

# Séquence VII – Régression linéaire

## Exercice 1

On dispose de cinq spécimens fossiles d'un animal disparu pour lesquels on possède les mesures de la longueur en cm de leur humérus  $x$  et de leur fémur  $y$ .

1. Compléter le tableau suivant et en déduire les valeurs des variances et covariance :

|   | $x_i$ | $y_i$ | $x_i^2$ | $y_i^2$ | $x_i y_i$ |
|---|-------|-------|---------|---------|-----------|
| 1 | 44    | 40    |         |         |           |
| 2 | 65    | 60    |         |         |           |
| 3 | 71    | 59    |         |         |           |
| 4 | 75    | 65    |         |         |           |
| 5 | 87    | 77    |         |         |           |

2. Déterminer, par la méthode des moindres carrés, l'équation de la droite de régression de  $y$  en  $x$ .
3. Passe-t-elle par le point moyen  $G$  ?
4. Déterminer le coefficient de corrélation linéaire. Commenter.
5. Calculer la longueur, selon ce modèle, du fémur d'un spécimen dont l'humérus mesurerait 55 cm.

## Solution de l'exercice 1

1. Le tableau complété donne :

|   | $x_i$ | $y_i$ | $x_i^2$ | $y_i^2$ | $x_i y_i$ |
|---|-------|-------|---------|---------|-----------|
| 1 | 44    | 40    | 1936    | 1600    | 1760      |
| 2 | 65    | 60    | 4225    | 3600    | 3900      |
| 3 | 71    | 59    | 5041    | 3481    | 4189      |
| 4 | 75    | 65    | 5625    | 4225    | 4875      |
| 5 | 87    | 77    | 7569    | 5929    | 6699      |

2. L'équation de la droite de régression de  $y$  en  $x$  est

$$y = 0,8319x + 3,3$$

3. La droite de régression passe par le point moyen  $G(68,4; 60,2)$
4. Le coefficient de corrélation linéaire est

$$r \approx 0,9856$$

$r > 0,84$  donc la longueur en cm des humérus des cinq spécimens de fossiles est corrélée à la longueur de leur fémur. L'ajustement affine est raisonnable.

5.  $y = 0,8319 \times 55 + 3,3 = 49,05$ . Ainsi, selon ce modèle, un fémur d'un spécimen dont l'humérus mesurerait 55 cm aurait pour mesure environ 49,05 cm.

## Exercice 2

Pour étudier les problèmes de malnutrition dans un pays pauvre, on a calculé le poids moyen par âge d'un échantillon de 2400 enfants répartis uniformément en 12 classes d'âge.

On a obtenu le tableau suivant :

| $x_i =$<br>classe d'âge | $y_i =$<br>poids moyen<br>en kg | $x_i^2$ | $y_i^2$ | $x_i y_i$ |
|-------------------------|---------------------------------|---------|---------|-----------|
| 1                       | 3,5                             |         |         |           |
| 2                       | 3,5                             |         |         |           |
| 3                       | 3,3                             |         |         |           |
| 4                       | 4,4                             |         |         |           |
| 5                       | 4,4                             |         |         |           |
| 6                       | 4,2                             |         |         |           |
| 7                       | 5,1                             |         |         |           |
| 8                       | 5,3                             |         |         |           |
| 9                       | 5,5                             |         |         |           |
| 10                      | 5,5                             |         |         |           |
| 11                      | 6,2                             |         |         |           |
| 12                      | 5,7                             |         |         |           |

1. Compléter ce tableau avec les valeurs de  $x_i^2$ , de  $y_i^2$  et de  $x_i y_i$ . Déterminer l'équation de la droite de régression de  $y$  en  $x$ .
2. Déterminer le coefficient de corrélation linéaire. Interpréter.

## Solution de l'exercice 2

- 1.

| $x_i =$<br>classe d'âge | $y_i =$<br>poids moyen<br>en kg | $x_i^2$ | $y_i^2$ | $x_i y_i$ |
|-------------------------|---------------------------------|---------|---------|-----------|
| 1                       | 3,5                             | 1       | 12,25   | 3,5       |
| 2                       | 3,5                             | 4       | 12,25   | 7         |
| 3                       | 3,3                             | 9       | 10,89   | 9,9       |
| 4                       | 4,4                             | 16      | 19,36   | 17,6      |
| 5                       | 4,4                             | 25      | 19,36   | 22        |
| 6                       | 4,2                             | 36      | 17,64   | 25,2      |
| 7                       | 5,1                             | 49      | 26,01   | 35,7      |
| 8                       | 5,3                             | 64      | 28,09   | 42,4      |
| 9                       | 5,5                             | 81      | 30,25   | 49,5      |
| 10                      | 5,5                             | 100     | 30,25   | 55        |
| 11                      | 6,2                             | 121     | 38,44   | 68,2      |
| 12                      | 5,7                             | 144     | 32,49   | 68,4      |

L'équation de la droite de régression de  $y$  en  $x$  est

$$y = 0,2552x + 3,058$$

2. Le coefficient de corrélation linéaire est

$$r \approx 0,95$$

$r > 0,84$  donc l'ajustement affine est raisonnable. Il y a une corrélation entre le poids moyen des enfants de cet échantillon et leurs âges.

### Exercice 3

Le tableau ci-dessous représente l'évolution du revenu brut disponible et de la consommation des ménages en euros pour un pays donné sur la période 1992 – 2001. On arrondira les calculs à  $10^{-4}$  près.

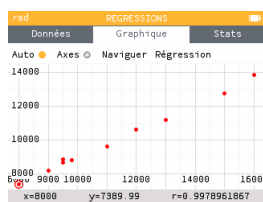
| Année | Revenu | Consommation |
|-------|--------|--------------|
| 1992  | 8000   | 7389,99      |
| 1993  | 9000   | 8169,65      |
| 1994  | 9500   | 8831,71      |
| 1995  | 9500   | 8652,84      |
| 1996  | 9800   | 8788,08      |
| 1997  | 11000  | 9616,21      |
| 1998  | 12000  | 10593,45     |
| 1999  | 13000  | 11186,11     |
| 2000  | 15000  | 12758,09     |
| 2001  | 16000  | 13869,62     |

On cherche à expliquer la consommation des ménages ( $y$ ) par le revenu ( $x$ ).

1. Tracer le nuage de points. Attention aux unités sur les axes !
2. Calculer les coordonnées du point moyen.
3. Déterminer l'équation de la droite de régression linéaire par la méthode des moindres carrés.
4. En 2002 et 2003, on prévoyait respectivement 16 800 et 17 000 euros pour la valeur du revenu. Déterminer les valeurs prévues par le modèle de la consommation pour ces deux années.
5. Déterminer le coefficient de corrélation. Interpréter.

### Solution de l'exercice 3

1.



2. Le point moyen est  $G(11280; 9985,575)$
3. L'équation de la droite de régression linéaire est

$$y = 0,781x + 1176,09$$

4. Avec le modèle :

- En 2002, on a

$$y = 0,781 \times 16800 + 1176,09 = 14296,89$$

Ainsi, en 2002, grâce au modèle, on prévoit une consommation de 14296,89 euros.

