

Chapitre II

Statistiques à 2 variables quantitatives

Contenus

- Nuage de points associé à une série statistique à deux variables quantitatives.
- Ajustement affine.

Capacités attendues

- 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 $u^2, \frac{1}{t}, \frac{1}{\sqrt{n}}, \log(y), \dots$) afin de conjecturer une relation de linéarité entre de nouvelles variables.

Table des matières

Rappel	2
Tracer une droite en connaissant son équation	2
I. Séries statistiques à 2 variables	2
I.1. Définition	2
I.2. Nuages de points.....	2
II. Ajustement affine	3
II.1. Définition	3
II.2. Méthode	4
a. Au jugé (méthode moins précise) :	4
b. À l'aide de la calculatrice :	4
c. Avec un tableur :	4
III. Nuage de points avec changement de variables	4

Rappel

Tracer une droite en connaissant son équation

Soit la droite d'équation $y = 3x - 2$ (ou $f(x) = 3x - 2$).

Pour la tracer, nous avons besoin des coordonnées de 2 points de cette droite.

Nous allons faire un tableau de valeurs, et choisir 2 valeurs de x (prenons 0 et 2).

x	0	2
y ou $f(x)$	$3 \times 0 - 2 = -2$	$3 \times 2 - 2 = 4$

On a ainsi les coordonnées de deux points : (0; -2) et (2; 4).

Vous les placez dans un repère et tracez la droite passant par ces 2 points.

I. Séries statistiques à 2 variables

Act. 1 p 102

I.1. Définition

Une série statistique à 2 variables est une série statistique étudiant simultanément deux caractères sur un même échantillon de n individus extraits d'une population.

Si les valeurs prises par le premier caractère sont x_1, x_2, \dots, x_n et celles prises par le second caractère y_1, y_2, \dots, y_n , alors on peut représenter cette série dans le tableau ci-dessous :

Valeurs du 1 ^{er} caractère	x_1	x_2	...	x_n
Valeurs du 2 nd caractère	y_1	y_2	...	y_n

Exemple :

Voici un tableau donnant la température moyenne dans l'atmosphère en fonction de l'altitude :

Altitude(km)	0	0,5	1	1,5	2	2,5
Température(°C)	15	11,8	8,9	5,1	1,8	-1,4

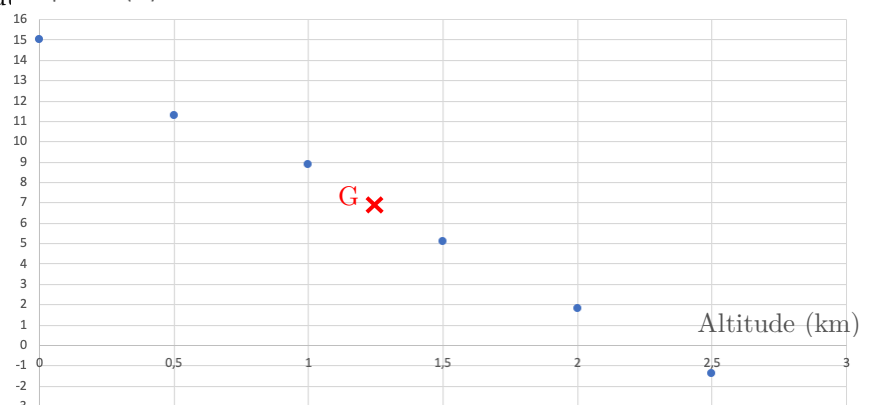
I.2. Nuages de points

Définition :

Dans un repère orthogonal du plan, l'ensemble des points $M(x_i; y_i)$ pour i allant de 1 à n est appelé le nuage de plan de la série statistique.

Exemple :

Avec le tableau précédent représentant l'évolution de la température en fonction de l'altitude, le nuage de points suivants :



Définition :

Soit \bar{x} la moyenne des valeurs prises par le 1^{er} caractère et \bar{y} celle prise par le 2nd caractère. On appelle point moyen du nuage de points de la série statistique, le point G de coordonnées $(\bar{x} ; \bar{y})$.

Exemple :

Pour notre exemple, la moyenne des altitudes est :

$$\frac{0 + 0,5 + 1 + 1,5 + 2 + 2,5}{6} = 1,25 \text{ km}$$

La moyenne des températures est :

$$\frac{15 + 11,8 + 8,9 + 5,1 + 1,8 + (-1,4)}{6} \approx 6,9 \text{ }^{\circ}\text{C}$$

Ainsi le point moyen est le point G de coordonnées $(1,25 ; 6,9)$ (voir la croix sur le graphique ci-dessus).



Point calculatrice !

Ti : Pour rentrer les points, aller dans **stats** puis **ÉDIT**. Dans L1 rentrer les valeurs de votre 1^{er} caractère, et dans L2 celles du 2nd caractère. (Vous pouvez choisir d'autres colonnes).

