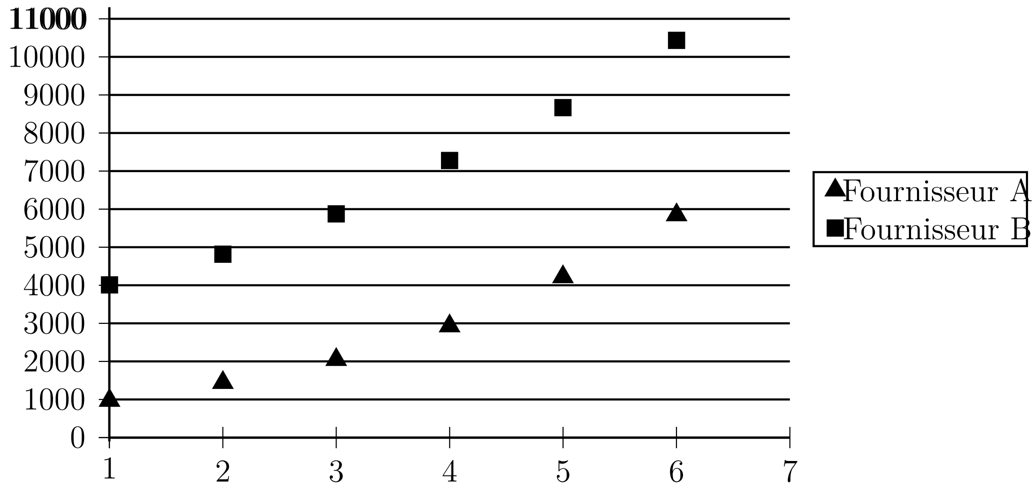
Statistique à deux variables : suppléments

Exercice I (Bac ES 2005)

Dans une ville, deux fournisseurs d'accès au réseau internet sont en concurrence.
Pour étudier l'évolution du nombre d'abonnés à ces deux fournisseurs A et B, on a reporté dans le tableau suivant, à la fin de chaque année, le nombre total d'abonnés déclaré par chacun des deux fournisseurs.

Année	1999	2000	2001	2002	2003	2004
Rang x_i de l'année	1	2	3	4	5	6
Nombre total y_i d'abonnés par le fournisseur A	975	1443	2049	2930	4220	5850
Nombre total t_i d'abonnés par le fournisseur B	4012	4813	5872	7281	8664	10432



1°) L'allure du nuage de points associé à la série statistique $(x_i\ ;\ y_i)$ permet d'envisager un ajustement exponentiel. On pose $Y_i = \ln(y_i)$.
Écrire une équation de la droite (d) d'ajustement de Y en x par la méthode des moindres carrés (arrondir les résultats au millième).

2°) En utilisant cet ajustement, donner une estimation du nombre d'abonnés au fournisseur A en 2006.

3°) L'allure du nuage de points associé à la série statistique $(x_i\ ;\ t_i)$ permet d'envisager un ajustement exponentiel.
En posant $T_i = \ln(t_i)$, on obtient, par la méthode des moindre carrés, une équation de la droite (Δ) d'ajustement de T en x sous la forme : $T = 0,193x + 8,102$ (ce résultat est admis).
En utilisant cet ajustement, donner une estimation du nombre d'abonnés au fournisseur B en 2006.
En supposant que les ajustements précédents restent pertinents, préciser l'année à partir de laquelle le nombre d'abonnés au fournisseur A dépassera le nombre d'abonnés au fournisseur B.
Justifier.

Exercice II

Le tableau suivant donne l'évolution de 1985 à 1994 du prix moyen en francs d'un kilogramme d'un produit fabriqué dans une entreprise.

Année	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
x_i : rang de l'année	1	2	3	4	5	6	7	8	9	10
y_i : prix en francs	12	17	18	22,5	26	27,5	32,5	33	36,5	37

1°) a) Représenter graphiquement par un nuage de points M_i $(x_i\,;\,y_i)$ cette série statistique dans un repère orthogonal $(O\,;\,\vec{\imath}\,,\,\vec{\jmath})$:

unités graphiques : 1 cm pour 1 an en abscisses
1 cm pour 5 francs en ordonnées.

b) Déterminer les coordonnées du point moyen G. Placer G.

2°) On considère les droites (D) et (D') d'équations respectives :

$$\text{(D)} \quad : \quad y = 2,8x + 10,8$$
$$\text{(D')} \quad : \quad y = 6x - 6,8$$

a) Démontrer que (D) et (D') passent par G.

b) Tracer (D) et (D') dans le même repère que le nuage.

c) Parmi ces deux droites, indiquer celle qui ajuste le mieux le nuage de points.

3°) Utilisation de l'ajustement précédent :

a) Les experts prévoient un prix supérieur à 47 F en 1996. Cette prévision vous semble-t-elle justifiée ?

b) Estimer l'année à partir de laquelle la valeur de y sera supérieure ou égale à 50 F.

