Le modèle de régression linéaire simple

A– 1 • Introduction :

Un modèle économique est l'expression mathématique d'une certaine théorie économique. L'exemple de la loi psychologique fondamentale de Keynes est assez pertinent à cet effet. D'après cette loi, en moyenne et la plupart du temps lorsque le revenu d'un individu augmente, il augmente aussi sa consommation, mais dans une proportion moindre à l'augmentation de son revenu. Mathématiquement, si on note la consommation par C_t *et le revenu par* Y_t *, cette loi peut être spécifiée comme suit :* $C_t = a_0 + a_1 Y_t$ *,*

(avec a_1 *: propension marginale à consommer ,* $0 < a_1 < 1$*)*

En général, le modèle spécifié par l'économiste est défini comme étant une maquette de la réalité ou d'un phénomène sous forme d'équations dont les variables sont des grandeurs économiques.

Un modèle économétrique n'est autre chose qu'un modèle économique qui contient les spécifications nécessaires pour son application empirique. C'est donc le modèle économique auquel on ajoute un terme d'erreur u_t *.*

$C_t = a_0 + a_1 Y_t + u_t$ *, (modèle spécifié par l'économètre)*

La première partie de ce modèle $(a_0 + a_1 Y_t)$ *constitue sa partie systématique et la deuxième* (u_t) *sa partie stochastique ou aléatoire.*

Il convient de noter également que le terme d'erreur u_t *(bruit, perturbation ou aléa) dénote de la différence entre l'économiste et l'économètre. Il synthétise l'influence sur* C_t *(variable expliquée) de toutes les autres variables oubliées et des erreurs éventuelles de spécification de la forme fonctionnelle dans le modèle spécifié par l'économiste.*

De plus, sa présence dans le modèle rend les paramètres a_0 et a_1 inconnus, on ne sait plus les calculer, il faut donc les estimer.

A– 2 • Les principaux modèles économétriques (spécifications courantes) :

a • La spécification Niveau-Niveau : $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$

Une augmentation d'une unité de X augmente Y de β_2 unités , pour $\beta_2 > 0$

b • La spécification Niveau-Log : $y_i = \beta_1 + \beta_2 ln(x_i) + \varepsilon_i$

Une augmentation de 1% de X augmente Y de $\dfrac{\beta_2}{100}$ unités, pour $\beta_2 > 0$

c • La spécification Log-Niveau : $ln(y_i) = \beta_1 + \beta_2 x_i + \varepsilon_i$

Une augmentation d'une unité de X augmente Y de $(\beta_2 \times 100)$ unités, pour $\beta_2 > 0$

β_2 mesure la semi-élasticité de y par rapport à x

d • La spécification Log-Log : $ln(y_i) = \beta_1 + \beta_2 ln(x_i) + \varepsilon_i$

Une augmentation de (points de pourcentage)% de x augmente y de β_2 % (points de pourcentage), pour $\beta_2 > 0$, pour $\beta_2 > 0$.

Ce qui nous permet d'interpréterβ_2 comme un paramètre d'élasticité

☑ ***Modèles logarithmiques et Élasticités*** : *Soit $\epsilon_{x_0}(y/x)$ l'élasticité de y par rapport à x*

$$\text{donc } \epsilon = \frac{dy/y}{dx/x} \text{ or } \begin{cases} \dfrac{d\ln(y)}{dy} = \dfrac{1}{y} \\ \dfrac{d\ln(x)}{dx} = \dfrac{1}{x} \end{cases} \Rightarrow \begin{cases} d\ln(y) = \dfrac{dy}{y} \\ d\ln(x) = \dfrac{dx}{x} \end{cases} \text{ et comme on a } \beta_2 = \frac{d\ln(y)}{d\ln(x)} \Rightarrow \beta_2 = \frac{dy/y}{dx/x} = \epsilon$$

e • Récapitulation :

Type de modèle	Variable dépendante	Variable explicative	Interprétation du du coefficient β_2
Niveau-Niveau	y_i	x_i	$\Delta y = \beta_2 \Delta x$
Niveau-Log	y_i	$ln(x_i)$	$\Delta y = (\beta_2/100)\%\Delta x$
Log-Niveau	$ln(y_i)$	x_i	$\Delta y = (100\beta_2)\%\Delta x$
Log-Log	$ln(y_i)$	$ln(x_i)$	$\%\Delta y = \beta_2 \%\Delta x$

A– 3 • Le modèle et ses hypothèses :

a • L'équation de régression : *Le modèle de régression linéaire simple considère que la variable à expliquer Y est une fonction affine de la variable explicative X.*

Mathématiquement, cette dépendance linéaire s'écrit de la sorte :

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_i + \hat{\varepsilon}_i \ ; i = 1, 2, \dots, n$$

☞ *L'indice i correspond à une observation particulière*

☞ *La variable y_i s'appelle indifféremment variable endogène, ou variable expliquée (la variable à expliquer), ou variable dépendante*

☞ *La variable x_i s'appelle indifféremment variable exogène, ou variable explicative, ou variable indépendante. On parle aussi de régresseur ou prédicteur*

☞ *Le terme ε_i est un terme d'erreur aléatoire inobservable*

☞ *β_1 et β_2 sont les paramètres du modèle (à estimer) qui permettent de caractériser la relationde dépendance linéaire qui existe à chaque observation i entre x_i et y_i*

b • Les hypothèses : *Les estimateurs $\widehat{\beta}_1$ et $\widehat{\beta}_2$ vont dépondre des y_i, donc des ε_i : ce seront des variables aléatoires, et nous aurons besoin des moments de leurs distribution. Il nous faut donc faire des hypothèses sur la distribution des ε_i*

$$(M): y_i = \beta_1 + \beta_2 x_i + \varepsilon_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_i + \hat{\varepsilon}_i \quad ; \ i = 1, 2, \dots, n$$

■ **$\mathcal{H}_1$(Hypothèse sur la partie systématique):** *Hypothèses sur Y et X :*

X et Y Sont deux grandeurs numériques mesurées sans erreur. X Est une donnée exogène dans le modèle. Elle est supposée non aléatoire (ou non stochastique). Y Est aléatoire par l'intermédiaire $\mathcal{E}$ c-à-d. la seule erreur qu'on a sur Y provient des insuffisances de X à expliquer ses valeurs dans le modèle.

■ **$\mathcal{H}_2$ (Hypothèse sur la partie stochastique):** *Hypothèses sur les termes aléatoires ε_i : Les ε_i sont i.i.d (indépendants et identiquement distribués).*

■ **$\mathcal{H}_3$(Hypothèse sur la partie stochastique):** $E(\varepsilon_i) = 0 \ \forall \ i = 1, 2, \dots, n$

En moyenne les erreurs s'annulent c-à-d. le modèle est bien spécifié.

■ **$\mathcal{H}_4$(Hypothèse sur la partie stochastique):** $Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_\varepsilon^2 \ \forall \ i = 1, 2, \dots, n$

La variance de l'erreur est constante et ne dépend pas de l'observation. C'est l'hypothèse d'homoscédasticité des erreurs.

Il s'agit d'assumer que les variables explicatives omises dans le modèle influent toutes pratiquement de façon constante sur la variable expliquée.

■ **$\mathcal{H}_5$(Hypothèse d'indépendance entre la partie systématique et la partie aléatoire):**

L'erreur est indépendante de la variable exogène c-à-d. $Cov(\varepsilon_i, x_i) = 0 \ \forall \ i = 1, 2, \dots, n$

Cette hypothèse signifie que l'erreur et les variables explicatives ont une influence séparée sur la variable endogène

- $\mathcal{H}_6$(*Hypothèse sur la partie stochastique*): *Indépendance des erreurs. Les erreurs relatives à 2 observations différentessont indépendantes c-à-d.*

$Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0 \;\; pour \;\; i \neq j$.

On parle de l'hypothèse de la non autocorrélation des erreurs

- $\mathcal{H}_7$(*Hypothèse sur la partie stochastique*): $\varepsilon_i \backsim \mathcal{N}(0, \sigma_\varepsilon^2)$.

L'hypothèse de normalité des erreurs est un élément clé pour l'inférence statistique. Elle est donc nécessaire pour mener les tests.

- $\mathcal{H}_8$(*Hypothèse sur la partie systématique*) : x_i *prend au moins deux valeurs différentes*

- $\mathcal{H}_9$(*Hypothèse sur la partie systématique*): $n > 2$, *c-à-d. le nombre d'observations n doit être supérieur au nombre des paramètres à estimer.*

☞ **Remarque 1** : *Lorsque les hypothèses $\mathcal{H}_9$, $\mathcal{H}_6$, $\mathcal{H}_4$ et $\mathcal{H}_3$ sont réalisées, on dit que les erreurs sont des bruits blancs. Et lorsqu'on y ajoute l'hypothèse $\mathcal{H}_7$, on parle des bruits blancs gaussiens.*

☞ **Remarque 2** : *Lorsque toutes les hypothèses sous-tendant la méthode des MCO sont remplies, le théorème de Gauss-Markov avance que ses estimateurs sont BLUE (Best Linear Unbiased Estimator), c'est-à-dire qu'ils sont les meilleurs estimateurs linéaires, non biaisés et à variance minimale.*

$\mathcal{A}$– 4 • Estimateur des moindres carrés ordinaires (MCO) "régression avec constante β_1" :

La différence : $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i$ *s'appelle résidu, et est une estimation de l'erreur* ε_i .

On peut écrire indifféremment : $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ *ou* $y_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\varepsilon}_i$

La méthode des moindres carrés ordinaire consiste à minimiser la somme des carrés des

$$\textit{résidus}:\min_{\widehat{\beta}_1,\widehat{\beta}_2}\sum_{i=1}^{n}\hat{\varepsilon}_i^2=\min_{\widehat{\beta}_1,\widehat{\beta}_2}\underbrace{\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i\right)^2}_{\varphi(\widehat{\beta}_1,\widehat{\beta}_2)}$$

• ***Condition nécessaire :***

$$\begin{cases}\dfrac{\partial\varphi(\widehat{\beta}_1,\widehat{\beta}_2)}{\partial\widehat{\beta}_1}=0\\ \dfrac{\partial\varphi(\widehat{\beta}_1,\widehat{\beta}_2)}{\partial\widehat{\beta}_2}=0\end{cases}\Leftrightarrow\begin{cases}-2\sum_{i=1}^{n}(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i)=0\\ -2\sum_{i=1}^{n}(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i)x_i=0\end{cases}\Leftrightarrow\begin{cases}\sum_{i=1}^{n}y_i-n\widehat{\beta}_1-\widehat{\beta}_2\sum_{i=1}^{n}x_i=0\\ \sum_{i=1}^{n}x_iy_i-\widehat{\beta}_1\sum_{i=1}^{n}x_i-\widehat{\beta}_2\sum_{i=1}^{n}x_i^2=0\end{cases}$$

$$\Leftrightarrow\begin{cases}n\bar{y}-n\widehat{\beta}_1-n\widehat{\beta}_2\bar{x}=0\\ n\widehat{\beta}_1\bar{x}\ +\widehat{\beta}_2\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_iy_i\end{cases}\Leftrightarrow\begin{cases}\widehat{\beta}_1=\bar{y}-\widehat{\beta}_2\bar{x}\\ n(\bar{y}-\widehat{\beta}_2\bar{x})\bar{x}\ +\widehat{\beta}_2\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_iy_i\end{cases}$$

$$\Leftrightarrow\begin{cases}\widehat{\beta}_1=\bar{y}-\widehat{\beta}_2\bar{x}\\ \widehat{\beta}_2\sum_{i=1}^{n}x_i^2-n\widehat{\beta}_2\bar{x}^2=\sum_{i=1}^{n}x_iy_i-n\overline{xy}\end{cases}\Leftrightarrow\begin{cases}\widehat{\beta}_1=\bar{y}-\widehat{\beta}_2\bar{x}\\ \widehat{\beta}_2\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)=\sum_{i=1}^{n}x_iy_i-n\overline{xy}\end{cases}$$

■ ***$\mathcal{H}_8$(Hypothèse sur la partie systématique) : x_i prend au moins deux valeurs différentes*** $\Rightarrow\sum_{i=1}^{n}(x_i-\bar{x})^2\neq 0$

$$\Leftrightarrow\begin{cases}\widehat{\beta}_1=\bar{y}-\widehat{\beta}_2\bar{x}\\ \widehat{\beta}_2=\sum_{i=1}^{n}x_iy_i-n\overline{xy}\Big/\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\Big/\sum_{i=1}^{n}(x_i-\bar{x})^2\end{cases}$$

• ***Condition suffisante :***

$$\frac{\partial^2\varphi}{\partial\widehat{\beta}_1^2}(\widehat{\beta}_1,\widehat{\beta}_2)=\frac{\partial\left(-2\sum_{i=1}^{n}(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i)\right)}{\partial\widehat{\beta}_1}=-2\sum_{i=1}^{n}-1=2n$$

$$\frac{\partial^2\varphi}{\partial\widehat{\beta}_2^2}(\widehat{\beta}_1,\widehat{\beta}_2)=\frac{\partial\left(-2\sum_{i=1}^{n}(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i)x_i\right)}{\partial\widehat{\beta}_2}=-2\sum_{i=1}^{n}-x_i^2=2\sum_{i=1}^{n}x_i^2$$

$$\frac{\partial^2\varphi}{\partial\widehat{\beta}_2\partial\widehat{\beta}_1}(\widehat{\beta}_1,\widehat{\beta}_2)=\frac{\partial^2\varphi}{\partial\widehat{\beta}_1\partial\widehat{\beta}_2}(\widehat{\beta}_1,\widehat{\beta}_2)=\frac{\partial\left(-2\sum_{i=1}^{n}(y_i-\widehat{\beta}_1-\widehat{\beta}_2x_i)\right)}{\partial\widehat{\beta}_2}=-2\sum_{i=1}^{n}-x_i=2\sum_{i=1}^{n}x_i$$

$$\left|H\left(\varphi(\widehat{\beta}_1,\widehat{\beta}_2)\right)\right|=\begin{vmatrix}\dfrac{\partial^2\varphi}{\partial\widehat{\beta}_1^2}(\widehat{\beta}_1,\widehat{\beta}_2) & \dfrac{\partial^2\varphi}{\partial\widehat{\beta}_2\partial\widehat{\beta}_1}(\widehat{\beta}_1,\widehat{\beta}_2)\\ \dfrac{\partial^2\varphi}{\partial\widehat{\beta}_1\partial\widehat{\beta}_2}(\widehat{\beta}_1,\widehat{\beta}_2) & \dfrac{\partial^2\varphi}{\partial\widehat{\beta}_2^2}(\widehat{\beta}_1,\widehat{\beta}_2)\end{vmatrix}=\begin{vmatrix}2n & 2\sum_{i=1}^{n}x_i\\ 2\sum_{i=1}^{n}x_i & 2\sum_{i=1}^{n}x_i^2\end{vmatrix}$$

$$\left|H\left(\varphi(\widehat{\beta}_1,\widehat{\beta}_2)\right)\right| = 4n\sum_{i=1}^{n}x_i^2 - 4\left(\sum_{i=1}^{n}x_i\right)^2 = 4n\sum_{i=1}^{n}x_i^2 - 4(n\overline{x})^2 = 4n^2\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2\right)$$

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2 \text{ étant la variance empirique-non corrigée- de } (x_1, \dots, x_n)$$

$$S_y^2 = \frac{1}{n}\sum_{i=1}^{n}y_i^2 - \overline{y}^2 \text{ étant la variance empirique biaisée de } (y_1, \dots, y_n)$$

$$S_{x,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) = \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{xy} \text{ étant la covariance empirique non biaisée}$$

$$\left|H\left(\varphi(\widehat{\beta}_1,\widehat{\beta}_2)\right)\right| = 4n^2S_x^2 > 0 \Rightarrow (\widehat{\beta}_1,\widehat{\beta}_2) \text{ est un minimum local}$$

$$\begin{cases} \widehat{\beta}_1 = \overline{y} - \widehat{\beta}_2\overline{x} \\ \widehat{\beta}_2 = \dfrac{\sum_{i=1}^{n}x_iy_i - n\overline{xy}}{\sum_{i=1}^{n}x_i^2 - n\overline{x}^2} = \dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \dfrac{Cov(x_i,y_i)}{Var(x_i)} = \dfrac{S_{x,y}}{S_x^2} \end{cases}$$

☞ **1ère conséquence** : $\widehat{y}_i - \overline{y} = \frac{S_{x,y}}{S_x^2}(x_i - \overline{x}) = \widehat{\beta}_2(x_i - \overline{x})$

On démontre aussi que la moyenne arithmétique de $\widehat{y}_i$ est égal à $\overline{y}$:

$$\overline{\widehat{y}_i} = \frac{1}{n}\sum_{i=1}^{n}(\widehat{\beta}_1 + \widehat{\beta}_2x_i) = \frac{1}{n}\underbrace{\sum_{i=1}^{n}\widehat{\beta}_1}_{n\widehat{\beta}_1} + \frac{\widehat{\beta}_2}{n}\sum_{i=1}^{n}x_i = \widehat{\beta}_1 + \widehat{\beta}_2\overline{x} = \overline{y} \Rightarrow \overline{\widehat{y}_i} = \overline{y}$$

☞ **2ème conséquence** : *La somme -et donc la moyenne arithmétique- des résidus est nulle dans une régression avec constante. En effet :*

$$\sum_{i=1}^{n}\widehat{\varepsilon}_i = \sum_{i=1}^{n}(y_i - \widehat{\beta}_1 - \widehat{\beta}_2x_i) = n\overline{y} - n\widehat{\beta}_1 - \widehat{\beta}_2\sum_{i=1}^{n}x_i = n\overline{y} - n(\overline{y} - \widehat{\beta}_2\overline{x}) - n\widehat{\beta}_2\overline{x} = 0 \Rightarrow \sum_{i=1}^{n}\widehat{\varepsilon}_i = 0$$

☞ **3ème conséquence** :

$$\sum_{i=1}^{n}\widehat{\varepsilon}_ix_i = \sum_{i=1}^{n}(y_i - \widehat{\beta}_1 - \widehat{\beta}_2x_i)\,x_i = \sum_{i=1}^{n}x_iy_i - \widehat{\beta}_1\sum_{i=1}^{n}x_i - \widehat{\beta}_2\sum_{i=1}^{n}x_i^2$$

$$\sum_{i=1}^{n}\widehat{\varepsilon}_ix_i = \sum_{i=1}^{n}x_iy_i - (\overline{y} - \widehat{\beta}_2\overline{x})(n\overline{x}) - \widehat{\beta}_2\sum_{i=1}^{n}x_i^2 = \left(\sum_{i=1}^{n}x_iy_i - n\overline{xy}\right) + n\widehat{\beta}_2\overline{x}^2 - \widehat{\beta}_2\sum_{i=1}^{n}x_i^2$$

$$= nS_{x,y} - \widehat{\beta}_2\left(\sum_{i=1}^{n}x_i^2 - n\overline{x}^2\right) = nS_{x,y} - \frac{S_{x,y}}{S_x^2}(nS_x^2) \Rightarrow \sum_{i=1}^{n}\widehat{\varepsilon}_ix_i = 0$$

☞ $4^{\text{ème}}$ ***conséquence*** : ***Il existe un lien entre la pente d'une régression linéaire simple et le coefficient de corrélation*** :

$$\hat{\beta}_2 = \frac{S_{x,y}}{S_x^2} = \frac{S_{x,y}}{S_x S_x} \times \frac{S_y}{S_y} = \frac{Cov(x_i, y_i)}{S_x S_y} \times \frac{S_y}{S_x}$$

$\Rightarrow \hat{\beta}_2 = \dfrac{S_y}{S_x}\rho_{x,y}$; ***$\hat{\beta}_2$ et $\rho_{x,y}$ auront toujours le même signe .***

☑ ***L'erreur quadratique moyenne minimale :***

L'équation de la droite des moindres carrées étant :