- En 2003, on a

$$y = 0,781 \times 17000 + 1176,09 = 14453,09$$

Ainsi, en 2003, grâce au modèle, on prévoit une consommation de 14453,09 euros.

5. Le coefficient de corrélation linéaire est

$$r \approx 0,9979$$

$r > 0,84$  donc l'ajustement affine est raisonnable. Il y a une corrélation entre la consommation des ménages et la revenu brut disponible.

### Exercice 4

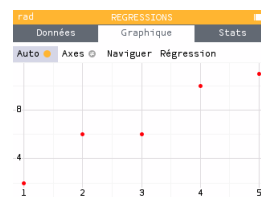
On considère la série statistique double suivante :

|       |   |   |    |   |    |
|-------|---|---|----|---|----|
| $x_i$ | 2 | 3 | 5  | 1 | 4  |
| $y_i$ | 6 | 6 | 11 | 2 | 10 |

1. Représenter graphiquement le nuage de points associé à cette série.
2. Déterminer l'équation de la droite de régression de  $y$  en  $x$ .
3. Déterminer l'équation de la droite de régression de  $x$  en  $y$ .
4. Déterminer le coefficient de corrélation. Interpréter.

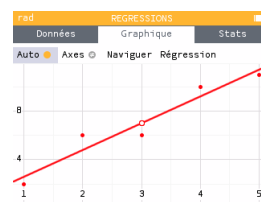
### Solution de l'exercice 4

1.



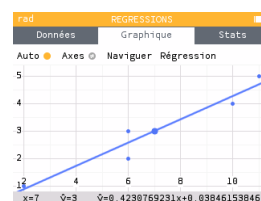
2. L'équation de la droite de régression de  $y$  en  $x$  est

$$y = 2,2x + 0,4$$



3. L'équation de la droite de régression de  $x$  en  $y$  est

$$y = 0,423x + 0,0385$$



4. Le coefficient de corrélation est  $r \approx 0,9648$ . Ainsi  $r > 0,84$  donc les données  $y_i$  sont bien corrélées aux données  $x_i$ .



### Exercice 5

Tous les résultats numériques seront arrondis à l'unité près sauf indication contraire.

Une machine est achetée 3000 euros.

Le prix de revente  $y$ , exprimé en euros, est donné en fonction du nombre  $x$  d'années d'utilisation par le tableau ci-contre

| $x_i$ | 0    | 1    | 2    | 3    | 4    | 5   |
|-------|------|------|------|------|------|-----|
| $y_i$ | 3000 | 2400 | 1920 | 1536 | 1229 | 983 |

1. Représenter le nuage de points associé à la série statistique  $(x_i ; y_i)$  dans un repère orthogonal du plan. Les unités graphiques seront de 2 cm pour une année sur l'axe des abscisses et de 1 cm pour 200 euros sur l'axe des ordonnées.
2. Calculer le pourcentage de dépréciation du prix de revente après les trois premières années d'utilisation.
3. Dans cette question, les calculs effectués à la calculatrice ne seront pas justifiés. Donner une équation de la droite de régression de  $y$  en  $x$  obtenue par la méthode des moindres carrés.  
Représenter la droite dans le repère précédent.

### Solution de l'exercice 5



## Exercice 6

Dans tout l'exercice, le détail des calculs n'est pas demandé. Les résultats seront arrondis à  $10^{-3}$ .

On rappelle que l'image d'un réel  $x$  par la fonction exponentielle peut être notée  $\exp(x) = e^x$ .

On veut étudier l'évolution des records de l'épreuve d'athlétisme du 100 mètres masculin.

Pour cela, on cherche un ajustement des records pour en prévoir l'évolution.

Certains records établis depuis 1900 sont donnés dans le tableau suivant.

|   | Année | Rang de l'année, $x_i$ | Temps en secondes, $y_i$ |
|---|-------|------------------------|--------------------------|
| 1 | 1900  | 0                      | 10,80                    |
| 2 | 1912  | 12                     | 10,60                    |
| 3 | 1921  | 21                     | 10,40                    |
| 4 | 1930  | 30                     | 10,30                    |
| 5 | 1964  | 64                     | 10,06                    |
| 6 | 1983  | 83                     | 9,93                     |
| 7 | 1991  | 91                     | 9,86                     |
| 8 | 1999  | 99                     | 9,79                     |

### Étude d'un modèle affine

1. Construire le nuage de points  $M_i(x_i; y_i)$  pour  $i$  compris entre 1 et 8, associé à cette série statistique double. On prendra comme unité graphique 1 cm pour dix ans en abscisse et 1 cm pour un dixième de seconde en ordonnées, en commençant les graduations au point de coordonnées (0; 9).
2. Peut-on envisager un ajustement affine à court terme ? Cet ajustement permet-il des prévisions pertinentes à long terme sur les records futurs ?

### Étude d'un modèle exponentiel

Après étude, on choisit de modéliser la situation par une autre courbe. On effectue les changements de variables suivants :  $X = e^{-0,00924x}$  et  $Y = \ln y$ . On obtient le tableau suivant :

|                         |       |       |       |       |       |       |       |       |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| $X_i = e^{-0,00924x_i}$ | 1     | 0,895 | 0,824 | 0,758 | 0,554 | 0,464 | 0,431 | 0,401 |
| $Y_i = \ln y_i$         | 2,380 | 2,361 | 2,342 | 2,332 | 2,309 | 2,296 | 2,288 | 2,281 |

1. Donner une équation de la droite de régression de  $Y$  en  $X$  obtenue par la méthode des moindres carrés.
2. En déduire qu'on peut modéliser une expression de  $y$  en fonction de  $x$  sous la forme suivante :

$$y = \exp \left( a e^{-0,00924x} + b \right)$$

où  $a$  et  $b$  sont deux réels à déterminer.

3. À l'aide de cet ajustement, quel record du 100 mètres peut-on prévoir en 2010 ?
4. Calculer la limite en  $+\infty$  de la fonction  $f$  définie sur  $\mathbb{R}$  par l'expression suivante :

$$f(t) = \exp \left( 0,154 e^{-0,00924t} + 2,221 \right)$$

Que peut-on en conclure, en utilisant ce modèle, concernant les records du 100 mètres masculin à très long terme ?