4°) a) Calculer le coefficient de corrélation linéaire de la série à deux variables $(x; y)$.

b) Déterminer l'équation de la droite de régression de y en x. Utiliser cette équation pour donner le prix d'un kilogramme de produit en 1988.

c) Donner l'équation de la droite de régression de x en y.

Exercice III (d'après BTS)

Un fabricant de produits manufacturés utilise un jeu de moules qui permet de fabriquer 10 000 pièces par jour. Avec un jeu de moules neuf on obtient des pièces dont la masse est de 18 kilogrammes. Cette masse augmente avec l'usure des moules. On considère toutefois que la production reste acceptable tant que la masse moyenne des pièces d'une palette choisie au hasard dans la production se situe dans l'intervalle $[16,5\,;\,20]$.

Pour évaluer l'usure d'un jeu de moules au cours de la production, on prélève chaque quinzaine, au hasard, une palette de pièces que l'on pèse (la mesure de la masse est le seul contrôle de qualité que l'on fait subir à la production de ces pièces). En commençant avec un moule neuf à la date 0, on obtient ainsi la série de mesures suivante :

Numéro de la quinzaine : x_i	0	1	2	3	4	5	6
Masse moyenne : y_i	18	18,1	18,2	18,4	18,5	18,7	18,9

1°) Représenter le nuage de points (N) associé à cette série statistique double, en adoptant pour unités : 2 cm pour une quinzaine sur l'axe des abscisses et 4 cm pour 1 kg sur l'axe des ordonnées.
On pourra prendre l'origine du repère au point de coordonnées $(0\,;\,18)$.

2°) Déterminer la valeur approchée à 10^{-3} près du coefficient de corrélation linéaire de la série statistique.

3°) Déterminer par la méthode des moindres carrés une équation de la droite de régression $\mathscr{D}$ de y en x de la forme $y = mx + p$. On donnera les valeurs approchées à 10^{-2} près des coefficients m et p.

4°) Utiliser l'ajustement précédent pour estimer au bout de combien de temps le remplacement des moules est nécessaire.

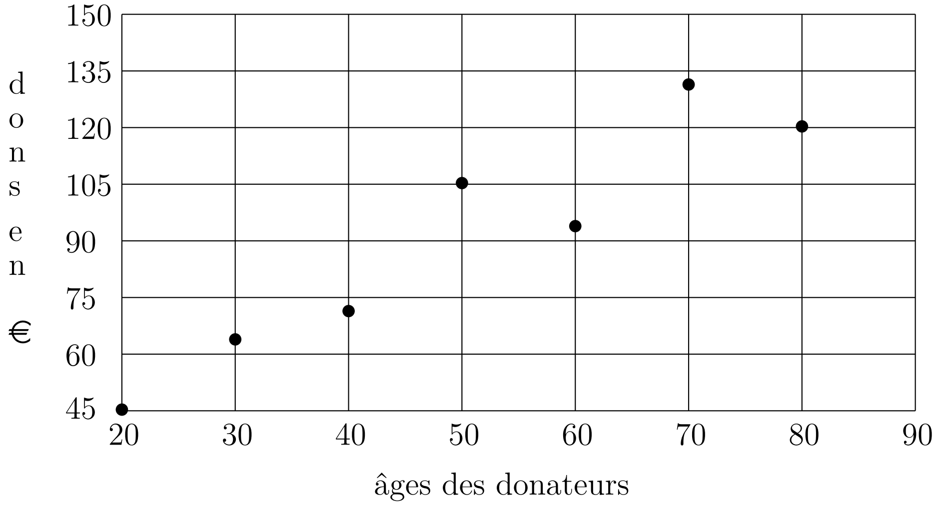
5°) Déterminer une équation de la droite de régression de x en y.

Exercice IV

Un sondage réalisé par la Fondation de France pendant une certaine année a permis de dresser le tableau suivant :

Âge moyen du donateur, x_i	20	30	40	50	60	70	80
Don, en euros, y_i	45	63,75	71,25	105	93,75	131,25	120

On a représenté ci-dessous le nuage de points M_i $(x_i\,;\,y_i)$ associé à la série statistique double $(X,\,Y)$.



1°) a) Tracez une droite ajustant approximativement ce nuage.

b) Trouvez l'équation de cette droite.

2°) a) On partage le nuage de points en deux sous-nuages : le premier, de point moyen G_1, est composé des trois points d'abscisses 20 ; 30 et 40, le second, de point moyen G_2, est composé des quatre autres points. Calculer les coordonnées des points G_1 et G_2.

b) Déterminer une équation de la droite d'ajustement affine (G_1G_2). La tracer sur le graphique.

3°) En utilisant l'équation de la droite de Mayer, estimer, en euros, le don d'une personne de 25 ans.