$\hat{y}_i = \hat{\beta}_1 - \hat{\beta}_2 x_i = \bar{y} + \dfrac{S_{x,y}}{S_x^2}(x_i - \bar{x})$. ***L'erreur quadratique moyenne minimale sera :***

$$\delta_{min}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y} - \frac{S_{x,y}}{S_x^2}(x_i - \bar{x})\right)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 - 2\frac{S_{x,y}}{S_x^2}\left(\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})\right) + \frac{S_{x,y}^2}{S_x^4}\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

$$= S_y^2 - 2\frac{S_{x,y}}{S_x^2}(S_{x,y}) + \frac{S_{x,y}^2}{S_x^4}(S_x^2) = S_y^2 - \frac{S_{x,y}^2}{S_x^2} = S_y^2 - \left(\frac{S_{x,y}^2}{S_x^2 S_y^2}\right)S_y^2 = S_y^2 - \rho_{x,y}^2 S_y^2$$

$$\delta_{min}^2 = S_y^2(1 - \rho_{x,y}^2)$$

On en déduit en particulier que $\delta_{min}^2 = 0$ si et seulement si $|\rho_{x,y}| = 1$

$\mathcal{A}$– 5 • Moments des estimateurs de moindres carrées :

Soient les identités suivantes :

- $w_i = \dfrac{x_i - \bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$
- $z_i = \dfrac{1}{n} - \bar{x} w_i$

Il est facile de vérifier que :

☞ $\hat{\beta}_1 = \sum_{i=1}^{n} z_i y_i$ ☞ $\hat{\beta}_2 = \sum_{i=1}^{n} w_i y_i$ ☞ $\sum_{i=1}^{n} w_i = 0$ ☞ $\sum_{i=1}^{n} w_i^2 = \dfrac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

☞ $\sum_{i=1}^{n} w_i x_i = 1$ ☞ $\sum_{i=1}^{n} z_i = 1$ ☞ $\sum_{i=1}^{n} z_i^2 = \dfrac{1}{n} + \dfrac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \dfrac{\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}$

☞ $\sum_{i=1}^{n} z_i x_i = 0$ ☞ $\sum_{i=1}^{n} z_i w_i = -\dfrac{\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$a \bullet$ ***Espérances mathématiques :*** *Nous allons vérifier que* $\widehat{\beta}_1$ *et* $\widehat{\beta}_2$ *sont des estimateurs sans biais de* β_1 *et* β_2 *. On a :*

$$\blacksquare\ \widehat{\beta}_1 = \sum_{i=1}^{n} z_i y_i = \sum_{i=1}^{n} z_i(\beta_1 + \beta_2 x_i + \varepsilon_i) = \beta_1 \underbrace{\sum_{i=1}^{n} z_i}_{1} + \beta_2 \underbrace{\sum_{i=1}^{n} z_i x_i}_{0} + \sum_{i=1}^{n} z_i \varepsilon_i$$

$$\widehat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} z_i \varepsilon_i \Rightarrow E(\widehat{\beta}_1) = \underbrace{E(\beta_1)}_{\beta_1} + \sum_{i=1}^{n} z_i \underbrace{E(\varepsilon_i)}_{\blacksquare\ \mathcal{H}_3:\ 0}$$

$E(\widehat{\beta}_1) = \beta_1$ ***donc*** $\widehat{\beta}_1$ ***est un estimateur sans biais de*** β_1

$$\blacksquare\ \widehat{\beta}_2 = \sum_{i=1}^{n} w_i y_i = \sum_{i=1}^{n} w_i(\beta_1 + \beta_2 x_i + \varepsilon_i) = \beta_1 \underbrace{\sum_{i=1}^{n} w_i}_{0} + \beta_2 \underbrace{\sum_{i=1}^{n} w_i x_i}_{1} + \sum_{i=1}^{n} w_i \varepsilon_i$$

$$\widehat{\beta}_2 = \beta_2 + \sum_{i=1}^{n} w_i \varepsilon_i \Rightarrow E(\widehat{\beta}_2) = \underbrace{E(\beta_2)}_{\beta_2} + \sum_{i=1}^{n} w_i \underbrace{E(\varepsilon_i)}_{\blacksquare\ \mathcal{H}_3:\ 0}$$

$E(\widehat{\beta}_2) = \beta_2$ ***donc*** $\widehat{\beta}_2$ ***est un estimateur sans biais de*** β_2

$b \bullet$ ***Variance :***

$$\blacksquare\ Var(\widehat{\beta}_2) = E\left[\left(\widehat{\beta}_2 - \underbrace{E(\widehat{\beta}_2)}_{\beta_2}\right)^2\right] = E\left[(\widehat{\beta}_2 - \beta_2)^2\right]$$

$$Or\ \widehat{\beta}_2 - \beta_2 = \sum_{i=1}^{n} w_i \varepsilon_i\ , donc\ , Var(\widehat{\beta}_2) = E\left[\left(\sum_{i=1}^{n} w_i \varepsilon_i\right)^2\right]$$

• Produits de sommes :

• Développement d'un carré : *Soit* $I = [m, n]$ *un intervalle d'entiers.*

$$\left(\sum_{i=m}^{n} x_i\right)^2 = \sum_{m \le i,j \le n} x_i x_j = \sum_{i=m}^{n}\sum_{j=m}^{n} x_i x_j = \sum_{i=m}^{n} x_i^2 + \sum_{\substack{m \le i,j \le n \\ i \ne j}} x_i x_j = \sum_{i=m}^{n} x_i^2 + 2\sum_{m \le i < j \le n} x_i x_j$$

On peut écrire aussi : $\displaystyle \left(\sum_{i=m}^{n} x_i\right)^2 = \sum_{i=m}^{n} x_i^2 + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} x_i x_j$

$$Ainsi\ \left(\sum_{i=1}^{n} w_i \varepsilon_i\right)^2 = \sum_{i=m}^{n} w_i^2 \varepsilon_i^2 + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} w_i w_j \varepsilon_i \varepsilon_j$$

$$Et\ Var(\hat{\beta}_2) = E\left[\sum_{i=m}^{n} w_i^2 \varepsilon_i^2 + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} w_i w_j \varepsilon_i \varepsilon_j\right] = \sum_{i=m}^{n} w_i^2 \underbrace{E(\varepsilon_i^2)}_{\mathcal{H}_4:\ \sigma_\varepsilon^2} + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} w_i w_j \underbrace{E(\varepsilon_i\varepsilon_j)}_{\mathcal{H}_6:\ 0}$$

$$= \sigma_\varepsilon^2 \underbrace{\sum_{i=m}^{n} w_i^2}_{1/\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$Var(\hat{\beta}_1) = E\left[\left(\hat{\beta}_1 - \underbrace{E(\hat{\beta}_1)}_{\beta_1}\right)^2\right] = E\left[(\hat{\beta}_1-\beta_1)^2\right]$$

$$Or\ \hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} z_i\varepsilon_i\ \ et\ \left(\sum_{i=1}^{n} z_i\varepsilon_i\right)^2 = \sum_{i=m}^{n} z_i^2\varepsilon_i^2 + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} z_i z_j \varepsilon_i\varepsilon_j$$

$$ainsi\ Var(\hat{\beta}_1) = E\left[\sum_{i=m}^{n} z_i^2 \varepsilon_i^2 + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} z_i z_j \varepsilon_i \varepsilon_j\right] = \sum_{i=m}^{n} z_i^2 \underbrace{E(\varepsilon_i^2)}_{\mathcal{H}_4:\ \sigma_\varepsilon^2} + 2\sum_{i=m}^{n-1}\sum_{j=i+1}^{n} z_i z_j \underbrace{E(\varepsilon_i\varepsilon_j)}_{\mathcal{H}_6:\ 0}$$

$$= \sigma_\varepsilon^2 \underbrace{\sum_{i=m}^{n} z_i^2}_{\frac{1}{n}+\frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}}$$

$$Var(\hat{\beta}_1) = \sigma_\varepsilon^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right] = \sigma_\varepsilon^2\left[\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]$$

c • Covariance :

$$Cov(\hat{\beta}_1,\hat{\beta}_2) = E\left[\left(\hat{\beta}_1 - \underbrace{E(\hat{\beta}_1)}_{\beta_1}\right)\left(\hat{\beta}_2 - \underbrace{E(\hat{\beta}_2)}_{\beta_2}\right)\right] = E\left[\underbrace{(\hat{\beta}_1-\beta_1)}_{\sum_{i=1}^{n} z_i\varepsilon_i}\underbrace{(\hat{\beta}_2-\beta_2)}_{\sum_{i=1}^{n} w_i\varepsilon_i}\right]$$

$$= E\left[\left(\sum_{i=1}^{n} z_i\varepsilon_i\right)\left(\sum_{i=1}^{n} w_i\varepsilon_i\right)\right] = E\left[\sum_{i=1}^{n} z_i w_i \varepsilon_i^2 + \sum_{i=1}^{n}\sum_{j\neq i}(\varepsilon_i z_i)(\varepsilon_j w_j)\right]$$

$$= \sum_{i=1}^{n} z_i w_i \underbrace{E(\varepsilon_i^2)}_{\mathcal{H}_4:\ \sigma_\varepsilon^2} + \sum_{i=1}^{n}\sum_{j\neq i} z_i w_j \underbrace{E(\varepsilon_i\varepsilon_j)}_{\mathcal{H}_6:\ 0} = \sigma_\varepsilon^2 \underbrace{\sum_{i=1}^{n} z_i w_i}_{-\frac{\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}$$

$$Cov(\hat{\beta}_1,\hat{\beta}_2) = -\,\sigma_\varepsilon^2\left[\frac{\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]$$

$\mathcal{A}-6$ • ***Convergence en probabilité*** : *Sous la condition* $\lim\limits_{n\to+\infty}\left[\frac{\sum_{i=1}^{n}x_i^2}{n}\right]$ *existe , on a* :

- $\begin{cases}\lim\limits_{n\to+\infty}E(\widehat{\beta}_1)=\beta_1\\ \lim\limits_{n\to+\infty}Var(\widehat{\beta}_1)=\lim\limits_{n\to+\infty}\left[\frac{\sigma_\varepsilon^2\sum_{i=1}^{n}x_i^2}{Var(x_i)}\right]\frac{1}{n^2}=0\end{cases}\Leftrightarrow\left(\widehat{\beta}_1\xrightarrow[n\to+\infty]{mq}\beta_1\right)\Rightarrow \widehat{\beta}_1\xrightarrow[n\to+\infty]{p}\beta_1$

- $\begin{cases}\lim\limits_{n\to+\infty}E(\widehat{\beta}_2)=\beta_2\\ \lim\limits_{n\to+\infty}Var(\widehat{\beta}_2)=\lim\limits_{n\to+\infty}\left[\frac{\sigma_\varepsilon^2}{Var(x_i)}\right]\frac{1}{n}=0\end{cases}\Leftrightarrow\left(\widehat{\beta}_2\xrightarrow[n\to+\infty]{mq}\beta_2\right)\Rightarrow \widehat{\beta}_2\xrightarrow[n\to+\infty]{p}\beta_2$

$\mathcal{A}-7$ • ***Interprétation matricielle*** :

En réunissant toutes les observations sur l'équation de régression : $y_i=\beta_1+\beta_2x_i+\varepsilon_i$;

$i=1,2,\dots,n$ *On obtient* :

$$\underbrace{Y}_{(n\times1)}=\underbrace{X}_{(n\times2)}\underbrace{\beta}_{(2\times1)}+\underbrace{\mathcal{E}}_{(n\times1)}\ ;\ avec\ Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},X=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix},\beta=\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}\ et\ \mathcal{E}=\begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix}$$

$$\Rightarrow(X'X)^{-1}=\begin{pmatrix}n & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2\end{pmatrix}^{-1}=\frac{1}{n\sum_{i=1}^{n}(x_i-\overline{x})^2}\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & -\sum_{i=1}^{n}x_i\\ -\sum_{i=1}^{n}x_i & n\end{pmatrix}$$

$$\widehat{\beta}=(X'X)^{-1}(X'Y)\ minimise\ \sum_{i=1}^{n}\varepsilon_i^2$$

$$X'X=\begin{pmatrix}n & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2\end{pmatrix}\Rightarrow(X'X)^{-1}=\frac{1}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\begin{pmatrix}\sum_{i=1}^{n}x_i^2/n & -\overline{x}\\ -\overline{x} & 1\end{pmatrix}$$

$$\Rightarrow\sigma_\varepsilon^2(X'X)^{-1}=\begin{pmatrix}Var(\widehat{\beta}_1) & Cov(\widehat{\beta}_1,\widehat{\beta}_2)\\ Cov(\widehat{\beta}_1,\widehat{\beta}_2) & Var(\widehat{\beta}_2)\end{pmatrix}$$

$\mathcal{A}-8$ • ***Théorème de Gauss-Markov*** :

$\widehat{\beta}_1$ *et* $\widehat{\beta}_2$ ***sont les estimateurs de*** β_1 *et* β_2 ***sans biais et de variance minimale*** (*ESBVM*)

parmi tous les estimateurs sans biais linéaires (*qui s'écrivent comme des combinaisons linéaires des* y_i)

Soit $\widetilde{\beta}$ *un autre estimateur linéaire* ($\widetilde{\beta}=AY$) *sans biais de* β *alors* : $Var(\widetilde{\beta})\geq Var(\widehat{\beta})$

$\mathcal{A}$– 9 • Décomposition de la variance totale et coefficient de détermination :

a • Estimation de la variance des erreurs :

- *Les variables aléatoires* $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i$ *sont appelées les résidus empiriques*

- *La variance empirique biaisée des résidus est notée* $S_{\hat{\varepsilon}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\varepsilon}_i - \bar{\hat{\varepsilon}}\right)^2$

comme on a démontré que $\sum_{i=1}^{n}\hat{\varepsilon}_i = 0$, *donc* $\bar{\hat{\varepsilon}} = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i = 0$, *ainsi* $S_{\hat{\varepsilon}}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$

or l'erreur quadratique moyenne minimale est : $\delta_{min}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$

Donc la variance résiduelle n'est rien d'autre que l'erreur quadratique moyenne minimale et l'expression de la variance empirique biaisée des résidus sera :

$$S_{\hat{\varepsilon}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\varepsilon}_i - \bar{\hat{\varepsilon}}\right)^2 = \delta_{min}^2 = S_y^2\left(1 - \rho_{x,y}^2\right)$$

Nous avons : $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i = \beta_1 + \beta_2 x_i + \varepsilon_i - \left(\bar{y} - \hat{\beta}_2\bar{x}\right) - \hat{\beta}_2 x_i$

$$= \beta_1 + \beta_2 x_i + \varepsilon_i - \left[(\beta_1 + \beta_2\bar{x} + \bar{\varepsilon}) - \hat{\beta}_2\bar{x}\right] - \hat{\beta}_2 x_i$$

$$= \beta_1 + \beta_2 x_i + \varepsilon_i - \beta_1 - \beta_2\bar{x} - \bar{\varepsilon} + \hat{\beta}_2\bar{x} - \hat{\beta}_2 x_i$$

$$= (\varepsilon_i - \bar{\varepsilon}) + \beta_2(x_i - \bar{x}) - \hat{\beta}_2(x_i - \bar{x}) = (\varepsilon_i - \bar{\varepsilon}) + \left(\beta_2 - \hat{\beta}_2\right)(x_i - \bar{x})$$

Alors $\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \sum_{i=1}^{n}\left[(\varepsilon_i - \bar{\varepsilon}) + \left(\beta_2 - \hat{\beta}_2\right)(x_i - \bar{x})\right]^2$

$$= \sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2 + \left(\beta_2 - \hat{\beta}_2\right)^2\sum_{i=1}^{n}(x_i - \bar{x})^2 + 2\left(\beta_2 - \hat{\beta}_2\right)\sum_{i=1}^{n}(x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon})$$

Or $\sum_{i=1}^{n}(x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon}) = \sum_{i=1}^{n}(x_i - \bar{x})^2\left[\sum_{i=1}^{n}\left(\underbrace{\left(\frac{x_i - \bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}_{w_i}(\varepsilon_i - \bar{\varepsilon})\right)\right]$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2\left[\sum_{i=1}^{n}\left(w_i(\varepsilon_i-\bar{\varepsilon})\right)\right] = \sum_{i=1}^{n}(x_i-\bar{x})^2\left[\underbrace{\sum_{i=1}^{n} w_i\varepsilon_i}_{\hat{\beta}_2-\beta_2} - \bar{\varepsilon}\underbrace{\sum_{i=1}^{n} w_i}_{0}\right]$$

$$\sum_{i=1}^{n}(x_i-\bar{x})(\varepsilon_i-\bar{\varepsilon}) = -(\beta_2-\hat{\beta}_2)\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$Donc \sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2 + (\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + 2(\beta_2-\hat{\beta}_2)\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})(\varepsilon_i-\bar{\varepsilon})}_{-(\beta_2-\hat{\beta}_2)\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$= \sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2 + (\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2 - 2(\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2 - (\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$E\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\right) = E\left(\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2\right) - E\left((\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

Calculons les espérances de $\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2$ *et* $(\beta_2-\hat{\beta}_2)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$

- $$E\left[\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2\right] = E\left[\sum_{i=1}^{n}\varepsilon_i^2 - n\bar{\varepsilon}^2\right] = \sum_{i=1}^{n}\underbrace{E(\varepsilon_i^2)}_{\mathcal{H}_4:\,\sigma_\varepsilon^2} - nE(\bar{\varepsilon}^2) = \sum_{i=1}^{n}\sigma_\varepsilon^2 - n\left[Var(\bar{\varepsilon}) + \left(E(\bar{\varepsilon})\right)^2\right]$$

$$= n\,\sigma_\varepsilon^2 - n\left[Var\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right) + \left(E\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right)\right)^2\right]$$

$$= n\,\sigma_\varepsilon^2 - n\left[\frac{1}{n^2}Var\left(\sum_{i=1}^{n}\varepsilon_i\right) + \left(\frac{1}{n}E\left(\sum_{i=1}^{n}\varepsilon_i\right)\right)^2\right]$$

$$= n\,\sigma_\varepsilon^2 - n\left[\frac{1}{n^2}\sum_{i=1}^{n}\underbrace{Var(\varepsilon_i)}_{\substack{\mathcal{H}_4:\,Var(\varepsilon_i)=\sigma_\varepsilon^2\\ \mathcal{H}_6:\,Cov(\varepsilon_i,\varepsilon_j)=0}} + \left(\frac{1}{n}\sum_{i=1}^{n}\underbrace{E(\varepsilon_i)}_{\mathcal{H}_3:\,0}\right)^2\right] = n\,\sigma_\varepsilon^2 - n\left[\frac{1}{n^2}\sum_{i=1}^{n}\sigma_\varepsilon^2\right]$$

$$= n\,\sigma_\varepsilon^2 - n\left[\frac{n\,\sigma_\varepsilon^2}{n^2}\right]$$

$$E\left[\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2\right]=(n-1)\,\sigma_\varepsilon^2$$

■ $E\left[\left(\beta_2-\hat{\beta}_2\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right]=\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]E\left[\left(\hat{\beta}_2-\beta_2\right)^2\right]=\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]E\left[\left(\hat{\beta}_2-E(\hat{\beta}_2)\right)^2\right]$

$$=\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]Var(\hat{\beta}_2)=\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\sigma_\varepsilon^2$$

$D'\text{où } E\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\right)=E\left(\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2\right)-E\left(\left(\beta_2-\hat{\beta}_2\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right)=(n-1)\,\sigma_\varepsilon^2-\sigma_\varepsilon^2$

$$E\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\right)=(n-2)\,\sigma_\varepsilon^2$$

Dans le cas d'un échantillon, la variance empirique est un estimateur biaisée de la variance de l'échantillon. Pour la débiaiser, on la multiplie par $\frac{n}{n-1}$