## Solution de l'exercice 6





## Exercice 7

Craignant une propagation de grippe infectieuse, un service de santé d'une ville de 50 000 habitants a relevé le nombre de consultations hebdomadaires concernant cette grippe pendant 7 semaines. Ces semaines ont été numérotées de 1 à 7. On note  $x_i$  les rangs des semaines et  $y_i$  le nombre de consultations :

|                               |     |     |     |      |      |      |      |
|-------------------------------|-----|-----|-----|------|------|------|------|
| Rang de la semaine $x_i$      | 1   | 2   | 3   | 4    | 5    | 6    | 7    |
| Nombre de consultations $y_i$ | 540 | 720 | 980 | 1320 | 1800 | 2420 | 3300 |

1. Tracer le nuage de points sur papier millimétré en prenant 2 cm pour une unité en  $x$  et 1 cm pour 200 en  $y$ . Pourquoi un modèle d'ajustement affine a-t-il été rejeté ?
2. Pour un ajustement exponentiel, on utilise  $z_i = \ln y_i$ . Compléter le tableau suivant (en arrondissant les  $z_i$  à 0,01 près) et déduire une équation de la droite de régression de  $z$  en  $x$  par la méthode des moindres carrés. Exprimer  $y$  en fonction de  $x$  sous la forme  $y = e^{ax+b}$ .

|                          |   |   |   |   |   |   |   |
|--------------------------|---|---|---|---|---|---|---|
| Rang de la semaine $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $z_i = \ln y_i$          |   |   |   |   |   |   |   |

3. À partir du modèle exponentiel, calculer :

- Une estimation du nombre de consultations à la 10ème semaine (arrondie à l'unité).
- La semaine à partir de laquelle le nombre de consultations dépassera le quart de la population.

## Solution de l'exercice 7





## Exercice 8

Le glacier d'Aletsch, classé à l'UNESCO, est le plus grand glacier des Alpes ; situé dans le sud de la Suisse, il alimente la vallée du Rhône.

Pour étudier le recul de ce glacier au fil des années, une première mesure a été effectuée en 1900 : ce glacier mesurait alors 25,6 km.

Des relevés ont ensuite été effectués tous les 20 ans. Le recul du glacier est mesuré par rapport à la position où se trouvait initialement le pied du glacier en 1900. Les mesures successives sont indiquées dans le tableau ci-dessous. On note  $t$  la durée, en années, écoulée depuis 1900, et  $r$  le recul correspondant, mesuré en kilomètres.

| Année de mesure                 | 1900 | 1920 | 1940 | 1960 | 1980 | 2000 |
|---------------------------------|------|------|------|------|------|------|
| Durée $t$ écoulée (depuis 1900) | 0    | 20   | 40   | 60   | 80   | 100  |
| Recul $r$ (en km)               | 0    | 0,3  | 0,6  | 1    | 1,6  | 2,3  |

Par exemple, en 1940 ( $t = 40$ ), le recul du glacier par rapport à 1900 a été de 0,6 km : la longueur du glacier était donc de  $25,6 - 0,6 = 25$  km.

Dans cet exercice, les résultats seront arrondis, si nécessaire, à  $10^{-3}$  près.

### Partie A : Étude d'un modèle affine

- Tracer le nuage de points dans le repère donné en annexe (durée  $t$  en abscisse, distance  $r$  en ordonnée).
- À l'aide de la calculatrice, donner l'équation de la droite d'ajustement affine par la méthode des moindres carrés de  $r$  en fonction de  $t$ , puis tracer cette droite dans le repère précédent.
- À partir du modèle affine obtenu précédemment, estimer par le calcul :
  - Le recul puis la longueur du glacier en 2011.
  - L'année de disparition du glacier (arrondir à l'unité).

### Partie B : Utilisation d'un modèle exponentiel

Le résultat du 3.b de la partie A étant peu en accord avec la plupart des autres études, les glaciologues considèrent un autre modèle : le modèle exponentiel.

On pose  $y = \ln(r)$ . On rappelle que  $\ln(r)$  désigne le logarithme népérien du recul  $r$ .

- Recopier puis compléter le tableau suivant sur votre copie (pour permettre le calcul de  $y$ , la durée 0 de l'année 1900 a été exclue du tableau).

| Durée $t$ (à partir de 1900) | 20 | 40 | 60 | 80 | 100 |
|------------------------------|----|----|----|----|-----|
| $y = \ln(r)$                 |    |    |    |    |     |

- À l'aide de la calculatrice, donner l'équation de la droite d'ajustement affine par la méthode des moindres carrés de  $y$  en fonction de  $t$ .
  - En déduire l'expression de  $r$  en fonction de  $t$ .
- En utilisant le modèle obtenu précédemment, estimer par le calcul :
  - Le recul puis la longueur du glacier en 2011.
  - L'année de disparition du glacier (arrondir à l'unité).

## Solution de l'exercice 8





## Exercice 9

Le tableau suivant donne la population d'une ville nouvelle entre les années 1970 et 2000.

| Année                                  | 1970 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 |
|----------------------------------------|------|------|------|------|------|------|------|
| Rang de l'année $x$                    | 0    | 5    | 10   | 15   | 20   | 25   | 30   |
| Population en milliers d'habitants $y$ | 18   | 21   | 25   | 30   | 36   | 42   | 50   |

Représenter le nuage de points associé à ce tableau avec le rang  $x$  de l'année en abscisse et la population  $y$  en ordonnée.

### Partie A : Un ajustement affine

1. À l'aide de la calculatrice, déterminer une équation de la droite d'ajustement affine de  $y$  en  $x$  par la méthode des moindres carrés (les coefficients seront arrondis au centième).
2. Tracer cette droite sur le graphique.
3. Dédire de cet ajustement une estimation de la population en 2003, à un millier près.

### Partie B : Un ajustement exponentiel

L'allure du nuage incite à chercher un ajustement par une fonction  $f$  définie sur  $[0; +\infty[$  par  $f(x) = ae^{bx}$  où  $a$  et  $b$  sont des réels.

1. Déterminer  $a$  et  $b$  tels que  $f(0) = 18$  et  $f(30) = 50$ . On donnera une valeur arrondie de  $b$  au millième.
2. Dédire de cet ajustement une estimation de la population en 2003, à un millier près.
3. Tracer la courbe représentative de  $f$  sur le graphique.
4. La population en 2003 était de 55 milliers. Lequel des deux ajustements vous semble le plus pertinent ? Justifier votre choix.

### Partie C : Calcul d'une valeur moyenne

On considère maintenant que, pour une année, la population est donnée en fonction du rang  $x$  par  $f(x) = 18e^{0,034x}$ .

1. Calculer la valeur moyenne de la fonction  $f$  sur  $[0; 30]$  ; on donnera le résultat arrondi au dixième.
2. À l'aide d'une lecture graphique, déterminer l'année au cours de laquelle la population atteint cette valeur moyenne.

## Solution de l'exercice 9



## Exercice 10

Une entreprise de services d'une ville cherche à modéliser la consommation des ménages sur les dernières années.

Le rang  $x_i = 1$  correspond à 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  $P_i(x_i; y_i)$  dans un repère orthogonal du plan. On prendra 1 cm comme unité en abscisse et 1 cm pour 10 000 € en ordonnée.
2. Déterminer les coordonnées du point moyen  $G$  de ce nuage et le placer dans le repère précédent.
3. On réalise un ajustement affine de ce nuage par la droite  $D$  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 le résultat sous forme 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  $10^{-2}$ . En déduire que :  $y = 20,49e^{0,23x}$ .
- c. Estimer alors, à l'aide de ce nouvel ajustement, la consommation des ménages de cette ville en 2007 à 100 € près.

