EXERCICES CORRIGES STATISTIQUES A DEUX VARIABLES

1) Introduction

Dans une population on fait deux mesures notés *x* et *y* sur chaque individu statistique. Le but de l'ajustement consiste s'il y a une relation entre ces deux mesures.

En classe de Terminale Technologique, on recherche s'il existe une relation affine, donc de la forme **y = ax + b** entre ces deux variables

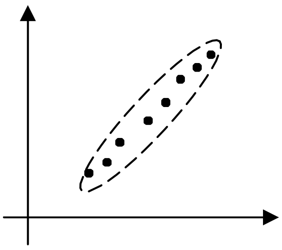
Exemples

On mesure la taille et le poids d'un ensemble de personnes. On veut vérifier si le poids *y* augmente avec la taille *x*

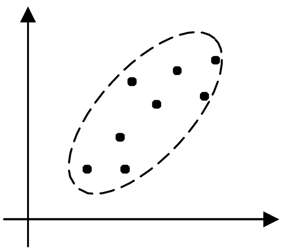
Une entreprise veut étudier l'évolution des son chiffre d'affaire mensuel.
x est le n° du mois et *y* est le chiffre d'affaire

2) Justification d'un ajustement affine

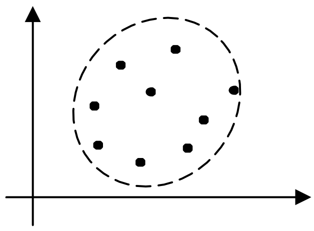
Chaque individu statistique est donc représenté par un couple de nombres (x ;y). On peut donc représenter graphiquement ces couples de nombres par un ensemble de points. Cet ensemble de points est appelé nuage de points. La forme de ce nuage permet de justifier ou non un ajustement affine.



Nuage très allongé
L'ajustement affine est justifié



Nuage allongé
L'ajustement affine reste justifié



Nuage non allongé
L'ajustement affine n'est pas justifié

3) Les deux méthodes d'ajustement affine

Point moyen

Soient *n* points de coordonnes $(x_1 ; y_1)$; $(x_2 ; y_2)$; $(x_3 ; y_3)$; ; $(x_n ; y_n)$

On appelle point moyen le point $G(\,(\bar{x} ; \bar{y})\,)$

La méthode de Mayer

On range la série suivant l'ordre croissant de *x*
On considère deux sous nuages égaux ou différents d'une unité
Pour chaque sous nuage on calcule les coordonnées
du point moyen. On obtient donc deux points G_1 **et** G_2
La droite de Mayer est alors la droite $(G_1 \; G_2)$

La droite de Mayer passe par le point moyen du nuage

La méthode des moindres carrés

Cette méthode de calcul se fait uniquement à la calculatrice. On obtient deux nombres *a et b*.
a et *b* sont les coefficients de la droite d'ajustement.

La droite des moindres carrés passe par le point moyen du nuage

4) Exercice1

Dans une maternité on a relevé le poids et la taille de 10 nouveaux nés.
Les résultats sont consignés dans le tableau suivant :

Enfant	1	2	3	4	5	6	7	8	9	10
Masse en kg	**2,5**	**2,6**	**2,7**	**3**	**3,2**	**3,3**	**3,4**	**3,6**	**3,8**	**3,9**
Taille en cm	**45**	**46**	**48**	**50**	**51**	**52**	**53**	**54**	**54**	**57**

On veut savoir si connaissant le poids d'un nouveau né ont peut avoir une idée de sa taille.

1. L'ajustement affine vous parait il justifié ?
2. Faire un ajustement affine de la taille en fonction du poids par la méthode de Mayer:
3. Vérifier que le point moyen est sur la droite d'ajustement
4. Si un bébé pèse 4,2 kg quelle sera sa taille probable ?
5. Refaire les calculs avec la droite des moindres carrés

CORRIGE :

On pose *x* le poids en kg et *y* la taille en cm

1) Les points du nuage sont pratiquement alignés, un ajustement affine est donc bien justifié.

2) Le tableau est dans l'ordre croissant des *x,* on fait donc deux sous nuages, le premier avec les nouveaux nés n° 1 à 5 et le second avec les nouveaux nés n° 6 à 10.

