TP

Régression linéaire
Méthode des moindres carrés

L'objectif de ce TP est de calculer, tracer, et exploiter la méthode des moindres carrés avec un tableur (LibreOffice, Excel, ...).

1) Ajustement affine

Pour un ensemble de points de coordonnées $(x_i; y_i)$ on rappelle que la droite des moindres carrés a pour équation $y = ax + b$ avec

$$a = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} \qquad \text{et} \qquad b = \overline{y} - a\overline{x}$$

en notant $\overline{x}$ la moyenne des valeurs x_i, $\overline{y}$ la moyenne des valeurs y_i, $\overline{xy}$ la moyenne des valeurs x_iy_i, et $\overline{x^2}$ la moyenne des valeurs x_i^2.

1) À partir d'une feuille de calcul créer le tableau des coordonnées de 7 points :

$$(1, 3),\ (2, 4),\ (3, 9),\ (4, 8),\ (5, 12),\ (6, 11),\ (7, 18).$$

Calculer les coefficients a et b de la droite des moindres carrés (on utilisera la fonction *SOMME*).

2) Tracer sur votre feuille la droite des moindres carrés et l'ensemble des 7 points $(x_i; y_i)$.

3) Vérifier votre tracé et votre calcul précédent en traçant les points sur le tableur et en rajoutant une courbe de tendance ad'hoc.

2) Ajustement exponentiel : durée de vie et maintenance d'équipements.

On s'interesse à la durée de vie d'appareils mécaniques, entre autre en vu de la planification de la maitenance / remplacement des appareils.

Les pourcentages $R(t_i)$ des appareils mécaniques encore en service après un nombre t_i d'heures de fonctionnement ont été relevés et notés dans le tableau suivant :

t_i	100	200	300	400	500	600	750	1000	1500
$R(t_i)$	0,80	0,64	0,52	0,40	0,32	0,28	0,20	0,12	0,04

1. Saisir les coordonnées des points correspondants, les tracer dans une feuille de calcul.
   Un ajustement affine semble-t'il pertinent ? Justifier précisément.
2. On pose $y_i = \ln R(t_i)$. Représenter graphiquement le nuage de points M_i de coordonnées $(t_i; y_i)$.
3. Peut-on envisager un ajustement affine de ce nuage de points ?
   Donner l'équation de la droite de régression de y en t.
   En déduire une expression de la forme $R(t) = ke^{-\lambda t}$, avec k et λ des constantes.
4. Déterminer à l'aide du modèle précédent, le nombre d'équipements encore en service au bout de 900 heures de fonctionnement.
5. À l'aide de ce modèle, déterminer à partir de quand le pourcentage d'appareils encore en fonctionnement sera inférieur à 1%.