## Solution de l'exercice 10





## Exercice 11

Un centre d'appel comptait en 2001 soixante-six employés. Le tableau ci-dessous donne l'évolution du nombre d'employés en fonction du rang de l'année.

| Année                   | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|-------------------------|------|------|------|------|------|------|------|
| Rang de l'année $x_i$   | 1    | 2    | 3    | 4    | 5    | 6    | 7    |
| Nombre d'employés $y_i$ | 66   | 104  | 130  | 207  | 290  | 345  | 428  |

On cherche à étudier l'évolution du nombre  $y$  d'employés en fonction du rang  $x$  de l'année. Une étude graphique montre qu'un ajustement affine ne convient pas.

On pose alors  $z = \sqrt{y} - 3$ .

1. Recopier et compléter le tableau suivant (on donnera les résultats sous forme décimale, arrondis au centième).

|                       |      |   |   |   |   |   |   |
|-----------------------|------|---|---|---|---|---|---|
| Rang de l'année $x_i$ | 1    | 2 | 3 | 4 | 5 | 6 | 7 |
| $z_i$                 | 5,12 |   |   |   |   |   |   |

2. Représenter le nuage de points  $M_i(x_i; z_i)$  associé à cette série statistique dans le plan muni d'un repère orthonormal d'unité graphique 1 cm.

Un ajustement affine vous paraît-il approprié ? Justifier la réponse.

3. Déterminer, à l'aide de la calculatrice, une équation de la droite d'ajustement affine de  $z$  en  $x$  par la méthode des moindres carrés (on donnera les coefficients sous forme décimale, arrondis au centième). Tracer cette droite sur le graphique précédent.

En utilisant cet ajustement, à partir de quelle année peut-on prévoir que l'effectif de ce centre d'appel dépassera 900 employés ?

## Solution de l'exercice 11



## Exercice 12

### Partie 1

Sachant qu'il y avait 13 millions de cotisants au régime général de retraites en France métropolitaine en 1975 et 16,6 millions de cotisants en 2005, calculer le pourcentage d'augmentation du nombre de cotisants entre 1975 et 2005. On arrondira le résultat à 0,1% près.

### Partie 2

Le tableau ci-dessous donne le nombre de retraités en France métropolitaine entre 1975 et 2005 :

| Année                                   | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2005 |
|-----------------------------------------|------|------|------|------|------|------|------|
| Rang de l'année $x_i$                   | 0    | 1    | 2    | 3    | 4    | 5    | 6    |
| Nombre de retraités (en millions) $y_i$ | 4,1  | 5,0  | 5,9  | 7,4  | 8,3  | 9,7  | 10,7 |

Source : INSEE / Caisse Nationale d'Assurance Vieillesse 2007

1. Sur une feuille de papier millimétré, représenter le nuage de points  $M_i(x_i; y_i)$  associé à la série statistique dans un repère orthogonal d'unités graphiques 2 cm en abscisse (pour les rangs d'année) et 1 cm en ordonnée (pour 1 million de retraités).
2.
  - a. Calculer les coordonnées du point moyen  $G$  de cette série statistique.
  - b. Donner, à l'aide de la calculatrice, l'équation réduite de la droite  $d$  d'ajustement de  $y$  en  $x$  par la méthode des moindres carrés (on arrondira les coefficients au dixième).
  - c. Placer le point  $G$  et tracer la droite  $d$  dans le repère construit à la première question.
  - d. En utilisant l'ajustement trouvé à la question 2, déterminer par un calcul une estimation du nombre de retraités en 2010.

### Partie 3

On utilisera les données des parties 1 et 2. Dans cette partie, les résultats seront donnés sous forme de pourcentage, arrondis au dixième.

On appelle rapport démographique de l'année  $n$  le rapport

$$R_n = \frac{\text{nombre de cotisants de l'année } n}{\text{nombre de retraités de l'année } n}.$$

1. Calculer le taux d'évolution de  $R_n$  entre 1975 et 2005.
2. Entre 2005 et 2010, une étude montre que le nombre de cotisants devrait augmenter de 6,4% et que le nombre de retraités devrait augmenter de 12,1%. Calculer le taux d'évolution du rapport démographique entre 2005 et 2010.

Toute trace de recherche, même incomplète, sera prise en compte dans l'évaluation.

## Solution de l'exercice 12



### Exercice 13

Le responsable d'un site Internet s'intéresse au nombre de pages visitées sur son site chaque semaine.

#### Partie A

Le tableau ci-dessous donne le nombre de pages visitées, exprimé en milliers, durant chacune des quatre semaines suivant l'ouverture du site.

|                                              |    |    |    |    |
|----------------------------------------------|----|----|----|----|
| Semaine : $x_i$                              | 1  | 2  | 3  | 4  |
| Nombre de pages visitées en milliers : $y_i$ | 40 | 45 | 55 | 70 |

Ainsi, au cours de la deuxième semaine après l'ouverture du site, 45 000 pages ont été visitées.

1. a. Représenter le nuage de points  $M_i(x_i; y_i)$  associé à cette série statistique dans un repère orthogonal. L'allure de ce nuage suggère-t-elle un ajustement affine ?  
b. Déterminer les coordonnées du point moyen  $G$  de ce nuage, puis placer ce point sur le graphique.  
c. On appelle  $(d)$  la droite d'ajustement de  $y$  en  $x$  obtenue par la méthode des moindres carrés. Parmi les deux propositions ci-dessous, une seule correspond à l'équation réduite de la droite  $(d)$ . Préciser laquelle, en utilisant le point moyen  $G$  :

$$y = 9x + 29 \quad \text{ou} \quad y = 10x + 27,5$$

- d. Tracer la droite  $(d)$  sur le graphique.
2. En supposant que cet ajustement reste valable pendant les deux mois qui suivent l'ouverture du site, donner une estimation du nombre de pages visitées au cours de la huitième semaine suivant l'ouverture du site.

#### Partie B

Le responsable décide de mettre en place, au cours de la quatrième semaine suivant l'ouverture du site, une vaste campagne publicitaire afin d'augmenter le nombre de visiteurs du site.

|                                              |    |    |    |    |    |     |     |
|----------------------------------------------|----|----|----|----|----|-----|-----|
| Semaine $x_i$                                | 1  | 2  | 3  | 4  | 5  | 6   | 7   |
| Nombre de pages visitées en milliers : $y_i$ | 40 | 45 | 55 | 70 | 95 | 125 | 175 |

1. Compléter le nuage de points en ajoutant les trois nouveaux points définis dans le tableau précédent. Compte tenu de l'allure du nuage, un ajustement exponentiel semble approprié. Pour cela, on pose  $z = \ln y$ .
2. On donne ci-dessous les valeurs de  $z_i = \ln(y_i)$  pour  $1 \leq i \leq 7$ , les résultats étant arrondis au centième.

|                  |      |      |      |      |      |      |      |
|------------------|------|------|------|------|------|------|------|
| Semaine $x_i$    | 1    | 2    | 3    | 4    | 5    | 6    | 7    |
| $z_i = \ln(y_i)$ | 3,69 | 3,81 | 4,01 | 4,25 | 4,55 | 4,83 | 5,16 |

- a. À l'aide de la calculatrice, déterminer une équation de la droite d'ajustement de  $z$  en  $x$  obtenue par la méthode des moindres carrés. On donnera la réponse sous la forme  $z = ax + b$ , en arrondissant les coefficients  $a$  et  $b$  au centième.
- b. En déduire la relation  $y = \alpha e^{\beta x}$ , où 27,94 et 0,25 sont des valeurs approchées au centième des réels  $\alpha$  et  $\beta$ , respectivement.
- c. À l'aide de ce nouvel ajustement, donner une estimation du nombre de pages visitées au cours de la huitième semaine suivant l'ouverture du site.  
Combien de semaines auraient été nécessaires pour atteindre ce résultat sans campagne publicitaire ? (On utilisera l'ajustement obtenu dans la partie A).



