Chapitre 09 - Corrélation et causalité

Pour tous le cours, on utilisera la série statistique suivante, qui correspond au nombre de personnes en emploi selon l'année (données INSEE):

Année	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**	**1986**
Nombre de personnes employées (en milliers)	22 193	22 365	22 538	22 557	22 814	22 985	22 854	22 842	22 806	22 612	22 646	22 850

Année	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
Nombre de personnes employées (en milliers)	22 857	22 946	23 278	23 359	23 268	23 231	23 104	23 031	23 319	23 448	23 363	23 612

Année	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Nombre de personnes employées (en milliers)	23 892	24 548	24 906	25 166	25 200	25 304	25 491	25 671	26 115	26 468	26 204	26 282

Année	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Nombre de personnes employées (en milliers)	26 374	26 350	26 377	26 442	26 580	26 803	27 022	27 133	26 995

I - Série statistique à deux variables

Définition

Lorsqu'on étudie conjointement deux caractères (ou variables) x et y sur une même population de taille n, on associe à chaque individu de la population un couple $(x_i; y_i)$, où x_i et y_i sont les valeurs respectives des variablse x et y prises par l'individu numéro i (où i est un nombre entier entre 1 et n, ou parfois entre 0 et $n-1$).

On appelle série statistique double (x, y) l'ensemble des couples $(x_i; y_i)$ associés à chaque individu de la population. On la présente en général dans un tableau.

Valeurs x_i de la variable x	x_1	x_2	**...**	x_i	**...**	x_n
Valeurs y_i de la variable y	y_1	y_2	...	y_i	...	y_n

Exemple

La série statistique de l'introduction est-elle une série statistique double?

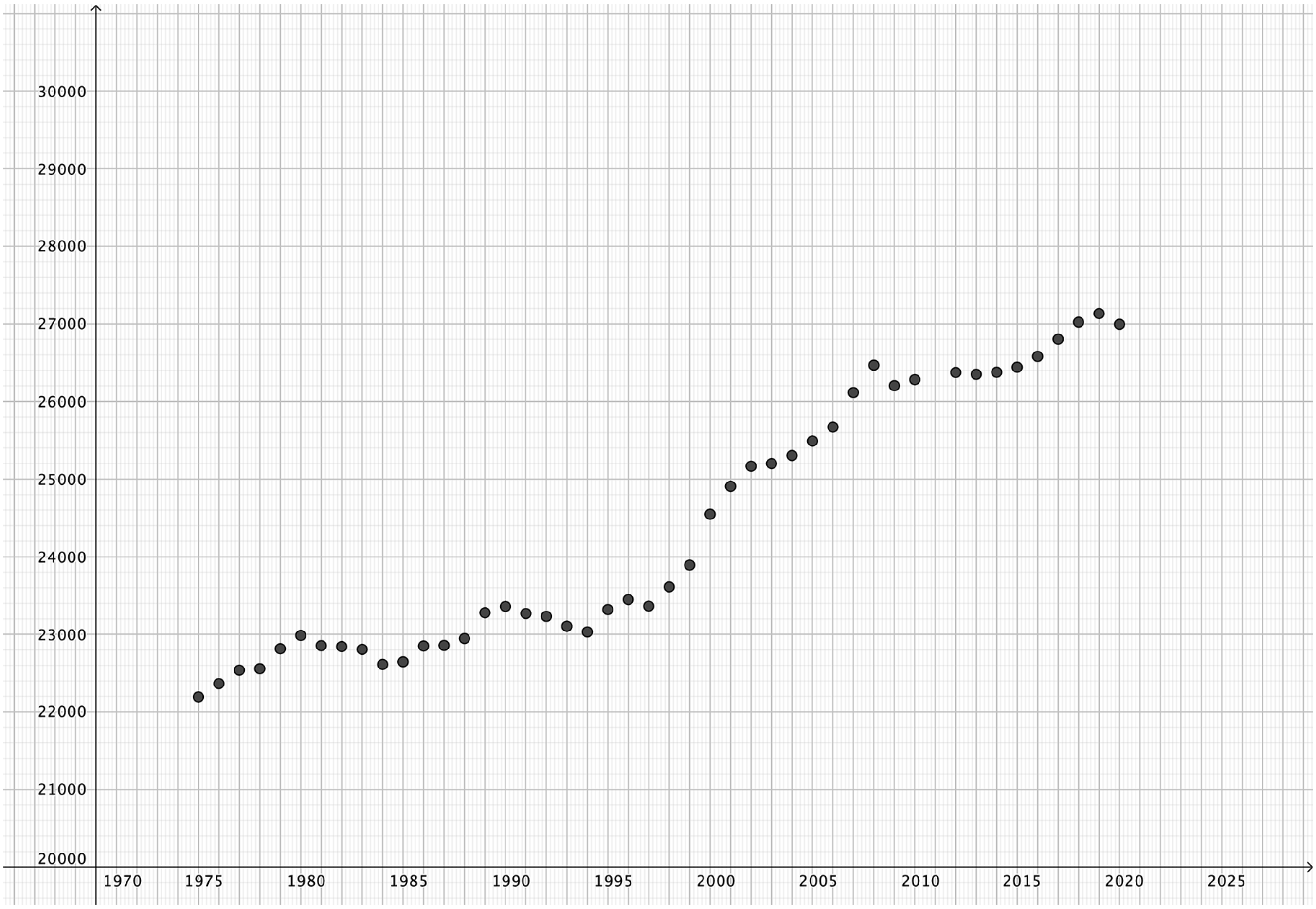
> En posant x la variable correspondant à l'année et y la varible correspondant au nombre de personnes employées (en milliers), on peut considérer que la série statistique est une série statistique double.