Ici, on a deux échantillons. On peut montrer qu'alors la variance résiduelles est un estimateur biaisée de σ_ε^2 *, et que, pour la débiaiser, il faut la multiplier par* $\frac{n}{n-2}$

D'où finalement : $\hat{\sigma}_\varepsilon^2=\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}=\frac{SCR}{n-2}=\frac{S_{\hat{\varepsilon}}^2}{n-2}=\frac{\delta_{min}^2}{n-2}=\frac{S_y^2\left(1-\rho_{x,y}^2\right)}{n-2}$ *est un ESB de* σ_ε^2

$SCR=\sum_{i=1}^{n}\hat{\varepsilon}_i^2$ *: est la somme des carrés résiduels. Elle indique la variabilité de Y non expliquée par le modèle.*

• *L'équation d'analyse de la variance :*

☑ ***Lemme :*** $SCR=\sum_{i=1}^{n}\hat{\varepsilon}_i^2=\sum_{i=1}^{n}(y_i-\bar{y})^2-\hat{\beta}_2^2\sum_{i=1}^{n}(x_i-\bar{x})^2=n\left(S_y^2-\hat{\beta}_2^2S_x^2\right)$

■ $SCT=\sum_{i=1}^{n}(y_i-\bar{y})^2=nS_y^2$ *: est la somme des carrés totaux. Elle indique la variabilité totale de Y.*

■ $SCE=\sum_{i=1}^{n}\left(\hat{y}_i-\bar{\hat{y}}\right)^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$ *: est la somme des carrés expliqués. Elle indique la* *variation de Y due à sa régression linéaire sur X.*

- $SCR = \sum_{i=1}^{n}(\hat{\varepsilon}_i - \bar{\hat{\varepsilon}})^2 = \sum_{i=1}^{n}\hat{\varepsilon}_i^2$: *est la somme des carrés résiduels. Elle indique la variabilité de Y non expliquée par le modèle.*

Pour prouver que $SCT = SCE + SCR$ *, il suffit de montrer que* $\hat{\beta}_2^2\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2$

On a : $\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2 = \sum_{i=1}^{n}[(\hat{\beta}_1+\hat{\beta}_2x_i)-(\hat{\beta}_1+\hat{\beta}_2\bar{x})]^2 = \sum_{i=1}^{n}\left(\hat{\beta}_2(x_i-\bar{x})\right)^2 = \hat{\beta}_2^2\sum_{i=1}^{n}(x_i-\bar{x})^2$

Ainsi on a :
$$\begin{cases} SCE = \sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2 = \hat{\beta}_2^2\sum_{i=1}^{n}(x_i-\bar{x})^2 = n\hat{\beta}_2^2S_x^2 = \hat{\beta}_2\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) = n\hat{\beta}_2S_{x,y} \\ SCT = \sum_{i=1}^{n}(y_i-\bar{y})^2 = nS_y^2 \\ SCR = \sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \sum_{i=1}^{n}(y_i-\bar{y})^2 - \hat{\beta}_2^2\sum_{i=1}^{n}(x_i-\bar{x})^2 = n(S_y^2-\hat{\beta}_2^2S_x^2) \end{cases}$$

D'où $\underbrace{\sum_{i=1}^{n}(y_i-\bar{y})^2}_{SCT} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2}_{SCE} + \underbrace{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}_{SCR}$ *ou encore* $SCT = SCE + SCR$ $\left(\textit{seulement pour } \beta_1 \neq 0\right)$

c • Tableau d'analyse de variance :

Source de la variabilité	Somme des carrés	Degré de liberté	Carrés moyens
Régression	SCE	1	SCE
Résidus	SCR	$n-2$	$\hat{\sigma}^2 = SCR/n-2$
Total	SCT	$n-1$	$S^2 = SCT/n-1$

d • Coefficient de détermination : *La part de variance de Y expliquée par le modèle est toujours traduite par le coefficient de détermination :*

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} \in [0,1] (\textit{seulement pour } \beta_1 \neq 0)$$

- *Plus le* R^2 *est proche de 1, meilleur est l'ajustement, la connaissance des valeurs de X permet de deviner avec précision celles de Y.*
- *Plus le* R^2 *est proche de 0, mauvais est l'ajustement, X n'apporte pas d'informations utiles sur Y.*
- *Il faut tout de même faire attention quant au crédit à accorder au* R^2 *, il doit toujours être accompagné d'autres tests (Student et Fisher essentiellement) avant de*

trancher sur sur la bonté d'un modèle, mais il reste un critère non négligeable pour la prévision.

■ *Les résultats précédentes ne sont pas nécessairement valables dès qu'ils'agit d'un modèle sans terme constant* (β_1). *En effet, on a peut avoir* $\sum_{i=1}^{n} \hat{\varepsilon}_i \neq 0$, *comme on peut avoir* $SCT \neq SCE + SCR$ *ou encore* $R^2 \notin [0,1]$

■ *Mentionnons dès à présent une interprétation statistique de* R^2. *Nous démontrons, en régression multiple, que si* $\beta_2 = 0$, *alors,*

$$\frac{(n-2)R^2}{1-R^2} \backsim \mathcal{F}(1,(n-2)) \text{ ou encore } \sqrt{\frac{(n-2)R^2}{1-R^2}} \backsim \mathcal{T}(n-2)$$

Avec un seuil de signification α, *le* R^2 *sera donc bon si :* $\left[\frac{(n-2)R^2}{1-R^2}\right]_{obs} > t^2_{1-\frac{\alpha}{2}}(n-2)$

Au meilleur des cas	*Au pire des cas*
$SCR = 0$	$SCE = 0$
$SCT = SCE$	$SCT = SCR$
$R^2 = 1$	$R^2 = 0$
Le modèle est parfait, la droite de régression passe par tous les points du nuage.	*Le modèle est mauvais, la meilleure prédiction de Y est sa propre moyenne.*

e • Coefficient de détermination corrigé ou ajusté: *Le* R^2 *est un indicateur de qualité, mais il présente un défaut ennuyeux : plus nous augmentons le nombre de variables explicatives, même non pertinentes, n'ayant aucun rapport avec le problème que l'on cherche à résoudre, plus grande sera sa valeur.*

$$\bar{R}^2 = 1 - \frac{SCR/n-2}{SCT/n-1} = 1 - \left(\frac{n-1}{n-2}\right)(1-R^2), \text{ avec}, \bar{R}^2 < R^2 \text{ et si } n \gg \Rightarrow \bar{R}^2 \approx R^2$$

En augmentant le nombre d'explicatives, nous augmentons de manière mécanique la valeur du R^2 *mais, dans le même temps, nous diminuons le degré de liberté. Il faudrait donc intégrer cette dernière notion pour contrecarrer l'évolution du* R^2.

C'est exactement ce que fait-le R^2*-ajusté (ou* R^2*-corrigé)*

Le $\bar{R}^2$ *en tant que tel n'est pas d'une grande utilité. Son principal avantage est qu'il permet de comparer des modèles imbriqués.*

f • **Relation entre coefficient de corrélation $\rho_{x,y}$ et coefficient de détermination R^2 :**

Pour une régression linéaire simple, et seulement dans ce cas, le R^2 n'est rien d'autre que le carré du coefficient de corrélation

$$\widehat{\beta}_2 = \frac{S_y}{S_x}\rho_{x,y} \Rightarrow \widehat{\beta}_2^2 = \frac{S_y^2}{S_x^2}\rho_{x,y}^2 \Rightarrow \rho_{x,y}^2 = \frac{\widehat{\beta}_2^2 S_x^2}{S_y^2} \Rightarrow \rho_{x,y}^2 = \frac{\frac{1}{n}\widehat{\beta}_2^2\sum_{i=1}^{n}(x_i-\overline{x})^2}{\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2} \Rightarrow \rho_{x,y}^2 = \frac{\widehat{\beta}_2^2\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$$

$$\Rightarrow \rho_{x,y}^2 = \frac{SCE}{SCT} = R^2 \text{ . Par conséquent : } \rho_{x,y} = signe(\widehat{\beta}_2)\sqrt{R^2}$$

g • **Signification du coefficient de corrélation :** *Le travail d'interprétation d'un coefficient de corrélation linéaire se fait toujours en deux temps : une interprétation par rapport au signe/sens de la liaison et une interprétation par rapport au degré de dépendance.*

Interprétation par rapport au signe	■ Si $\rho_{x,y} > 0$, X et Y sont positivement corrélées (la relation linéaire entre X et Y est positive) ■ Si $\rho_{x,y} < 0$, X et Y sont négativement corrélées (la relation linéaire entre X et Y est négative) ■ Si $\rho_{x,y} = 0$, X et Y sont non corrélées (pas de liaison linéaire, mais possibilité d'une liaison d'un autre type)
Interprétation par rapport à l'intensité	■ Si $\rho_{x,y} = \pm 1$, le lien linéaire entre X et Y est parfait. Dans ce cas, l'une des variables est fonction affine de l'autre, les n points (x_i, y_i) sont alignés. ■ Si $0,8 < \rho_{x,y} < 1$, le lien linéaire est très fort. ■ Si $0,65 < \rho_{x,y} < 0,8$, le lien linéaire est fort (élevé) ■ Si $0,5 < \rho_{x,y} < 0,65$, le lien linéaire est modéré ■ Si $0,25 < \rho_{x,y} < 0,5$, le lien linéaire est faible. ■ Si $0,025 < \rho_{x,y} < 0,25$, le lien linéaire est très faible ■ Si $\rho_{x,y}$, proche de 0, alors il y a absence de lien entre X et Y

$\mathcal{A}$ – 10 • Intervalles de confiance et Tests de significativité :

a • **Variances et covariance estimées des estimateurs des paramètres :**

■ $$\begin{cases} \widehat{\sigma}_\varepsilon^2 = \dfrac{SCR}{n-2} \\ \sigma_{\widehat{\beta}_1}^2 = \sigma_\varepsilon^2\left[\dfrac{1}{n} + \dfrac{\overline{x}^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right] \end{cases} \Rightarrow \widehat{\sigma}_{\widehat{\beta}_1}^2 = \widehat{\sigma}_\varepsilon^2\left[\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right]$$

■ $$\begin{cases} \widehat{\sigma}_\varepsilon^2 = \dfrac{SCR}{n-2} \\ \sigma_{\widehat{\beta}_2}^2 = \dfrac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2} \end{cases} \Rightarrow \widehat{\sigma}_{\widehat{\beta}_2}^2 = \frac{\widehat{\sigma}_\varepsilon^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

- $\begin{cases} \hat{\sigma}_\varepsilon^2 = \dfrac{SCR}{n-2} \\ Cov(\hat{\beta}_1, \hat{\beta}_2) = -\sigma_\varepsilon^2 \left[\dfrac{\bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}\right] \end{cases} \Rightarrow \widehat{Cov}(\hat{\beta}_1, \hat{\beta}_2) = -\hat{\sigma}_\varepsilon^2 \left[\dfrac{\bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}\right]$

b • *Distributions des statistiques de décision :*

Sous l'hypoyhèse de la normalité $\mathcal{H}_7 : \varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, *nous aurons :*

- $\begin{cases} \hat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} z_i \varepsilon_i \\ E(\hat{\beta}_1) = \beta_1 \\ Var(\hat{\beta}_1) = \sigma_{\hat{\beta}_1}^2 = \sigma_\varepsilon^2 \left[\dfrac{1}{n} + \dfrac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right] \end{cases} \Rightarrow \hat{\beta}_1 \sim \mathcal{N}\left(\beta_1, \sigma_{\hat{\beta}_1}^2\right) \Rightarrow \dfrac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim \mathcal{N}(0,1)$

En pratique σ_ε^2 *est inconnue, en la remplaçant par son estimateur non biaisé* $\hat{\sigma}_\varepsilon^2 = \dfrac{SCR}{n-2}$,

on obtient un estimateur sans biais de $\hat{\sigma}_{\hat{\beta}_1}^2$ *qui sera :* $\hat{\sigma}_{\hat{\beta}_1}^2 = \hat{\sigma}_\varepsilon^2 \left[\dfrac{1}{n} + \dfrac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right]$

Ainsi $\dfrac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim \mathcal{T}(n-2)$ *ou encore* $\dfrac{\left(\hat{\beta}_1 - \beta_1\right)^2}{\hat{\sigma}_{\hat{\beta}_1}^2} \sim \mathcal{F}\big(1, (n-2)\big)$,

du fait que : $\dfrac{(n-2)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} = \dfrac{(n-2)\hat{\sigma}_{\hat{\beta}_1}^2}{\sigma_{\hat{\beta}_1}^2} \sim \chi^2(n-2)$

- $\begin{cases} \hat{\beta}_2 = \beta_2 + \sum_{i=1}^{n} w_i \varepsilon_i \\ E(\hat{\beta}_2) = \beta_2 \\ Var(\hat{\beta}_2) = \sigma_{\hat{\beta}_2}^2 = \dfrac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{cases} \Rightarrow \hat{\beta}_2 \sim \mathcal{N}\left(\beta_2, \sigma_{\hat{\beta}_2}^2\right) \Rightarrow \dfrac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} \sim \mathcal{N}(0,1)$

En pratique σ_ε^2 *est inconnue, en la remplaçant par son estimateur non biaisé* $\hat{\sigma}_\varepsilon^2 = \dfrac{SCR}{n-2}$

on obtient un estimateur sans biais de $\hat{\sigma}_{\hat{\beta}_2}^2$ *qui sera :* $\hat{\sigma}_{\hat{\beta}_2}^2 = \dfrac{\hat{\sigma}_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$

Ainsi $\dfrac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}_2}} \sim \mathcal{T}(n-2)$ *ou encore* $\dfrac{\left(\hat{\beta}_2 - \beta_2\right)^2}{\hat{\sigma}_{\hat{\beta}_2}^2} \sim \mathcal{F}\big(1, (n-2)\big)$,

du fait que : $\dfrac{(n-2)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} = \dfrac{(n-2)\hat{\sigma}_{\hat{\beta}_2}^2}{\sigma_{\hat{\beta}_2}^2} \sim \chi^2(n-2)$

- *On démontre aussi que* $\dfrac{SCE/1}{SCR/n-2} = \dfrac{R^2/1}{1-R^2/n-2} \sim \mathcal{F}(1,(n-2))$

c • Intervalles de confiance :

☞ *Intervalle de confiance de seuil α pour* β_1: $IC_{1-\alpha}(\beta_1) = \left[\hat{\beta}_1 \pm t_{1-\frac{\alpha}{2}}(n-2)\hat{\sigma}_{\hat{\beta}_1}\right]$

☞ *Intervalle de confiance de seuil α pour* β_2: $IC_{1-\alpha}(\beta_2) = \left[\hat{\beta}_2 \pm t_{1-\frac{\alpha}{2}}(n-2)\hat{\sigma}_{\hat{\beta}_2}\right]$

☞ *Intervalle de confiance de seuil α pour* σ_ε^2: $IC_{1-\alpha}(\sigma_\varepsilon^2) = \left[\dfrac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{1-\frac{\alpha}{2}}(n-2)}, \dfrac{(n-2)\hat{\sigma}_\varepsilon^2}{\chi^2_{\frac{\alpha}{2}}(n-2)}\right]$

d • Tests sur les coefficients individuels :

Pour tester : $\begin{cases} H_0 : \beta_j = \beta_j^0 \\ \quad contre \\ H_1 : \beta_j \neq \beta_j^0 \end{cases}$, *On rejette* H_0 *si* $\left|\dfrac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}}\right| > t_{1-\frac{\alpha}{2}}(n-2)$; $j \in \{1,2\}$

Pour tester : $\begin{cases} H_0 : \beta_j = \beta_j^0 \\ \quad contre \\ H_1 : \beta_j > \beta_j^0 \end{cases}$, *On rejette* H_0 *si* $\dfrac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}} > t_{1-\alpha}(n-2)$; $j \in \{1,2\}$

Pour tester : $\begin{cases} H_0 : \beta_j = \beta_j^0 \\ \quad contre \\ H_1 : \beta_j < \beta_j^0 \end{cases}$, *On rejette* H_0 *si* $\dfrac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}} < t_{\alpha}(n-2)$; $j \in \{1,2\}$

☞ *Test de significativité individuelle* : *Les hypothèses du test sont* :

$$\begin{cases} H_0 : \beta_j = 0 \ \text{(le paramètre est statistiquement nul, non significatif)} \\ \qquad\qquad contre \\ H_1 : \beta_j \neq 0 \ \text{(le paramètre est statistiquement non nul, significatif)} \end{cases} ;\ j \in \{1,2\}$$

On rejette H_0 *si* $\left|t_{\hat{\beta}_j}\right| > t_{1-\frac{\alpha}{2}}(n-2)$ *avec* $t_{\hat{\beta}_j} = \dfrac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}$ *le t-ratio de Student*

En rejetant H_0 *le paramètre* β_j *est dit statistiquement non nul, la variable x lui associée a une influence significatif dans l'explication de y*

- *Significativité du paramètre* β_1: $\begin{cases} H_0 : \beta_1 = 0 \\ \quad contre \\ H_1 : \beta_1 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \Delta x/x = \Delta y/y \\ \quad contre \\ H_1 : \Delta x/x \neq \Delta y/y \end{cases}$

Autrement dit, tester la significativité de la constante β_1 *revient à tester l'hypothèse d'un transfert intégral d'une variation relative de la variable explicative x vers la variable à expliquer y*

- **Significativité du paramètre β_2:** $\begin{cases} H_0 : \beta_2 = 0 \text{ , (le modèle n'est pas bon)} \\ \text{contre} \\ H_1 : \beta_2 \neq 0 \text{ , (le modèle est bon)} \end{cases}$

$$\Leftrightarrow \begin{cases} H_0 : R^2 = 0 \\ \text{contre} \\ H_1 : R^2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \rho_{x,y} = 0 \text{ , (hypothèse d'absence de corrélation)} \\ \text{contre} \\ H_1 : \rho_{x,y} \neq 0 \text{ , (hypothèse d'absence de décorrélation)} \end{cases}$$

On vérifie bien que $t^2_{\hat\beta_2} = \dfrac{(n-2)R^2}{1-R^2}$ ou encore $R^2 = \dfrac{t^2_{\hat\beta_2}}{t^2_{\hat\beta_2} + (n-2)}$

- **Test de rendement d'échelle constant :** $\begin{cases} H_0 : \beta_2 = 1 \\ \text{contre} \\ H_1 : \beta_2 \neq 1 \end{cases} \Leftrightarrow \begin{cases} H_0 : \Delta y = \Delta x \\ \text{contre} \\ H_1 : \Delta y = \Delta x \end{cases}$

Autrement dit, tester l'hypothèse d'un report intégral d'une variation absolue de la variable explicative x vers la variable à expliquer y $H_0 : \Delta y = \Delta x$ contre $H_1 : \Delta y = \Delta x$ revient à tester le rendement d'échelle constant $H_0 : \beta_2 = 1$ contre $H_1 : \beta_2 \neq 1$

e • Test sur une combinaison linéaire des coefficients : Soit $\hat\gamma$ un estimateur sans biais d'une combinaison linéaire $\hat\gamma = a\hat\beta_1 + b\hat\beta_2$

$$V(\hat\gamma) = V(a\hat\beta_1 + b\hat\beta_2)$$

$$= a^2 \underbrace{V(\hat\beta_1)}_{\sigma_\varepsilon^2\left(\frac{1}{n}+\frac{\bar X^2}{\sum_{i=1}^n (X_i-\bar X)^2}\right)} + b^2 \underbrace{V(\hat\beta_2)}_{\sigma_\varepsilon^2/\sum_{i=1}^n (X_i-\bar X)^2} + 2ab \underbrace{Cov(\hat\beta_1,\hat\beta_2)}_{-\sigma_\varepsilon^2(\bar X/\sum_{i=1}^n (X_i-\bar X)^2)}$$

$$= a^2\sigma_\varepsilon^2\left(\frac{1}{n} + \frac{\bar X^2}{\sum_{i=1}^n (X_i-\bar X)^2}\right) + \frac{b^2\sigma_\varepsilon^2}{\sum_{i=1}^n (X_i-\bar X)^2} - \frac{2ab\sigma_\varepsilon^2 \bar X}{\sum_{i=1}^n (X_i-\bar X)^2}$$

$$= \sigma_\varepsilon^2\left(\frac{a^2}{n} + \frac{(a\bar X)^2}{\sum_{i=1}^n (X_i-\bar X)^2} + \frac{b^2}{\sum_{i=1}^n (X_i-\bar X)^2} - \frac{2ab\bar X}{\sum_{i=1}^n (X_i-\bar X)^2}\right) = \sigma_\varepsilon^2\left(\frac{a^2}{n} + \frac{(a\bar X)^2 + b^2 - 2ab\bar X}{\sum_{i=1}^n (X_i-\bar X)^2}\right)$$

$$V(\hat\gamma) = \sigma_\varepsilon^2\left(\frac{a^2}{n} + \frac{(b - a\bar X)^2}{\sum_{i=1}^n (X_i-\bar X)^2}\right)$$

Ainsi $\dfrac{(a\hat\beta_1 + b\hat\beta_2) - (a\beta_1 + b\beta_2)}{\sigma_\varepsilon\sqrt{\dfrac{a^2}{n} + \dfrac{(b-a\bar X)^2}{\sum_{i=1}^n (X_i-\bar X)^2}}} \backsim \mathcal{T}(n-2)$ ou encore $\dfrac{\left[(a\hat\beta_1 + b\hat\beta_2) - (a\beta_1 + b\beta_2)\right]^2}{\sigma_\varepsilon^2\left(\dfrac{a^2}{n} + \dfrac{(b-a\bar X)^2}{\sum_{i=1}^n (X_i-\bar X)^2}\right)} \backsim \mathcal{F}(1,(n-2))$

$\mathcal{A}$– 11 • Prévision dans le modèle de régression linéaire simple :

En cherchant à trouver un intervalle de confiance sur une valeur future y_θ pour $\theta > n$ où n est la période pour laquelle le modèle (M): $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$; $i = 1, 2, \dots, n$ est *vérifié.*

On parlerait alors d'intervalle de prévision.

