

1) a) $X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ $X^T Y = \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = n \begin{pmatrix} \bar{y} \\ \bar{xy} \end{pmatrix} = 90 \begin{pmatrix} -19,607 \\ -709,9204 \end{pmatrix} = \dots$
 (On pose $n=90$) $X^T \hat{Y} = \dots = n \begin{pmatrix} \bar{\hat{y}} \\ \bar{\hat{xy}} \end{pmatrix}$

b) $\bar{Y} = \frac{1}{n} (X^T Y)_1$ $\bar{\hat{Y}} = \frac{1}{n} (X^T \hat{Y})_1$
 Puis utiliser $\hat{Y} = X \hat{\beta}$ et $\hat{\beta} = (X^T X)^{-1} X^T Y$ comme 6 : c'est l'estimateur des moindres carrés.

Donc $X^T \hat{Y} = X^T X (X^T X)^{-1} X^T Y = X^T Y$ donc $\frac{1}{n} (X^T \hat{Y})_1 = \frac{1}{n} (X^T Y)_1$
 d'où $\bar{\hat{Y}} = \bar{Y}$. (Accessoirement, la 2^e coordonnée fournit $\bar{X\hat{Y}} = \bar{XY}$)

c) $Y^T \hat{Y} = \sum_i y_i \hat{y}_i = n \bar{Y\hat{Y}}$. Comme $\hat{Y} = X \hat{\beta} = X (X^T X)^{-1} X^T Y$, on n'a pas le droit d'écrire $(X^T X)^{-1} = X^{-1} (X^T)^{-1}$: ces matrices ne sont pas inversibles ! (indice : elles ne sont pas carrées)
 $n \bar{Y\hat{Y}} = Y^T \hat{Y} = Y^T X (X^T X)^{-1} X^T Y$ et $n \bar{\hat{Y}^2} = \hat{Y}^T \hat{Y} = (X (X^T X)^{-1} X^T Y)^T X (X^T X)^{-1} X^T Y$
 $= Y^T X (X^T X)^{-1} X^T X (X^T X)^{-1} X^T Y = Y^T X (X^T X)^{-1} X^T Y$

Donc $Y^T \hat{Y} = \hat{Y}^T \hat{Y}$, c'est-à-dire $\bar{Y\hat{Y}} = \bar{\hat{Y}^2}$.
 $(A^T)^{-1} = (A^{-1})^T$
 $(AB)^T = B^T A^T$
 $(A^T A)^T = A^T (A^T)^T = A^T A$

2) $e = \underset{\text{observés}}{Y} - \underset{\text{prédits}}{\hat{Y}} = Y - X \hat{\beta}$ (dans le meilleur des mondes $\hat{\beta} = \beta$ et donc $e = \varepsilon$)

indice	1	...	23	24	...
valeur	-97,434	...	-35,701	-31,741	...

Residuels :
 Min -97,434 1Q -34,711 (= 1^{re} quantile)

$23,25 = \frac{3}{4} 23 + \frac{1}{4} 24$
 $-34,711 = \frac{3}{4} (-35,701) + \frac{1}{4} (-31,741)$

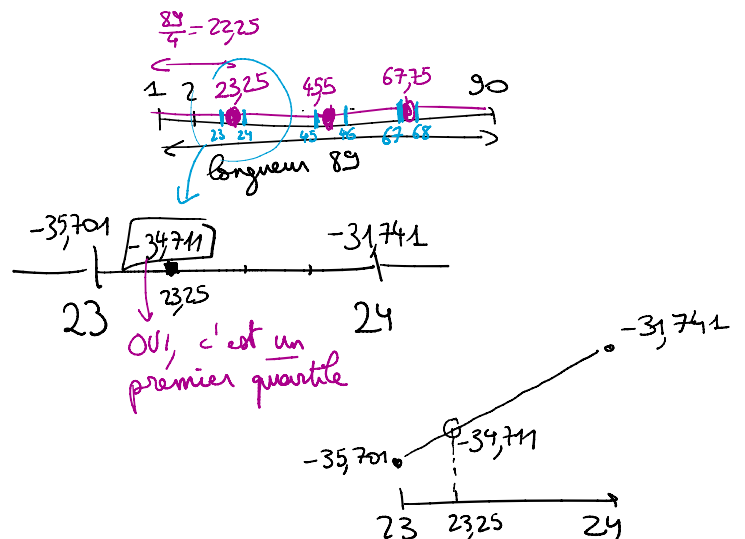
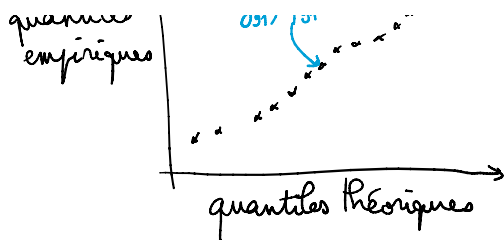