Pour accéder au calcul pour votre série, aller dans **stats** puis **CALC**, et choisissez **2 :Stats 2 Var**.

Sélectionnez les bonnes colonnes pour Xliste et Yliste, calculer.

Vous avez accès ainsi à notamment les coordonnées du point moyen.

Numworks : Choisissez le menu Régression, puis rentrez vos valeurs en X1 et Y1 par exemple. En allant dans l'onglet graphique, vous pourrez visualiser vos points, voir les coordonnées du point moyen.

II. Ajustement affine

Act. 2 p 102

II.1. Définition

Si les points d'une série statistique à 2 variables semblent plus ou moins alignés, il est possible de construire une droite passant au plus près de ces points.

Cette droite réalise alors un ajustement affine du nuage de points.

Mais pourquoi ???

Cela permet de dégager une tendance et aussi de pouvoir calculer des valeurs (avec précaution !) en dehors de celles que l'on nous avait fournies.

Définition :

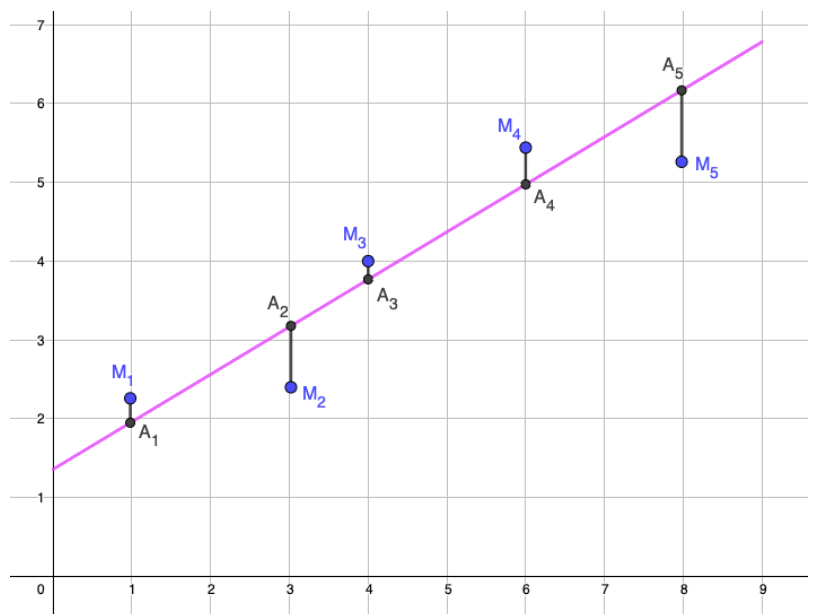
La droite obtenue par la méthode des moindres carrés est appelée la droite d'ajustement de y en x par la méthode des moindres carrés.

Explication :

Nous avons ci-contre un nuage de points M_1, \dots, M_5 .

La méthode des moindres carrés consiste à trouver la droite qui rend la somme $M_1A_1^2 + M_2A_2^2 + \dots + M_5A_5^2$ la plus petite possible.

Cette droite existe et est unique.



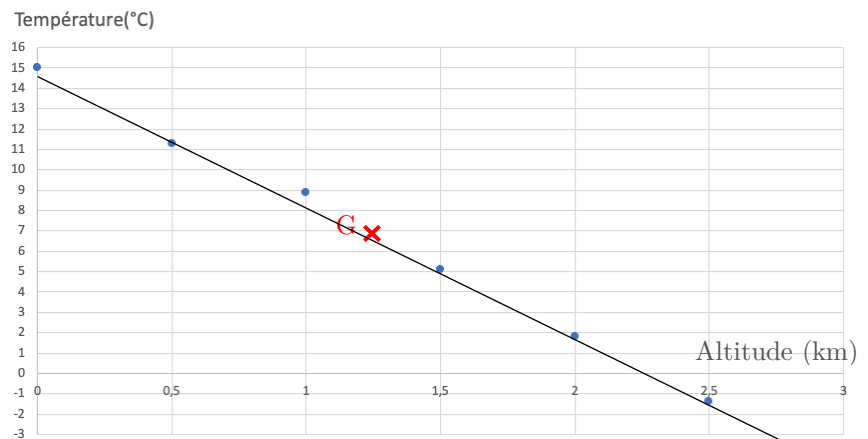
Remarque :

La droite d'ajustement passe par le point moyen G.

