Université de Sidi Bel Abbès *2022/2023*
Faculté de Médecine *Module :Biostatistique.*
1[ère] année Médecine

TD-N° 08 : Régression linéaire

Exercice 1 *Pour mesurer la dépendance entre l'âge et le risque cardio-vasculaire, on a observé 12 patients, pour lesquels on dispose de l'âge en années (variable X), et du logarithme du dosage en d-dimères (variable Y). On donne les quantités suivantes :* $\sum x_i = 596$; $\sum x_i^2 = 32435$; $\sum y_i = 5.2$; $\sum y_i^2 = 4.3$; $\sum x_i y_i = 188.58$.

- *Estimer l'équation de la droite de régression linéaire de Y sur X.*
- *Calculer le coefficient de détermination entre X et Y. Que peut-on conclure.*
- *Tester la significativite globale du modèle.*

Exercice 2 *On veut étudier la relation entre le QI du père (X) et le QI du fils (Y). Sur un échantillon aléatoire de 12 couples (père, fils), on a relevé le QI du père et le QI du fils. On dispose de 2 échantillons appariés de mesures $(x_i; y_i)$:*

couple (père,fils) Num i	1	2	3	4	5	6	7	8	9	10	11	12
QI du père x_i	123	144	105	110	98	138	131	90	119	109	125	100
QI du fils y_i	102	138	126	133	95	146	115	100	142	105	130	120

- *Représenter graphiquement le nuage des points.*
- *Un ajustement linéaire semble-t-il justifié ? Si oui, calculer les paramètres de la droite de régression de Y en X.*
- *Calculer le coefficient de détermination et le coefficeient de corrélation linéaire de X et Y.*
- *Si $\sum \varepsilon_i^2 = 2131.836$, représenter l'analyse de variance et le test de Fisher de corrélation. Que peut-on conclure*
- *Assurer à l'aide d'un test de Student que β est significativement différent de 0.*
- *Déterminer l'intervalle de confiance du paramètre β.*
- *Exploiter la droite d'ajustement pour déterminer le QI d'un fils dont le QI de père est 120.*

On donne les quantités suivantes : $\sum x_i = 1392$; $\sum x_i^2 = 164566$; $\sum y_i = 1452$; $\sum y_i^2 = 179068$; $\sum x_i y_i = 170394$.

$\mathscr{B}.\ \mathscr{T}.$

Rappels

- *La variance de X est* **estimé** *par (dans le cas d'un échantillon) :* $S_X'^2 = \dfrac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n-1} = \dfrac{\sum_{i=1}^{n}(x_i)^2 - n(\overline{X})^2}{n-1}$
- *La covaiance de X et Y est* **estimé** *par :* $\widehat{Cov}(X,Y) = S'_{X,Y} = \dfrac{\sum_{i=1}^{n}(x_i y_i) - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n-1}$
- *La pente β est donnée par la formule suivante :*

$$b = \frac{S'_{X,Y}}{S_X'^2}$$

- $a = m_Y - b m_X = \overline{Y} - b\,\overline{X}$ *où* $m_Y = \overline{Y} = \frac{\sum y_i}{n}$ *et* $m_X = \overline{X} = \frac{\sum x_i}{n}$
- *Estimation du cœfficient de corrélation entre X et Y :* $\hat{r} = c\hat{o}r(X,Y) = \frac{S'_{X,Y}}{S'_X\, S'_Y}$.