Devoir surveillé n°3

Statistiques à deux variables

EXERCICE 3.1 (6 points).

Le tableau suivant donne, en milliers, le nombre de Pactes civils de solidarité (PACS) signés chaque année en France :

Années	2 000	2 001	2 002	2 003	2 004
Rang de l'année, x_i	0	1	2	3	4
Nombres de PACS en milliers, y_i	22,1	19,4	25	31,1	39,6

Source INSEE.

1. **On envisage un ajustement affine**
   (a) À l'aide de la calculatrice, donner l'équation de la droite d'ajustement de y en x par la méthode des moindres carrés, sous la forme $y = ax + b$. Par la suite, on pose $f(x) = ax + b$.
   (b) En supposant que cet ajustement affine est valable jusqu'en 2 007, donner une estimation du nombre de milliers de Pactes civils de solidarité signés en 2 007.
2. **On envisage un autre type d'ajustement**
   On modélise le nombre de milliers de Pactes civils de solidarité signés durant l'année $2000 + x$ (x entier) à l'aide de la fonction g définie par
   $$g(x) = 1,6x^2 - 1,8x + 21,4.$$
   (a) En utilisant ce second modèle, calculer le nombre de milliers de Pactes civils de solidarité signés en 2 007.
   (b) On suppose que l'évolution se poursuit selon ce modèle.
   Déterminer, à l'aide de cet ajustement et par le calcul, en quelle année le nombre de Pactes civils de solidarité signés sera supérieur à 100 000.
3. **Comparaison des deux ajustements**
   On a calculé ci-dessous le tableau des carrés des écarts entre les valeurs réelles et les valeurs calculées avec le premier ajustement.

x_i	0	1	2	3	4
$\left[(y_i - f(x_i)\right]^2$	16	11,36	5,95	1,02	7,95

   (a) Compléter le tableau 3.1, page 53, les valeurs étant arrondies au centième.
   (b) Lequel de ces deux ajustements semble le plus proche de la réalité ? Justifier.

EXERCICE 3.2 (8,5 points).

Dans une région de 1 000 km^2, la superficie moyenne des terrains urbanisés entre 1970 et 1998 est donnée par le tableau suivant

année a_i	1970	1974	1978	1982	1986	1990	1994	1998
rang de l'année x_i	0	4	8	12	16	20	24	28
superficie en km^2 y_i	80	94	110	129	152	178	205	236

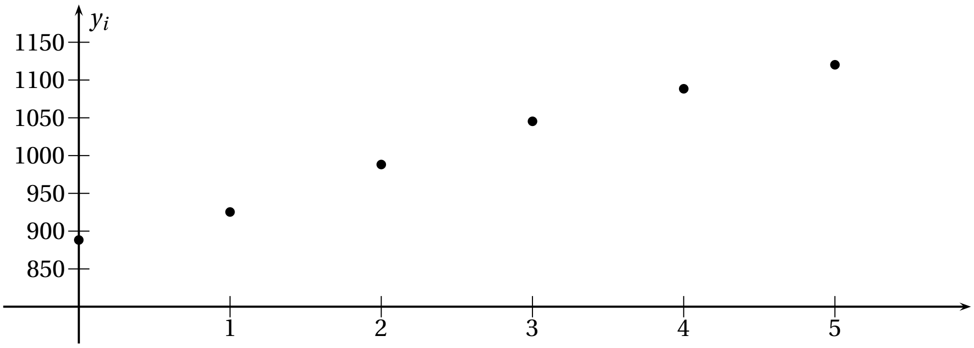
On a représenté, en annexe, dans le repère orthogonal de la figure 3.2, page 53, la série chronologique double $(x_i ; y_i)$.

1. Déterminer les coordonnées de G, point moyen du nuage $(x_i ; y_i)$, et le placer sur la figure 3.2.
2. On propose de chercher un ajustement affine.
   (a) Déterminer l'équation réduite de la droite Δ de régression de y en x par la méthode des moindres carrés. *On arrondira les coefficients à* 10^{-2}.
   (b) À l'aide de cet ajustement, estimer la superficie des terrains urbanisés en 2010. *On arrondira à l'unité.*
   (c) Représenter, sur la figure 3.2, la droite Δ.
   Que peut-on penser alors de l'estimation précédente ?
3. À cause de la légère courbure de la forme du nuage de points, un statisticien pense que les superficies sont plutôt liées au carré du rang de l'année.
   Pour vérifier son hypothèse, nous allons étudier la série $(x_i ; z_i)$ où $z_i = \sqrt{y_i}$.
   (a) Compléter sur l'annexe, page 53, le tableau 3.2 (*on arrondira les résultats à* 10^{-2}).
   (b) Représenter dans le repère orthogonal de la figure 3.3, donnée en annexe page 53, la série chronologique double $(x_i ; z_i)$.
   (c) On admettra pour la suite que l'équation réduite de $\mathscr{D}$, la droite de régression de z en x obtenue par la méthode des moindres carrés, est $z = 0,23x + 8,75$.
      i. Représenter $\mathscr{D}$ dans le repère de la figure 3.3.
      ii. Cela donne-t-il raison au statisticien ? *Justifier brièvement.*
      iii. À l'aide de ce nouvel ajustement, estimer la superficie des terrains urbanisés en 2010. *On arrondira à l'unité.*

EXERCICE 3.3 (5,5 points)**.**
Pour les élèves ***ne suivant pas*** *l'enseignement de spécialité.*
Toutes les données de cet exercice sont issues du site Observatoire des inégalités.
Le tableau suivant donne le nombre de titulaires du RMI, en milliers, des années 1994 à 1999.

