

STATISTIQUES A DEUX VARIABLES

Objectifs :

- Représenter un nuage de points.
- Déterminer et utiliser un ajustement affine pour interpoler ou extrapoler des valeurs inconnues.
- Représenter un nuage de points en effectuant un changement de variable donné (par exemple u^2 , $\frac{1}{t}$, $\frac{1}{\sqrt{n}}$, $\log(y)$...) afin de conjecturer une relation de linéarité entre de nouvelles variables.

1. Vocabulaire

Définition 1. Nuage de points

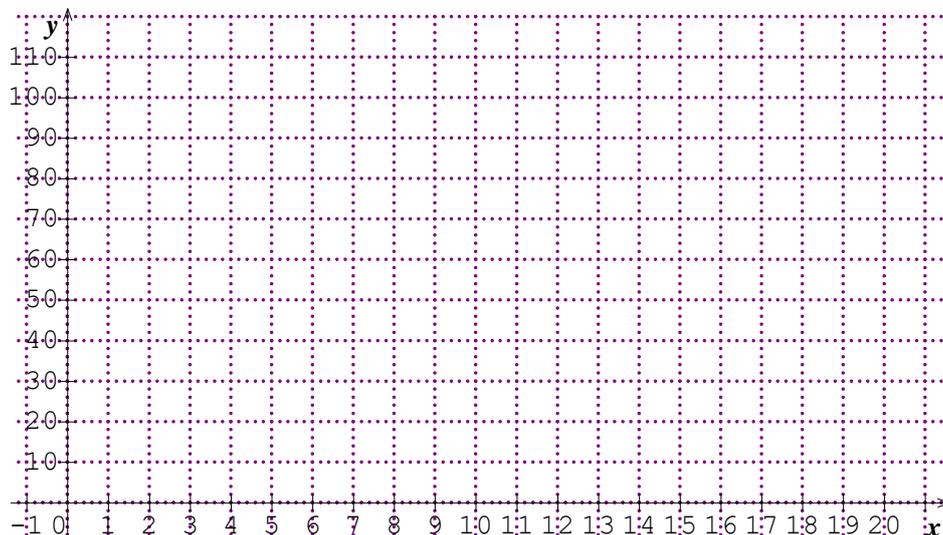
Soit une série statistique à deux variables x et y chacune de n valeurs x_1, x_2, \dots, x_n et y_1, y_2, \dots, y_n .

Soit un repère orthogonal $(O ; I, J)$, on appelle l'ensemble des n points (non reliés) M_i de coordonnées $(x_i ; y_i)$ pour i entier variant de 1 à n .

Exemple : Le tableau suivant présente l'évolution du budget publicitaire et du chiffre d'affaire d'une société au cours des 6 dernières années :

Budget publicitaire en milliers d'euros x_i	8	10	12	14	16	18
Chiffre d'affaire en milliers d'euros y_i	40	55	55	70	75	95

Dans un repère, on a représenté le nuage de points $(x_i ; y_i)$.



Définition 2. Point moyen

Soit une série statistique à deux variables x et y chacune de n valeurs x_1, x_2, \dots, x_n et y_1, y_2, \dots, y_n , et \bar{x} et \bar{y} les moyennes arithmétiques respectives des n valeurs x_i et des n valeurs y_i .

On appelle point moyen G du nuage, le point de coordonnées

Exemple : Déterminer les coordonnées du point moyen G du nuage de points de l'exemple précédent.

$$\bar{x} = \frac{\dots + \dots + \dots + \dots + \dots + \dots}{\dots} = \dots \quad \text{et} \quad \bar{y} = \frac{\dots + \dots + \dots + \dots + \dots + \dots}{\dots} = \dots$$

Le point moyen G du nuage de points a pour coordonnées (..... ;). On peut placer ce point dans le repère précédent.