## Exercice 14

Le tableau ci-dessous donne l'évolution de la population d'un pays de 1950 à 1985.

$t_i$  désigne le rang de l'année et  $p_i$  la population en millions d'habitants.

| Année                 | 1950 | 1955 | 1960 | 1965 | 1970 | 1975 | 1980 | 1985 |
|-----------------------|------|------|------|------|------|------|------|------|
| Rang de l'année $t_i$ | 0    | 5    | 10   | 15   | 20   | 25   | 30   | 35   |
| $p_i$ (en millions)   | 8    | 8,9  | 9,9  | 11   | 12   | 13,5 | 15   | 16,6 |

### Partie A : Exploitation des données - Recherche d'un modèle

1. Représenter le nuage de points  $M_i(t_i; p_i)$  associé à la série statistique dans un repère orthogonal.

Sur l'axe des abscisses, choisir 2 cm pour 5 unités (5 ans).

Sur l'axe des ordonnées, placer 8 à l'origine, puis choisir 2 cm pour une unité (1 million d'habitants).

2. Les experts cherchent à modéliser cette évolution par une fonction dont la courbe est voisine du nuage de points.

- On pose  $y_i = \ln p_i$ . Donner une valeur approchée à  $10^{-3}$  près par défaut du coefficient de corrélation linéaire  $r$  de la série  $(t_i; y_i)$ .
- Déterminer une équation de la droite de régression de  $y$  en  $t$ . Les coefficients seront arrondis à  $10^{-3}$  près.
- En déduire l'expression de la population  $p$  en fonction du rang  $t$  de l'année.

### Partie B : Étude du modèle exponentiel

On admet que la fonction  $f$  définie sur  $[0; 35]$  par :  $f(t) = 8e^{0,02t}$  est une modélisation satisfaisante de l'évolution de la population (en millions d'habitants) de 1950 à 1985.

- Étudier le sens de variation de  $f$  sur  $[0; 35]$  et dresser le tableau de variation complet de  $f$  sur cet intervalle.
- Construire soigneusement la courbe représentative de  $f$ , notée  $(C)$ , dans le repère de la partie A. Qu'observe-t-on ?
- On pose  $I = \int_0^{35} f(t) dt$ . Donner une valeur approchée de  $I$ , arrondie à  $10^{-2}$  près.  
En déduire la population moyenne  $m$  du pays durant ces 35 années et la représenter sur le graphique.
- Calculer le rapport  $\frac{f(t+1)-f(t)}{f(t)}$  et en donner une interprétation en termes de pourcentage.
- Si le modèle exponentiel étudié dans la partie B restait valable après 1985, en quelle année la population aurait-elle dépassé les 19 millions d'habitants ?

## Solution de l'exercice 14



## Exercice 15

Le tableau suivant, publié en août 1999 dans une revue économique, donne la part du temps partiel au sein de la population active (les valeurs pour 2000 et 2004 sont le résultat d'une estimation).

| Année $x_i$                        | 1980 | 1985 | 1990 | 1995 | 1997 | 2000 | 2004 |
|------------------------------------|------|------|------|------|------|------|------|
| Part du temps partiel en % : $y_i$ | 8,3  | 11   | 12   | 15,6 | 16,8 | 18   | 20   |

On étudie la série statistique  $(x_i, y_i)$  pour  $1980 \leq x_i \leq 1997$ . Les calculs seront effectués à la calculatrice.

1. Représenter dans un repère orthogonal le nuage de points de coordonnées  $(x_i, y_i)$  pour  $1980 \leq x_i \leq 1997$ .

On prendra 1 cm pour une part de 2% en ordonnée, 2 cm pour 5 ans en abscisse en prenant pour origine le point (1980; 0).

2. Déterminer les coordonnées de  $G$ , point moyen de la série statistique  $(x_i, y_i)$ . Le placer sur le graphique.

3. a. Donner la valeur arrondie à  $10^{-3}$  près du coefficient de corrélation linéaire de la série  $(x_i, y_i)$ . Un ajustement affine est-il justifié ?

b. Déterminer une équation de la droite d'ajustement affine de  $y$  en  $x$  par la méthode des moindres carrés (avec  $a$  et  $b$  arrondis à  $10^{-3}$  près). Dessiner cette droite sur le graphique.

c. Peut-on considérer que les estimations pour 2000 et 2004 faites par la revue ont été réalisées en utilisant l'équation obtenue à la question 3.b. ?

## Solution de l'exercice 15





## Exercice 16

Le prix de vente des terrains à bâtir dans la même commune rurale est donné par le tableau suivant :

|                                        |      |      |      |      |      |      |      |
|----------------------------------------|------|------|------|------|------|------|------|
| Année                                  | 1980 | 1985 | 1987 | 1990 | 1995 | 1997 | 2000 |
| Rang de l'année $x_i$                  | 0    | 5    | 7    | 10   | 15   | 17   | 20   |
| Prix du m <sup>2</sup> en francs $y_i$ | 58,8 | 60,9 | 62,1 | 67,5 | 71,7 | 73   | 73,8 |

1. Quelle est, en pourcentage, l'augmentation du prix du m<sup>2</sup> entre 1980 et 2000 ?
2. Représenter le nuage de points  $M_i(x_i, y_i)$  dans un repère orthogonal où 5 cm représentent 10 ans en abscisses et 5 cm représentent 10 francs en ordonnées.
3. Déterminer le point moyen  $G$  du nuage et le placer sur le graphique.
4. Donner le coefficient de corrélation linéaire de la série  $(x_i, y_i)$  à 0,01 près.

On considère que ce coefficient justifie un ajustement affine par la méthode des moindres carrés. Écrire une équation de la droite d'ajustement affine de  $y$  en  $x$ , notée  $(D)$  (les coefficients sont arrondis à 0,01). Tracer  $(D)$ .

5. Estimer à 1 millier de francs près le prix d'un terrain de 1500 m<sup>2</sup> en 2003.

## Solution de l'exercice 16



## Exercice 17

Les résultats numériques seront obtenus à l'aide de la calculatrice ; aucun détail des calculs statistiques n'est demandé.