Figure 2, a-a-plot



quantile empirique d'ordre $\frac{23}{91}$ 91 intervalles
 i 1 2 3 ... 90
 $e(i)$ -97,434 ...
 quantile empirique d'ordre $\frac{2}{91} = \frac{2}{91}$ etc.
 $e(i)$ est la i -ème résidu quand on les a triés par ordre croissant : $e_{(1)} \leq e_{(2)} \leq \dots \leq e_{(n)}$



$e(i)$ est la i -ème résidu quand on les a triés par ordre croissant: $e_{(1)} \leq e_{(2)} \leq \dots \leq e_{(n)}$

Notons z_α le quantile d'ordre α de la loi normale $\mathcal{N}(0,1)$

Si les e_i étaient iid $\sim \mathcal{N}(\mu, \sigma^2)$, alors pour n grand, $\hat{q}_{\frac{i}{n}} \approx \mu + \sigma z_{\frac{i}{n}}$
autrement dit ces points s'alignent suivant une droite (celle tracée en bleu)

3) a) (Cours 6) $\hat{S}^2 = \frac{1}{n} \|Y - X\hat{\beta}\|^2$, alors $\frac{n\hat{S}^2}{\sigma^2} \sim \chi^2(n-2)$
(donc $\mathbb{E}[\hat{S}^2] = \frac{n-2}{n} \sigma^2$) $\xrightarrow{1 + \text{nombre de } X = 2}$ $\xrightarrow{\frac{n-2}{n} \approx 1}$ $\xrightarrow{\text{biaisé!}}$

$$\hat{\sigma}^2 = \frac{1}{n-2} \|Y - \hat{Y}\|^2$$

b) 88 degree of freedom: $\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$
 $\xrightarrow{\text{nombre de degrés de liberté de la loi du chi-deux}} \xrightarrow{= 90 - 2 = 88}$

c) Residual standard error
écart-type des résidus

$$(n-2)\hat{\sigma}^2 = \|Y - \hat{Y}\|^2 = \sum_i (Y_i - \hat{Y}_i)^2 = \sum_i Y_i^2 - 2 \sum_i Y_i \hat{Y}_i + \sum_i \hat{Y}_i^2 = n(\bar{Y}^2 - 2\bar{Y}\bar{\hat{Y}} + \bar{\hat{Y}}^2)$$

$$= n(\bar{Y}^2 - \bar{\hat{Y}}^2) \quad \text{16) : } \bar{\hat{Y}} = \bar{Y}$$

Donc $\hat{\sigma}^2 = \frac{n}{n-2} (\bar{Y}^2 - \bar{\hat{Y}}^2) = \frac{90}{88} (2852 - 532) = 2373$

D'où $\hat{\sigma} = 48,71$

4) a) On peut utiliser $\hat{\beta}_1 = \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{\overline{X^2} - \bar{X}^2}$
 $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$

b) $\hat{\beta} \sim \mathcal{N}(\beta, \hat{\sigma}^2 (X^T X)^{-1})$

(par le modèle gaussien homosédastique $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$)

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \begin{pmatrix} \text{Var}(\hat{\beta}_0) & \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) \\ \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \text{Var}(\hat{\beta}_1) \end{pmatrix} \right)$$