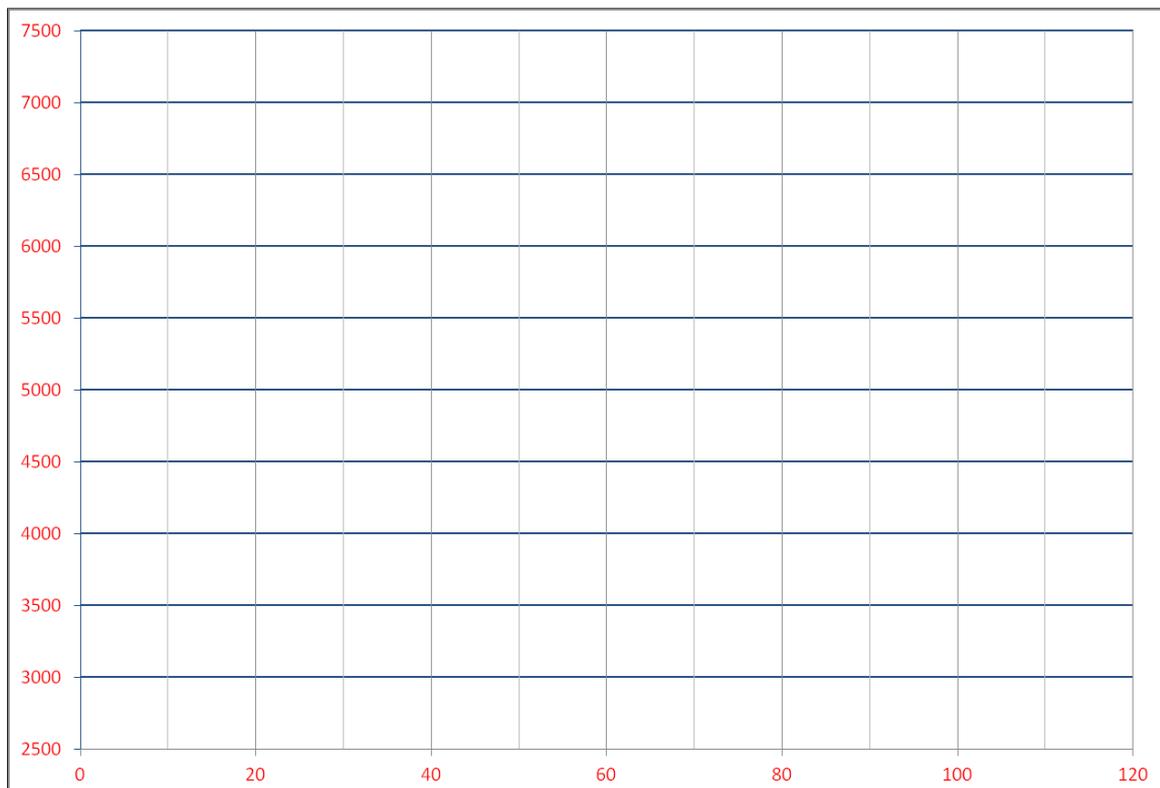
[corrigé en vidéo](#)

Exercice 1

Le tableau ci-dessous indique le poids (en grammes) et l'âge d'un nouveau-né.

Date	27/06	07/07	18/07	07/08	25/08	17/09	25/09	08/10
Rang x_i	1	11	22	42	60	83	91	...
Poids y_i	3 150	3 440	3 830	4 680	5 530	6 400	6 545	7 000

- 1) Construire le nuage de points dans le repère suivant.
- 2) Déterminer le point moyen G de ce nuage de points et le placer dans le repère.

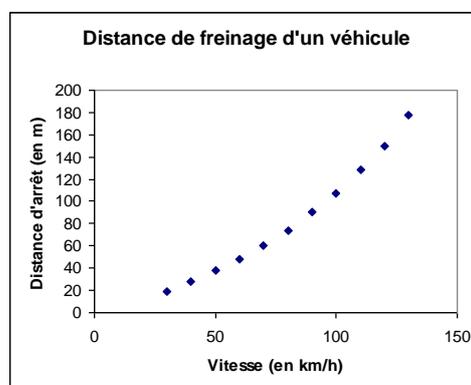
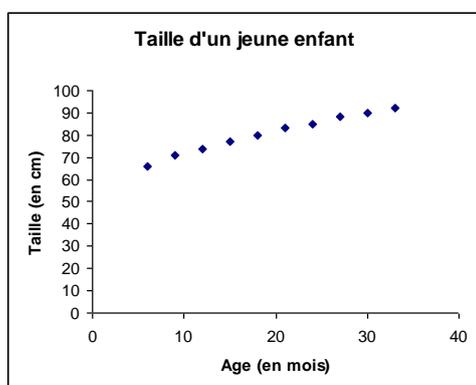