Le tableau suivant donne la dépense, en millions d'euros, des ménages en produits informatiques (matériels, logiciels, réparations) de 1990 à 1998.

| Année                               | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 |
|-------------------------------------|------|------|------|------|------|------|------|------|------|
| Rang de l'année $x_i$               | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    |
| Dépense $y_i$ (en millions d'euros) | 398  | 451  | 423  | 501  | 673  | 956  | 1077 | 1255 | 1427 |

### Source INSEE

- Représenter le nuage de points associé à la série statistique  $(x_i, y_i)$  et le point moyen dans un repère orthogonal tel que 2 cm représentent une année en abscisse et 1 cm représente 100 millions d'euros en ordonnée (ainsi 398 sera représenté par 3,98 cm).
- Donner la valeur arrondie à  $10^{-3}$  du coefficient de corrélation linéaire de la série  $(x_i, y_i)$ . Un ajustement affine paraît-il justifié ?
  - Écrire une équation de la droite d'ajustement affine  $D$  de  $y$  en  $x$  par la méthode des moindres carrés (les coefficients seront arrondis à  $10^{-3}$ ). Représenter  $D$  dans le repère précédent.
  - En utilisant cet ajustement affine, donner une estimation de la dépense des ménages (arrondie à un million d'euros) en produits informatiques en 2000.
- L'allure du nuage permet d'envisager un ajustement exponentiel. On pose  $z_i = \ln y_i$ .
  - Recopier et compléter le tableau suivant où  $z_i$  est arrondi à  $10^{-3}$  :

|       |       |       |       |       |   |   |   |   |   |
|-------|-------|-------|-------|-------|---|---|---|---|---|
| $x_i$ | 0     | 1     | 2     | 3     | 4 | 5 | 6 | 7 | 8 |
| $z_i$ | 5,986 | 6,111 | 6,047 | 6,217 |   |   |   |   |   |

- Donner la valeur arrondie à  $10^{-3}$  du coefficient de corrélation linéaire de la série  $(x_i, z_i)$ . Écrire une équation de la droite d'ajustement affine de  $z$  en  $x$  par la méthode des moindres carrés (les coefficients seront arrondis à  $10^{-3}$ ).
  - En utilisant cet ajustement, donner une estimation de la dépense des ménages (arrondie à un million d'euros) en produits informatiques en 2000.
- En 2000, les ménages ont dépensé 68,9 milliards d'euros pour la culture, les loisirs et les sports, dont 3,1% pour les produits informatiques.  
Avec lequel des deux ajustements l'estimation faite est-elle la meilleure ?

## Solution de l'exercice 17



## Exercice 18

Les guichets d'une agence bancaire d'une petite ville sont ouverts au public cinq jours par semaine : mardi, mercredi, jeudi, vendredi et samedi.

Le tableau ci-dessous donne la répartition journalière des 250 retraits d'argent liquide effectués aux guichets une certaine semaine.

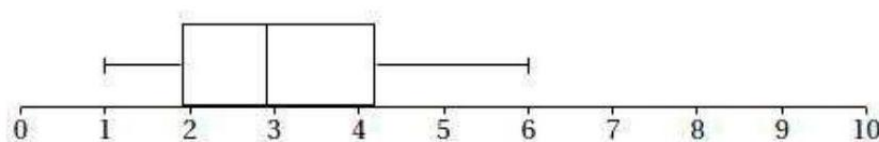
| Jour de la semaine | mardi | mercredi | jeudi | vendredi | samedi |
|--------------------|-------|----------|-------|----------|--------|
| Rang $i$ du jour   | 1     | 2        | 3     | 4        | 5      |
| Nombre de retraits | 37    | 55       | 45    | 53       | 60     |

On veut tester l'hypothèse « le nombre de retraits est indépendant du jour de la semaine ». On suppose donc que le nombre des retraits journaliers est égal à  $\frac{1}{5}$  du nombre des retraits de la semaine.

On pose  $d_{\text{obs}}^2 = \sum_{i=1}^5 \left(f_i - \frac{1}{5}\right)^2$ , où  $f_i$  est la fréquence des retraits du  $i$ -ième jour.

### Partie A

1. Calculer les fréquences des retraits pour chacun des cinq jours de la semaine.
2. Calculer alors la valeur de  $1000 d_{\text{obs}}^2$  (la multiplication par 1000 permet d'obtenir un résultat plus lisible).
3. En supposant qu'il y a équiprobabilité des retraits journaliers, on a simulé 2000 séries de 250 retraits hebdomadaires. Pour chaque série, on a calculé la valeur du  $1000 d_{\text{obs}}^2$  correspondant. Ces valeurs ont permis de construire le diagramme en boîte ci-dessous où les extrémités des « pattes » correspondent respectivement au premier décile et au neuvième décile.



4. Lire sur le diagramme une valeur approchée du neuvième décile.  
En argumentant soigneusement la réponse, dire si pour la série observée au début, on peut affirmer, avec un risque d'erreur inférieur à 10%, que « le nombre de retraits est indépendant du jour de la semaine ».

**Solution de l'exercice 18** Des éléments de solutions sont disponibles à la première page de ce document :





## Exercice 19

En 2004, une caisse de retraite propose à ses adhérents un barème de rachat d'un trimestre de cotisation des années antérieures selon le tableau suivant :

| Âge de l'adhérent en années                                   | 54   | 55   | 56   | 57   | 58   |
|---------------------------------------------------------------|------|------|------|------|------|
| Rang $x_i$                                                    | 0    | 1    | 2    | 3    | 4    |
| Montant $y_i$ du rachat d'un trimestre de cotisation en euros | 2229 | 2285 | 2340 | 2394 | 2449 |

(Source : CARMF Mai 2004)

1. Calculer l'augmentation en pourcentage du montant du rachat d'un trimestre entre un salarié de 54 ans et un salarié de 58 ans. On donnera le résultat arrondi à l'unité.
2. Sur votre copie, représenter le nuage de points associé à la série statistique  $(x_i; y_i)$  dans un repère orthogonal :
  - sur l'axe des abscisses, on placera 0 à l'origine et on choisira 2 cm pour une unité ;
  - sur l'axe des ordonnées, on placera 2200 à l'origine et on choisira 1 cm pour 20 euros.
3. Dans cette question, les calculs effectués à la calculatrice ne seront pas justifiés. Le nuage de points permet de penser qu'un ajustement affine est justifié.
  - a. Donner une équation de la droite de régression ( $D$ ) de  $y$  en  $x$ , obtenue par la méthode des moindres carrés.
  - b. Représenter la droite ( $D$ ) dans le repère précédent.
  - c. Quel serait, avec cet ajustement affine, le montant du rachat d'un trimestre pour un salarié âgé de 60 ans ?
  - d. En fait, le montant du rachat d'un trimestre pour un salarié âgé de 60 ans est de 2555 euros, et le montant du rachat d'un trimestre après 60 ans est calculé de la façon suivante : à partir de 60 ans, le montant du rachat baisse de 3% par an. Calculer le montant du rachat d'un trimestre pour un salarié ayant 65 ans.

