Pondichéry – Avril 2007 - Exercice

Une entreprise de services d'une ville cherche à modéliser la consommation des ménages sur les dernières années.
Le rang $x_1 = 1$ est donné pour l'année 1998. La consommation est exprimée en milliers d'euros.

Année	1998	2000	2001	2002	2004
Rang de l'année x_i	1	3	4	5	7
Consommation en milliers d'euros y_i	28.5	35	52	70.5	100.5

1. Représenter le nuage de points $P_i\left(x_i ; y_i\right)$ dans un repère orthogonal du plan (on prendra 1 cm comme unité en abscisses et 1 cm pour 10 000€ en ordonnées) ;
2. Déterminer les coordonnées du point moyen G de ce nuage ; le placer dans le repère précédent ;
3. On réalise un ajustement affine de ce nuage par la droite d'équation $y = 12,5x + b$ qui passe par le point G.
   a) Déterminer la valeur de b ;
   b) Tracer la droite D dans le repère précédent.
4. Déterminer, à l'aide de l'ajustement précédent, la consommation estimée des ménages de cette ville en 2005 ;
5. En réalité, un relevé récent a permis de constater qu'en 2005 la consommation réelle des ménages de cette ville était de $y_8 = 140\ 000€$.
   Déterminer, en pourcentage, l'erreur commise par l'estimation précédente par rapport à la valeur exacte (on donnera un résultat à l'aide d'un nombre entier en effectuant un arrondi) ;
6. Un nouvel ajustement, de type exponentiel semble alors plus adapté.
   a) Recopier et compléter le tableau suivant sachant que $z = \ln y$.
   Les résultats seront arrondis au centième.

x_i	1	3	4	5	7	8
$z_i = \ln y_i$	3,35	…	…	…	…	4,94

b) Déterminer l'équation réduite de la droite de régression de z en x obtenue par la méthode des moindres carrés à l'aide de la calculatrice ; cette équation est de la forme $z = cx + d$; on donnera les arrondis des coefficients c et d à 10^{-2}.

c) En déduire que : $y = 20,49e^{0,23x}$;

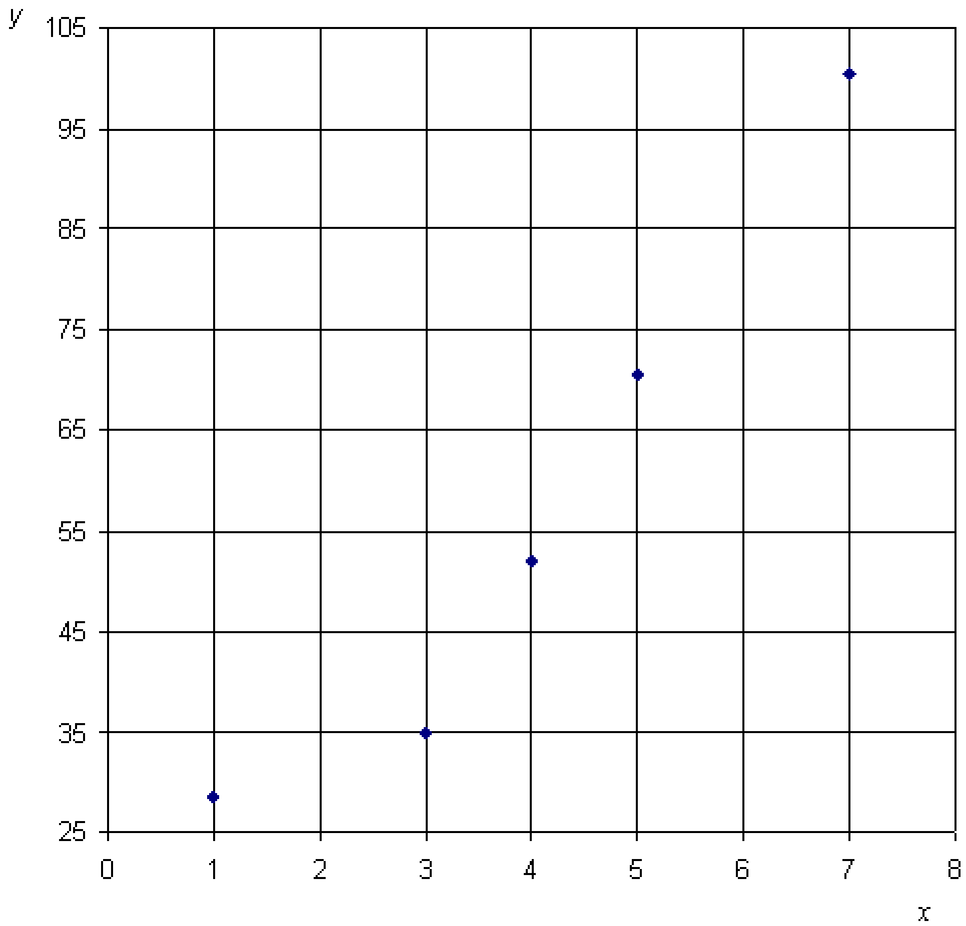
d) Estimer alors, à l'aide de ce nouvel ajustement, la consommation des ménages de cette ville en 2007 à 100€ près.