Sous l'hypothèse que le modèle reste inchangé, nous aurons : $y_\theta = \beta_1 + \beta_2 x_\theta + \varepsilon_\theta$ *et*

$\hat{y}_\theta = \hat{\beta}_1 + \hat{\beta}_2 x_\theta$ *sera sans biais.*

La variable $y_\theta - \hat{y}_\theta = \varepsilon_\theta - (\hat{\beta}_1 - \beta_1) - (\hat{\beta}_2 - \beta_2)x_\theta$ *est normale de paramètres* $E(y_\theta - \hat{y}_\theta)$ *et* $V(y_\theta - \hat{y}_\theta)$

or $E(y_\theta - \hat{y}_\theta) = E[\varepsilon_\theta - (\hat{\beta}_1 - \beta_1) - (\hat{\beta}_2 - \beta_2)x_\theta] = \underbrace{E(\varepsilon_\theta)}_{0} - E(\hat{\beta}_1 - \beta_1) - E[(\hat{\beta}_2 - \beta_2)x_\theta]$

$E(y_\theta - \hat{y}_\theta) = \underbrace{\left(E(\hat{\beta}_1) - \beta_1\right)}_{0} - x_\theta \underbrace{\left(E(\hat{\beta}_2) - \beta_2\right)}_{0} = 0 \Rightarrow \hat{y}_\theta$ *sans biais de* y_θ

$$V(y_\theta - \hat{y}_\theta) = E\left((y_\theta - \hat{y}_\theta)^2\right) - \underbrace{[E(y_\theta - \hat{y}_\theta)]^2}_{0} = E\left((y_\theta - \hat{y}_\theta)^2\right) = E\left(\left(\varepsilon_\theta - \left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)\right)^2\right)$$

$$V(y_\theta - \hat{y}_\theta) = E(\varepsilon_\theta^2) - 2\underbrace{E\left(\varepsilon_\theta\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)\right)}_{0} + E\left(\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)^2\right)$$

Or $\hat{\beta}_1$ *et* $\hat{\beta}_2$ *ne dépendent que de* $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ *et tant que* $E(\varepsilon_i \varepsilon_\theta) = 0 \ \forall\, i = 1, 2, \dots, n$:

On a donc $E(\varepsilon_\theta \hat{\beta}_1) = E(\varepsilon_\theta \hat{\beta}_2) = 0$

$$E\left(\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)^2\right) = V\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right) \text{ Car } E\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right) = 0$$

$$E\left(\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)^2\right) = V(\hat{\beta}_1 + \hat{\beta}_2 x_\theta - \beta_1 - \beta_2) = V\left(\underbrace{\hat{\beta}_1 + x_\theta \hat{\beta}_2}_{\hat{y} = 1 \times \hat{\beta}_1 + x_\theta \hat{\beta}_2}\right)$$

$$= \sigma_\varepsilon^2 \left(\frac{1^2}{n} + \frac{\left(x_\theta - (1.\bar{x})\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) = \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{(x_\theta - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)$$

$$\text{or } V(y_\theta - \hat{y}_\theta) = E(\varepsilon_\theta^2) + E\left(\left((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)x_\theta\right)^2\right) = \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{(x_\theta - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)$$

$$V(y_\theta - \hat{y}_\theta) = \sigma_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{(x_\theta - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) \Rightarrow \hat{V}(y_\theta - \hat{y}_\theta) = \hat{\sigma}_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{(x_\theta - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)$$

La statistique de décision sera : $\dfrac{\hat{y}_\theta - y_\theta}{\hat{\sigma}_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{(x_\theta - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)} \backsim \mathcal{T}(n-2)$

$D'où\ l'intervalle\ de\ prévision : IP_{1-\alpha}(Y_\theta) = \left[\widehat{Y}_\theta \pm t_{1-\frac{\alpha}{2}}(n-2)\widehat{\sigma}_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(X_\theta - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}}\right]$

Le modèle linéaire général

$\mathcal{B}-1$ • Formulation et hypothèses de base :

En économétrie, on ne considère pas simplement que les variables soient observées sur des unités statistiques. On postule l'existence d'un modèle qui régit les relations entre les variables. La relation la plus simple est une relation linéaire, entre les variables explicatives et la variable dépendante.

Le modèle linéaire général s'écrit : $(M): y_i = \beta_1 + \sum_{j=2}^{k} \beta_j x_{ij} + \varepsilon_i \ ;\ i = 1, 2 \dots, n$

a • Données observées, et formulation matricielle : En observant n réalisations du modèle, on peut l'écrire sous sa forme matricielle :

$$Y = X\beta + \mathcal{E} \text{ avec } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1k} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \text{ et } \mathcal{E} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Où X est une matrice de constantes (non-aléatoire) de plein rang de dimension $n \times k$ d'observations sur les variables explicatives

Y est un vecteur $n \times 1$ d'observations sur la variable dépendante

β est un vecteur $k \times 1$ de paramètres inconnus

Et $\mathcal{E}$ un vecteur $n \times 1$ d'erreursaléatoires inobservables

☑ **Remarque :** La somme des termes d'erreur $1_n'\mathcal{E} = \mathcal{E}'1_n = \sum_{i=1}^{n} \varepsilon_i$ n'est pas nécessairement nulle

b • Les hypothèses :

- $\mathcal{H}_1$ **(Hypothèse structurelle):** $Y = X\beta + \mathcal{E}$: Le modèle est linéaire ou linéarisable *en* X *(ou sur ses paramètres)*

■ $\mathcal{H}_2$**(Hypothèse structurelle)**: *X est non aléatoire c.-à-d. les exogènes x_{ij} et la variable endogène Y sont sont observés sans erreur. Y est aléatoire par l'intermédiaire de $\mathcal{E}$.*

■ $\mathcal{H}_3$**(Hypothèse stochastique)**: $E(\mathcal{E}) = 0$ *En moyenne les erreurs s'annulent c.-à-d. le modèle est bien spécifié.*

■ $\mathcal{H}_4$**(Hypothèse stochastique)**: $E(\mathcal{E}\mathcal{E}') = \sigma_\varepsilon^2 I_n = \Omega_\varepsilon$

où $\begin{cases} Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0 \text{ , si } i \neq j : (\text{non-autocorrélation des erreurs}) \\ E(\varepsilon_i \varepsilon_j) = \sigma_\varepsilon^2 \text{ , si } i = j : (\text{erreurs homoscédastiques}) \end{cases}$

■ $\mathcal{H}_5$**(Hypothèse inférentielle)** : $\mathcal{E} \rightsquigarrow \mathcal{N}(0, \sigma_\varepsilon^2 I_n)$: *Les erreurs ε_i et la variable endogène Y sont gaussiennes et stationnaires en niveau.*

L'hypothèse de normalité des erreurs est un élément clé pour l'inférence statistique.

Elle est donc nécessaire pour mener les tests.

■ $\mathcal{H}_6$**(Hypothèse structurelle)** : *La matrice $(X'X)$ est régulière c.-à-d. $det(X'X) \neq 0$ et $(X'X)^{-1}$ existe. Elle indique l'absence de colinéarité entre les variables exogènes,*

autrement dit les différents vecteurs $X_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix}$ *sont linéairement indépendants.*

En cas de multicolinéarité, la méthode des MCO devient défaillante.

Nous pouvons aussi voir cette hypothèse structurelle sous l'angle :

$rang(X) = rang(X'X) = k$

■ $\mathcal{H}_7$**(Hypothèse structurelle)** : $rang(X) = k < n$ *, c.-à-d. le nombre d'observation est supérieur au nombre de paramètres à estimer (les β_j).*

Lorsque $k > n$, la matrice $(X'X)$ n'est plus inversible.

■ $\mathcal{H}_8$**(Hypothèse structurelle)** : *$X'X/n$ tend vers une matrice finie non singulière lorsque $n \to +\infty$*

■ $\mathcal{H}_9$**(Hypothèse stochastique)** : $E(X\mathcal{E}) = 0$ *ou encore* $Cov(\varepsilon_i, x_{ij}) = 0$:

Il y a indépendance entre la partie systématique et la partie stochastique

$\mathcal{B}$– 2 • Estimateur des moindres carrés ordinaires (MCO):

La régression de Y en X au sens des moindres carrés consiste à chercher l'ajustement qui minimise en $\hat{\beta}$ la somme des carrés des résidus $\hat{\varepsilon}'\hat{\varepsilon} = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})$

☑ **Remarque :** *$\hat{\varepsilon}$ est orthogonal avec toutes les colonnes de X c.-à-d.* $\hat{\varepsilon}'X = X'\hat{\varepsilon} = 0$

$\hat{Y} = X\hat{\beta}$ est une combinaison linéaire des colonnes de X donc $\hat{Y}$ et $\hat{\varepsilon}$ sont aussi orthogonaux c.-à-d. $\hat{\varepsilon}'Y = Y'\hat{\varepsilon} = 0$

☑ **Condition pour que la somme des résidus soit nulle :** *La matrice X peut contenir une variable constante de manière explicite, c.-à-d. qu'une des colonnes de cette matrice contient une variable constante. La constante peut également être définie de manière implicite, ce qui signifie qu'il existe une combinaison linéaire des colonnes de X qui permet d'obtenir une colonne de uns. Formellement, on suppose qu'il existe un vecteur $\lambda = (\lambda_1 \ \cdots \ \lambda_k)'$ de $\mathbb{R}^k$ tel que : $X\lambda = 1_n = (1 \ \cdots \ 1)'$*

■ **Théorème :** *Si la matrice X contient une variable constante définie de manière explicite ou implicite, alors la somme des résidus est nulle :* $1_n'\hat{\varepsilon} = \sum_{i=1}^{n} \hat{\varepsilon}_i = 0$

☞ **Démonstration :**

On a $\hat{\varepsilon}'X = 0$. Or, s'il existe un vecteur $\lambda = (\lambda_1 \ \cdots \ \lambda_k)'$ de $\mathbb{R}^k$ tel que $X\lambda = 1_n = (1 \ \cdots \ 1)'$, alors $\hat{\varepsilon}'X\lambda = \hat{\varepsilon}'1_n = 0'\lambda = 0$

■ *Si la somme des résidus est nulle, la moyenne des valeurs ajustées est égale à la moyenne des valeurs observées, autrement dit :*

$$On\ a \begin{cases} \bar{\hat{Y}} = \dfrac{1_n'\hat{Y}}{n} \\ \bar{Y} = \dfrac{1_n'Y}{n} \end{cases}. \ Or\ 1_n'\hat{\varepsilon} = 0 \Leftrightarrow 1_n'(Y - \hat{Y}) = 0 \Leftrightarrow 1_n'Y = 1_n'\hat{Y} \Leftrightarrow \bar{\hat{Y}} = \bar{Y}$$

$$\hat{\beta} = (X'X)^{-1}(X'Y)\ minimise \sum_{i=1}^{n} \hat{\varepsilon}_i^2$$

$$X'X = \begin{pmatrix} n & \sum_{i=1}^{n} x_{i2} & \cdots & \sum_{i=1}^{n} x_{i,p-1} & \sum_{i=1}^{n} x_{i,p} \\ \sum_{i=1}^{n} x_{i2} & \sum_{i=1}^{n} x_{i2}^2 & \cdots & \sum_{i=1}^{n} x_{i,2}x_{i,p-1} & \sum_{i=1}^{n} x_{i,2}x_{i,p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{i=1}^{n} x_{i,p-1} & \sum_{i=1}^{n} x_{i,2}x_{i,p-1} & \cdots & \sum_{i=1}^{n} x_{i,p-1}^2 & \sum_{i=1}^{n} x_{i,p-1}x_{i,p} \\ \sum_{i=1}^{n} x_{i,p} & \sum_{i=1}^{n} x_{i,2}x_{i,p} & \cdots & \sum_{i=1}^{n} x_{i,p-1}x_{i,p} & \sum_{i=1}^{n} x_{i,p}^2 \end{pmatrix} \text{ et } X'Y = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i2}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{i,p}y_i \end{pmatrix}$$

$\mathcal{B}$– 3 • *Moments des estimateurs de moindres carrées :*

a • Espérance de $\widehat{\beta}$:

$$E(\widehat{\beta}) = E[(X'X)^{-1}(X'Y)] = E\left[(X'X)^{-1}\left(X'(X\beta+\mathcal{E})\right)\right] = E[(X'X)^{-1}(X'X)\beta + (X'X)^{-1}X'\mathcal{E}]$$

$$= E[\beta + (X'X)^{-1}X'\mathcal{E}] = \beta + (X'X)^{-1}X' \underbrace{E(\mathcal{E})}_{\blacksquare\ \mathcal{H}_3:\ 0}$$

$E(\widehat{\beta}) = \beta$ ***donc*** $\widehat{\beta}$ ***est un estimateur sans biais de*** β

b • Matrice de covariance de $\widehat{\beta}$:

$$Var(\widehat{\beta}) = E\left[\left(\widehat{\beta} - E(\widehat{\beta})\right)\left(\widehat{\beta} - E(\widehat{\beta})\right)'\right] = E\left[(\widehat{\beta} - \beta)(\widehat{\beta} - \beta)'\right] = E[((X'X)^{-1}X')(\mathcal{E}\mathcal{E}')(X(X'X)^{-1})]$$

$$= ((X'X)^{-1}X')\left(\underbrace{E(\mathcal{E}\mathcal{E}')}_{\blacksquare\ \mathcal{H}_4:\ \sigma_\varepsilon^2 I_n}\right)(X(X'X)^{-1}) = \sigma_\varepsilon^2 \underbrace{(X'X)^{-1}(X'X)}_{I_k}(X'X)^{-1}$$

$$Var(\widehat{\beta}) = \Omega_{\widehat{\beta}} = \sigma_{\widehat{\beta}}^2 = \sigma_\varepsilon^2 (X'X)^{-1} = \begin{pmatrix} \sigma_{\widehat{\beta}_1}^2 & Cov(\widehat{\beta}_1, \widehat{\beta}_2) & \ldots & Cov(\widehat{\beta}_1, \widehat{\beta}_k) \\ Cov(\widehat{\beta}_2, \widehat{\beta}_1) & \sigma_{\widehat{\beta}_2}^2 & \ldots & Cov(\widehat{\beta}_2, \widehat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\widehat{\beta}_k, \widehat{\beta}_1) & Cov(\widehat{\beta}_k, \widehat{\beta}_2) & \ldots & \sigma_{\widehat{\beta}_k}^2 \end{pmatrix}$$

c • Le théorème de Gauss-Markov : ***Si*** $\widetilde{\beta}$ ***est un autre estimateur linéaire sans biais de*** β ***,***

c.-à-d. si $E(\widetilde{\beta}) = \beta$ ***et*** $\widetilde{\beta} = AY$ ***, les variances de ses composantes ne peuvent être***

inférieures à celle des composantes de $\widehat{\beta}$ ***:*** $Var(\widetilde{\beta}_j) \geq Var(\widehat{\beta}_j)$ ***, pour*** $j = 1, 2, \ldots, k$

☞ *Démonstration :* ***Soit*** $\widetilde{\beta} = AY$ ***un autre estimateur de*** β ***.***

En posant $A = (X'X)^{-1}X' + C$ ***, on obtient :***

$$\widetilde{\beta} = [(X'X)^{-1}X' + C][X\beta + \mathcal{E}] = (X'X)^{-1}(X'X)\beta + (X'X)^{-1}X'\mathcal{E} + CX\beta + C\mathcal{E}$$

$$= \beta + CX\beta + (X'X)^{-1}X'\mathcal{E} + C\mathcal{E} = (I_k + CX)\beta + ((X'X)^{-1}X' + C)\mathcal{E} = (I_k + CX)\beta + A\mathcal{E}$$

$$(\widetilde{\beta} \text{ est un estimateur sans biais de } \beta) \Leftrightarrow E(\widetilde{\beta}) = \beta \Leftrightarrow (I_k + CX)\beta + A \underbrace{E(\mathcal{E})}_{\blacksquare\ \mathcal{H}_3:\ 0}$$

$$\Leftrightarrow (I_k + CX)\beta = \beta \Leftrightarrow CX = 0$$

En imposant $CX = 0$, on obtient $\widetilde{\beta} = \beta + A\mathcal{E}$

La matrice de covariance de $\widetilde{\beta}$ est alors :

$$Var(\widetilde{\beta}) = E\left[\left(\widetilde{\beta} - E(\widetilde{\beta})\right)\left(\widetilde{\beta} - E(\widetilde{\beta})\right)'\right] = E\left[(\widetilde{\beta} - \beta)(\widetilde{\beta} - \beta)'\right] = E[(A\mathcal{E})(A\mathcal{E})'] = E[A\mathcal{E}\mathcal{E}'A']$$

$$= A \underbrace{E(\mathcal{E}\mathcal{E}')}_{\blacksquare\ \mathcal{H}_4:\ \sigma_\varepsilon^2 I_n} A' = \sigma_\varepsilon^2 AA' = \sigma_\varepsilon^2[(X'X)^{-1}X' + C][(X'X)^{-1}X' + C]'$$

$$= \sigma_\varepsilon^2[(X'X)^{-1}X' + C][X(X'X)^{-1} + C']$$

$$= \sigma_\varepsilon^2[(X'X)^{-1}(X'X)(X'X)^{-1} + (X'X)^{-1}X'C' + CX(X'X)^{-1} + CC']$$

$$= \sigma_\varepsilon^2\left[(X'X)^{-1} + (X'X)^{-1}\underbrace{X'C'}_{0} + \underbrace{CX}_{0}(X'X)^{-1} + CC'\right] = \sigma_\varepsilon^2[(X'X)^{-1} + CC']$$

$$= \underbrace{\sigma_\varepsilon^2(X'X)^{-1}}_{Var(\widehat{\beta})} + \sigma_\varepsilon^2 CC'$$

$$Var(\widetilde{\beta}) - Var(\widehat{\beta}) = \sigma_\varepsilon^2 CC'$$

Mais les éléments de la diagonale de CC' sont des sommes de carrés, donc non négatives