## Solution de l'exercice 19



## Exercice 20

Une enquête menée pour le compte d'une entreprise a permis d'établir le nombre d'acheteurs d'un produit X selon le montant de son prix de vente. Les résultats de l'enquête sont résumés dans le tableau ci-dessous dans lequel :

- $x_i$  désigne le prix de vente unitaire (en euros) du produit X ;
- $y_i$  le nombre d'acheteurs en milliers.

|       |      |      |   |   |     |
|-------|------|------|---|---|-----|
| $x_i$ | 1    | 1,50 | 2 | 3 | 4   |
| $y_i$ | 3,75 | 2,8  | 2 | 1 | 0,5 |

1. Représenter sur votre copie le nuage de points associé à la série  $(x_i; y_i)$  dans un repère orthogonal  $(O; \vec{i}, \vec{j})$  du plan avec les unités graphiques : 4 cm pour 1 euro en abscisse et 2 cm pour 1000 acheteurs en ordonnée.
2. On recherche un ajustement affine de la série  $(x_i; y_i)$ .
  - a. Donner l'équation de la droite d'ajustement de  $y$  en  $x$  obtenue par la méthode des moindres carrés.

Les calculs seront faits à la calculatrice, et les valeurs cherchées seront arrondies au centième ; aucune justification n'est demandée.

2.
  - b. Tracer cette droite dans le même repère que précédemment.
  - c. Utiliser cet ajustement pour estimer le nombre d'acheteurs potentiels pour un produit vendu 2,50 euros.

## Solution de l'exercice 20



## Exercice 21

### Partie A

Dans un pays européen, le montant des recettes touristiques, exprimées en millions d'euros, est donné dans le tableau ci-dessous :

| Année                                                       | 2000  | 2001  | 2002  | 2003  | 2004  | 2005  |
|-------------------------------------------------------------|-------|-------|-------|-------|-------|-------|
| Rang de l'année $x_i$                                       | 0     | 1     | 2     | 3     | 4     | 5     |
| Montant des recettes touristiques $y_i$ en millions d'euros | 24495 | 26500 | 29401 | 33299 | 33675 | 34190 |

1. On utilise un ajustement affine. Donner, à l'aide de la calculatrice, l'équation de la droite d'ajustement de  $y$  en  $x$ , obtenue par la méthode des moindres carrés. Les coefficients, obtenus à l'aide de la calculatrice, seront arrondis au centième.
2. En supposant que cet ajustement est valable jusqu'en 2007, calculer le montant que l'on peut prévoir pour les recettes touristiques de l'année 2007, arrondi au million d'euros.

### Partie B

On considère la fonction  $f$  définie pour tout nombre entier  $n$  par  $f(n) = e^{10,13+0,07n}$ .

On utilise cette fonction pour modéliser l'évolution des recettes touristiques de ce pays européen. Ainsi,  $f(n)$  représente le montant des recettes touristiques (exprimé en millions d'euros) de ce pays européen pour l'année  $2000 + n$ .

1. Selon ce modèle, calculer le montant des recettes touristiques que l'on peut prévoir pour l'année 2007. Arrondir le résultat au million d'euros.
2. a. Déterminer le nombre entier  $n$  à partir duquel  $f(n) > 45000$ .  
b. En déduire l'année à partir de laquelle, selon ce modèle, le montant des recettes touristiques dépasserait 45000 millions d'euros.

## Solution de l'exercice 21



## Exercice 22

On se propose d'étudier l'évolution des ventes d'un modèle de voiture de gamme moyenne depuis sa création en 1999. Les parties I et II peuvent être traitées indépendamment l'une de l'autre.

### Partie I

Le tableau suivant donne le nombre annuel, exprimé en milliers, de véhicules vendus les cinq premières années de commercialisation :

| Année                                                 | 1999 | 2000 | 2001  | 2002  | 2003  |
|-------------------------------------------------------|------|------|-------|-------|-------|
| Rang de l'année : $x_i$                               | 0    | 1    | 2     | 3     | 4     |
| Nombre annuel de véhicules vendus en milliers : $y_i$ | 81,3 | 92,3 | 109,7 | 128,5 | 131,2 |

- Dans le plan (P) muni d'un repère orthogonal d'unités graphiques 1 cm pour une année sur l'axe des abscisses et 1 cm pour 10 milliers de véhicules vendus sur l'axe des ordonnées, représenter le nuage de points associé à la série statistique  $(x_i; y_i)$  pour  $i$  entier variant de 0 à 4.
- L'allure du nuage de points permet d'envisager un ajustement affine.
  - Déterminer les coordonnées du point moyen  $G$  de ce nuage.
  - Déterminer l'équation  $y = ax + b$  de la droite  $(D)$  d'ajustement affine de  $y$  en  $x$  obtenue par la méthode des moindres carrés.
  - Placer le point  $G$  et tracer la droite  $(D)$  sur le graphique précédent.
  - En utilisant l'ajustement affine de b., donner une estimation du nombre de véhicules vendus en 2007.
- Le tableau suivant donne le nombre annuel de véhicules vendus, exprimé en milliers, de 2003 à 2007 :

| Année                                                 | 2003  | 2004  | 2005  | 2006 | 2007 |
|-------------------------------------------------------|-------|-------|-------|------|------|
| Rang de l'année : $x_i$                               | 4     | 5     | 6     | 7    | 8    |
| Nombre annuel de véhicules vendus en milliers : $y_i$ | 131,2 | 110,8 | 101,4 | 86,3 | 76,1 |

- Compléter le nuage de points précédent à l'aide de ces valeurs.
- L'ajustement précédent est-il encore adapté ? Justifier la réponse.
- On décide d'ajuster le nuage de points associé à la série statistique  $(x_i; y_i)$ , pour  $i$  entier variant de 4 à 8, par une courbe qui admet une équation de la forme  $y = e^{cx+d}$ .  
Déterminer les réels  $c$  et  $d$  pour que cette courbe passe par les points  $A(4; 131,2)$  et  $B(8; 76,1)$ . On donnera la valeur exacte, puis l'arrondi au millième de chacun de ces nombres réels.

### Partie II

Soit  $f$  la fonction définie sur l'intervalle  $[4; 10]$  par  $f(x) = e^{-0,136x+5,421}$ .

On suppose que  $f$  modélise en milliers l'évolution du nombre annuel de véhicules vendus à partir de l'année 2003.

- Déterminer le sens de variation de la fonction  $f$  sur l'intervalle  $[4; 10]$ .
- Tracer la courbe  $(C)$  représentative de la fonction  $f$  dans le même repère que le nuage de points.
- L'entreprise décide d'arrêter la fabrication du modèle l'année où le nombre annuel de véhicules vendus devient inférieur à 65 000.
  - Résoudre algébriquement dans l'intervalle  $[4; 10]$  l'inéquation  $f(x) < 65$ . En quelle année l'entreprise doit-elle prévoir cet arrêt ?
  - Retrouver graphiquement le résultat précédent en laissant apparents les traits de construction nécessaires.