Analyse

Un exercice classique sur les statistiques à deux variables et l'ajustement affine. Dans un premier temps on utilise un ajustement affine. Puis, à l'aide d'un changement de variable, on détermine un ajustement exponentiel, mieux adapté aux données disponibles.

Résolution

→ *Question 1.*



→ *Question 2.*

Les coordonnées du point moyen G sont les moyennes $\overline{x}$ et $\overline{y}$ des séries des x_i et des y_i respectivement. L'effectif total étant de $n=5$, il vient :

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{5}(1+3+4+5+7)=\frac{20}{5}=4$$

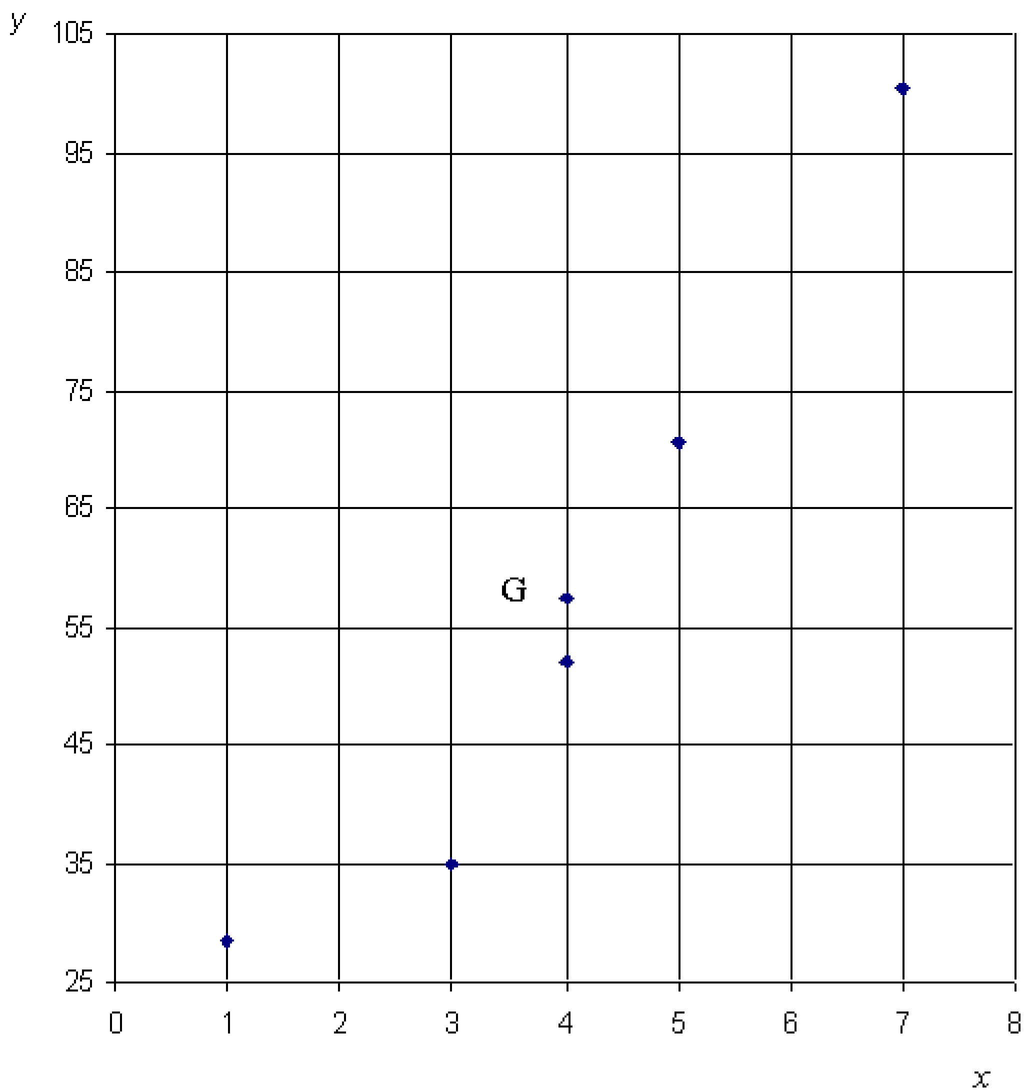
Et :