Point moyen du premier sous nuage

1 $x = \frac{2,5 + 2,6 + 2,7 + 3 + 3,2}{5} = 2,8$

1 $y = \frac{45 + 46 + 48 + 50 + 51}{5} = 48$

On a donc le point

G_1 (2,8 ; 48)

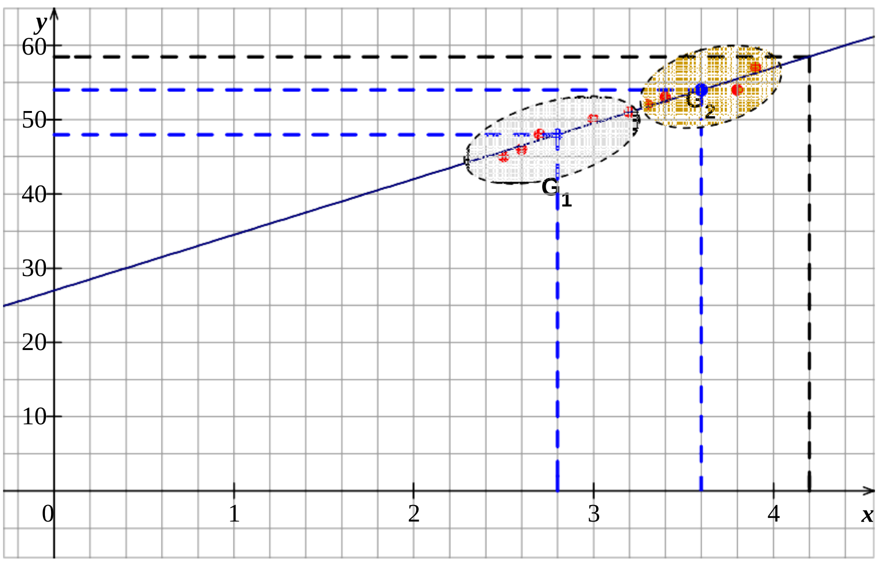
Point moyen du second sous nuage

2 $x = \frac{3,3 + 3,4 + 3,6 + 3,8 + 3,9}{5} = 3,6$

2 $y = \frac{52 + 53 + 54 + 54 + 57}{5} = 54$

On a donc le point

G_2(3,6 ; 54)



Calculons l'équation de la droite (G_1 ; G_2)
Son coefficient directeur est :

$$\mathbf{a = \frac{54 - 48}{3,6 - 2,8} = 7,5}$$

L'équation de cette droite est y = 7,5x + b

Le point G_1(2,8 ; 48) est un point de cette droite donc :
48 = 7,5 x 2,8 + b   48 – 7,5 x 2,8 = b   Donc, b = 27
La droite d'ajustement est donc :

y = 7,5x + 27

3)*Calcul du point moyen *G
En faisant la moyenne des 10 poids on trouve x_G = 3,2
Et la moyenne des 10 tailles donne y_G= 54

Le point moyen est donc *G***(3,2 ; 54)**
Dans l'équation, pour x = 3,2 on a y = 7,5 x 3,2 + 27 = 51 = y_G *G* est donc un point de la droite

4)*Si un bébé pèse 4,2 kg alors *x = 4,2
On remplace *x* par 4,2 dans l'équation :
y = 7,5 x 4,2 + 27 = 58,8
Un bébé pesant 4,2 kg mesurera probablement 58 cm

5)Ajustement par les moindres carrés
A la calculatrice on trouve (précision 2 décimales) a = 7,55   b = 26,85

La droite des moindre carrés est donc : **y = 7,55x + 26,85**

Pour *x* = 3,2 on a y = 7,55 x 3,2 + 26,85 = 51,01 *(la différence avec 51 s'explique par les arrondis)*
Le point moyen *G* est bien sur la droite des moindres carrés

Pour *x* = 4,2 on a ; y = 7,55 x 4,2 + 26,85 = 58,51
On retrouve un résultat équivalent à celui obtenu par la méthode de Mayer

EXERCICE :

À partir des recensements effectués tous les dix ans, on a établi le tableau suivant qui donne l'évolution de la population française en millions d'individus entre 1851 et 1911. Peu de données sont disponibles pour l'année 1871

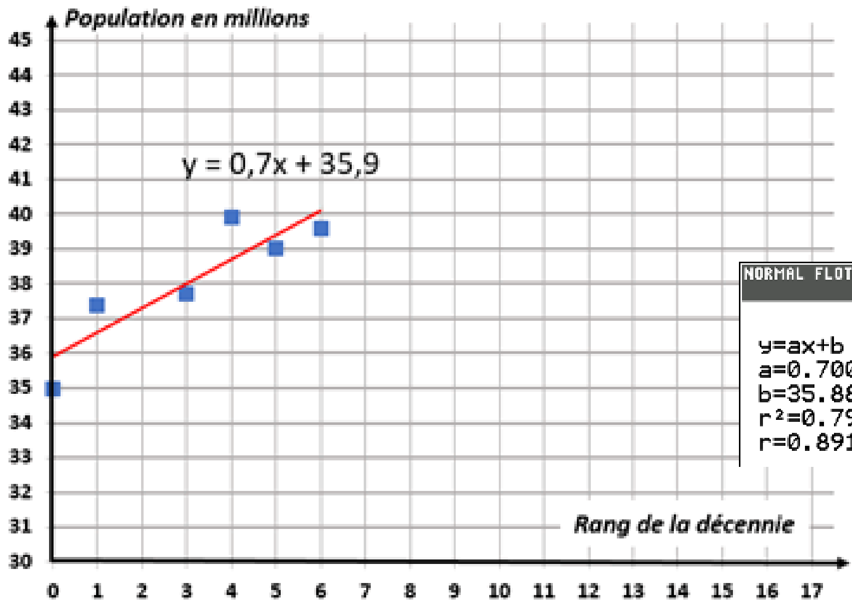
	Population en 1851	*Population en 1861*	*Population en 1881*	*Population en 1891*	*Population en 1901*	*Population en 1911*
Rang de la décennie : x_i	**0**	**1**	**3**	**4**	**5**	**6**
Population en millions : y_i	**35**	**37,4**	**37,7**	**39,9**	**39**	**39,6**