Donc $\text{Var}(\hat{\beta}_0) = \hat{\sigma}^2 \frac{\bar{X}^2}{n(\bar{X}^2 - \bar{X}^2)}$

$$\approx \hat{\sigma}^2 \frac{\bar{X}^2}{n(\bar{X}^2 - \bar{X}^2)} = 48,71^2 \cdot \frac{1118}{90(118 - 3318^2)} = 1725$$

$$X^T X = n \begin{pmatrix} 1 & \bar{X} \\ \bar{X} & \bar{X}^2 \end{pmatrix}$$

donc $(X^T X)^{-1} = \frac{1}{\det(X^T X)} \times n \begin{pmatrix} \bar{X}^2 & -\bar{X} \\ -\bar{X} & 1 \end{pmatrix}$
 $= \frac{1}{n(\bar{X}^2 - \bar{X}^2)} \begin{pmatrix} \bar{X}^2 & -\bar{X} \\ -\bar{X} & 1 \end{pmatrix}$

Donc $\sqrt{\text{Var}(\hat{\beta}_0)} \approx 41,5$

De même $\text{Var}(\hat{\beta}_1) \approx \hat{\sigma}^2 \frac{1}{n(\bar{X}^2 - \bar{X}^2)} \approx \frac{48,71^2}{90(118 - 3318^2)} \approx 1,543$

De même $\overline{\text{Var}(\hat{\beta}_1)} \approx \hat{\sigma}^2 \frac{1}{n(\bar{X}^2 - \bar{X}^2)} \approx \frac{48,71^4}{90(1118 - 3318^2)} \approx 1,543$
 donc $\boxed{\sqrt{\text{Var}(\hat{\beta}_1)} \approx 1,24}$

Ces valeurs correspondent à celles de la colonne « Std. Error ».

(d) [Speller: cette notation signifie $P(|T| \geq |t^{obs}|)$ pour $T \sim ?$
 et $t^{obs} = ?$]

(e) $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2(X^T X)^{-1})$ indépendant de $\hat{\sigma}^2$!

$$\frac{\hat{\beta} - \beta}{\sigma} \sim \mathcal{N}(0, (X^T X)^{-1}) = \mathcal{N}\left(0, \frac{1}{n(\bar{X}^2 - \bar{X}^2)} \begin{pmatrix} \bar{X}^2 & -\bar{X} \\ -\bar{X} & 1 \end{pmatrix}\right)$$

Donc $\frac{\hat{\beta}_1 - \beta_1}{\sigma} \sim \mathcal{N}\left(0, \frac{1}{n(\bar{X}^2 - \bar{X}^2)}\right)$ et $\frac{\hat{\beta}_0 - \beta_0}{\sigma} \sim \mathcal{N}\left(0, \frac{\bar{X}^2}{n(\bar{X}^2 - \bar{X}^2)}\right)$

⚠ La formule de test de Student vue en cours ne marche que pour tester $\beta_1 = 0$ vs $\beta_1 \neq 0$!

Test $H_0: \beta_0 = 0$ vs. $H_1: \beta_0 \neq 0$.

$$\frac{\hat{\beta}_0 - \beta_0}{\sigma} \sim \mathcal{N}\left(0, \frac{\bar{X}^2}{n(\bar{X}^2 - \bar{X}^2)}\right) \text{ donc } \frac{\hat{\beta}_0 - \beta_0}{\sigma} \sqrt{\frac{n(\bar{X}^2 - \bar{X}^2)}{\bar{X}^2}} \sim \mathcal{N}(0, 1).$$

Par ailleurs, $\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$ est indépendant de $\hat{\beta}_0$, donc de

D'où
$$\frac{[(\hat{\beta}_0 - \beta_0) / \sigma] \sqrt{\frac{n(\bar{X}^2 - \bar{X}^2)}{\bar{X}^2}}}{\sqrt{\frac{(n-2)\hat{\sigma}^2}{\sigma^2} / (n-2)}} \sim \mathcal{C}(n-2)$$

Sous H_0 , $T = \frac{\hat{\beta}_0}{\hat{\sigma}} \sqrt{\frac{n(\bar{X}^2 - \bar{X}^2)}{\bar{X}^2}} \sim \mathcal{C}(n-2) = \mathcal{C}(88)$

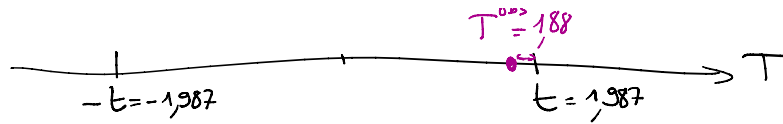
A.N. $T = \frac{77,563}{48,71} \sqrt{90 \left(1 - \frac{3318^2}{1118}\right)} \approx 1,88$

Région de rejet (test bilatéral, la loi de T sous H_0 est symétrique): $]-\infty; -t] \cup [t; +\infty[$.