Définition

A chaque couple $(x_i; y_i)$ de la série statistique double (x, y), on peut associer le point M_i de coordonnées $(x_i; y_i)$ dans un repère. L'ensemble de ces points est appelé nuage de points associé à la série statistique double (x, y).

Exemple

Tracer le nuage de points correspondant à la série statistique de l'introduction.

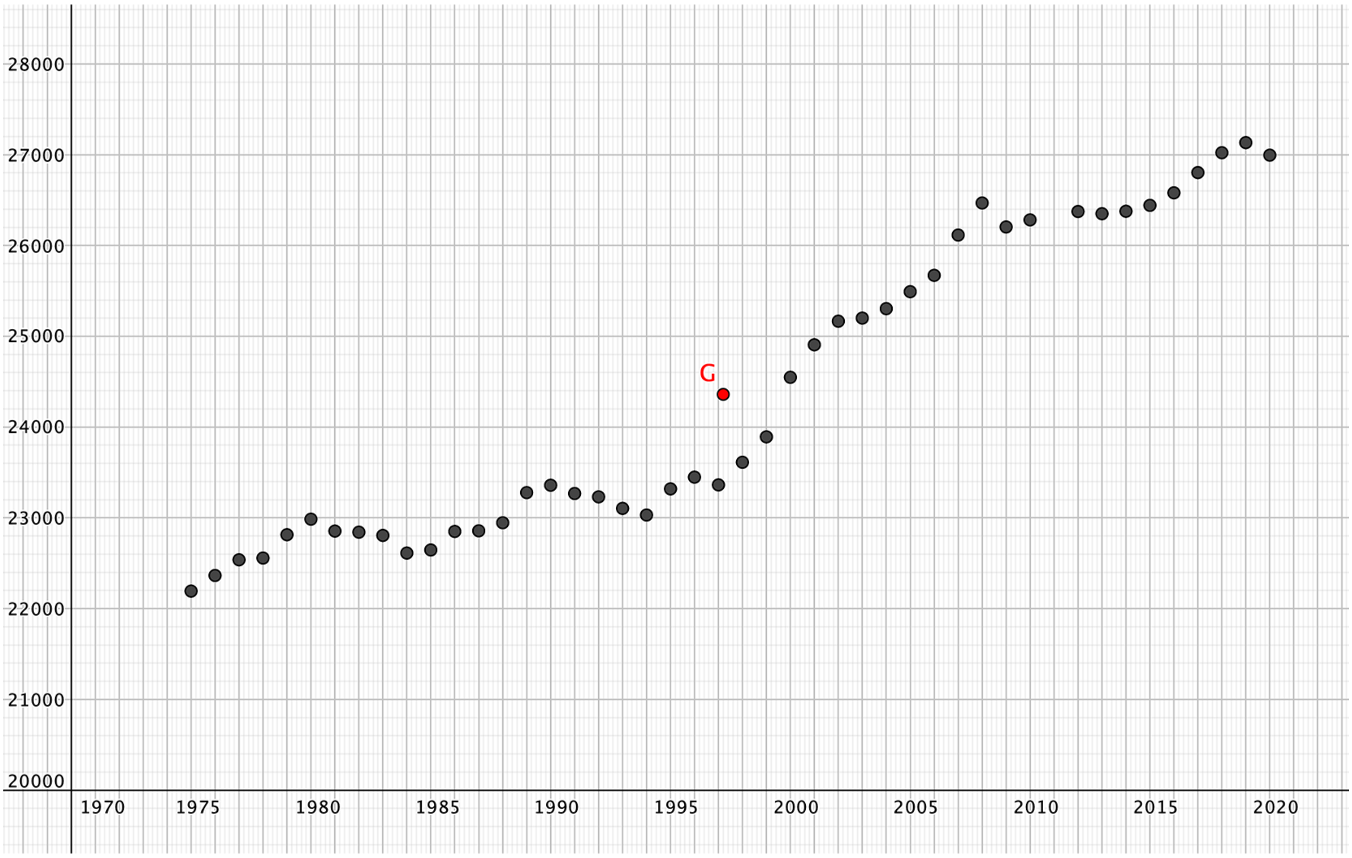


Définition

On appelle point moyen d'un nuage de point, le point G de coordonnées $\left(\overline{x}; \overline{y}\right)$ où $\overline{x}$ est la moyenne des abscisses des points du nuage et $\overline{y}$ est la moyenne des ordonnées des points du nuage.

Exemple

Déterminer les coordonnées du point moyen du nuage de points. Vous arrondirez les résultats au dixième

> $$\overline{x} = \frac{\sum_{i=1}^{46} x_i}{46} = \frac{1975 + 1976 + \ldots + 2020}{46} = 1997,2$$
>
> $$\overline{y} = \frac{\sum_{i=1}^{46} y_i}{46} = \frac{22193 + 22365 + \ldots + 26995}{46} = 24360,02$$
>
> Donc le point moyen G a pour coordonnées $G(1997,2; 24360,02)$.
>
> 
>

Propriété

On appelle covariance des séries statistiques x et y le nombre, noté $var(x, y)$ ou σ_{xy} défini par:

$$cov(x, y) = \sigma_{xy} = \frac{(x_1 - \overline{x})(y_1 - \overline{y}) + (x_2 - \overline{x})(y_2 - \overline{y}) + \ldots + (x_n - \overline{x})(y_n - \overline{y})}{n} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y})$$

Exemple

Calculer la covariance de x et y.