2. Ajustement d'un nuage de points

Définitions 3. Ajustement d'un nuage de points

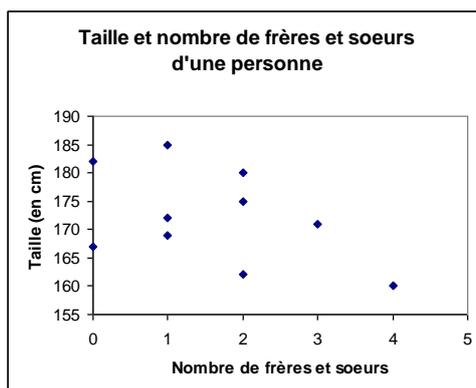
- On appelle ajustement du nuage de points toute courbe « »
..... » le nuage de points.
- Toute droite « résumant approximativement » le nuage est appelée
..... du nuage de points.

Remarque : Il existe d'autres types d'ajustement : dans certains cas, on peut observer que visiblement une droite ne convient pas mais que le nuage de points semble être approché par un autre type de courbe, parabole par exemple. En outre, certains nuages peuvent ne pas sembler être approchables par une quelconque courbe auquel cas les deux variables ne sont pas reliées entre elles.



On peut tracer une droite \mathcal{D} au voisinage des 10 points ; on dit alors que l'on a un ajustement affine.

Un ajustement affine ne convient pas ; on peut penser à « approcher » le nuage à l'aide d'une parabole.



Les points sont dispersés de façon quelconque ; cela veut dire qu'il n'existe aucun lien entre x_i et y_i , un ajustement n'est pas possible.

3. Détermination d'une équation d'une droite d'ajustement affine

Méthode graphique « au jugé »

On trace « au jugé » une droite qui « semble résumer » le nuage de points.
C'est une méthode simple mais qui dépend de la droite tracée.

Exemple : Le tableau ci-dessous propose les tailles et poids de 10 embryons durant la 14^{ème} semaine de grossesse (16^{ème} d'aménorrhée), les données étant ordonnées par ordre croissant de taille.

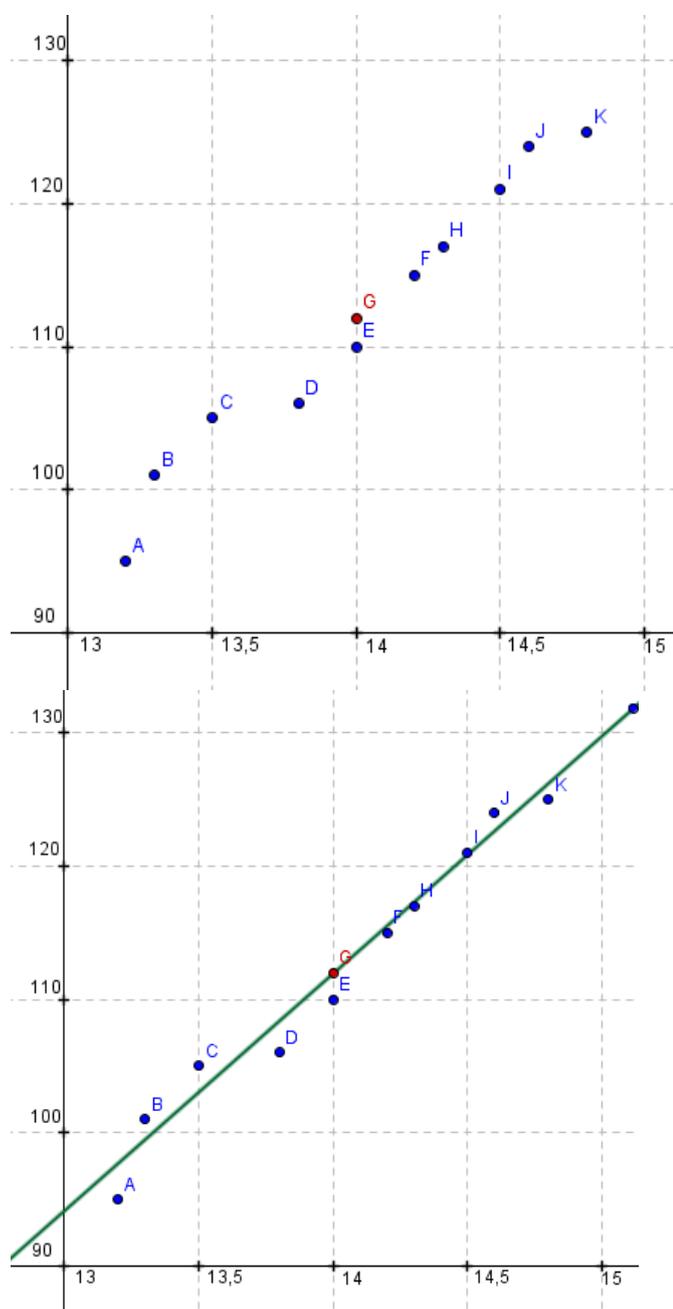
											Moyenne
Taille x_i (en cm)	13,2	13,3	13,5	13,8	14	14,2	14,3	14,5	14,6	14,8	14
Poids y_i (en g)	95	101	105	106	110	115	117	121	124	125	112