Les variances des composantes de $\widetilde{\beta}$ sont donc supérieures ou égales aux variances des composantes de $\widehat{\beta}$

☑ ***Conséquences du théorème de Gauss-Markov :***

☞ ***1ère conséquence : Sous l'hypothèse de normalité des erreurs, non seulement que l'estimateur des MCO est BLUE par le théorème de Gauss-Markov, mais il devient le meilleur estimateur sans biais de β. La variance des estimateurs des MCO atteint la borne de l'inégalité de Cramer-Rao, borne inférieure pour tous les estimateurs.***

☞ ***2ème conséquence : Sous l'hypothèse de normalité, on obtient des tests exacts. Sachant que $\widehat{\beta} \sim \mathcal{N}(\beta, \sigma_\varepsilon^2(X'X)^{-1})$, cela revient à dire que l'on connaît les distributions exactes des tests. On peut donc construire les tests de Student et de Fisher dans les petits échantillons.***

$d \bullet$ Estimation de la variance des erreurs σ_ε^2 :

$$On\ a\ \hat{\mathcal{E}} = Y - X\hat{\beta} = X\beta + \mathcal{E} - X(X'X)^{-1}X'Y = X\beta + \mathcal{E} - X(X'X)^{-1}X'(X\beta + \mathcal{E})$$

$$= X\beta + \mathcal{E} - X\underbrace{(X'X)^{-1}(X'X)}_{I_k}\beta - X(X'X)^{-1}X'\mathcal{E} = X\beta + \mathcal{E} - X\beta - X(X'X)^{-1}X'\mathcal{E}$$

$$= \mathcal{E} - X(X'X)^{-1}X'\mathcal{E} = (I_n - X(X'X)^{-1}X')\mathcal{E} = M\mathcal{E}$$

On a vérifié que $M = I_n - X(X'X)^{-1}X'$ *est à la fois idempotente et symétrique*

Ainsi : $\hat{\mathcal{E}}'\hat{\mathcal{E}} = \mathcal{E}'M'M\mathcal{E} = \mathcal{E}'M\mathcal{E}$

$E(\hat{\mathcal{E}}'\hat{\mathcal{E}}) = E(\mathcal{E}'M\mathcal{E}) = E[Tr(\mathcal{E}'M\mathcal{E})]$, *car* $\mathcal{E}'M\mathcal{E}$ *est un scalaire*

$$= E[Tr(M\mathcal{E}\mathcal{E}')] = Tr[E(M\mathcal{E}\mathcal{E}')] = Tr\left[M\ \underbrace{E(\mathcal{E}\mathcal{E}')}_{\blacksquare\ \mathcal{H}_4:\ \sigma_\varepsilon^2 I_n}\right] = Tr(M\sigma_\varepsilon^2 I_n) = \sigma_\varepsilon^2 Tr(M)$$

$$= \sigma_\varepsilon^2 Tr(I_n - X(X'X)^{-1}X') = \sigma_\varepsilon^2[Tr(I_n) - Tr(X(X'X)^{-1}X')]$$

$$= \sigma_\varepsilon^2\left[n - Tr\left(\underbrace{(X'X)^{-1}(X'X)}_{I_k}\right)\right] = \sigma_\varepsilon^2[n - Tr(I_k)] = (n-k)\sigma_\varepsilon^2$$

D'où $\hat{\sigma}_\varepsilon^2 = \dfrac{\hat{\mathcal{E}}'\hat{\mathcal{E}}}{n-k} = \dfrac{SCR}{n-k}$ *est un estimateur sans biais de* σ_ε^2

Ainsi on obtient :

$$\widehat{Var}(\hat{\beta}) = \hat{\Omega}_{\hat{\beta}} = \hat{\sigma}_{\hat{\beta}}^2 = \hat{\sigma}_\varepsilon^2(X'X)^{-1} = \begin{pmatrix} \hat{\sigma}_{\hat{\beta}_1}^2 & \widehat{Cov}(\hat{\beta}_1,\hat{\beta}_2) & \dots & \widehat{Cov}(\hat{\beta}_1,\hat{\beta}_k) \\ \widehat{Cov}(\hat{\beta}_2,\hat{\beta}_1) & \hat{\sigma}_{\hat{\beta}_2}^2 & \dots & \widehat{Cov}(\hat{\beta}_2,\hat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{Cov}(\hat{\beta}_k,\hat{\beta}_1) & \widehat{Cov}(\hat{\beta}_k,\hat{\beta}_2) & \dots & \hat{\sigma}_{\hat{\beta}_k}^2 \end{pmatrix}$$

Lorsque le nombre d'observations tend vers l'infini, l'expression $\hat{\Omega}_{\hat{\beta}}$ *tend vers zéro.*

Par conséquent, l'estimateur $\hat{\beta}$ *est convergent . Toutefois, la condition suffisante serait que les variables exogènes ne tendent pas à devenir colinéaires lorsque n tend vers l'infini* $\blacksquare$ $\mathcal{H}_8$: $\lim\limits_{n\to+\infty}(X'X/n) = N$, *avec N une matrice finie non singulière*

$\mathcal{B}$– 4 • Régression avec les données centrées :

Supposons que la première colonne de la matrice X *soit composée de constantes* :

$$X = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1k} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{pmatrix}.$$

$Dans\ ce\ cas\ la\ régression\ multiple\ s'écrit$: $(M): y_i = \beta_1 + \sum_{j=2}^{k} \beta_j x_{ij} + \varepsilon_i = \widehat{\beta}_1 + \sum_{j=2}^{k} \widehat{\beta}_j x_{ij} + \hat{\varepsilon}_i$

$On\ peut\ aussi\ travailler\ avec\ les\ données\ centrées\ en\ sommant\ sur\ les\ i\ et\ en\ divisant$

$par\ n\ l'équation\ (M). On\ obtient$: $(\overline{M}): \overline{y} = \beta_1 + \sum_{j=2}^{k} \beta_j \overline{x}_j + \overline{\varepsilon} = \widehat{\beta}_1 + \sum_{j=2}^{k} \widehat{\beta}_j \overline{x}_j$

$et\ donc\ en\ soustrayant\ (\overline{M})$ à (M) , $on\ a\ finalement$:

$$(M_c): y_i - \overline{y} = \sum_{j=2}^{k} \beta_j (x_{ij} - \overline{x}_j) + (\varepsilon_i - \overline{\varepsilon}) = \sum_{j=2}^{k} \beta_j (x_{ij} - \overline{x}_j) + \hat{\varepsilon}_i \text{ , ou encore}$$

$$(M_c) : y_{i,c} = \sum_{j=2}^{k} \beta_j x_{ij,c} + \varepsilon_{i,c} = \sum_{j=2}^{k} \widehat{\beta}_j x_{ij,c} + \hat{\varepsilon}_i$$

$Définissons\ maintenant$:

- $\widetilde{\beta} = (\widehat{\beta}_2 \quad \cdots \quad \widehat{\beta}_k)'$, $de\ \mathbb{R}^{k-1}$ $composé\ des\ k-1\ dernières\ composantes\ de\ \widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_k \end{pmatrix}$
- $\widetilde{X} = \begin{pmatrix} x_{12} & \cdots & x_{1k} \\ \vdots & \cdots & \vdots \\ x_{n2} & \cdots & x_{nk} \end{pmatrix}$, $la\ matrice\ n \times (k-1) composé\ des\ k-1\ dernières\ colonnes\ de\ X$
- $1_n = (1 \quad \cdots \quad 1)'$, $le\ vecteur\ colonne\ de\ n$ "uns"
- $La\ matrice\ idempotente\ et\ symétrique\ qui\ centre\ les\ valeurs$:

$$P_c = I_n - \frac{1_n 1_n'}{n} = \begin{pmatrix} (1-1/n) & -1/n & \cdots & -1/n & -1/n \\ -1/n & (1-1/n) & \cdots & -1/n & -1/n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -1/n & -1/n & \cdots & (1-1/n) & -1/n \\ -1/n & -1/n & \cdots & -1/n & (1-1/n) \end{pmatrix}$$

- $Y_c = P_c Y = Y - 1_n \overline{Y} = \begin{pmatrix} y_1 - \overline{y} \\ \vdots \\ y_n - \overline{y} \end{pmatrix}$
- $X_c = P_c \widetilde{X}$, $la\ matrice\ \widetilde{X}\ centrée$: $X_c = \begin{pmatrix} (x_{12} - \overline{x}_2) & \cdots & (x_{1k} - \overline{x}_k) \\ \vdots & \cdots & \vdots \\ (x_{n2} - \overline{x}_2) & \cdots & (x_{nk} - \overline{x}_k) \end{pmatrix}$

$La\ régression\ multiple\ peut\ maintenant\ s'écrire$: $(M_c) : Y_c = X_c \widetilde{\beta} + \widehat{\mathcal{E}}$

$Le\ vecteur\ \widetilde{\beta}\ est\ évidemment\ définie\ par$: $\widetilde{\beta} = (X_c' X_c)^{-1} (X_c' Y_c) = (X_c' X_c / n)^{-1} (X_c' Y_c / n)$

Cette présentation est intéressante à plus d'un titre. En effet : $\frac{X'_cX_c}{n}$ *n'est autre que la matrice*

variance-covatiance $\Sigma = \frac{X'_cX_c}{n} = \begin{pmatrix} S_2^2 & S_{2,3} & \cdots & S_{2,k-1} & S_{2,k} \\ S_{2,3} & S_3^2 & \cdots & S_{3,k-1} & S_{3,k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{2,k-1} & S_{3,k-1} & \cdots & S_{k-1}^2 & S_{k-1,k-1} \\ S_{2,k} & S_{3,k} & \cdots & S_{k-1,k} & S_k^2 \end{pmatrix}$,

avec $S_j^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2$ *et* $S_{pq} = \frac{1}{n}\sum_{i=1}^{n}(x_{ip}-\overline{x}_p)(x_{i,q}-\overline{x}_q)$; j, p *et* $q = 2, 3, \dots, k$

Et $\frac{X'_cY_c}{n}$ *est le vecteur des covariances entre les variables explicatives et la variable*

dépendante: $\frac{X'_cY_c}{n} = (S_{2,y} \quad S_{3,y} \quad \cdots \quad S_{k-1,y} \quad S_{k,y})'$ *avec* $S_{jy} = \frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)(y_i-\overline{y})$; $j = 2, \dots, k$

B– 5 • Décomposition de la variance-coefficients de détermination :

a • Lemme : $\widehat{\mathcal{E}}'\widehat{\mathcal{E}} = Y'Y - \widehat{\beta}'X'Y$

☞ *Démonstration :*

$\widehat{\mathcal{E}}'\widehat{\mathcal{E}} = (Y - X\widehat{\beta})'(Y - X\widehat{\beta})Y'Y - 2\widehat{\beta}'X'Y + \widehat{\beta}'(X'X)\widehat{\beta}$, *or* $(X'X)\widehat{\beta} = X'Y$ *d'où* $\widehat{\mathcal{E}}'\widehat{\mathcal{E}} = Y'Y - \widehat{\beta}'X'Y$

b • L'équation d'analyse de la variance :

- $SCT = \sum_{i=1}^{n}(y_i-\overline{y})^2 = nS_y^2$: *est la somme des carrés totaux. Elle indique la variabilité totale de Y.*

- $SCE = \sum_{i=1}^{n}(\widehat{y}_i-\overline{\widehat{y}})^2 = \sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2$: *est la somme des carrés expliqués. Elle indique la variation de Y due à sa régression linéaire sur* $x_{i2}, x_{i3}, \dots, x_{ik}$.

- $SCR = \sum_{i=1}^{n}(\hat{\varepsilon}_i-\overline{\hat{\varepsilon}})^2 = \sum_{i=1}^{n}\hat{\varepsilon}_i^2$: *est la somme des carrés résiduels. Elle indique la variabilité de Y non expliquée par le modèle.*

On a : $SCT = \sum_{i=1}^{n}(y_i-\overline{y})^2 = \sum_{i=1}^{n}y_i^2 - n\overline{y}^2 = Y'Y - \frac{(1'_nY)^2}{n} = Y'_cY_c$

$$D'autre\ part\ :\ SCE = \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2 = (X\hat{\beta})'(X\hat{\beta}) - \frac{(1'_n X\hat{\beta})^2}{n}$$

$$= \hat{\beta}'(X'X)\hat{\beta} - \frac{(1'_n Y)^2}{n}, puisque\ 1'_n Y = 1'_n(X\hat{\beta} + \hat{\varepsilon}) = 1'_n X\hat{\beta} + \underbrace{1'_n\hat{\varepsilon}}_{0} = 1'_n X\hat{\beta}$$

$$= \hat{\beta}'(X'X)(X'X)^{-1}X'Y - \frac{(1'_n Y)^2}{n}$$

$$SCE = \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \hat{\beta}'X'Y - \frac{(1'_n Y)^2}{n} = \hat{Y}'_c\hat{Y}_c = \tilde{\beta}'X'_cY_c = \tilde{\beta}'(X'X)\tilde{\beta}$$

$$SCR = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \hat{\varepsilon}'\hat{\varepsilon} = Y'Y - \hat{\beta}'X'Y = \left(SCT + \frac{(1'_n Y)^2}{n}\right) - \left(SCE + \frac{(1'_n Y)^2}{n}\right) = SCT - SCE$$

$$D'où \underbrace{\sum_{i=1}^{n} (y_i - \bar{y})^2}_{SCT} = \underbrace{\sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2}_{SCE} + \underbrace{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}_{SCR}\ ou\ encore\ SCT = SCE + SCR\ (seulement\ pour\ \beta_1 \neq 0)$$

$$SCR = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \hat{\varepsilon}'\hat{\varepsilon} = Y'Y - \hat{\beta}'X'Y = Y'_cY_c - \hat{Y}'_c\hat{Y}_c$$

c • Tableau d'analyse de variance :

Source de la variabilité	Somme des carrés	Degré de liberté	Carrés moyens
Régression	SCE	$k-1$	$SCE/k-1$
Résidus	SCR	$n-k$	$\hat{\sigma}^2 = SCR/n-k$
Total	SCT	$n-1$	$S^2 = SCT/n-1$

d • Coefficient de détermination (dans un modèle avec terme constant β_1) :

La part de variance de Y expliquée par le modèle est toujours traduite par le coefficient de détermination.

Connaissant l'équation d'analyse de la variance : $SCT = SCE + SCR$, *le* R^2 *correspond au rapport :*

$$R^2 = \frac{SCE}{SCT} = \frac{\hat{Y}'_c\hat{Y}_c}{Y'_cY_c} = 1 - \frac{SCR}{SCT} = 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{Y'_cY_c} \in [0,1]\ (seulement\ pour\ \beta_1 \neq 0)$$

Il faut noter que, comme pour le modèle simple, le coefficient de détermination reste un indicateur du caractère explicatif de l'équation de régression à bien modéliser Y. Il mesure ainsi la part de variance de la variable endogène attribuable à sa régression sur les X. Ceci est confirmé par le fait que le coefficient de détermination n'est rien d'autre que le carré du coefficient de corrélation entre les valeurs observées et les *valeurs prédites de Y* : $R^2 = \frac{Cov(\widehat{y}_i, y_i)}{\sqrt{Var(\widehat{y}_i)Var(y_i)}} = \rho^2_{\widehat{Y},Y}$

☑ **Remarque** : *Le coefficient de corrélation linéaire entre $\widehat{Y}$ et Y $(\rho_{\widehat{Y},Y})$ est appelé coefficient de corrélation multiple. Cela suggère d'ailleurs de construire le graphique nuage de points confrontant $\widehat{Y}$ et Y pour évaluer la qualité de la régression. Si le modèle est parfait, les points seraient parfaitement alignés.*

Bien évidement $R^2 \in [0, 1]$, plus R^2 est proche de 1, plus le caractère explicatif du modèle est important.

Le R^2 est certes un indicateur de qualité, mais il présente l'inconvénient d'être mécanique c-à-d. que sa valeur augmente avec l'augmentation des variables explicatives, mêmes non pertinentes à l'explication du phénomène étudié.

A l'extrême, si on augmente le nombre de variables explicatives, mêmes impertinentes, tels que le nombre de paramètres devienne égal au nombre d'observations, on aurait un $R^2 = 1$. Ainsi, en tant que tel, le R^2 n'est pas l'outil approprié pour juger de l'apport des des variables supplémentaires lors de la comparaison de plusieurs modèles. Lorsqu'il augmente de manière mécanique, de l'autre côté l'on perd en degrés de liberté.

e • Coefficient de détermination corrigé ou ajusté: *Le R^2 est un indicateur de qualité, mais il présente un défaut ennuyeux : plus nous augmentons le nombre de variables explicatives, même non pertinentes, n'ayant aucun rapport avec le problème que l'on cherche à résoudre, plus grande sera sa valeur.*

La mesure alternative, plus robuste à l'ajout des variables, qui corrige ce problème associé aux degrés de liberté est le R^2-ajusté (ou R^2-corrigé). Elle se définit comme suit :

$$\overline{R}^2 = 1 - \frac{SCR/n-k}{SCT/n-1} = 1 - \left(\frac{n-1}{n-k}\right)(1 - R^2) \text{ , avec } \quad \overline{R}^2 < R^2 \text{ et } \lim_{n\to+\infty}(\overline{R}^2) = R^2$$

☑ **Remarque** : *Il faut faire attention de ne pas interpréter le $\overline{R}^2$ en termes de part de variance expliquée. Son seul avantage est qu'il permet de comparer plusieurs modèles. De plus, le $\overline{R}^2$ peut prendre des valeurs négatives. Dans ce dernier cas, il faut l'assimiler à zéro.*

$\oint$ • ***Coefficient de détermination partielle**:* *Le coefficient de détermination partielle associé à une variable quelconque est la part de la variance laissée inexpliquée par les autres variables qui est expliquée grâce à l'introduction de la variable étudiée.*

Il mesure l'influence de X_j sur Y si on élimine l'effet des autres variables explicatives.

- *Une première régression en l'absence de la variable étudiée X_j :*

$$SCT_1 = SCE_1 + SCR_1 \;\; (M_1)$$

- *Une seconde régression en prenant toutes les variables, la variable X_j comprise :*

$$SCT_2 = SCE_2 + SCR_2 \;\; (M_2)$$

- SCR_1 : *Ce qui n'est pas expliqué par la première régression (en l'absence de la variable explicative X_j)*

- $SCE_2 - SCE_1$: *Le supplément d'explication apporté par X_j*

Le coefficient de détermination partielle associé à la variable X_j est défini par la formule suivante: $r^2_{Y,X_j} = \frac{SCE_2 - SCE_1}{SCR_1} = \frac{SCR_1 - SCR_2}{SCR_1} = \frac{R_2^2 - R_1^2}{1 - R_1^2} = \frac{t^2_{\hat{\beta}_j}}{t^2_{\hat{\beta}_j} + (n-k)}$ où $t_{\hat{\beta}_j} = \frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}$

$$r^2_{Y,X_j} \in [0,1] : \begin{cases} si\;\; r^2_{Y,X_j} = 0 \;\Rightarrow \text{la variable } X_j \text{ n'apporte rien de plus à l'explication de } Y \\ si\;\; r^2_{Y,X_j} = 1 \;\Rightarrow \text{la variable } X_j \text{ explique tout ce qui restait à expliqué} \end{cases}$$

La notion de corrélation partielle est importante dans la mesure où elle permet de juger de la pertinence d'introduire une variable exogène dans le modèle. Plus élevé sera le coefficient de corrélation partielle d'une variable, plus importante sera sa contribution à l'explication globale du modèle.