> $$cov(x, y) = \frac{(1975-1997,2)(22193-24360,02)+(1976-1997,2)(22365-24360,02)+\ldots+(2020-1997,2)(26995-24360,02)}{46} \approx 20787,17$$

Propriété

On a aussi $cov(x, y) = \frac{x_1y_1 + x_2y_2 + \ldots + x_ny_n}{n} - \overline{x}\overline{y} = \frac{1}{n}\left(\sum_{k=1}^{n} x_ky_k\right) - \overline{x}\overline{y}$

Démonstration

> $$cov(x, y) = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y}) = \frac{1}{n}\sum_{k=1}^{n}(x_ky_k - x_k\overline{y} - \overline{x}y_k + \overline{x}\overline{y}) = \frac{1}{n}\sum_{k=1}^{n}(x_ky_k) - \frac{1}{n}\sum_{k=1}^{n}(x_k\overline{y}) - \frac{1}{n}\sum_{k=1}^{n}(\overline{x}y_k) + \frac{1}{n}\sum_{k=1}^{n}(\overline{x}\overline{y})$$

> $cov(x,y) = \frac{1}{n}\sum_{k=1}^{n}(x_k y_k) - \overline{y} \times \frac{1}{n}\sum_{k=1}^{n}(x_k) - \overline{x} \times \frac{1}{n}\sum_{k=1}^{n}(y_k) + \overline{x}\overline{y} \times \frac{1}{n}\sum_{k=1}^{n} 1 = \frac{1}{n}\sum_{k=1}^{n}(x_k y_k) - \overline{x}\overline{y} - \overline{x}\overline{y} + \overline{x}\overline{y} \times \frac{n}{n} = \frac{1}{n}\sum_{k=1}^{n}(x_k y_k) - \overline{x}\overline{y}$

Exemple

Vérifier avec cette formule la covariance de x et y.

> > $cov(x,y) = \frac{1975 \times 22193 + 1976 \times 22365 + ... + 2020 \times 26995}{46} - 1997,2 \times 24360,02 \approx 20787,17$

II - Ajustement affine

Définition

Lorsque les points d'un nuage sont "presque" alignés, on peut imaginer une droite qui passe au plus près des points du nuage. Cette droite, appelée droite d'ajustement, réalise un ajustement affine du nuage de points.

