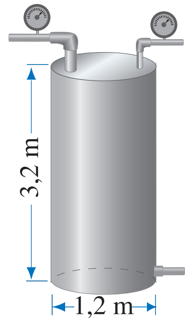
11. Une moto se déplace à une vitesse de 15 m/s. Si pendant quatre secondes la moto a une accélération de 6 m/s^2, quelle sera sa vitesse finale ?

12. On laisse tomber une pierre du haut d'un édifice et on mesure la durée de la chute. Sachant que celle-ci a été de 2,7 secondes et que l'accélération due à l'attraction terrestre est de 9,8 m/s^2, trouver la vitesse de la pierre au moment de l'impact au sol.

13. Une conduite sert à acheminer un liquide dans un réservoir vide pouvant contenir 375 m^3. Le débit dans cette conduite est de 15 m^3/min.
   a) Représenter graphiquement la fonction décrivant le débit au temps t en minutes pendant la période de remplissage.
   b) Quel est le domaine de validité du modèle ?
   c) Définir la fonction décrivant le volume de liquide dans le réservoir au temps t pendant la période de remplissage.
   d) Représenter graphiquement la fonction volume et donner son domaine de validité.
   e) Quelle relation peut-on établir entre la fonction débit et la fonction volume ?

14. Un réservoir vide de 2 200 litres peut être alimenté par deux conduites. L'une de ces conduites a un débit de 40 litres à la minute et la deuxième, un débit de 20 litres à la minute. La valve de la première conduite est ouverte dix minutes avant celle de la deuxième conduite.
   a) Représenter graphiquement la fonction décrivant le débit total au temps t pendant la période de remplissage.
   b) Quel est le domaine de validité du modèle ?
   c) Définir la fonction décrivant le volume de liquide dans le réservoir au temps t pendant la période de remplissage.
   d) Représenter graphiquement la fonction volume et donner son domaine de validité.
   e) Quelle relation peut-on établir entre la fonction débit et la fonction volume ?
   f) Quel est le volume de liquide dans le réservoir au bout de 5 minutes ? au bout de quinze minutes ?

15. Un réservoir de 2 200 litres est rempli à capacité. On décide de procéder à la vidange du réservoir pour éliminer les résidus accumulés dans le fond du réservoir. On met en marche le système de vidange à un débit de 60 L/min. Au bout de 30 minutes, on diminue le débit à 20 L/min.

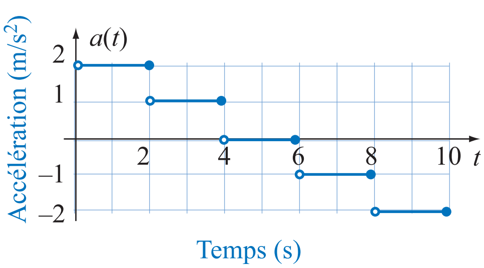


a) Représenter graphiquement la fonction décrivant le débit total au temps *t* pendant la période de vidange.
b) Quel est le domaine de validité du modèle ?
c) Définir la fonction décrivant le volume de liquide dans le réservoir au temps *t* pendant la période de vidange.
d) Représenter graphiquement la fonction volume et donner son domaine de validité.
e) Quelle relation peut-on établir entre la fonction débit et la fonction volume ?

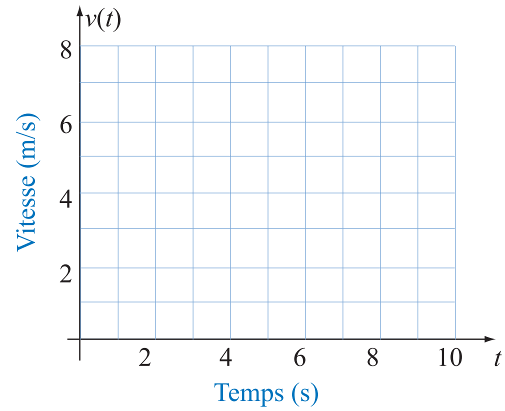
16. Un mobile, initialement au repos, accélère à un taux de 0,9 m/s² pour atteindre une vitesse finale de 32 m/s. Quelle était sa vitesse 8 secondes avant d'atteindre la vitesse finale ?

17. Une voiture roule à 28 m/s et prend 12 secondes pour s'arrêter. Quelle est son accélération pendant le temps de freinage ?