Coefficient de corrélation partielle du premier ordre	Coefficient de corrélation partielle du deuxième ordre
Soit Y une variable endogène, et X_i, X_j, X_k des exogènes, le coefficient de corrélation partielle mesure le lien entre Y et chaque X, l'influence juste d'une troisième variable exogène étant exclue. Partant de l'exemple choisi, on peut calculer ainsi six coefficients de corrélation partielle du premier ordre : $r_{YX_i \to X_j}$; $r_{YX_i \to X_k}$; $r_{YX_j \to X_i}$; $r_{YX_j \to X_k}$; $r_{YX_k \to X_i}$; $r_{YX_k \to X_j}$ *Dans ce cas, le coefficient de corrélation partielle du premier ordre peut être calculé à partir des coefficients de corrélation comme suit :*	*En considérant l'exemple ci-contre, le coefficient de corrélation partielle du deuxième ordre sert à quantifier le lien entre Y et chaque X, l'influence de deux autres étant exclue. Ainsi, partant du même exemple, il est possible de calculer trois coefficients de corrélation partiels du deuxième ordre, soit :* $r_{YX_i \to X_jX_k}$; $r_{YX_j \to X_iX_k}$; $r_{YX_k \to X_iX_j}$
$r_{YX_i \to X_j} = \dfrac{r_{YX_i} - \left(r_{YX_j} r_{X_iX_j}\right)}{\sqrt{\left(1 - r^2_{YX_j}\right)\left(1 - r^2_{X_iX_j}\right)}}$	*Etapes de calcul de $r_{YX_i \to X_jX_k}$:* *Supposons que l'on veut mesurer le lien entre Y et X_i, l'influence de X_j et de X_k étant neutralisée, soit*
$r_{YX_i \to X_k} = \dfrac{r_{YX_i} - \left(r_{YX_k} r_{X_iX_k}\right)}{\sqrt{\left(1 - r^2_{YX_k}\right)\left(1 - r^2_{X_iX_k}\right)}}$	1 *Calcul des résidus $\mathcal{E}_1$ issus de la régression de Y sur X_j et X_k*
$r_{YX_j \to X_i} = \dfrac{r_{YX_j} - \left(r_{YX_i} r_{X_jX_i}\right)}{\sqrt{\left(1 - r^2_{YX_i}\right)\left(1 - r^2_{X_jX_i}\right)}}$	2 *Calcul des résidus $\mathcal{E}_2$ issus de la régression de X_i sur X_j et X_k*
$r_{YX_j \to X_k} = \dfrac{r_{YX_j} - \left(r_{YX_k} r_{X_jX_k}\right)}{\sqrt{\left(1 - r^2_{YX_k}\right)\left(1 - r^2_{X_jX_k}\right)}}$	3 *$r_{YX_i \to X_jX_k}$ correspondra au carré du coefficient de corrélation linéaire calculé entre $\mathcal{E}_1$ et $\mathcal{E}_2$:* $r_{YX_i \to X_jX_k} = r^2_{\mathcal{E}_1\mathcal{E}_2}$
$r_{YX_k \to X_i} = \dfrac{r_{YX_k} - \left(r_{YX_i} r_{X_kX_i}\right)}{\sqrt{\left(1 - r^2_{YX_i}\right)\left(1 - r^2_{X_kX_i}\right)}}$	
$r_{YX_k \to X_j} = \dfrac{r_{YX_k} - \left(r_{YX_j} r_{X_kX_j}\right)}{\sqrt{\left(1 - r^2_{YX_j}\right)\left(1 - r^2_{X_kX_j}\right)}}$	

*g • **Modèle sans terme constant(β_1)** : La somme des résidus n'est pas nécessairement nulle, en effet, on a peut avoir* $\sum_{i=1}^{n} \hat{\varepsilon}_i \neq 0$ *, comme on peut avoir* $SCT \neq SCE + SCR$ *ou encore* $R^2 \notin [0,1]$

Néanmoins, on a toujours, en vertu du lemme : $Y'Y = \hat{\mathcal{E}}'\hat{\mathcal{E}} + \hat{\beta}'X'Y = \hat{\mathcal{E}}'\hat{\mathcal{E}} + \hat{Y}'\hat{Y}$

On peut définir $R^2_* = \dfrac{\hat{Y}'\hat{Y}}{Y'Y} = 1 - \dfrac{\hat{\mathcal{E}}'\hat{\mathcal{E}}}{Y'Y} \in [0,1]$

Ce coefficient R_*^2 peut être utilisé dans tous les cas, tant dans un modèle sans constante que dans un modèle avec constante.

Comme précédemment, nous pouvons ajuster ce dernier coefficient de détermination aux nombres de degrés de liberté, comme suit : $\overline{R}_*^2 = 1 - \dfrac{\hat{\varepsilon}'\hat{\varepsilon}/(n-k)}{Y'Y/(n-1)} = \left(\dfrac{n-1}{n-k}\right)R_*^2 - \left(\dfrac{k-1}{n-k}\right)$

☑ ***Interprétation des coefficients de détermination :*** *R^2 est une fonction monotone de la statistique de Fisher à employer pour tester la nullité de tous les coefficients de la régression sauf la constante.*

Nous verrons aussi que R_*^2 est une fonction monotone de l'autre statistique de Fisher à employer pour tester la nullité de tous les coefficients, constante comprise

B-6 • Problèmes particuliers : Multicolinéarité, Variables muettes, Variables Proxy :

a • Multicolinéarité :

■ *L'existence d'une relation linéaire exacte entre les colonnes de X nous empêche de déterminer l'estimateur $\widehat{\beta}$ de manière unique. Ce cas est un cas extrême de multicolinéarité. Mais il arrive souvent que certaines des colonnes de X présentent une dépendance linéaire approximative. Les conséquences de ce problème sont les suivantes:*

☞ *Un manque de précision dans les estimateurs des β_j, se traduisant par de fortes variances*

☞ *Les estimateurs des β_j présenteront souvent des distorsions importantes, dues à des raisons numériques. Le nombre de chiffres significatifs des emplacements mémoire d'un ordinateur est en effet limité, ce qui se traduit par un manque de stabilité des programmes d'inversion matricielle, pour des matrices qui sont presque singulières.*

Pour illustrer le premier point, prenons le modèle de régression simple :

$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$. *Nous avons vu que :* $Var(\widehat{\beta}_2) = \dfrac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \overline{x})^2}$.

La multicolinéarité se traduira dans ce cas par une série d'observations (x_i) presque *constante, c-à-d.* $x_i \approx \overline{x}$ *pour tout i . On a alors* $\sum_{i=1}^{n}(x_i - \overline{x})^2 \approx 0$, *ce qui se traduit par une forte variance de* $\widehat{\beta}_2$

- *La multicolinéarité peut être mesuré en calculant le rapport* $\frac{\lambda_{max}}{\lambda_{min}}$ *de la plus grande à la plus petite valeur propre de* $X'X$
- *Pour corriger le problème de multicolinéarité, on peut :*

☞ *Soit ajouter des observations à l'échantillonquand la chose est possible ; il faut néanmoins que les observations supplémentaires ne présentent pas de multicolinéarité*

☞ *Soit introduire une information a priori . Supposons par exemple que dans la fonction de production :* $\ln(Q_t) = A + \alpha \ln(K_t) + \beta \ln(L_t) + \varepsilon_t$

Les variables $\ln(K_t)$ *et* $\ln(L_t)$ *sont fortement colinéaires. Si l'on sait que les rendement d'échelleconstant* $(\alpha + \beta = 1)$, *on peut transformer le modèle comme suit :*

$\ln(Q_t) = A + \alpha \ln(K_t) + (1 - \alpha) \ln(L_t) + \varepsilon_t$ *ou* $[\ln(Q_t) - \ln(L_t)] = A + \alpha[\ln(K_t) - \ln(L_t)] + \varepsilon_t$

Ce qui a donc pour effet de supprimer un régresseur. Ceci peut résoudre le problème . Essentiellement , l'informationa priori $\alpha + \beta = 1$ *supplée au défaut d'information présent dans l'échantillon (tentative d'estimer trop de paramètres avec trop peut de données)*

• Variables muettes :

- ***Synonyme :*** *Variables indicatrices, qualitatives , binaires , dummy , dichotomiques, auxiliaires, artificielles*

Une variable indicatrice est une variable spéciale qui ne prend que deux valeurs, à savoir : $\begin{cases} 1 \text{ , pour indiquer que le phénomène (ou l'événement) a lieu} \\ 0 \text{ , pour indiquer que le phénomène (ou l'événement) n'a pas lieu} \end{cases}$

Elle est utilisée en économétrie pour saisir les facteurs qualitatifs- comme la race, le sexe, la religion ou même un événement tel qu'une guerre, une grève, un tsunami, etc. - que l'on désire intégrer dans les modèles. Comme variable explicative, on la note généralement par la lettre D, pour dire dummy.

Il est également important de noter que les variables binaires peuvent intervenir dans le modèle de deux manières, soit comme endogène (modèle de probabilité linéaire, modèles Logit, Probit, Tobit, Gombit) soit comme exogène (modèles ANOVA et ANCOVA). Nous ne nous intéressons qu'au cas où la variable muette entre comme explicative dans le modèle.

Aussi, l'utilisation de ces variables dépend fortement du problème posé. Comme exogènes, les variables dummy sont utilisées pour répondre à un triple objectif :

- *Corriger les écarts aberrants (ou déviants)*
- *Capter la présence de la discrimination*
- *Capter les variations saisonnières*

❶ Corriger les valeurs singulières (ou anormales) :

Lorsque la variable endogène comporte, à certaines dates, des valeurs atypiques c-à-d. des valeurs anormalement élevées ou anormalement basses -associées en général à la survenance de chocs ou d'événement rares, il y a lieu d'incorporer une dummy dans le modèle afin d'en tenir compte. La démarche consisterait simplement à détecter les valeurs anormales et à les corriger, en mettant 1 à ces dates là et 0 ailleurs, afin que les déviants ne perturbent pas l'estimation statistique des autres variables.

☑ **Remarques :**

- *La correction effectuée n'est valable que si le coefficient associé à la variable dummy est statistiquement significatif.*
- *Après estimation, le signe affecté à la variable binaire est proportionnelle à l'anomalie constatée dans les données. S'il s'agit d'une observation anormalement basse, le signe affecté à la dummy sera (−), ce qui indique que l'écart criant avait tendance à ramener la droite de régression vers le bas. En revanche, s'il est plutôt question d'une observation anormalement élevée, le signe affecté à la dummy sera (+), ce qui indique que le déviant avait tendance à tirer la droite de régression vers le haut.*
- *Attention à ne pas saisir les écarts anormalement élevés et anormalement bas par une même une variable muette. Lorsque la série présente à la fois les deux types*

d'écarts, il convient de les capter par deux variables auxiliaires différentes, l'une pour les observations exceptionnellement élevées et l'autre pour celles exceptionnellement basses.

❷ Capter la présence de la discrimination :

L'explication d'un phénomène peut parfois nécessiter la présence des variables qualitatives. Supposons que l'on souhaite expliquer, pour dix étudiants de première licence en Economie échantillonnés, le phénomène (note obtenue en macroéconomie : NM_i); tout naturellement les variables comme présence au cours PC_i, nombre d'heures d'étude consacrées à la macroéconomie HE_i … s'avèrent pertinentes. Mais il est tout à fait aussi possible que des variables comme la religion de l'étudiant RE_i, ou sa tribu TE_i, soient déterminantes dans l'explication du phénomène étudié. Dans ce cas, l'utilisation d'une variable binaire permet de segmenter les individus en deux groupes et de déterminer si le critère de segmentation est réellement discriminant.

Dans notre exemple la note obtenue en macroéconomie, si l'on assume que l'appartenance ou non à la religion catholique est déterminante dans la réussite, ce qui revient à dire que la religion est un facteur de discrimination, le modèle à estimer sera :

$$NM_i = \beta_0 + \beta_1 PC_i + \beta_2 HE_i + \beta_3 RE_i + \varepsilon_i \text{ ; où } RE_i = \begin{cases} 1 \text{ , si l'étudiantest catholique} \\ 0 \text{ , si autre religion} \end{cases}$$

Puis estimer, on applique les MCO. Après estimation, si β_3 est statistiquement significatif, on en conclurait que la religion (catholique) a joué sur la note en macroéconomie, elle est donc bien un facteur discriminant de la note obtenue en macroéconomie. A l'opposé, si β_3 est statistiquement non significatif, on en conclurait que la religion (catholique) n'a pas joué sur la réussite en macroéconomie.

☑ Remarques :

- *Dans le cas de variables dummy à plusieurs modalités, par exemple l'état civil (célibataire, marié, divorcé, autres), il est convenable de coder alors autant de variables indicatrices qu'il y a de modalités moins une. Ainsi, pour l'état civil, on* *définira trois variables binaires :* $Célibataire_i = \begin{cases} 1\text{ , si l'individu est célibataire} \\ 0\text{ , si non} \end{cases}$

$Marié_i = \begin{cases} 1\text{ , si l'individu est marié} \\ 0\text{ , si non} \end{cases}$; $Divorcé_i = \begin{cases} 1\text{ , si l'individu est divorcé} \\ 0\text{ , si non} \end{cases}$

, la modalité autres étant implicitement contenue dans le terme constant et ne serait donc spécifiée à part que dans un modèle sans terme constant.

■ *La codification dépend du modélisateur et doit être prise en compte dans l'interprétation des résultats. A titre exemplatif, si l'on considère la variable qualitative sexe, le modélisateur est libre de coder 1; si femme et 0; si homme et inversement. Il doit seulement en tenir compte lors de l'interprétation.*

❸ *Capter les variations saisonnières :*

Les variables indicatrices sont aussi utilisées pour prendre en compte les mouvements saisonniers qui caractérisent certaines variables comme les dépenses de publicité, qui sont généralement plus importantes en certaines périodes de l'année qu'en d'autres.
Supposons que l'on s'intéresse à la relation entre le chiffre d'affaires (Ch_t) *et les dépenses de publicité* ($Dpub_t$). *On peut écrire :* $(M_1) : Ch_t = \beta_0 + \beta_1 Dpub_t + \varepsilon_t$
En utilisant les données trimestrielles, il ne serait pas correct d'estimer directement le modèle (M_1) *, parce qu'on n'aurait pas tenu compte de l'effet saisonnier, les dépenses de publicité ne sont pas les mêmes tous les trois mois (trimestre).*
On peut capter l'effet saisonnier en introduisant dans (M_1) *une variable dummy. Pour notre cas, on aura autant de variables dummy qu'il y a de trimestres, soit quatre dummy. Sachant qu'on compte quatre trimestres par année, l'introduction des variables dummy se fera comme suit :*

	Trimestre	D_{1t}	D_{2t}	D_{3t}	D_{4t}	$\sum$
2005	1er trimestre	1	0	0	0	1
	2ème trimestre	0	1	0	0	1
	3ème trimestre	0	0	1	0	1
	4ème trimestre	0	0	0	1	1
2006	1er trimestre	1	0	0	0	1
	2ème trimestre	0	1	0	0	1
	3ème trimestre	0	0	1	0	1
	4ème trimestre	0	0	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Et le modèle (M_1) *devient* $Ch_t = \beta_0 + \beta_1 Dpub_t + \beta_2 D_{1t} + \beta_3 D_{2t} + \beta_4 D_{3t} + \varepsilon_t \quad (M_2)$

ou encore $Ch_t = \beta_1 Dpub_t + \beta_2 D_{1t} + \beta_3 D_{2t} + \beta_4 D_{3t} + \beta_5 D_{4t} + \varepsilon_t \quad (M_3)$

Si le modèle contient un terme constant, celui-ci joue d'office le rôle de l'une des quatre variables dummy. Dans (M_2) *par exemple,* β_0 *joue le rôle de* D_{4t} *(on a le choix pour la variable binaire à écarter). En revanche, en absence du terme constant, il convient de prendre en compte, comme dans la relation* (M_3) *de toutes les variables dummy.*

Une fois cette gymnastique terminée, on peut alors, sans difficulté normalement, appliquer les MCO soit sur le modèle (M_2)*, soit sur le modèle* (M_3).

c • Variables Proxy :

■ **Présentation du problème :**

☑ **Estimer le rendement de l'éducation :**

Soit la fonction mincérienne de salaire : $ln(W_t) = \beta_0 + \beta_1 EDUC_t + \beta_2 EXPER_t + u_t$

Avec : W*: salaire,* $EDUC$*: nombre d'années d'études et* $EXPER$*: nombre d'années d'expérience.*

Les capacités intellectuelles intrinsèques (CII) de l'individu (ability) sont :

- *Inobservées*
- *Positivement corrélées avec le niveau d'éducation :* $EDUC_t = \delta_1 CII_t + v_t$
- *Positivement corrélées avec le niveau de salaire :* $ln(W_t) = \delta_2 CII_t + w_t$

Le rendement de l'éducation estimé par : $ln(W_t) = \beta_0 + \beta_1 EDUC_t + \beta_2 EXPER_t + u_t$

La question qui se pose : Est-t-il surestimé ou sous-estimé ?

■ **Étendue du problème :**

Ce problème est inhérent à toute analyse économétrique :

- *Certaines variables sont par nature inobservables ou non mesurables (Dynamisme, charisme, capital social d'un individu, esprit d'équipe dans une entreprise, etc.)*
- *Certaines variables sont non disponibles pour l'économètre (questions non posées, réponses biaisées surdes sujets sensibles, etc.)*
- *Il faut juste rester conscient que certaines variables peuvent capter des effets plus larges que ce pourquoi elles sont incluses dans le modèle*

■ **Formalisation :**

• *Lorsque l'on est face au problème de variable omise, il est possible d'utiliser des variables « Proxy » qui permettent de corriger les biais de spécification*

• *Une variable proxy est une variable corrélée à la variable observable mais non directement explicative du modèle étudié*

• *La variable proxy est substituée à la variable inobservée et le signe du coefficient estimé est interprété en fonction de la relation théorique reliant variable proxy et variable inobservée*

Reprenons le modèle général : $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2}^* + u_t \quad (1)$

x_2^* *inobservable, corrélée avec* x_1*, avec* $x_{t2}^* = \delta_0 + \delta_1 x_{t2} + v_t \quad (2)$ *où* x_2 *est observable*

Sous quelles conditions, la variable proxy x_2 *permet une estimation sans biais et convergente de* β_1 *par* $y_t = \pi_0 + \beta_1 x_{t1} + \pi_2 x_{t2} + w_t$?