Remarque

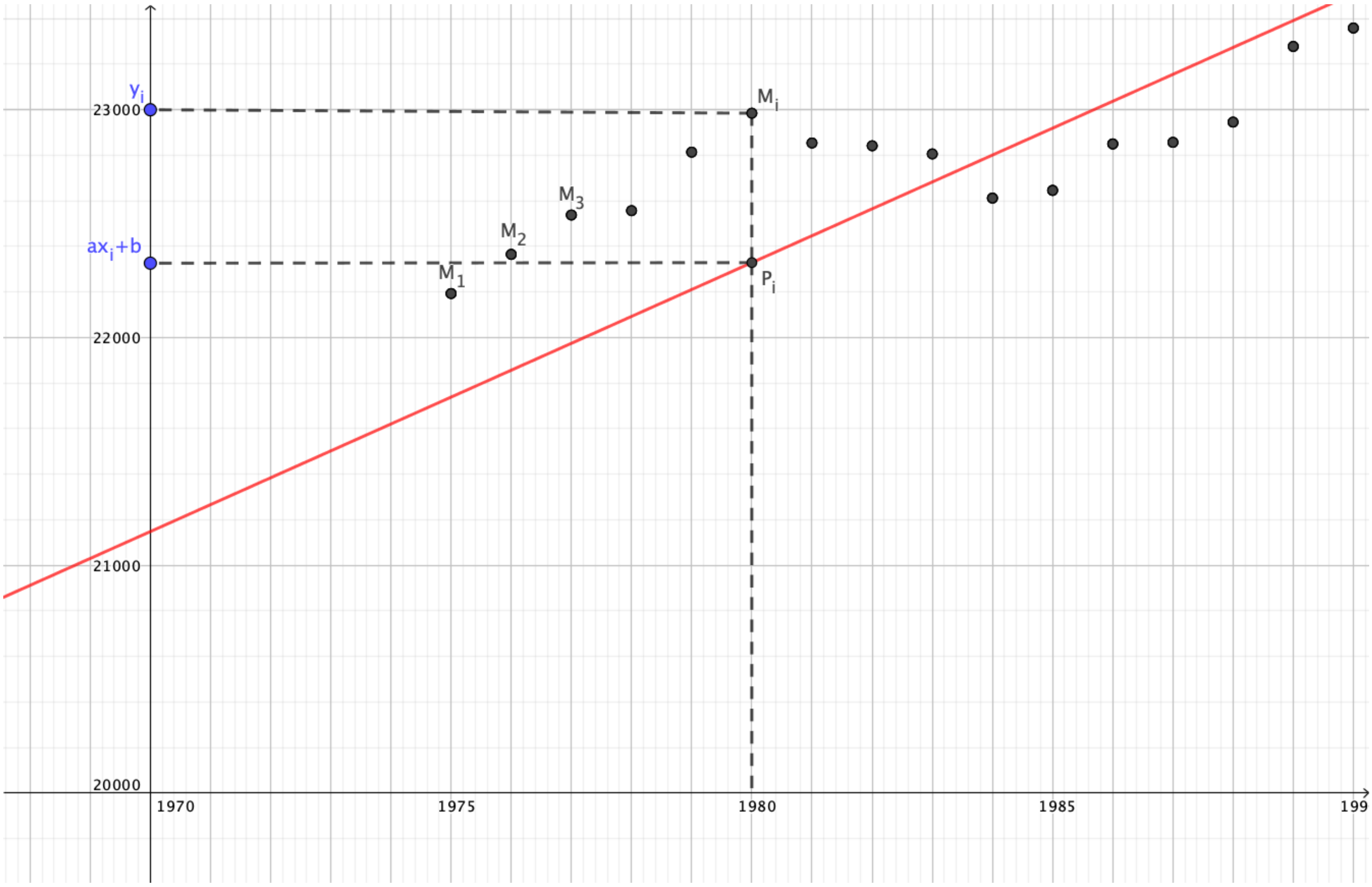
Lorsqu'on a fait un ajustement affine, on peut exprimer y en fonction de x ou x en fonction de y.

En général, la droite d'ajustement passe par le point moyen G du nuage.

Méthode des moindres carrés

Soit un nuage de points $M_i(x_i; y_i)_{1 \leq i \leq n}$, une droite d'équation $y = ax + b$ qui ajuste le nuage, et les points $P_i\ (x_i; ax_i + b)$ sur la drotie (d). Il existe deux réels a et b tels que la somme $M_1P_1^2 + M_2P_2^2 + ... + M_nP_n^2$ soit minimale.

Pour ces valeurs de a et b, la droite d'équation $y = ax + b$ est appelée droite des moindres carrés associée au nuage de points $M_i(x_i; y_i)_{1 \leq i \leq n}$, ou droite de régression de y en x.



Propriété

Soit un nuage de point $M_i(x_i; y_i)_{1 \leq i \leq n}$ de point moyen $G\left(\overline{x}; \overline{y}\right)$.

La droite des moindres carrés de ce nuage de points a pour équation $y = ax + b$ avec $a = \frac{cov(x,y)}{var(x)}$ et $b = \overline{y} - a\overline{x}$.

