Feuille de TD n°6 : correction

Exo 1.

1) Si l'on suppose être dans le cadre du modèle linéaire multiple avec constante, il est possible d'écrire une équation du type analyse de la variance :

$$\underbrace{\sum (y_i-\bar{y})^2}_{\text{variation totale}} = \underbrace{\sum (y_i-\hat{y}_i)^2}_{\text{variation résiduelle}} + \underbrace{\sum (\hat{y}_i-\bar{y})^2}_{\text{variation expliquée}}$$

Par conséquent :

variation expliquée = variation totale - variation résiduelle

$= 1743,281 - 442,292$

$= 1300,989$

2) la proportion recherchée = $\dfrac{\text{variation expliquée}}{\text{variation totale}} = \dfrac{1300,989}{1743,281}$

$= 74,6\ \%$

Rq : dans le modèle avec constante, cette quantité n'est autre que R^2.

3°) Rq : auparavant, il faudrait vérifier, pour répondre à cette question que toutes les hypothèses du modèle sont bien vérifiées, à savoir :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$
avec $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$

Comme nous ne disposons pas des données, c'est ici impossible et donc nous admettons que c'est bien le cas.

hypothèse nulle : H_0 : $\beta_1 = \beta_2 = \beta_3 = 0$

contre

hypothèse alternative : H_1 $\exists\, i \in \{1,2,3\}$ tq $\beta_i \neq 0$

La statistique de test à utiliser est :

(voir rappel) $$F = \frac{\|\bar{y}-\hat{y}\|^2/(p-p_0)}{\|y-\hat{y}\|^2/(n-p)}$$ avec ici $n = 22$, $p = 4$, $p_0 = 1$

$$\Rightarrow F = \frac{1300,989/3}{442,292/18} \approx 17,649$$