Les équations (1) et (2) impliquent :

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2(\delta_0 + \delta_1 x_{t2} + v_t) + u_t = (\beta_0 + \beta_2\delta_0) + \beta_1 x_{t1} + \beta_2\delta_1 x_{t2} + (\beta_2 v_t + u_t)$$

$$= \pi_0 + \beta_1 x_{t1} + \pi_2 x_{t2} + w_t$$

w_t *ne doit pas être corrélé avec les variables explicatives du modèle, donc* u_t *et* v_t *doivent être non corrélés avec* x_1 *et* x_2

• **NB** : *Les valeurs des coefficients estimés pour le terme constant et la proxy ne sont pas directement interprétables.*

Si en revanche on a, $x_{t2}^* = \delta_0 + \delta_1 x_{t1} + \delta_2 x_{t2} + v_t$ *avec* w_t *non-corrélé à* x_1 *et* x_2*, alors :*

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2(\delta_0 + \delta_1 x_{t1} + \delta_2 x_{t2} + v_t) + u_t$$

$$= (\beta_0 + \beta_2\delta_0) + (\beta_1 + \beta_2\delta_1)x_{t1} + \beta_2\delta_2 x_{t2} + (\beta_2 v_t + u_t)$$

• *Le signe et la valeur du biais dépendent des valeurs des paramètres* β_2 *et* δ_1

• *Le biais reste en général inférieur au biais initial (variable omise)*

■ **Biais d'omission :**

Supposons deux modèles : $\begin{cases} y_i = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\varepsilon}_i & (\text{vrai modèle}) \\ y_i = \tilde{\beta}_1 + \tilde{\beta}_2 x_{2i} + \tilde{\varepsilon}_i & (\text{modèle sous-dimensionné}) \end{cases}$

• *Les conséquences de l'omission de* x_3 *sont*

☞ *L'omission d'une variable explicative importante conduit à un biais d'omission*

☞ *L'importance du biais dépend de la dépendance entre la variable omise et les variables explicatives incluses dans la régression*

☞ *Si la variable omise x_3 est corrélée avec la variable x_2, c.-à-d., si r_{23}, le coefficient de corrélation entre les deux variables n'est pas nul $\tilde{\beta}_1$ et $\tilde{\beta}_2$ sont biaisées et incohérents. Autrement dit $E(\tilde{\beta}_1) \neq \beta_1$ et $E(\tilde{\beta}_2) \neq \beta_2$. Le biais ne disparaît pas lorsque la taille de l'échantillon s'accroît*

☞ *Même si x_2 et x_3 ne sont pas corrélés, $\tilde{\beta}_1$ est biaisé, bien que $\tilde{\beta}_2$ ne le soit plus*

☞ *La variance des erreurs σ_ε^2 est incorrectement estimée*

☞ *La mesure habituelle de la variance de $\tilde{\beta}_2$ est un estimateur biaisé de la variance du véritable estimateur $\hat{\beta}_2$*

☞ *En conséquence, l'intervalle de confiance standard et les procédés de test d'hypothèse fourniront probablement des conclusions erronées sur la signification statistique des paramètres estimés*

☞ *Une autre conséquence est que les prévisions basées sur le modèle incorrect et les intervalles de prévision ne seront pas fiable*

- *Le biais est d'autant négligeable que*

☞ *L'effet partiel de x_3 sur y est négligeable*

☞ *Les variables x_2 et x_3 sont faiblement corrélées*

☞ *La variance de x_2 est élevée*

- *Le biais du paramètre est donné par :* $E(\tilde{\beta}_2) - \beta_2 = \beta_3 \dfrac{S_{x_2,x_3}}{S_{x_2}^2}$

	$S_{x_2,x_3} > 0$	$S_{x_2,x_3} < 0$
$\beta_3 > 0$	*Biais positif*	*Biais négatif*
$\beta_3 < 0$	*Biais négatif*	*Biais positif*

- *Erreur quadratique moyenne :* $\dfrac{EQM(\tilde{\beta}_2)}{EQM(\hat{\beta}_2)} = 1 + r_{23}^2(t_3^2 - 1)$

où r_{23}^2 *le coefficient de corrélation entre les deux variables* x_2 *et* x_3 *et* $t_3 = \dfrac{\beta_3}{\hat{\sigma}_{\hat{\beta}_3}}$

BEN AHMED MOHSEN Téléphone: (+216) 97 619191 / 54 619191
omega.center.cp@gmail.com https://web.facebook.com/OMEGACENTER2014

A partir de ce rapport on peut conclure que $\tilde{\beta}_2$ *a une erreur quadratique i, férieure à celle de* $\hat{\beta}_2$ *si* $|t_3| < 1$

Comme t_3 *est inconnue, on doit l'estimer à partir de l'équation complète*

Comme estimateur de β_2 *, on peut utiliser l'estimateur défini par :*

$$\breve{\beta}_2 = \begin{cases} \hat{\beta}_2 : \text{estimateur des MCO de l'équation du vrai modèle , n si } |\hat{t}_3| \geq 1 \\ \tilde{\beta}_2 : \text{estimateur de la variable omise , si } |\hat{t}_3| < 1 \end{cases}$$

☞ **Remarque :** *Au lieu d'utiliser* $\hat{\beta}_2$ *et* $\tilde{\beta}_2$ *en fonction de* $\hat{t}_3$ *on peut , on peut considérer une combinaison linéaire des deux,* $\lambda\hat{\beta}_2 + (1-\lambda)\tilde{\beta}_2$ *appelé estimateur moyen pondéré et il y a une erreur quadratique moyenne* $\lambda = \dfrac{t_3^2}{1+t_3^2}$

Moindres carrées sous contraintes linéaires

C-1 • L'estimateur de β sous contraintes :

On s'intéresse maintenant à l'estimateur $\hat{\beta}_c$ *du vecteur* β *sous un système de* J *contraintes indépendantes, qui peut s'écrire sous la forme :* $\underbrace{R}_{J\times k}\underbrace{\hat{\beta}_c}_{k\times 1} = \underbrace{r}_{J\times 1}$

R *étant une matrice* $J \times k$ *de rang* J *,* r *est un vecteur* $J \times 1$ *et* $\hat{\beta}_c$ *est le vecteur des estimateurs de* β *sous contraintes.*

Nous minimisons la somme des carrées des résidus sous les contraintes du système $R\hat{\beta}_c = r$*. On écrira ce système sous la forme* $2(R\hat{\beta}_c - r) = 0$ *et on formera le Lagrangien :*

$$\phi = (Y - X\hat{\beta}_c)'(Y - X\hat{\beta}_c) - 2\lambda'(R\hat{\beta}_c - r).$$

Où λ' *est un vecteur ligne de* J *multiplicateurs de Lagrange.*

■ **Condition de premier ordre :**

$$\begin{cases} \dfrac{\partial\phi}{\partial\hat{\beta}_c} = -2X'Y + 2(X'X)\hat{\beta}_c - 2R'\lambda = 0 & (1) \\ \dfrac{\partial\phi}{\partial\lambda} = -2(R\hat{\beta}_c - r) = 0 & (2) \end{cases}$$

$(1) \Rightarrow \hat{\beta}_c = \hat{\beta} + (X'X)^{-1}R'\lambda$ (3) ; où $\hat{\beta} = (X'X)^{-1}(X'Y)$ *l'estimateur sans contraintes*

$En\ multipliant\ par\ R : R\widehat{\beta}_c = R\widehat{\beta} + R(X'X)^{-1}R'\lambda$, $or\ (2) \Rightarrow R\widehat{\beta}_c = r$

$ceci\ implque : \lambda = [R(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$

D'où $\widehat{\beta}_c = \widehat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$ (4)

$On\ constate\ que\ \widehat{\beta}_c$ ($l'estimateur\ contraint$) $diff$ère $de\ \widehat{\beta}$ ($l'estimateur\ non\ contraint$)

$par\ une\ fonction\ lin$éaire $du\ vecteur\ r - R\widehat{\beta}$. $Ce\ dernier\ sera\ nul\ si\ le\ vecteur\ \widehat{\beta}\ v$érifie

$les\ restrictions\ a\ priori$

C– 2 • $Efficacit$é $de\ l'estimateur\ de\ \beta\ sous\ contraintes$:

$On\ se\ propose\ de\ d$émontrer $que\ si\ les\ restrictions\ a\ priori\ sont\ v$érifiés $par\ le\ vecteur\ \beta$

(c-à-$d.$ $par\ les\ vraies\ valeurs\ des\ param$ètres à $estimer$), $l'estimateur\ \widehat{\beta}_c\ sera\ au\ moins$

$aussi\ efficace\ que\ l'estimateur\ \widehat{\beta}$; $en\ particulier$: $\begin{cases} E\left(\widehat{\beta}_c\right) = \beta \\ Var\left(\widehat{\beta}_{jc}\right) \le Var\left(\widehat{\beta}_j\right) \ , \forall j = 1, 2, \dots, k \end{cases}$

$En\ substituant\ \widehat{\beta} = \beta + (X'X)^{-1}X'\mathcal{E}\ dans\ (4) : \widehat{\beta}_c = \widehat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$

$on\ obtient$:

$$\widehat{\beta}_c = \beta + (X'X)^{-1}X'\mathcal{E} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\beta - R(X'X)^{-1}X'\mathcal{E})$$

$$\widehat{\beta}_c = \beta + \underbrace{[I_n - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R]}_{A}(X'X)^{-1}X'\mathcal{E}\ sous\ l'hypoth\text{è}se\ r - R\beta = 0$$

$\widehat{\beta}_c = \beta + A(X'X)^{-1}X'\mathcal{E}$ où $A\ est\ non\ stochastique, donc\ E\left(\widehat{\beta}_c\right) = \beta$

$$et\ Var\left(\widehat{\beta}_c\right) = E\left[\left(\widehat{\beta}_c - E\left(\widehat{\beta}_c\right)\right)'\left(\widehat{\beta}_c - E\left(\widehat{\beta}_c\right)\right)\right] = E\left[\left(\widehat{\beta}_c - \beta\right)'\left(\widehat{\beta}_c - \beta\right)\right]$$

$$= A(X'X)^{-1}X'E(\mathcal{E}\mathcal{E}')X(X'X)^{-1}A' = \sigma_\varepsilon^2 A \underbrace{(X'X)^{-1}(X'X)}_{I_k}(X'X)^{-1}A'$$

$$Var\left(\widehat{\beta}_c\right) = \sigma_\varepsilon^2 A(X'X)^{-1}A'$$

$On\ v$érifie $bien\ que\ si\ V = \sigma_\varepsilon^2(X'X)^{-1} = Var\left(\widehat{\beta}\right)$, $alors$

$$\sigma_\varepsilon^2 A\left(X'X\right)^{-1}A' = V - VR'\left(RVR'\right)^{-1}RV$$

$Var\left(\widehat{\beta}_c\right) - Var\left(\widehat{\beta}\right) = -VR'(RVR')^{-1}RV$, $matrice\ semi$– dé$finie\ n$égative , les éléments $de\ sa$

$diagonale\ sont\ non\ positifs\ et\ Var\left(\widehat{\beta}_{jc}\right) \le Var\left(\widehat{\beta}_j\right) \ , \forall j = 1, 2, \dots, k$

C– 3 • Décomposition $de\ la\ somme\ des\ r$ésidus $contraints$:

$Soit\ \widehat{\mathcal{E}}_c = Y - X\widehat{\beta}_c$, $le\ vecteur\ des\ r$ésidus $contraints.$ $On\ a$:

$$\hat{\varepsilon}_c\hat{\varepsilon}'_c = (Y - X\hat{\beta}_c)'(Y - X\hat{\beta}_c) = (Y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}_c)'(Y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}_c)$$

$$= (\hat{\varepsilon} + X[\hat{\beta} - \hat{\beta}_c])'(\hat{\varepsilon} + X[\hat{\beta} - \hat{\beta}_c]) = \hat{\varepsilon}'\hat{\varepsilon} + 2(\hat{\beta} - \hat{\beta}_c)'X'\hat{\varepsilon} + (\hat{\beta} - \hat{\beta}_c)'(X'X)(\hat{\beta} - \hat{\beta}_c)$$

$$\hat{\varepsilon}_c\hat{\varepsilon}'_c = \hat{\varepsilon}'\hat{\varepsilon} + (\hat{\beta} - \hat{\beta}_c)'(X'X)(\hat{\beta} - \hat{\beta}_c)$$

$(\hat{\beta} - \hat{\beta}_c)'(X'X)(\hat{\beta} - \hat{\beta}_c)$ ***est une matrice semi-définie positive. On a donc :*** $\hat{\varepsilon}_c\hat{\varepsilon}'_c \geq \hat{\varepsilon}'\hat{\varepsilon}$

Et comme $R_c^2 = 1 - \dfrac{\hat{\varepsilon}_c\hat{\varepsilon}'_c}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ ***et*** $R^2 = 1 - \dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \Rightarrow R_c^2 \leq R^2$

Cas particuliers de l'inférence en régression classique

D-1 • Distributions des statistiques de décision :

- $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 I_n)$
- $Y \sim \mathcal{N}(X\beta, \sigma_\varepsilon^2 I_n)$
- $\hat{\beta} \sim \mathcal{N}(\beta, \sigma_\varepsilon^2 (X'X)^{-1})$
- $\dfrac{(n-k)\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} = \dfrac{SCR}{\sigma_\varepsilon^2} = \dfrac{(n-k)\hat{\sigma}_{\hat{\beta}}^2}{\sigma_{\hat{\beta}}^2} \sim \chi^2(n-k)$
- $\dfrac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} =\sim \mathcal{T}(n-k)$, ***la loi de Student de d.d.l*** $(n-k)$
- $\dfrac{(\hat{\beta}_j - \beta_j)^2}{\hat{\sigma}_{\hat{\beta}_j}^2} =\sim \mathcal{F}(1, (n-k))$, ***la loi de Fisher dedegrés de liberté respectifs 1 et*** $(n-k)$
- $\dfrac{\hat{\beta}'(X'X)\hat{\beta}/k}{\hat{\sigma}_\varepsilon^2} = \dfrac{\hat{\beta}'(X'X)\hat{\beta}/k}{SCR/n-k} = \left(\dfrac{R_*^2}{1-R_*^2}\right)\left(\dfrac{n-k}{k}\right) \sim \mathcal{F}(k, (n-k))$

D-2 • Test de significativité individuelle :

Comme pour le cas simple, le test de significativité individuelle, qui porte sur chaque paramètre, est mené en calculant les ratios de Student. Pour un test bilatéral, les hypothèses du test sont: $\begin{cases} H_0 : \beta_j = \beta_j^0 \\ \text{contre} \\ H_1 : \beta_j \neq \beta_j^0 \end{cases}$ ***On rejette*** H_0 ***si*** $\left|\dfrac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}}\right| > t_{1-\frac{\alpha}{2}}(n-k)$; $j \in \{1, 2, \dots, k\}$

$\mathcal{D}-3$ • Test de nullité de tous les coefficients (la constante β_1 faisant partie) :

les hypothèses du test sont: $\begin{cases} H_0 : \beta = 0 \\ \text{contre} \\ H_1 : \beta \neq 0 \end{cases}$ On rejette H_0 si $F_{obs} > f_{1-\alpha}(k,(n-k))$

$$\text{où } F_{obs} = \frac{\hat{\beta}'(X'X)\hat{\beta}/k}{SCR/n-k} = \left(\frac{R_*^2}{1-R_*^2}\right)\left(\frac{n-k}{k}\right), \text{sous } H_0$$

On vérifie que $R_*^2 = \dfrac{kF_{obs}}{(n-k)+kF_{obs}}$

De manière équivalente on rejette H_0 si $R_*^2 > \dfrac{kf_{1-\alpha}(k,(n-k))}{(n-k)+kf_{1-\alpha}(k,(n-k))}$

$\mathcal{D}-4$ • Test de nullité de tous les coefficients (sauf la constante β_1) :

$$\begin{cases} H_0 : \breve{\beta} = 0 \\ \text{contre} \\ H_1 : \breve{\beta} \neq 0 \end{cases} \text{où } \breve{\beta} = \begin{pmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{pmatrix}$$

On rejette H_0 si $F_{obs} > f_{1-\alpha}((k-1),(n-k))$

$$\text{où } F_{obs} = \frac{\hat{Y}_c'\hat{Y}_c/k-1}{(Y_c'Y_c-\hat{Y}_c'\hat{Y}_c)/n-k} = \frac{SCE/k-1}{SCR/n-k} = \frac{\tilde{\beta}'(X'X)\tilde{\beta}/k-1}{SCR/n-k} = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-k}{k-1}\right)$$

$$\text{sous } H_0; \text{avec} \tilde{\beta} = \begin{pmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{pmatrix}$$

On vérifie que $R^2 = \dfrac{(k-1)F_{obs}}{(n-k)+(k-1)F_{obs}}$

De manière équivalente on rejette H_0 si $R^2 > \dfrac{(k-1)f_{1-\alpha}((k-1),(n-k))}{(n-k)+(k-1)f_{1-\alpha}((k-1),(n-k))}$

$\mathcal{D}-5$ • Test sur une combinaison linéaire des coefficients :

$$\begin{cases} H_0 : c'\beta - r = 0 \\ \text{contre} \\ H_1 : c\beta - r \neq 0 \end{cases} \text{où } c \text{ est un vecteur } k \times 1 \text{ de constantes et } r \text{ un scalaire}$$

Statistique de décision :

$$F = \frac{(c'\hat{\beta}-r)^2}{\hat{\sigma}_\varepsilon^2[c'(X'X)^{-1}c]} \leadsto \mathcal{F}(1,(n-k)) \text{ ou encore } T = \frac{c'\hat{\beta}-r}{\hat{\sigma}_\varepsilon\sqrt{c'(X'X)^{-1}c}} \leadsto \tau(T-k)$$

BEN AHMED MOHSEN Téléphone: (+216) 97 619191 / 54 619191
omega.center.cp@gmail.com https://web.facebook.com/OMEGACENTER2014

$\mathcal{D}-6$ • Test de stabilité structurelle (Chow):

On va diviser la période de l'échantillon en 2 sous-périodes de nombres d'observations $n_1 > k$ et $n_2 > k$, et étudier la stabilité des coéfficients de régression d'une sous période à l'autre. Sous l'hypothèse nulle (stabilité structurelle), les coefficients sont les même, sous l'hypothèse alternative, ils sont différents

Sous H_1 (pas de stabilité structurelle), le modèle s'écrit :

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \end{pmatrix}$$

Y_1 et $\mathcal{E}_1$ sont $n_1 \times 1$ et Y_2 et $\mathcal{E}_2$ sont $n_2 \times 1$, X_1 est $n_1 \times k$ et X_2 est $n_2 \times k$; β_1 et β_2 sont $k \times 1$

On a ici $2k$ régresseurs

$$\begin{cases} H_0 : \beta_1 = \beta_2 \\ \quad \text{contre} \\ H_1 : \beta_1 \neq \beta_2 \end{cases}$$

Sous H_0 (stabilité structurelle), le modèle s'écrit :

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \beta + \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \end{pmatrix} \text{ où } \beta = \beta_1 = \beta_2$$

On a ici k régresseurs

Statistique de décision : $F = \dfrac{(\hat{\mathcal{E}}_c'\hat{\mathcal{E}}_c - \hat{\mathcal{E}}'\hat{\mathcal{E}})/k}{\hat{\mathcal{E}}'\hat{\mathcal{E}}/(n-2k)} \rightsquigarrow \mathcal{F}(1,(n-k))$

*Le modèle **contraint** correspond aux hypothèses de régression classique avec* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

$$\hat{\mathcal{E}}_c'\hat{\mathcal{E}}_c = Y'Y - \hat{\beta}_c'X'Y = Y'[I_n - X(X'X)^{-1}X']Y = Y'MY$$

*Le modèle **non contraint**, on a comme matrice de régresseurs* $X_* = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$

$$\hat{\mathcal{E}}'\hat{\mathcal{E}} = Y'Y - \hat{\beta}'X_*'Y = Y'[I_n - X_*(X_*'X_*)^{-1}X_*']Y = Y'M_*Y$$

On vérifie facilement que : $Y'M_*Y = Y'M_1Y + Y'M_2Y$, avec $\begin{cases} M_1 = I_{n_1} - X_1(X_1'X_1)^{-1}X_1' \\ M_2 = I_{n_2} - X_2(X_2'X_2)^{-1}X_2' \end{cases}$

ainsi, on obtient : $F_{obs} = \dfrac{[Y'MY - (Y'M_1Y + Y'M_2Y)]/k}{Y'M_1Y + Y'M_2Y/(n-2k)}$,

on rejette H_0 (stabilité structurelle), si $F_{obs} > f_{1-\alpha}(k,(n-2k))$

Adresse : 2, Rue Ibn Arafa, av de l'environnement Manouba Centre, Manouba

$\mathcal{D}-7$ • *Prévision dans le modèle de régression linéaire multiple*:

Supposons que nous observons k valeurs futures des k régresseurs à une période $\boldsymbol{\theta}$ ***suivant la dernière période de l'échantillon. Ces valeurs forment un vecteur de dimension*** $\mathbf{1}\times \boldsymbol{k}$ ***, soit*** $\boldsymbol{x'_\theta}$.

