Série de TD $n1$:

Exercice 0.1. *On dispose de n points* $(x_i, y_i), i = 1, ..., n$ *et on sait qu'il existe une relation de la forme* $y_i = ax_i + b + \xi_i$*, où* ξ_i *sont centrées décorrélées et de même variance* σ^2.

1)− *On suppose connaître b mais pas a.*

(a)− *On revenant à la définition des MCO calculer l'estimateur* $\tilde{a}$ *des MC de a.*

(b)− *Calculer* $var(\tilde{a})$*, montrer qu'elle est inférieure à celle de* $\hat{a}$.

2)− *On suppose connaître a, mais pas b :*

(a)− *La même première question mais pour b.*

(b)− *La même deuxième question mais pour b.*

Exercice 0.2. *Rappeler la formule définissent qu'il est égal au carré du coefficient de détermination* R^2 *et la développer pour montrer qu'il égal au carré du coefficient de corrélation empirique entre* X *et* Y *noté* $\varrho^2_{x,y}$ *c'est-à-dire :*

$$R^2 = \varrho^2_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{(\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2)^{\frac{1}{2}}}$$

Exercice 0.3. *L'étude statistique ci-dessus porte sur les poids respectifs des pères et de leurs fils aînés :*

pères $= (65, 63, 67, 64, 68, 62, 70, 66, 68, 67, 69, 71)$.

fils $= (68, 66, 58, 65, 69, 66, 68, 65, 71, 67, 68, 70)$.

1)− *Calculer la droite des moindres carrés du poids des fils en fonction du poids des pères.*

2)− *Calculer la droite des moindres carrés du poids des pères en fonction du poids des fils.*

3)− *Montrer que le produit des pentes des deux droites est égal au carré du coefficient du corrélation empirique entre les* p_i *et les* f_i.

Exercice 0.4. *Nous souhaitons exprimer la hauteur* Y *(en pieds) d'un arbre d'une essence donnée en fonction de son diamètre* X *(en pouces) à* $1m30$ *du sol. Pour ce faire, nous avons mesuré* 20 *couples (diamètre, hauteur) et effectué les calculs suivants :*

$\overline{x} = 4.53$ *et* $\overline{y} = 8.65$.

$\frac{1}{20}\sum_{i=1}^{20}(x_i-\overline{x})^2 = 10.97$, $\frac{1}{20}\sum_{i=1}^{20}(y_i-\overline{y})^2 = 2.24$ *et* $\frac{1}{20}\sum_{i=1}^{20}(x_i-\overline{x})\frac{1}{20}\sum_{i=1}^{20}(y_i-\overline{y}) = 3.77$.

On note $y = \widehat{\beta}_0 + \widehat{\beta}_1 x$ *la droite de régression. Calculer* $\widehat{\beta}_0$ *et* $\widehat{\beta}_1$.

Exercice 0.5. *On examine l'évaluation d'une variable* y *en fonction de deux variables exogènes* x *et* z*, on dispose de* n *observations des ces variables.*

On note $X = (1, x, z)$*, où* 1*est le vecteur constant et* x *et* z *sont les vecteurs des variables explicatives.*

$I)-$ *Nous avons obtenus les résultats suivants :*

$$X'X = \begin{pmatrix} 25 & 0 & 0 \\ ? & 9.3 & 5.4 \\ ? & ? & 12.7 \end{pmatrix}$$

1)− Donner les valeurs manquantes.

2)− Que vaut n.

c)− Calculer le coefficient de corrélation linéaire empirique entre x et z.

II)− La régression linéaire de Y sur $(1, x, z)$ donne :

$$y = -1.6 + 0.61x + 0.46z + \widehat{\xi}, \quad SCR = 0.3$$

1)− Déterminer la moyenne empirique $\overline{y}$.

2)− Calculer SCE, SCT et R^2. Expliquer.

3)− Calculer R_a^2, $\widehat{\sigma}^2$ et un estimateur de $var(\widehat{\beta})$.