Démonstration

> Soit un nuage de point $M_i(x_i; y_i)_{1 \leq i \leq n}$ de point moyen $G\left(\overline{x}; \overline{y}\right)$.
>
> Soit une droite d'équation $y = ax + b$ qui ajuste le nuage, et les points $P_i\ (x_i; ax_i + b)$ sur la drotie (d).
>
> Soit $i \in [\![1; n]\!]$.
>
> $M_iP_i^2 = (y_i - (ax_i + b))^2 = (y_i - ax_i - b)^2$
>
> $\sum_{i=1}^{n} M_iP_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$
>
> Il faut donc trouver a et b pour que $\sum_{i=1}^{n} M_iP_i^2$ soit minimal.
>
> 1ère étape:
>
> Posons $f(b) = \sum_{i=1}^{n} M_iP_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$.

On cherche donc le minimum de f.

Si f admet un minimum en b_0 alors $f'(b_0) = 0$.

$$f'(b) = -2\sum_{i=1}^{n}(y_i - ax_i - b)$$

Résolvons donc $-2\sum_{i=1}^{n}(y_i - ax_i - b) = 0$ soit encore $\sum_{i=1}^{n}(y_i - ax_i - b) = 0$

On a alors:

$$\sum_{i=1}^{n}(y_i) - a\sum_{i=1}^{n}(x_i) - \sum_{i=1}^{n}(b) = 0$$

$$\sum_{i=1}^{n}(y_i) - a\sum_{i=1}^{n}(x_i) - bn = 0$$

$$bn = \sum_{i=1}^{n}(y_i) - a\sum_{i=1}^{n}(x_i)$$

$$b = \tfrac{1}{n}\sum_{i=1}^{n}(y_i) - a\frac{1}{n}\sum_{i=1}^{n}(x_i) = \overline{y} - a\overline{x}$$

2ème étape:

$$\sum_{i=1}^{n}M_iP_i^2 = \sum_{i=1}^{n}(y_i - ax_i - b)^2 = \sum_{i=1}^{n}\left(y_i - ax_i - \left(\overline{y} - a\overline{x}\right)\right)^2$$

$$\sum_{i=1}^{n}M_iP_i^2 = \sum_{i=1}^{n}\left(y_i - \overline{y} - ax_i + a\overline{x}\right)^2 = \sum_{i=1}^{n}\left(\left(y_i - \overline{y}\right) - a\left(x_i - \overline{x}\right)\right)^2$$

Posons $g(a) = \sum_{i=1}^{n}\left(\left(y_i - \overline{y}\right) - a\left(x_i - \overline{x}\right)\right)^2$

On cherche donc le minimum de g.

Si g admet un minimum en a_0 alors $g'(a_0) = 0$.

$$g'(a) = -2\sum_{i=1}^{n}\left[\left(x_i - \overline{x}\right)\left(\left(y_i - \overline{y}\right) - a\left(x_i - \overline{x}\right)\right)\right]$$

Résolvons $g'(a) = 0$

$$-2\sum_{i=1}^{n}\left[\left(x_i - \overline{x}\right)\left(\left(y_i - \overline{y}\right) - a\left(x_i - \overline{x}\right)\right)\right] = 0$$

$$\sum_{i=1}^{n}\left[\left(x_i - \overline{x}\right)\left(\left(y_i - \overline{y}\right) - a\left(x_i - \overline{x}\right)\right)\right] = 0$$

> $$\sum_{i=1}^{n}\left[\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)-a\left(x_i-\overline{x}\right)\left(x_i-\overline{x}\right)\right]=0$$
>
> $$\sum_{i=1}^{n}\left[\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)\right]-a\sum_{i=1}^{n}\left[\left(x_i-\overline{x}\right)\left(x_i-\overline{x}\right)\right]=0$$
>
> $$\frac{1}{n}\sum_{i=1}^{n}\left[\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)\right]-a\frac{1}{n}\sum_{i=1}^{n}\left[\left(x_i-\overline{x}\right)^2\right]=0$$
>
> $cov(x,y)-a\times var(x)=0$
>
> $a\times var(x)=cov(x,y)$
>
> $a=\frac{cov(x,y)}{var(x)}$
>
> Remarque: Le point $G\left(\overline{x};\overline{y}\right)$ appartient-il à la droite (d)?
>
> Calculons $a\overline{x}+b$.
>
> $a\overline{x}+b=a\overline{x}+\overline{y}-a\overline{x}=\overline{y}$
>
> Donc le point $G\left(\overline{x};\overline{y}\right)$ appartient à la droite (d)

Exemple

Déterminer l'équation de la droite de régression de y en x de notre nuage de points.