Source : site web Netenviesdesbebes.com / décembre 2009

Avec un logiciel (géométrie dynamique ou tableur), on peut afficher le nuage de points de cette série.

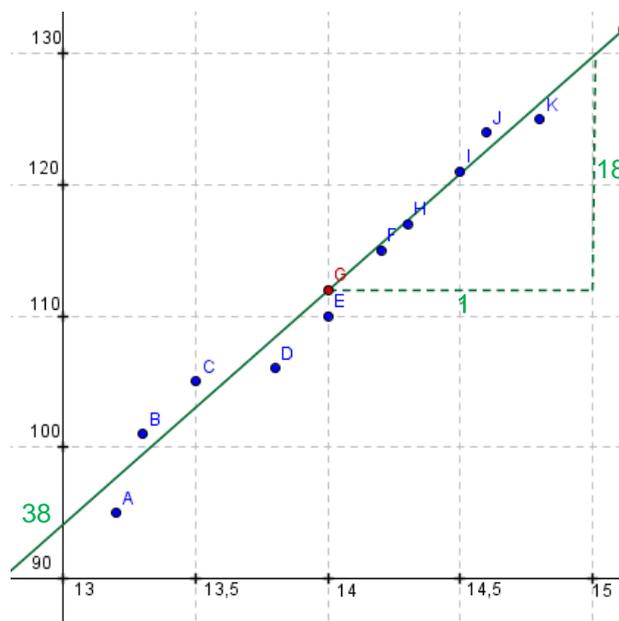
On n'oublie pas de placer le point moyen G après en avoir calculé ses coordonnées.

On trace « au jugé » et à la règle une droite passant par G et au plus près des autres points du nuage.



On lit graphiquement le coefficient directeur ($m \approx \dots$) et l'ordonnée à l'origine ($p \approx \dots$).

La droite d'ajustement affine a donc pour équation : $y = \dots x + \dots$.



Méthode des moindres carrés

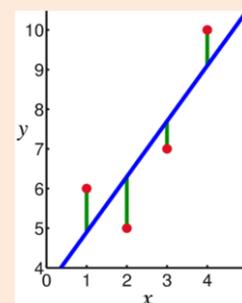
Cette méthode porte le nom de « moindre carrés » car elle consiste à rechercher la position de la droite d'ajustement telle que la somme des carrés des longueurs donnant les distances respectives (en vert)

entre la droite et les points soit

Pour cela, on utilise la calculatrice qui va donner l'équation de la droite cherchée.

Cette droite est appelée de y en x .

C'est une méthode simple mais qui dépend de la droite tracée.



Exercice 2 : On considère la série statistique à deux variables donnée dans le tableau suivant :

x_i	5	10	15	20	25	30	35	40
y_i	13	23	34	44	50	65	75	90

- 1) Dans un repère, représenter le nuage de points $(x_i ; y_i)$.
- 2) a) Déterminer, à l'aide de la calculatrice, une équation de la droite d'ajustement par la méthode des moindres carrés.
b) Représenter la droite d'ajustement de y en x .
- 3) Estimer graphiquement la valeur de x pour $y = 70$. Retrouver ce résultat par calcul. S'agit-il d'une interpolation ou d'une extrapolation ?



[corrigé en vidéo](#)

Définitions 4. Interpolation, extrapolation

L'interpolation et l'extrapolation sont des méthodes qui consistent à estimer une valeur inconnue dans une série statistique.

- Pour une interpolation, le calcul est réalisé dans le domaine d'étude fourni par les valeurs de la série.
- Pour une extrapolation, le calcul est réalisé en dehors du domaine d'étude.

Remarque : La droite d'ajustement par la méthode des moindres carrés passe par le point moyen G.