$$\overline{y}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{5}(28,5+35+52+70,5+100,5)=\frac{286,5}{5}=57,3$$

Finalement :

$$\boxed{\mathbf{G}(4;57,3)}$$

On obtient alors le graphique suivant :



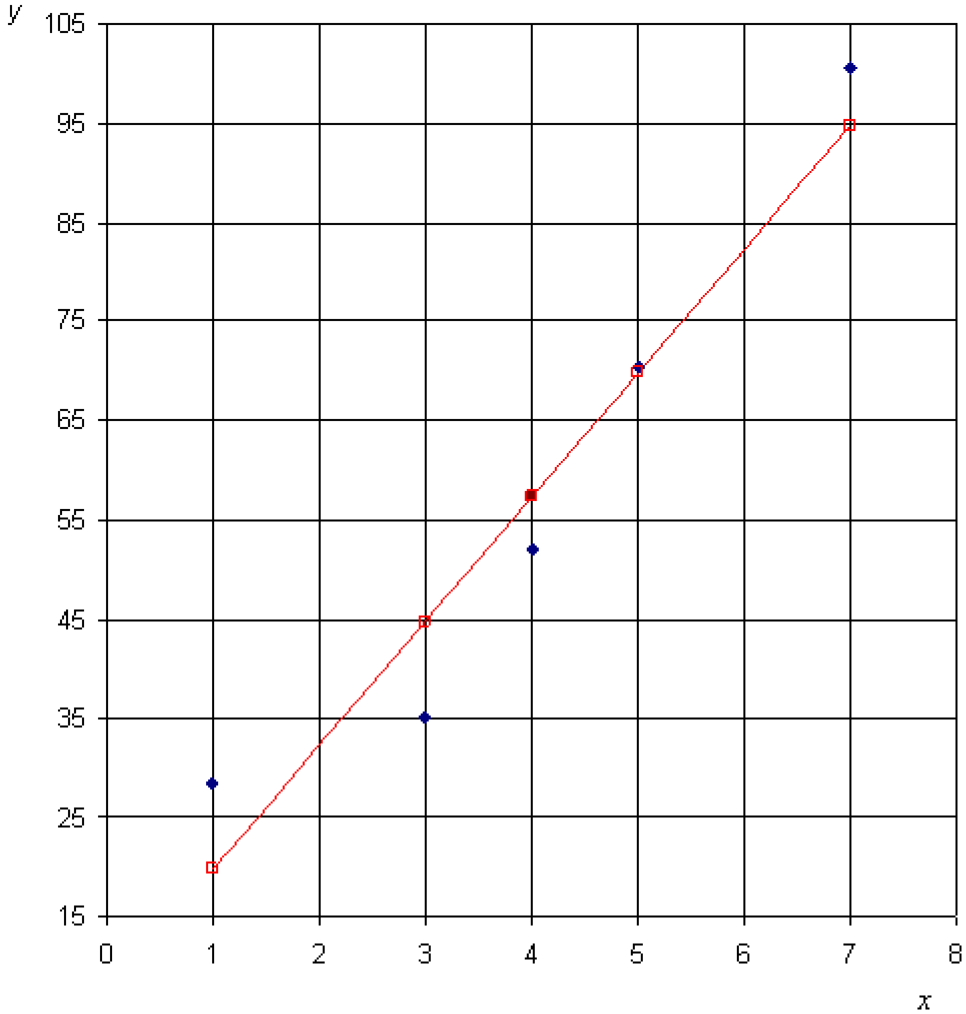
→ *Question 3.a.*

Les coordonnées du point G vérifiant l'équation de la droite D, on a : $57,3 = 12,5 \times 4 + b$.
Soit : $b = 57,3 - 50 = 7,3$.

On a $b = 7,3$ et l'équation réduite de la droite D s'écrit : $y = 12,5x + 7,3$.

→ *Question 3.b.*

On a représenté la droite D en rouge sur le graphique ci-après.



→ *Question 4.*

L'année 2005 correspond au rang $x = 8$. A l'aide de l'équation précédente, on peut effectuer une extrapolation. On obtient :

$$y_8 = 12,5 \times 8 + 7,3 = 100 + 7,3 = 107,3$$

A l'aide de l'ajustement affine obtenu, la consommation des ménages en 2005 peut être estimée à 107 300 €.

→ *Question 5.*

L'erreur relative commise en utilisant l'ajustement affine par rapport à la valeur exacte vaut :

$$\frac{140-107,3}{140}=\frac{32,7}{140}\simeq 0,23=23\%$$

A l'aide de l'ajustement affine obtenu, le pourcentage d'erreur commise par l'estimation obtenue à la question précédente par rapport à la valeur exacte est de 23% (valeur arrondie au pourcent).