Année : a_i	1994	1995	1996	1997	1998	1999
Rang de l'année : x_i	0	1	2	3	4	5
Nombre d'allocataires du RMI, en milliers : y_i	888	925	988	1045	1088	1120

La figure ci-dessous représente le nuage de points $(x_i\,;\,y_i)$.



Le tableau ci-dessous présente quelques calculs faits au tableur à partir des données précédentes :

	x_i	y_i	$X_i = x_i - \overline{x}$	$Y_i = y_i - \overline{y}$	X_i^2	$Y_i X_i$
	0	888	$-2,5$	-121	6,25	302,5
	1	925	$-1,5$	-84	2,25	126
	2	988	$-0,5$	-21	0,25	10,5
	3	1045	0,5	36	0,25	18
	4	1088	1,5	79	2,25	118,5
	5	1120	2,5	111	6,25	277,5
Moyenne	2,5	1009			2,92	142,17

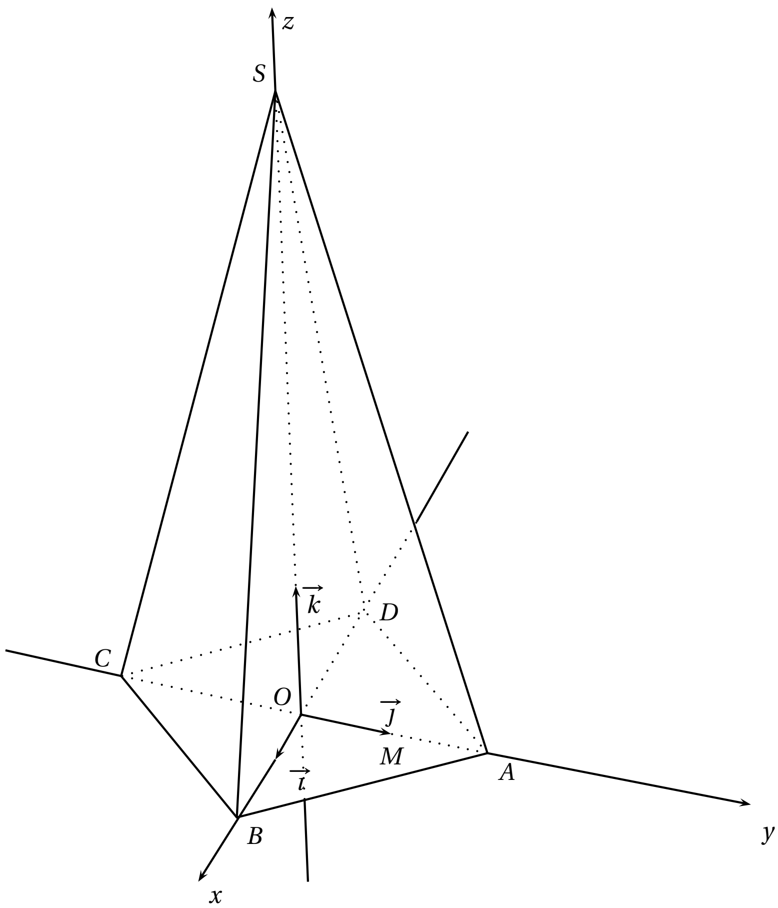
1. À l'aide du second tableau :
   (a) Déterminer les coordonnées de G, point moyen de la série double $(x_i\,;\,y_i)$.
   (b) Déterminer $V(x) = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2$ où N est le nombre de données.
   (c) Déterminer $\operatorname{cov}(x;y) = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)$ où N est le nombre de données.
   (d) En déduire le coefficient directeur de la droite de régression $\mathscr{D}$ de y en x par la méthode des moindres carrés.
   (e) En déduire l'équation réduite de $\mathscr{D}$.
   (f) À l'aide de cet ajustement, estimer le nombre d'allocataires du RMI qu'on pouvait prévoir pour 2008.
2. (a) La forme du nuage de points justifie-t-elle un ajustement affine ? *Justifier brièvement.*
   (b) D'après la même source, le nombre d'allocataires du RMI en 2008 était de 1 121 milliers. Qu'en penser ?

EXERCICE 3.3 (5,5 points)**.**
*Pour les élèves **suivant** l'enseignement de spécialité.*
La figure 3.1 de la présente page représente une pyramide $SABCD$ de sommet S.
On donne les coordonnées des points suivants dans un repère orthonormal $\left(\text{O}\,;\vec{\imath},\vec{\jmath},\vec{k}\right)$:

- $S(0\,;0\,;5)$;
- $A(0\,;2\,;0)$;
- $B(2\,;0\,;0)$;
- $C(0\,;-2\,;0)$;
- $D(-2\,;0\,;0)$;
- $M(0\,;1\,;0)$.