>> $Var(x)=\frac{1975^2+1976^2+...+2020^2}{46}-1997,5^2=176,03$
>>
>> $a=\frac{cov(x,y)}{var(x)}=\frac{20787,17}{176,03}=118,09$
>>
>> On rappelle que $G(1997,2;24402,7)$
>>
>> Donc $b=24360,02-118,09\times 1997,2=-211491,42$
>>
>> Ainsi la droite des moindres carrés du nuage de points à pour équation:
>>
>> $y=118,09x-211491,42$

Remarque

La calculatrice permet de trouver l'équation de la droite de régression de y en x:

- Effacer les listes L_1 et L_2 : `stats / EffListe` L_1 `,` L_2 `Enter`
- Rentrer les valeurs dans les listes L_1 et L_2 : stats / Edit / Modifier
- Faire les calculs: `stats / CALC / RégLin(ax+b)`
    - XListe : L_1
    - YListe : L_2
    - Calculer

S'affichent alors le coefficient directeur a et l'ordonnée à l'origine b de la droite de régression.

Définition

Le coefficient de corrélation linéaire d'une série statistique double de variable x et y est le nombre r défini par:

$r = \frac{cov(x,y)}{\sigma(x)\sigma(y)}$

Exemple

Calculer le coefficient de corrélation linéaire de notre série statistique.

> $Var(x) = \frac{1975^2+1976^2+...+2020^2}{46} - 1997,5^2 = 176,03$
>
> Donc $\sigma(x) = \sqrt{176,03} \approx 13,27$
>
> $Var(y) = \frac{22193^2+22365^2+...+26995^2}{46} - 24402,63^2 = 2665523,93$
>
> Donc $\sigma(y) = \sqrt{2665523,93} \approx 1632,64$
>
> $r = 20787,1713,27 \times 1632,64 \approx 0,96$

Remarque

On a toujours $-1 \leq r \leq 1$.

- Plus $|r|$ est proche de 1, plus l'ajustement affine est un bon modèle de corrélation entre les variables x et y.
- Plus $|r|$ se rapproche de 0, moins l'ajustement affine n'a de sens.
- Si $|r| = 1$ alors la droite de régression passe par tous les points du nuage.

Exemple

> Dans notre exemple, le coefficient de corrélation est proche de 1 donc la droite des moindres carrés est un bon modèle de corrélation entre les variables les années (x) et le nombre de personnes employées (y).

Ajustement exponentiel/Changement de variable pour se ramener à un ajustement affine

Parfois, le nuage de points $M_i(x_i; y_i)_{1\leq i\leq n}$ a une forme qui se rapproche d'une courbe exponentielle.

On fait alors un changement de variable en posant $Y = ln(y)$. Ainsi on étudie le nuage de points $N_i(x_i; Y_i)_{1\leq i\leq n}$. Ce nouveau nuage de points se rapprocha alors d'une forme rectiligne. On peut donc l'ajuster par la droite de régression de Y en x d'équation $Y = ax + b$.
On a alors $ln(y) = ax + b$, soit $y = e^{ax+b}$.

Cette équation est celle d'une courbe exponentielle qui ajuste au mieux le nuage de points $M_i(x_i; y_i)_{1\leq i\leq n}$.

On dit qu'on a réalisé un ajustement exponentiel du nuage de points.

Définition

Lorsqu'on a fait un ajustement affine ou exponentiel d'un nuage de points $M_i(x_i; y_i)_{1\leq i\leq n}$, on peut le considéré comme un modèle et l'utiliser pour faire une estimation de la quantité de la valeur y correspondant à une valeur x donnée.

Dans le cas où la valeur donnée de x reste dans l'intervalle des $[min(x_i); max(x_i)]$, cette estimation est une **interpolation**. Dans le cas contraire, c'est une **extrapolation**.

Exemple

Toujours dans notre exemple:
1. Estimer le nombre de personnes employées en 2010, puis en 2025.
2. A partir de quelle année pourrait-on avoir plus de 30 millions de personnes employées?

> > 1. La droite de régression de y en x a pour équation $y = 118,09x - 211491,42$.
> > Si $x = 2010$, alors $y = 118,09 \times 2010 - 211491,42 = 25869,48$.
> > Par cette interpolation, nous pouvons estimer que le nombre de personnes employées en 2010 était de 25869,48 milliers de personnes.
> > Si $x = 2025$, alors $y = 118,09 \times 2025 - 211491,42 = 27640,42$.
> > Par cette extrapolation, nous pouvons estimer que le nombre de personnes employées en 2025 sera de 27640,42 milliers de personnes.
> > 2. Résolvons $y \geq 30000$.
> > $118,09x - 2111491,42 \geq 30000$
> > $118,09x \geq 30000 + 211491,42$
> > $118,09x \geq 241491,42$ $x \geq \frac{241491,42}{118,09}$
> > $x \geq 2044,97$
> > A partir de 2045, on peut supposer qu'il y aura plus de 30 millions de personnes employées.