II.2. Méthode

a. Au jugé (méthode moins précise) :

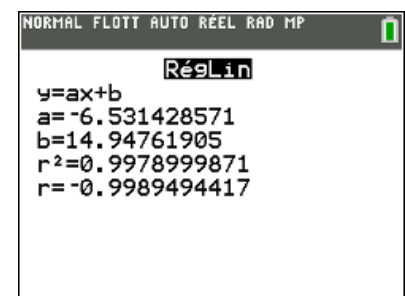
Après avoir tracé votre nuage de points, vous tracez la droite qui vous semble être la meilleure représentante pour l'ajustement affine.



b. À l'aide de la calculatrice :

TI : Après avoir rentré vos valeurs, aller dans stats, CALC, puis choisir 4 :RégLin(ax+b).

Vous obtenez le résultat ci-contre avec nos valeurs d'altitudes et de températures. Ainsi, l'équation de la droite d'ajustement est : $y = -6,617x + 15,138$.



Numworks :

Après avoir rentré vos valeurs dans le menu Régression, allez dans l'onglet graphique, puis sélectionnez Régression, puis modèle linéaire. Votre droite se trace et vous avez son équation.

c. Avec un tableur :

Rentrez vos valeurs, sélectionnez-les puis sélectionnez insertion/Nuage de points (Excel) ou Diagramme puis XY(dispersion) (LibreOffice et OpenOffice).

Pour avoir accès à la droite d'ajustement faites clic droit sur un des points puis choisir **Ajouter une courbe de tendance** (Excel) ou **Insérer une courbe de tendance** (LibreOffice), puis sélectionnez **Linéaire** et **Afficher l'équation**.

III. Nuage de points avec changement de variables

Parfois, un ajustement affine du nuage de points de notre série n'est pas pertinent. Dans ce cas, on peut faire ce que l'on appelle un changement de variable.

Exemple :

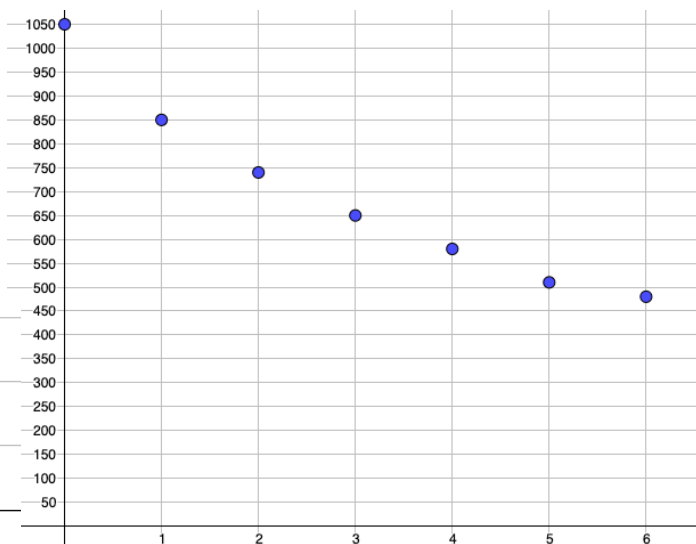
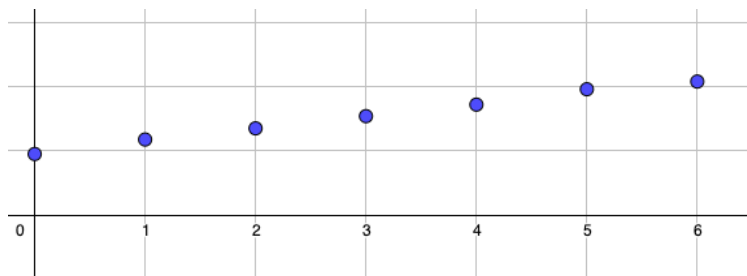
Dans le tableau ci-dessous, on a relevé le prix y , en euro, d'un smartphone chaque année de 2016 à 2022. On pose :

$$z = \frac{1}{y}.$$

Année	2016	2017	2018	2019	2020	2021	2022
Rang de l'année x_i	0	1	2	3	4	5	6
Prix y_i	1050	850	740	650	580	510	480
$z_i = \frac{1}{y_i}$	0,00095	0,00118	0,00135	0,00154	0,00172	0,00196	0,0021

S'il on représente le nuage des points $M_i(x_i; y_i)$, on voit que l'ajustement affine n'est pas pertinent (cf. ci-contre).

Par contre, à l'aide du changement de variable proposé : $z = \frac{1}{y}$, si l'on représente le nuage de points $N_i(x_i; z_i)$, on voit que l'ajustement affine devient possible (cf. graphique ci-dessous).



À l'aide de la calculatrice, on trouve que l'équation de la droite d'ajustement par la méthode des moindres carrés est : $z = 0,000192x + 0,000966$.

Pour retrouver, l'équation liant y à x , on utilise le changement de variable :

$$z = \frac{1}{y} = 0,000192x + 0,000966 \Leftrightarrow y = \frac{1}{0,000192x + 0,000966}$$

On obtient le graphique ci-dessous :