1. (a) Sans aucun calcul, donner une équation du plan contenant les points A, B, C et D.
   (b) Déterminer une équation du plan (ABS).
2. (a) Vérifier que le plan (BCS) admet pour équation : $5x-5y+2z=10$.
   (b) Placer le point $N(1\,;-1\,;1)$. Est-il dans le plan (BCS) ?
3. (a) Déterminer une équation du plan $\mathscr{R}$ parallèle au plan (BCS) passant par le point M.
   (b) Dessiner les traces du plan $\mathscr{R}$ sur les plans (xOy), (yOz) et (xOz).

FIGURE 3.1 – Figure de l'exercice 3.3 pour les spécialistes



Annexes

TABLE 3.1 – Tableau de la question 3a de l'exercice 3.1

x_i	0	1	2	3	4
$[(y_i - g(x_i))]^2$	0,49				

FIGURE 3.2 – Repère de l'exercice 3.2

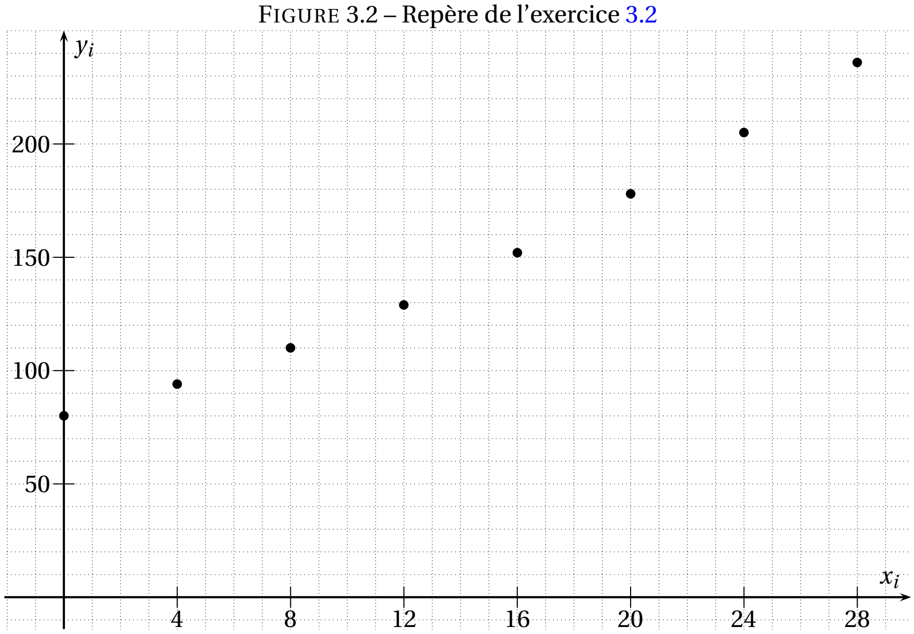


FIGURE 3.3 – Repère de la question 3b de l'exercice 3.2

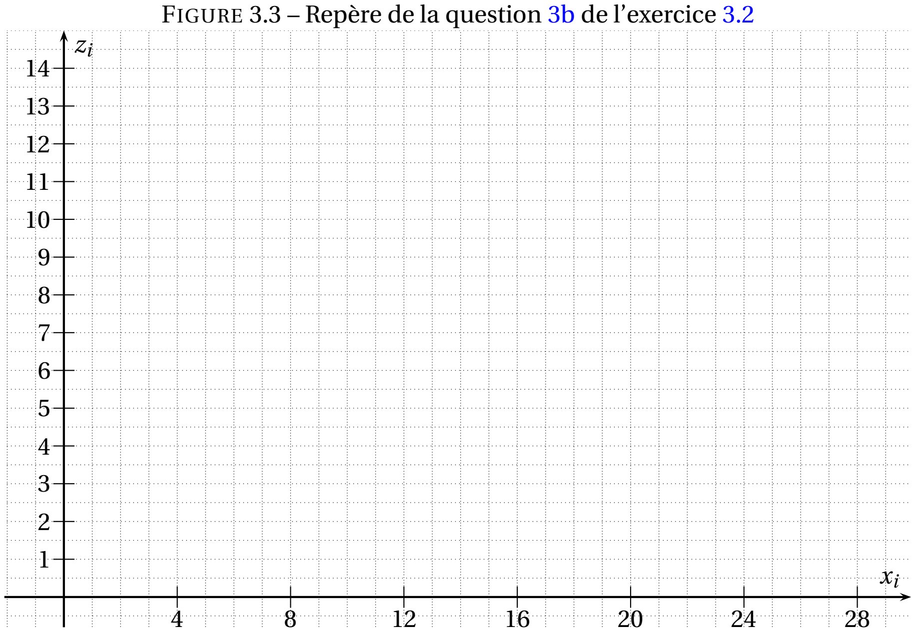


TABLE 3.2 – Tableau de la question 3a de l'exercice 3.2

rang de l'année x_i	0	4	8	12	16	20	24	28
z_i	8,94							