III - Quelques problèmes

1. Ajustement de Mayer (37 p 239)

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 rpésentés dans le tableau suivant.

Age 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
Age 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 commensera 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 les coordonnées du point moyen G_1 du premier sous-nuage.
   b. Déterminer les coordonnées du point moyen G_2 du deuxième sous-nuage.
   c. Placer les points G_1 et G_2 dans le repère et tracer la droite (G_1G_2). Cette droite est appelée droite de Mayer.
   d. Déterminer l'équation réduite de la droite (G_1G_2).
3. a. Déterminer les coordonnées du point moyen G du nuage de points.
   b. Le point G appartient-il à la droite de Mayeur?

2. Loi de désintégration radioactive

Suite à la catastrophe de Fukushima (2011), des aliments ont été contaminés par la radioactivité produite. Le taux Y; de césium 137, en becquerel par kg, a été mesuré dans le riz cultivé dans les communes proches de la centrale.

Ces mesures sont présentées dans le tableau ci-dessous.

Année	Rang x_i	y_i
2011	1	2000
2012	2	1960
2013	3	1900
2014	4	1870
2015	5	1825
2016	6	1785
2017	7	1740

On se propose de comparer les estimations de radioactivité en 2050 selon différents modèles utilisés.

1. ***Un ajustement affine***
   a. A l'aide du tableur, saisir cette série dans la plage A1:H2 et représenter le nuage de points $M_i\ (x_i; y_i)$ avec $1 \leq i \leq 7$.
   Afficher une courbe de tendance de type « linéaire » ainsi que son équation et le coefficient de détermination R^2 .
   b. En déduire le coefficient de corrélation linéaire rentre x et y. Interpréter ce coefficient. *Arrondir au millième.*
   c. Si l'évolution se poursuivait ainsi, estimer le taux de césium 137 en 2050. La région serait-elle alors considérée comme« légalement sûre»?
2. ***Le regard du physicien***
   En physique, on établit que la relation entre x et y est du type $y = ke^{-\lambda x}$ avec k et λ nombres réels, $\lambda > O$. On pose $z_i = ln(y_i)$.
   a. A l'aide du tableur, saisir les valeurs de z_i dans la plage A3:H3 et représenter le nuage de points $(x_i; z_i)$.
   b. Afficher la courbe de tendance de type «linéaire», ainsi que le coefficient R^2 . En déduire le coefficient de corrélation linéaire r' entre x et z. *Arrondir au millième.*
   c. Déterminer alors des valeurs approchées de k et λ.
   d. Si l'évolution se poursuivait ainsi, estimer le taux de césium 137 en 2050.
   Comparer avec la réponse obtenue à la question 1.c.

Info: les dernières mesures effectuées sur du riz récolté dans les aires des communes de Fukushima et Date ont révélé une teneur en césium radioactif supérieure à la limite légale de 500 becquerels par kilogramme fixée par le gouvernement central de Tokyo. Certains prélèvements ont affiché une teneur en césium 137 près de quatre fois supérieure au plafond, soit près de 2 000 becquerels par kilogramme.

3. Evolution de la température et des émissions de gaz à effet de serre dans le cadre du réchauffement climatique

Partie A - Etude de la concentration moyenne de CO_2 dans l'atmosphère

Le CO_2 est le principal gaz à « effet de serre ». Depuis 2015, les concentrations moyennes de CO_2 dans l'ensemble de l'atmosphère terrestre sont supérieures à 400 ppm(« partie par million »). Comme son nom l'indique, le ppm permet de savoir combien de molécules de CO_2 on trouve sur un million de molécules d'air. Il permet donc de rendre compte de la quantité de CO_2 dans une masse d'air donnée. Les prévisions estiment que d'ici 2030, le taux de CO_2 moyen dans l'atmosphère atteindra les 450 ppm.

1. Télécharger et ouvrir le fichier Excel suivant: lienmini.fr/maths-t09-01.
   Il donne la concentration de CO_2 dans l'atmosphère et la température moyenne mondiale chaque année de 1900 à 2015.
   Sélectionner les données des colonnes A et B, puis insérer un graphique **nuage de points**.
2. Ouvrir la fenêtre **Format de l'axe** à l'aide d'un double clic sur les valeurs de l'axe des abscisses, puis indiquer 1900 comme minimum et 2040 comme maximum, puis régler le « pas» des unités principales à 10.
3. Faire de même avec l'axe des ordonnées, en indiquant 280 comme minimum et un« pas» de 10 pour les unités principales.
4. a. À l'aide d'un clic droit sur les points du nuage, sélectionner Ajouter une courbe de tendance dans la fenêtre de dialogue. b. Dans la fenêtre d'affichage Format de courbe de tendance, faire afficher le coefficient de détermination r^2 sur le graphique. Puis sélectionner successivement **linéaire** puis **polynomiale degré 2** puis **polynomiale degré 3**, tout en observant la courbe et le coefficient r^2 . Que

constatez-vous? Quel est le meilleur ajustement possible? c. Toujours dans la fenêtre d'affichage **Format de courbe de tendance**, faire une prévision « en avant» de 15. Que remarque-t-on ? Cela confirme-t-il les prévisions données dans l'énoncé ?