Le tableau ci-dessous donne l'évolution de l'indice des prix de vente des appartements anciens à Paris au quatrième trimestre des années 2000 à 2007.

| Année                   | 2000 | 2001  | 2002  | 2003  | 2004  | 2005  | 2006  | 2007  |
|-------------------------|------|-------|-------|-------|-------|-------|-------|-------|
| Rang de l'année : $x_i$ | 0    | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
| Indice : $y_i$          | 100  | 108,5 | 120,7 | 134,9 | 154,8 | 176,4 | 193,5 | 213,6 |

(Source : INSEE)

1. Calculer le pourcentage d'augmentation de cet indice de l'année 2000 à l'année 2007.
2. Construire le nuage de points  $M_i(x_i; y_i)$  dans le plan  $(P)$  muni d'un repère orthogonal défini de la manière suivante :
  - sur l'axe des abscisses, on placera 0 à l'origine et on choisira 2 cm pour représenter une année ;
  - sur l'axe des ordonnées, on placera 100 à l'origine et on choisira 1 cm pour représenter 10 unités.
3. Déterminer les coordonnées du point moyen  $G$  de ce nuage. Placer le point  $G$  dans le plan  $(P)$ .
4. L'allure de ce nuage permet de penser qu'un ajustement affine est adapté.
  - a. À l'aide de la calculatrice, déterminer une équation de la droite  $(d)$  d'ajustement de  $y$  en  $x$ , obtenue par la méthode des moindres carrés. Les coefficients seront arrondis au centième.
  - b. Tracer la droite  $(d)$  dans le plan  $(P)$ .
5. En supposant que cet ajustement affine reste valable pour les deux années suivantes, estimer l'indice du prix de vente des appartements anciens de Paris au quatrième trimestre 2009. Justifier la réponse.

**Solution de l'exercice 22** Page 5 du document accessible sur ce lien :



## Exercice 23

Pour  $i$  nombre entier variant de 0 à 8, on définit le tableau suivant qui donne les valeurs du SMIC horaire brut, exprimé en euros, de 2001 à 2009 (source INSEE).

On se propose d'en étudier l'évolution :

| Année                               | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|-------------------------------------|------|------|------|------|------|------|------|------|------|
| Rang de l'année : $x_i$             | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    |
| SMIC horaire brut (en euros), $y_i$ | 6,67 | 6,83 | 7,19 | 7,61 | 8,03 | 8,27 | 8,44 | 8,71 | 8,82 |

Dans tout l'exercice, les pourcentages seront arrondis à 0,01% et les valeurs du SMIC horaire brut au centime d'euro.

### Partie A : Observation des données

1. Pour  $i$  entier variant de 0 à 8, représenter le nuage de points  $M_i(x_i; y_i)$  dans le plan muni d'un repère orthogonal défini de la façon suivante :

- sur l'axe des abscisses, on placera 0 à l'origine et on choisira 1 cm pour 1 année ;
- on graduera l'axe des ordonnées en commençant à 6 et on choisira 5 cm pour 1 euro.

2. Calculer le pourcentage d'augmentation de la valeur du SMIC horaire brut entre 2001 et 2009.

3. Démontrer qu'une valeur approchée du pourcentage annuel moyen d'augmentation de la valeur du SMIC horaire brut entre 2001 et 2005 est 4,75%.

On observe sur le graphique un changement de tendance à partir de 2005 : le pourcentage annuel moyen d'augmentation de la valeur du SMIC horaire brut est alors de 2,4% environ.

En supposant que cette nouvelle tendance se poursuive, on désire estimer la valeur du SMIC horaire brut en 2012.

Dans la suite de l'exercice, on ne s'intéresse qu'au sous-nuage constitué des cinq derniers points  $M_4, M_5, M_6, M_7$  et  $M_8$  du nuage précédent.

### Partie B : Modélisation de la série statistique $(x_i; y_i)_{4 \leq i \leq 8}$ par un ajustement exponentiel

En observant le pourcentage annuel moyen d'augmentation de la valeur du SMIC horaire brut entre 2005 et 2009, on estime à  $8,03 \times 1,024^n$  la valeur, exprimée en euros, du SMIC horaire brut pour l'année  $2005 + n$ ,  $n$  désignant un entier naturel.

On considère que ce nouveau modèle reste valable jusqu'à l'année 2016.

4. Calculer une estimation de la valeur du SMIC horaire brut en 2012.

5. À partir de quelle année la valeur du SMIC horaire brut dépassera-t-elle 10 euros ?

La Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (CNAMTS) publie chaque année des statistiques sur les accidents du travail en France. Celles-ci permettent d'obtenir divers indicateurs, notamment l'indice de fréquence (nombre moyen d'accidents du travail avec arrêt pour 1000 salariés).

Le tableau ci-dessous donne l'évolution de l'indice de fréquence pour le secteur du BTP (Bâtiment et Travaux Publics) en France, au cours des années 2001 à 2009 :

| Année                       | 2001  | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|-----------------------------|-------|------|------|------|------|------|------|------|------|
| Rang de l'année : $x_i$     | 1     | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
| Indice de fréquence : $y_i$ | 100,3 | 98,9 | 91,6 | 89,5 | 87,6 | 85,4 | 84,0 | 79,9 | 76,0 |

### Premier ajustement

Grâce à un logiciel, un élève a obtenu le nuage de points représentant la série statistique  $(x_i; y_i)$  et, par la méthode des moindres carrés, la droite d'ajustement de  $y$  en  $x$  dont une équation est  $y = -2,89x + 102,59$  (les coefficients sont arrondis à 0,01).

- En supposant que cet ajustement affine est valable jusqu'en 2012, déterminer une estimation de l'indice de fréquence en l'année 2012.
- Quel serait le pourcentage d'évolution entre 2007 et 2012 de l'indice de fréquence selon ce modèle ? On arrondira le résultat à  $10^{-2}$ .

## Deuxième ajustement

Un autre élève envisage un ajustement exponentiel de la série statistique  $(x_i; y_i)$ . On pose  $z_i = \ln y_i$ . Recopier et compléter le tableau ci-dessous (les valeurs de  $z_i$  seront arrondies à  $10^{-3}$ ).

| $x_i$           | 1     | 2     | 3     | 4 | 5 | 6 | 7 | 8 | 9 |
|-----------------|-------|-------|-------|---|---|---|---|---|---|
| $z_i = \ln y_i$ | 4,608 | 4,594 | 4,517 |   |   |   |   |   |   |

À l'aide de la calculatrice, déterminer, par la méthode des moindres carrés, une équation de la droite d'ajustement de  $z$  en  $x$  sous la forme  $z = ax + b$ , les coefficients  $a$  et  $b$  étant arrondis à  $10^{-4}$ . En déduire une expression de  $y$  en fonction de  $x$  sous la forme  $y = Ke^{-0,0328x}$ ,  $K$  étant une constante arrondie à  $10^{-1}$  près.

3. Dans cette question, toute trace de recherche, même incomplète, ou d'initiative même non fructueuse, sera prise en compte dans l'évaluation.

La stratégie européenne de santé au travail a fixé comme objectif une réduction de 25% de l'indice de fréquence entre 2007 et 2012. Peut-on prévoir d'atteindre cet objectif selon les deux ajustements précédents, que l'on suppose valables jusqu'en 2012 ?

**Solution de l'exercice 23** Page 4 du document accessible sur ce lien :