Source : INSEE

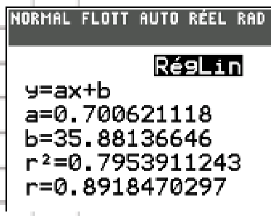
1. Placer sur le graphique donné en annexe le nuage de coordonnées $(x_i;y_i)$.
2. Donner une équation de la droite d'ajustement affine de y en fonction de x obtenue par la méthode des moindres carrés. Les coefficients seront arrondis au millième.
3. On décide d'ajuster ce nuage de points par la droite (d) d'équation y=0,7x+35,9. Tracer cette droite sur ce même graphique.
4. À l'aide de ce modèle, estimer la population en 1871.
5. À l'aide de ce modèle, estimer la population en 1971

CORRIGE

Annexe



2. L'équation fournie par calculatrice est la

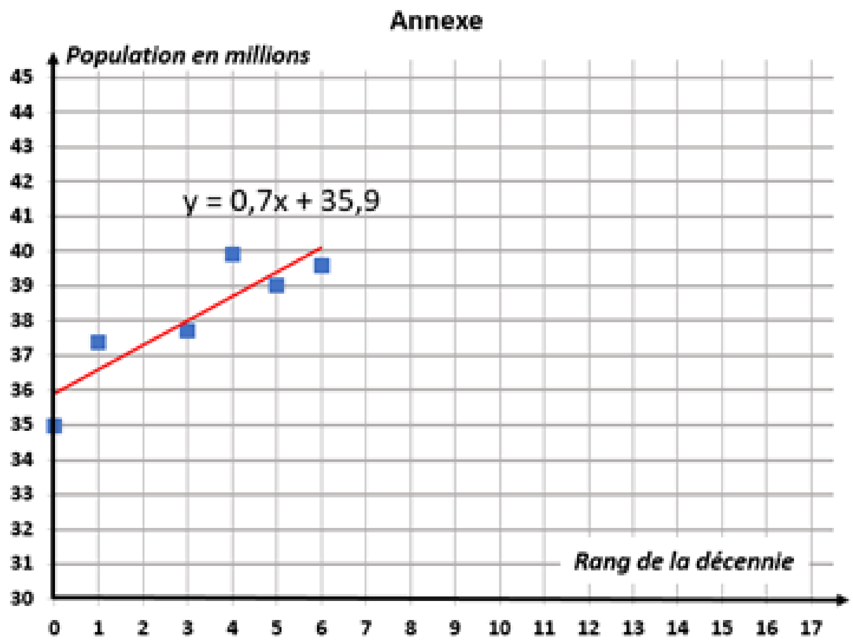


NORMAL FLOTT AUTO RÉEL RAD
RégLin
y=ax+b
a=0.700621118
b=35.88136646
r²=0.7953911243
r=0.8918470297

Sur cette fenêtre apparaissent le coefficient directeur de la droite a et l'ordonnée à l'origine b.

Ainsi l'équation de la droite d'ajustement est **y=0,701x+35,881** (coefficient arrondis au millième puisque nous sommes respectueux de l'énoncé).

3. Nuage de points et tracé de la droite



4. L'énoncé demande de procéder à une interpolation, c'est-à-dire d'estimer une valeur manquante entre d'autres qui sont connues
L'année 1871 correspond au rang 2. Pour cette valeur de x, nous obtenons
$y = 0{,}7\times2+35{,}9 = 37{,}3$. On estime la population française à 37,3 millions d'individus en 1871.

5. L'énoncé demande de procéder à une extrapolation, c'est-à-dire estimer une valeur au-delà de la série.
L'année 1971 correspond au rang 12. Pour cette valeur de x, nous obtenons
$y = 0{,}7\times12+35{,}9 = 44{,}3$. On estime la population française à 44,3 millions d'individus en 1971

REMARQUE : En 1971, la population française s'élevait à 51 millions d'individus

CONCLUSION : Le modèle mathématique établi pour la période 1851-1911 n'est plus valable en 1971