Exercice ③

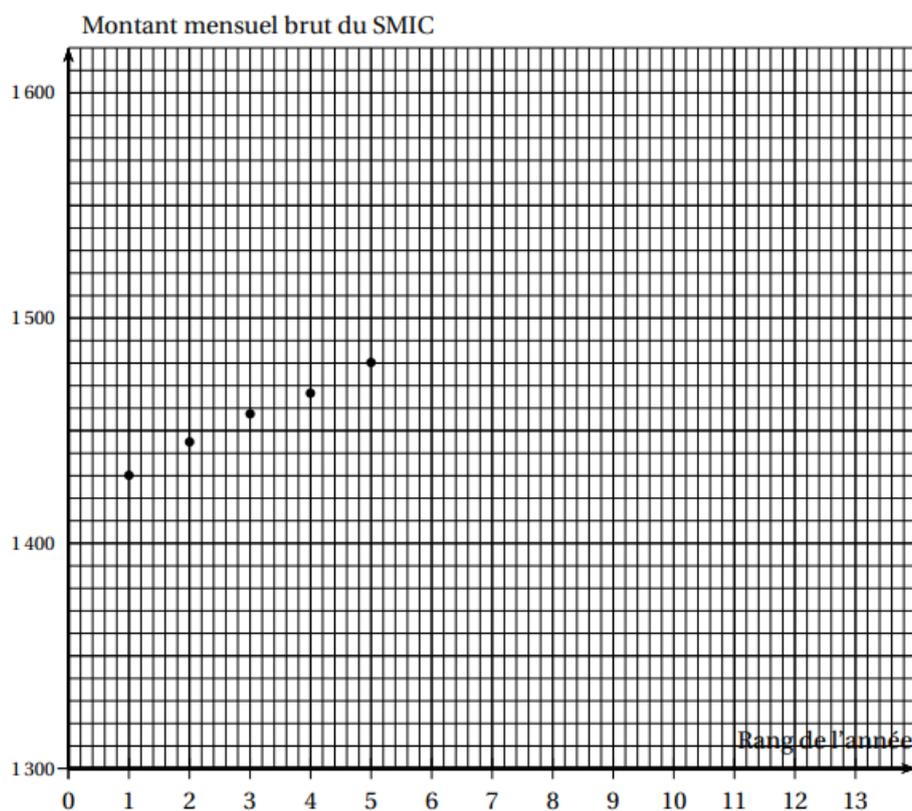
Le tableau suivant donne le montant mensuel brut, en euros, du SMIC pour 35 heures de travail hebdomadaire, entre 2013 et 2017 :

Année	2013	2014	2015	2016	2017
Rang de l'année : x_i	1	2	3	4	5
Montant mensuel brut du SMIC (en €) : y_i	1 430,22	1 445,38	1 457,52	1 466,62	1 480,27

Source I.N.S.E.E.

Une représentation graphique du nuage de points est donnée dans le repère ci-dessous.

- 1) À l'aide de la calculatrice, déterminer une équation de la droite d'ajustement de y en x obtenue par la méthode des moindres carrés.
- 2) a) Donner les coordonnées de deux points de cette droite, puis la tracer dans le repère ci-dessous.
b) En admettant que cet ajustement sera valide jusqu'en 2020, estimer la valeur du montant mensuel brut du SMIC en 2020.
En 2020, le smic s'élevait à 1 539,42 €. Que peut-on en déduire ?



4. Ajustement se ramenant à un ajustement affine

Exercice ④

Une entreprise vend des perles pour la fabrication de bijoux de fantaisie. Le tableau ci-dessous indique la quantité vendue y_i (en tonne), pour un prix au kilogramme fixé x_i (en euros).

x_i	3	4	5,5	7	8,5	9,3	10	10,8
y_i	4,926	3,773	2,773	2,197	1,820	1,669	1,554	1,430

- 1) Tracer le nuage de points. Peut-on envisager un ajustement affine ?
- 2) On pose $z = \frac{100}{y}$. Déterminer les valeurs de la série statistique z . Arrondir à 10^{-3} .
- 3) a) Déterminer, à l'aide de la calculatrice, l'équation de la droite des moindres carrés des séries x et z .
b) En déduire l'expression de y en fonction de x .
- 4) En extrapolant avec ce modèle, calculer la quantité de perles que vendrait l'entreprise si le prix montait à 24 € le kg.