→ *Question 6.a.*

On obtient :

x_i	1	3	4	5	7	8
$z_i=\ln y_i$	3,35	3,56	3,95	4,26	4,61	4,94

→ *Question 6.b.*

A l'aide de la calculatrice, on obtient, en arrondissant les coefficients à 10^{-2} :

$$\boxed{\boldsymbol{z=0,23x+3,02}}$$

→ *Question 6.c.*

Comme $z=\ln y$, l'égalité précédente se récrit : $\ln y=0,23x+3,02$.
En considérant l'exponentielle de chaque membre, il vient alors :

$$e^{\ln y}=y=e^{0,23x+3,02}=e^{3,02}\times e^{0,23x}$$

Or, $e^{3,02}\simeq 20,49$ (valeur approchée à 10^{-2}). On en déduit finalement :

$$\boxed{\boldsymbol{y=20,49e^{0,23x}}}$$

→ *Question 6.d.*

L'année 2007 correspond au rang $x_{10}=10$. On a alors, grâce à l'ajustement exponentiel obtenu à la question précédente :

$$y_{10}=20,49\times e^{0,23\times 10}=20,49\times e^{2,3}\simeq 204,4$$

(la quantité y désignant des milliers d'euros, le résultat précédent est fourni avec une précision de 10^{-1} correspondant aux 100€ de l'énoncé).

Conclusion :

A l'aide de l'ajustement exponentiel, la consommation des ménages de cette ville en 2007 peut être estimée à 204 400€ (à 100€ près).