Niveau 5% : $P(|T| \geq t) \leq 5\%$, donc $t =$ quantile d'ordre 97,5% de $\mathcal{C}(88)$.
 $t = 1,987$

Alors $T \notin R$: on conserve H_0 .

$T^{obs} = 1,88$



$$P_{T \sim \mathcal{L}(88)}(|T| \geq |T^{obs}|) = \alpha = p\text{-value.}$$

$\Pr(>|t|) \rightarrow p\text{-value} = 0.0641. > 0.05 : \text{on conserve } H_0 \text{ (OK)}$
 de pas beaucoup (T obs proche de t)

f) $\hat{\beta}_1 \sim \mathcal{N}(\beta_1, \sigma^2 \frac{1}{n(\bar{X}^2 - \bar{X}^2)})$

Donc $T = \frac{\hat{\beta}_1}{\hat{\sigma}} \sqrt{n(\bar{X}^2 - \bar{X}^2)} = \frac{-293}{48.7} \sqrt{30(1118 - 33^2)} = -2.37$

$$R =]-\infty; -1,987] \cup [1,987; +\infty[$$

$T^{obs} \in R \rightarrow \text{donc on rejette } H_0.$

5) a) $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{(\bar{\hat{y}}^2 - \bar{y}^2)}{(\bar{y}^2 - \bar{y}^2)} = \frac{\bar{\hat{y}}^2 - \bar{y}^2}{\bar{y}^2 - \bar{y}^2}$

$\bar{\hat{y}}^2 - \bar{y}^2 \approx \text{Var}(\hat{y})$
 $\bar{y}^2 - \bar{y}^2 \approx \text{Var}(y)$

$\frac{\bar{\hat{y}}^2 - \bar{y}^2}{\bar{y}^2 - \bar{y}^2} \approx \frac{\text{Var expliquée par best}}{\text{Var totale des } y}$

$\bar{\hat{y}}^2 - \bar{y}^2 = \frac{532 - 19.6^2}{2852 - 19.6^2} = 0.0600$
 cela confirme l'énoncé.

b) $R^2_{ajusté}$: calcul, OK.

Comme $R^2 \in [0, 1]$, $(1 - R^2) \geq 0$ donc $p \mapsto (1 - R^2) \frac{n-1}{n-p-1}$ est croissante
 donc $R^2_{ajusté}$ est décroissant en p .

c) $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^T X)^{-1})$

donc $\hat{\beta}_0 + \hat{\beta}_1 X_{g1} \sim \mathcal{N}(\beta_0 + \beta_1 X_{g1}, \frac{\sigma^2}{n(\bar{X}^2 - \bar{X}^2)} \begin{pmatrix} 1 & X_{g1} \\ \bar{X} & 1 \end{pmatrix})$

$= (\bar{X}^2 - \bar{X} X_{g1}, X_{g1} - \bar{X})$

D'où $\frac{(\hat{\beta}_0 + \hat{\beta}_1 X_{g1}) - (\beta_0 + \beta_1 X_{g1})}{\sigma} \sim \mathcal{N}(0, \frac{\bar{X}^2 - 2\bar{X} X_{g1} + X_{g1}^2}{n(\bar{X}^2 - \bar{X}^2)})$

Calculs: $\hat{\beta}_0 + \hat{\beta}_1 X_{g1} = 77,563 - 3,929 \cdot 50 = -68.9$

Calculo: $\hat{\beta}_0 + \hat{\beta}_1 x_{31} = 77,563 - 3,929 \cdot 50 = -68,9$