

Statistique en grande dimension et Apprentissage Série 2

Pr. Z. EN-NAIMANI



Exercice 1:

(Droite de régression linéaire)

L'analyse de la température de fonctionnement d'un procédé chimique sur le rendement du produit a donné les valeurs observées dressées dans cette table

Température en ${}^{\circ}C$ (x_i)	Rendement en $\%$ (y_i)
100	45
110	51
120	54
130	61
140	66
150	70
160	74
170	78
180	85
190	89

- 1. Déterminer la droite de régression du rendement en fonction de la température.
- 2. Calculer et interpréter le coefficient de détermination.
- 3. Deux points (105, 70) et (180, 70) ont été écartés de notre base de données. Déterminer la droite de régression pour les 12 points.
- 4. Comparer les deux situations via le coefficient de détermination.

Exercice 2: (Droite de régression linéaire)

1. Sachant que
$$SSReg = \sum_{i=1}^{n} (\hat{y}_i - \bar{y}_i)^2$$
, $SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ et $SST = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2$
Démontrer que

$$SST = SSReq + SSE$$

2. Démontrer que le coefficient de détermination
$$R^2 = \frac{SSReg}{SST}$$
 est le carré du coefficient de corrélation $R = \frac{Cov(x,y)}{2\pi r^2}$.