Partie B - Température moyenne globale

1. Sélectionner les données des colonnes A et C avec la touche `Ctrl` puis insérer un graphique nuage de points.
2. Régler le format de l'axe des abscisses comme dans la question A - 2.
3. Ajouter une courbe de tendance **polynomiale degré 3**.
4. Selon cette courbe de tendance, quelle température est prévue en 2030 ?

Partie C - Lien entre température et concentration de CO_2

1. Sélectionner les données des colonnes B et C, puis insérer un graphique« nuage de points».
2. Régler le format de l'axe des abscisses en prenant 280 comme minimum et 450 comme maximum, et 20 pour le« pas» des unités principales.
3. Faire de même avec l'axe des ordonnées, en indiquant 0, 1 pour les unités principales.
4. Ajouter une courbe de tendance linéaire.
5. Si la concentration de CO_2 atteint 450 ppm, quelle prévision de température donne la courbe de tendance de ce nuage de points ?

4. Loi de Moore

En 1965, Gordon E. Moore, l'un des cofondateurs de la société Intel, observe que le nombre de transistors qui composent un microprocesseur double tous les ans (première loi de Moore). En 1975, il précise qu'en fait ce nombre doit doubler tous les deux ans (deuxième loi de Moore) jusqu'en 2015, où on sera limité par la taille des atomes.

Partie A - Évolution du nombre de transistors selon la première loi de Moore

En 1971, un microprocesseur était constitué de 2300 transistors.

Soit (u_n) la suite qui donne, selon la première loi de Moore, le nombre de transistors dans un microprocesseur l'année de rang n, en prenant 1971 comme année de rang 0. Ainsi $u_0 = 2300$.

1. Pour tout entier naturel n, exprimer u_{n+1} en fonction de u_n .
2. Déterminer u_1 , u_2 et u_3 .
3. Quelle est la nature de cette suite?
4. En déduire une expression de u_n en fonction de n.
5. Selon la première loi de Moore combien un microprocesseur devrait contenir de transistors en 1974? Et en 1982?

Partie B - Évolution réelle du nombre de transistors par microprocesseur

Les données exactes de l'entreprise Intel concernant ses microprocesseurs depuis 1974 sont présentées dans le tableau suivant.

Année	Rang de l'année x_i	Nombre de transistors y_i (en millions)
1974	0	0,006
1979	5	0,029
1982	8	0,134
1985	11	0,275
1989	15	1,2
1993	19	3,1
1997	23	7,5
1999	25	9,5
2000	26	42
2004	30	125
2007	33	582
2008	34	820

1. Sur GeoGebra, représenter le nuage de points $M_i\ (x_i; y_i)$, pour i entier de 0 à 34, associé à la série statistique à deux variables « rang x de l'année » et « nombre y de transistors en millions ».
   Dans la barre de saisie, entrer successivement les coordonnées des points: `(0,0.006)` `(5,0.029)` `(8,0.134)` etc ..... Pour visualiser tous les points convenablement, changer l'échelle: clic droit sur le graphique Axe X: Axe Y puis 1 : 50
2. Soit (v_n) la suite définie sur $\mathbb{N}$, par: $\begin{cases} v_0 = 0,006 \\ v_{n+2} = 2v_n \end{cases}$
   a. Justifier que cette suite modélise l'évolution, selon la deuxième loi de Moore, du nombre de transistors, en millions, dans un microprocesseur l'année de rang n, en prenant 1974 comme année de rang 0.
   b. On admet que la suite (v_n) est géométrique. Démontrer que sa raison est \srqt2.
   c. Exprimer v_n en fonction de n.
   d. Déterminer v_{34} et comparer avec la valeur réelle.
3. On considère la fonction $f : x \mapsto 0,006 \times \sqrt{2}^x$ définie sur l'intervalle $[0; 34]$ telle que pour tout entier naturel n, $v_n = f(n)$. a. Représenter la courbe de la fonction f sur **GeoGebra**. Pour cela, dans la barre de saisie, entrer `f(x)=0.006*sqrt(2)^x`.
   b. La courbe obtenue coïncide-t-elle avec le nuage de points?
   c. Que peut-on dire de la deuxième loi de Moore pour la période 1974 à 2008?