

TD 3. Couples de variables quantitatives

Exercice 1.

Table 1 donne la température en degrés Celsius (C) et l'intensité lumineuse en kilolux (klx) relevées simultanément dans cinq villes françaises. \mathcal{D} désigne la droite de régression de l'intensité lumineuse en la température.

<i>ville</i>	<i>température (C)</i>	<i>intensité lumineuse (klx)</i>
1	3	10
2	0	9
3	2	8
4	3	11
5	-3	7

Table 1: *Température en degrés Celsius (C) et intensité lumineuse en kilolux (klx) relevées simultanément dans cinq villes*

1. Estimez la moyenne et la variance de chacune des variables
2. Estimez la covariance du couple : température \times intensité.
3. Que vaut l'ordonnée à l'origine de \mathcal{D} ?
4. Déterminez la pente de \mathcal{D} .
5. Que vaut la somme des carrés résiduels ?
6. Quelle part de la variabilité de l'intensité lumineuse peut-on attribuer à la température ?
7. A quelle valeur estime-t-on l'intensité lumineuse lorsque la température vaut un degré Celsius ?

Exercice 2.

Table 2 indique le nombre de salariés (x) et le chiffre d'affaire en millions d'euros (y) de dix entreprises sardes. \mathcal{D} désigne la droite de régression de y en x .

<i>entreprise (i)</i>	1	2	3	4	5	6	7	8	9	10
<i>nombre de salariés (x_i)</i>	5	5	1	5	4	8	3	9	4	9
<i>chiffre d'affaire (y_i)</i>	3,3	1,7	0,6	2,2	2,0	4,0	2,8	5,0	1,6	4,8

Table 2: *Nombre de salariés (x) et chiffre d'affaire en millions d'euros (y) de dix entreprises sardes*

1. Déterminez l'équation de \mathcal{D} ?
2. Que vaut la somme des carrés résiduels ?
3. Quelle part de la variabilité du chiffre d'affaire l'effectif de l'entreprise explique-t-il ?
4. A quelle valeur estime-t-on le chiffre d'affaire d'une entreprise de dix salariés ?

Exercice 3.

On considère toujours le modèle de régression linéaire expliquant le chiffre d'affaire d'une entreprise en millions d'euros (y) par le nombre de salariés (x) mais en Corse. Sur dix entreprises corses, on observe : $\sum_{i=1}^{10} x_i = 60$; $\sum_{i=1}^{10} y_i = 30,6$; $\sum_{i=1}^{10} x_i^2 = 430$; $\sum_{i=1}^{10} y_i^2 = 119,94$; $\sum_{i=1}^{10} x_i y_i = 213,4$.

1. Estimez la covariance du couple (x, y).

2. Déterminez l'équation de la droite de régression de y en x .
3. Que vaut la somme des carrés résiduels ?
4. Quelle part du chiffre d'affaire l'effectif de l'entreprise explique-t-il ?
5. A quelle valeur estime-t-on le chiffre d'affaire d'une entreprise de dix salariés ?

Exercice 4.

On considère le jeu de données `trees` de R. Il s'agit de trente et un arbres décrits par : la circonférence (`Girth`) de la taille (`Height`) et le volume (`Volume`). La commande `attach(trees)` rend autonomes les variables ; `n=nrow(trees)` définit la taille d'échantillon.

On s'intéresse au modèle de régression expliquant le volume par la circonférence : `myreg <- lm(Volume~Girth)` dont les statistiques sont dans : `mystat <- summary(myreg)`.

1. Que représente le graphique obtenu par : `plot(Girth,Volume)` ?
2. Comment interprétez-vous les résultats de : `myreg$coefficients` ?
3. Interprétez le résultat de : `abline(myreg$coefficients,col='blue')`.
4. Que représente le résultat de `myreg$residuals` ?
5. Déterminez et interprétez le coefficient R^2 .
6. Interprétez le résultat de : `predict(myreg,new=data.frame(Girth=20))`.

Exercice 5.

On considère les données `wine` de l'UCI. Ces données peuvent être téléchargées sous R grâce à :
`wine <- read.table('https://archive.ics.uci.edu/ml/machine-learning-databases/wine/wine.data',header=FALSE,sep=',',dec='.')`

Les variables qui décrivent ces données sont expliquées sous :

`https://archive.ics.uci.edu/ml/machine-learning-databases/wine/wine.names`

On cherche à expliquer la quantité d'alcool (y) par la quantité d'acide malique (x) dans le vin.

1. Représentez les vins dans le plan : Malic acid \times Alcohol.
 2. Estimez les paramètres du modèle de régression linéaire de y en x
 3. Tracez la droite de régression de y en x .
 4. Déterminez la part de la variabilité de la quantité d'alcool imputable à l'acide malique.
 5. A quelle quantité d'alcool doit-on s'attendre dans un vin dont l'acide malique vaut 1,8 ?
-