18. Le graphique ci-dessous représente l'accélération en m/s² d'un mobile au temps *t*.

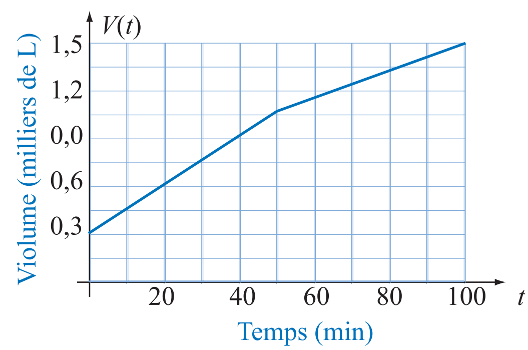


a) Sachant que la vitesse initiale est nulle, évaluer la vitesse du mobile à trois secondes.
b) Esquisser le graphique de la vitesse en fonction du temps *t*.

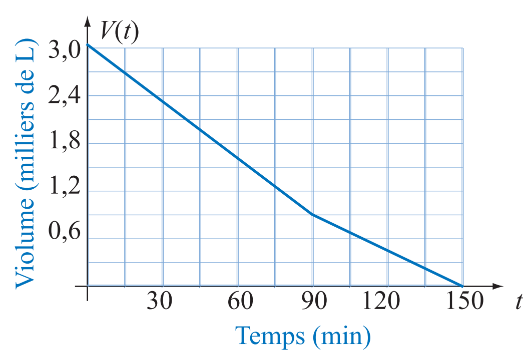


c) Quelle est l'accélération moyenne durant l'intervalle [0;4] ? durant l'intervalle [2;6] ?

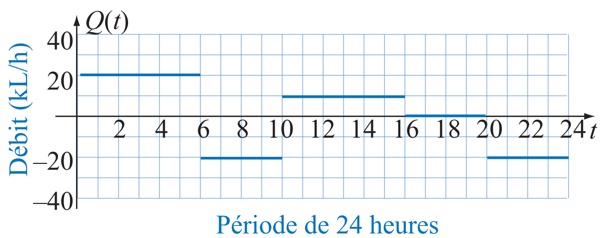
19. Le graphique ci-dessous décrit le volume de liquide dans un réservoir durant une période de remplissage. Exprimer algébriquement la fonction décrivant le volume de liquide dans le réservoir au temps *t*.



20. Le graphique ci-dessous décrit le volume de liquide dans un réservoir durant une période de vidange. Exprimer algébriquement la fonction décrivant le volume de liquide dans le réservoir au temps *t*.



21. Le graphique ci-dessous décrit le débit de la conduite d'alimentation d'un réservoir durant une journée.



a) Décrire algébriquement le débit en fonction du temps durant cette journée.

b) Sachant qu'au début de la journée le volume de liquide dans le réservoir était de 250 kL, exprimer algébriquement la fonction décrivant le volume de liquide dans le réservoir au temps t durant la journée.

22. Utiliser le programme pour déterminer si le lien entre les variables données dans les tableaux suivants est descriptible par un modèle affine et trouver les paramètres de ce modèle, le cas échéant.

a)

x	y
0	51,9
10	57,8
20	63,7
30	69,2
40	74,0
50	78,7
60	83,8
70	88,6

b)

x	y
2,4	0,85
6,1	2,11
7,2	2,51
11,6	4,01
17,1	5,86
21,8	7,35
30,4	10,08

23. On a réalisé une expérience qui consistait à plonger un manomètre dans un récipient rempli de liquide pour mesurer la pression à différentes profondeurs. Les données suivantes indiquant la pression absolue en fonction de la profondeur ont été recueillies.

h	p
2	121,54
3	131,61
4	141,70
5	151,80
6	161,89
7	171.98
8	182,08
9	192,17
10	202,26
11	212,36

a) Décrire mathématiquement le lien entre ces variables.

b) Montrer que ces données permettent de vérifier la loi fondamentale de l'hydrostatique qui s'énonce ainsi
La différence de pression entre deux niveaux d'un liquide est égale au poids d'une colonne de liquide ayant pour section l'unité de surface et pour hauteur la différence des niveaux.

24. On a mesuré la longueur d'une tige d'acier à différentes températures pour étudier son élongation. On a obtenu les données du tableau suivant.

T	L
10	25,64806
20	25,6512
30	25,65394
40	25,65688
50	25,65982
60	25,66276

a) Décrire mathématiquement la relation entre les variables.

b) Montrer que l'élongation de la tige est proportionnelle à la variation de température.

25. On a mesuré le volume d'un liquide à différentes températures pour en étudier la dilatation. On a obtenu les données du tableau suivant.

T	V
–20	1,5211
–10	1,5324
0	1,5437
10	1,5550
20	1,5663
30	1,5776
40	1,5889
50	1,6002
60	1,6115
70	1,6280
80	1,6341
90	1,6454

a) Décrire mathématiquement la relation entre les variables.

b) Montrer que la dilatation du volume du liquide est proportionnelle à la variation de température.