On se propose de calculer un intervalle de prévision centré sur la prévision $\hat{y}_\theta$ ***de la variable dépendante.***

Si le modèle reste inchangé à la période θ ***, on a :*** $y_\theta = x'_\theta\beta + \varepsilon_\theta$

Avec $E(\varepsilon_\theta\varepsilon_1) = \cdots = E(\varepsilon_\theta\varepsilon_n) = 0$ ***et*** $\hat{y}_\theta = x'_\theta\hat{\beta}$

Sous l'hypothèse $\varepsilon_\theta \sim \mathcal{N}(0,\ \sigma_\varepsilon^2)$***, déterminon la distribution de l'erreurde prévision :***

$y_\theta - \hat{y}_\theta = \varepsilon_\theta - x'_\theta(\hat{\beta}-\beta)$ ***est une variable normale de paramètres :*** $E(y_\theta - \hat{y}_\theta) = 0$ ***et***

$$Var(y_\theta - \hat{y}_\theta) = E(\varepsilon_\theta^2) + E\left[\left(x'_\theta(\hat{\beta}-\beta)\right)^2\right] - 2Cov\left(\varepsilon_\theta\, , x'_\theta(\hat{\beta}-\beta)\,\right)$$

$Cov\left(\varepsilon_\theta\, , x'_\theta(\hat{\beta}-\beta)\,\right) = 0$***, puisque*** $\hat{\beta}$ ***ne dépend que des erreurs*** $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ ***de l'échantillon qui sont indépendantes de*** ε_θ ***par hypothèse. On a alors :***

$$Var(y_\theta - \hat{y}_\theta) = \ \sigma_\varepsilon^2 + E\left(x'_\theta(\hat{\beta}-\beta)(\hat{\beta}-\beta)'x_\theta\right) = \ \sigma_\varepsilon^2 +\ \sigma_\varepsilon^2 x'_\theta(X'X)^{-1}x_\theta$$

$$\frac{y_\theta - \hat{y}_\theta}{\sigma_\varepsilon\sqrt{1 + x'_\theta(X'X)^{-1}x_\theta}} \sim \mathcal{N}(0,1) \text{ et } \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_\varepsilon^2} \sim\ \chi^2(n-k) \Rightarrow \frac{\hat{y}_\theta - y_\theta}{\hat{\sigma}_\varepsilon\sqrt{1 + x'_\theta(X'X)^{-1}x_\theta}} \sim \mathcal{T}(n-k)$$

On obtient ainsi l'intervalle de prévision :

$$IP_{1-\alpha}(y_\theta) = \left[\hat{y}_\theta \pm t_{1-\frac{\alpha}{2}}(n-k)\hat{\sigma}_\varepsilon\sqrt{1 + x'_\theta(X'X)^{-1}x_\theta}\right]$$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXVème PROMOTION
Juillet 2005

Exercice 1 : *(14 points : 1+1+1,5+1,5+1,5+1,5+1,5+1,5+1,5+1,5)*

ÉNONCÉ

Les observations concernant la quantité échangée d'un bien (Q) et son prix (P) sont résumées dans le tableau ci-dessous:

Q	101	93	72	81	62	57	74	103	81	69
P	22	24	23	26	27	28	25	20	25	27

1) *Calculer les moyennes et les variances empiriques des deux variables Q et P, sachant que :* $\sum_{i=1}^{10} Q_i^2 = 65095$, $\sum_{i=1}^{10} P_i^2 = 6157$ *et* $\sum_{i=1}^{10} Q_i P_i = 19284$.
2) *Calculer la covariance entre Q et P. En déduire le coefficient de corrélation linéaire entre ces deux variables.*
3) *Selon le résultat de la question 2, préciser la nature de la relation entre Q et P (s'agit-il d'une fonction d'offre ou d'une fonction de demande ?)*
4) *On considère à présent le modèle linéaire suivant :*
   $Q_i = \alpha + \beta P_i + \varepsilon_i,\ i = 1, 2, \dots, 10$
   Où $\varepsilon_i,\ i = 1, 2, \dots, 10$ sont des termes aléatoires identiquement et indépendamment distribués d'espérance mathématique zéro et de variance σ^2 et α et β sont des paramètres à estimer.
   a) *Donner les expressions des estimateurs de α et β obtenus par la méthode des moindres carrés ordinaires, notés $\hat{\alpha}$ et $\hat{\beta}$.*
   b) *Calculer les valeurs numériques des estimateurs $\hat{\alpha}$ et $\hat{\beta}$.*
   c) *Décomposer la variance de Q en variance expliquée et variance résiduelle.*
   d) *Déterminer la valeur numérique de l'estimation sans biais de la variance des erreurs, notée $\hat{\sigma}^2$*
   e) *Calculer la valeur de l'estimation de la variance de $\hat{\beta}$.*
   f) *En admettant la normalité des ε_i :*
      i. *Construire un intervalle de confiance au niveau 95% pour le paramètre β.*
      ii. *Tester l'hypothèse $H_0: \beta = -5$ contre l'hypothèse alternative $H_1: \beta \neq -5$ pour un niveau de confiance de 95%.*

Corrigé

1)

$\cdot\ \overline{Q} = \frac{1}{10}\sum_{i=1}^{10} Q_i = 79,3 \qquad \cdot\ \overline{P} = \frac{1}{10}\sum_{i=1}^{10} P_i = 24,7$

$\cdot\ S_Q^2 = \frac{1}{10}\sum_{i=1}^{10} Q_i^2 - \overline{Q}^2 = 221,01 \Rightarrow S_Q = 14,87 \qquad \cdot\ S_P^2 = \frac{1}{10}\sum_{i=1}^{10} P_i^2 - \overline{P}^2 = 5,61 \Rightarrow S_P = 2,37$

2)

$\cdot\ S_{P,Q} = \frac{1}{10}\sum_{i=1}^{10} Q_i P_i - \overline{P}\,\overline{Q} = -30,31 \qquad \cdot\ r_{P,Q} = \frac{S_{P,Q}}{S_P S_Q} = -0,86$

3) $r_{P,Q} < 0$, *P et Q sont négativement orrélées d'où une pente négative.*

En effet, il s'agit d'une fonction de demande vue sa pente négative due au fait qu'une baisse des prix entrainent des quantités demandées élevée

4)

a)

$\cdot\ \hat{\beta} = \frac{S_{P,Q}}{S_P^2} \qquad \cdot\ \hat{\alpha} = \overline{Q} - \hat{\beta}\overline{P}$

b)

$\cdot\ \hat{\beta} = \frac{-30,31}{5,61} = -5,4 \qquad \cdot\ \hat{\alpha} = 79,3 - \big((-5,4)\times 24,7\big) = 212,68$

c) $Q_i = \alpha + \beta P_i + \varepsilon_i = \hat{\alpha} + \hat{\beta}P_i + \hat{\varepsilon}_i \Rightarrow \hat{Q}_i = \hat{\alpha} + \hat{\beta}P_i$

$$S_Q^2 = \frac{1}{n}\sum_{i=1}^{n}(Q_i - \overline{Q})^2 = \frac{1}{n}\sum_{i=1}^{n}\left[(Q_i - \hat{Q}_i) - (\overline{Q} - \hat{Q}_i)\right]^2$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2 + \sum_{i=1}^{n}(\hat{Q}_i - \overline{Q})^2 - 2\underbrace{\sum_{i=1}^{n}(\overline{Q} - \hat{Q}_i)(Q_i - \hat{Q}_i)}_{0}\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2 + \sum_{i=1}^{n}\Big((\hat{\alpha} + \hat{\beta}P_i) - (\hat{\alpha} + \hat{\beta}\overline{P})\Big)^2\right] = \frac{1}{n}\left[\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2 + \hat{\beta}^2\sum_{i=1}^{n}(P_i - \overline{P})^2\right]$$

$$\underbrace{S_Q^2}_{VT} = \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2\right]}_{VR} + \underbrace{\left[\frac{1}{n}\hat{\beta}^2\sum_{i=1}^{n}(P_i - \overline{P})^2\right]}_{VE}$$

· $VT = S_Q^2 = 221,01$ · $VE = \frac{1}{n}\hat{\beta}^2 \sum_{i=1}^{n} (P_i - \bar{P})^2 = \hat{\beta}^2 S_P^2 = 163,59$

· $VR = \left[\frac{1}{n}\sum_{i=1}^{n} (Q_i - \hat{Q}_i)^2\right] = VT - VE = 57,42$

d)

$$\hat{\sigma}^2 = \frac{n}{n-2} VR = \frac{10 \times 57,42}{8} = 71,78 \Longrightarrow \hat{\sigma} = 8,47$$

e)

$$\hat{\sigma}_{\hat{\beta}}^2 = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(P_i - \bar{P})^2} = \frac{71,78}{56,1} = 1,33 \Longrightarrow \hat{\sigma}_{\hat{\beta}} = 1,13$$

f)

i.

· $\frac{\hat{\beta} - \beta}{\hat{\sigma}_{\hat{\beta}}} \backsim \mathcal{T}(n-2)$ · $t_{1-\frac{\alpha}{2}}(n-2) = t_{0,975}(8) = 2,306$

· $L'intervalle\ de\ confiance\ de\ seuil\ 5\%\ pour\ \beta\ sera :\ IC_{1-\alpha}(\beta) = \left[\hat{\beta} \pm t_{1-\frac{\alpha}{2}}(n-2)\hat{\sigma}_{\hat{\beta}}\right]$

$$IC_{95\%}(\beta) = [(-5,4) \pm (2,306 \times 1,13)] = [-8,11, -2,8]$$

ii. $H_0: \beta = -5$ *contre* $H_a: \beta \neq -5$, *étant un test bilatéral par la suite la région du non-rejet coïncide avec l'intervalle de confiance :*

$-5 \in [-8,11, -2,8]$, *ou encore* $-5 \in IC_{95\%}(\beta)$

D'où H_0 *n'est pas rejetée et on pourra accepter la valeur* (-5) *pour* β, *avec un risque de première espèce* $\alpha = 5\%$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXVII^{ème} PROMOTION (BANQUE)
AOÛT 2017

Exercice 2-Première Partie : (11 points : 1,5+1,5+1,5+1,5+1,5+1,5+2)

Les deux parties de l'exercice sont indépendantes.

ÉNONCÉ

On considère la régression entre le PIB en logarithme y et le niveau des dépenses publiques en logarithme x observés sur un ensemble de $T = 200$ périodes :

$$y_t = ax_t + b + \epsilon_t \text{ , pour } t = 1, 2, \dots, 200$$

où ϵ_t sont des termes aléatoires identiquement et indépendamment distribués selon une loi normale d'espérance mathématique nulle et de variance σ^2, a et b sont des paramètres à estimer. On dispose des statistiques suivantes :

$\sum_{t=1}^{200} y_t = 1610$; $\sum_{t=1}^{200} x_t = 1027$ *avec la matrice de variance-covariance du vecteur* $\begin{bmatrix} y \\ x \end{bmatrix}$

égale à :

	Y	X
Y	$8,53$	$6,12$
X	$6,12$	$6,22$

Première Partie :

1) *Utiliser les valeurs numériques exposées ci-dessous pour évaluer les estimations de a et b. Quelle est l'interprétation économique de l'estimation de a ?*
2) *Calculer le coefficient de corrélation linéaire entre x et y. Interpréter.*
3) *En déduire que le coefficient de détermination (connu également sous le nom coefficient d'ajustement) de la régression est approximativement égal à 0,7 . Justifier les étapes de calcul*
4) *En déduire la somme des carrés des résidus*
5) *Proposer une estimation sans biais de la variance σ^2 . Justifier votre réponse*
6) *Déterminer la valeur de l'estimation de la variance de $\hat{a}$*
7) *Tester la significativité de la variable des dépenses publiques x à un niveau 95% de significativité. Interpréter ce résultat.*

On rappelle que pour S une variable de Student, la probabilité $P[-2 < S < 2]$ est

$approximativement\ \text{é}gale\ \text{à}\ 0,95$

Corrigé

1)

$y_t = ax_t + b + \epsilon_t$

$$\boxed{\begin{pmatrix} S_y^2 & S_{x,y} \\ S_{x,y} & S_x^2 \end{pmatrix} = \begin{pmatrix} 8,53 & 6,12 \\ 6,12 & 6,22 \end{pmatrix}}$$

· $\hat{a} = \dfrac{S_{x,y}}{S_x^2} = \dfrac{6,12}{6,22} = 0,98$

· $\hat{b} = \bar{y} - \hat{a}\bar{x} = \dfrac{1}{T}\sum_{t=1}^{T} y_t - \hat{a}\dfrac{1}{T}\sum_{t=1}^{T} x_t = \dfrac{1}{T}\left(\sum_{t=1}^{T} y_t - \hat{a}\sum_{t=1}^{T} x_t\right) = \dfrac{1610\ - (0,98 \times 1027)}{200} = 3,02$

$Notons\ NDP, le\ niveau\ des\ d\text{é}penses\ publiques\ donc\ , x = \ln(NDP)\ \ et\ y = \ln(PIB)$

$Il\ s'agitdonc\ d'une\ sp\text{é}cification\ Log\text{-}Log : ln(PIB_t) = a\ln(NDP_t) + b + \epsilon_t$

$par\ la\ suite\ \hat{a} = 0,98\ est\ une\ estimation\ de\ l'\text{élasticité}\ du\ PIB\ par\ rapport\ au\ niveau$ $des\ d\text{é}penses\ publiques\ (NDP).$

· $Le\ PIB\ \text{é}volue\ dans\ le\ m\text{ê}me\ sens\ et\ au\ m\text{ê}me\ rythme\ que\ le\ niveau\ des\ d\text{é}penses$ $publiques\ (puisque\ l'\text{élasticité}\ est\ positive\ et\ proche\ de\ 1\ )$

· $Une\ augmentation\ de\ 1\%\ de\ niveau\ des\ d\text{é}penses\ publiques\ entraine\ une\ augmente$ $du\ PIB\ de\ l'ordre\ de\ \ 0,98\%$

2)

$r_{x,y} = \dfrac{S_{x,y}}{S_x S_y} = \dfrac{6,12}{\sqrt{6,22 \times 8,53}} = 0,84$

$Il\ s'agit\ d'une\ corr\text{é}lation\ lin\text{é}aire\ forte\ entre\ le\ PIB\ et\ le\ niveau\ des\ d\text{é}penses\ publiques$

$Ce\ qui\ justifie\ bien\ la\ r\text{é}gression\ lin\text{é}aire\ entre\ logarithme\ du\ PIB\ et\ celui\ du\ niveau\ des$ $d\text{é}penses\ publiques$

3)

$$R^2 = \frac{SCE}{SCT} = \frac{\hat{a}^2 \sum_{t=1}^{T}(x_t - \bar{x})^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2} = \frac{\hat{a}^2 \sum_{t=1}^{T}(x_t - \bar{x})^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

$$or\ S_x^2 = \frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2 \Rightarrow \sum_{t=1}^{T}(x_t - \bar{x})^2 = TS_x^2\ , d'autrepart, \hat{a} = \frac{S_{x,y}}{S_x^2}$$

$$et\ S_y^2=\frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2 \Rightarrow \sum_{t=1}^{T}(y_t-\bar{y})^2=TS_y^2$$

$$R^2=\frac{\left(S_{x,y}/S_x^2\right)^2TS_x^2}{TS_y^2}=\frac{\left(S_{x,y}/S_x^2\right)^2S_x^2}{S_y^2}=\frac{S_x^2S_{x,y}^2/S_x^4}{S_y^2}=\frac{S_{x,y}^2/S_x^2}{S_y^2}=\frac{S_{x,y}^2}{S_x^2S_y^2}=\left(\frac{S_{x,y}}{S_xS_y}\right)^2=r_{x,y}^2=0,84^2\cong 0,7$$

4)

$$R^2=\frac{SCE}{SCT}=1-\frac{SCR}{SCT}\Leftrightarrow\frac{SCR}{SCT}=1-R^2\Leftrightarrow SCR=\underbrace{SCT}_{TS_y^2}(1-R^2)=TS_y^2(1-R^2)$$

$$SCR=200\times 8,53\times 0,3=511,8$$

5)

$$E\left(\sum_{t=1}^{T}\hat{\epsilon}_t^2\right)=(T-2)\,\sigma^2 \Longrightarrow E\left(\frac{1}{T-2}\sum_{t=1}^{T}\hat{\epsilon}_t^2\right)=\ \sigma^2$$

$$\hat{\sigma}^2=\frac{\sum_{t=1}^{T}\hat{\epsilon}_t^2}{T-2}=\frac{SCR}{T-2}\text{est un estimateur sans biais de } \sigma^2$$

$$\hat{\sigma}^2=\frac{511,8}{198}=2,58\Rightarrow\hat{\sigma}=1,6$$

6)

$$\begin{cases}\hat{\sigma}^2=\dfrac{SCR}{T-2}\\ \sigma_{\hat{a}}^2=\dfrac{\hat{\sigma}^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\end{cases}\Rightarrow\hat{\sigma}_{\hat{a}}^2=\frac{\hat{\sigma}^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}=\frac{\hat{\sigma}^2}{TS_x^2}=\frac{2,58}{200\times 6,22}=0,002\Rightarrow\hat{\sigma}_{\hat{a}}=0,046$$

7)

Tester la significativité de la variable des dépenses publiques (x) ***à un niveau*** 95% ***revient au même de tester la significativité du paramètre*** a ***, à un même niveau de signification.***

Il s'agit donc de confronter $H_0: a=0$ ***(le paramètre est statistiquement nul, non significatif)***

Contre l'alternative $H_1: a\neq 0$ ***(le paramètre est statistiquement non nul, significatif)***

$$T=\frac{\hat{a}-a}{\hat{\sigma}_{\hat{a}}}\ \backsim \mathcal{T}(T-2), \text{la loi de Student de d.d.l } (198)$$

Règle de décision: on rejette H_0 ***avec un risque de première espèce*** α***, si*** $|T_0|\geq t_{1-\frac{\alpha}{2}}(T-2)$

$$or\ |T_0|=\left|\frac{\hat{a}-a_{sous\ H_0}}{\hat{\sigma}_{\hat{a}}}\right|=\left|\frac{0,98-0}{0,046}\right|=21,3\ et\ t_{0,975}(198)\cong 2\ ainsi\ |T_0|\geq t_{0,975}(198)$$

Le niveau des dépenses publiques contribue à l'explication du PIB

la variable x (Le niveau des dépenses publiques) est significative avec un risque de première espèce $\alpha = 5\%$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XL$^{\text{ème}}$ PROMOTION (ASSURANCE)
Mai 2024

Exercice 3 : *(7 points : 1,5+1,5+2+2)*

ÉNONCÉ

On s'intéresse à la régression simple reliant la variable endogène y_t représentant le logarithme du PIB avec z_t représentant le logarithme des crédits alloués à l'économie avec des observations annuelles durant la période 2011 à 2020 : $y_t = az_t + b + u_t$ où les termes d'erreurs u_t sont des lois normales indépendantes, avec $E(u_t) = 0$ et $E(u_t) = \sigma^2$ avec $t = 1, 2, \dots, 10$.

L'estimation des Moindres Carrées Ordinaires du coefficient de z_t est $\hat{a} = 1,24$ avec un écart-type estimé $\hat{\sigma}_{\hat{a}}$ de l'estimateur $\hat{a}$ égale à : $\hat{\sigma}_{\hat{a}} = 0,19$.

Par ailleurs, la somme des carrées des résidus est égale à 38,27 et $\sum_{t=1}^{T=10} (y_t - \bar{y})^2 = 240$ où $\bar{y}$ est la moyenne empirique de y_t

1- Interpréter économiquement cette relation ainsi que les coefficients a et b
2- Etudier la significativité statistique de la variable z_t au niveau de 95% .
On rappelle que pour une loi de Student ST, la probabilité $P[-2 \le St \le 2]$ est voisine de 0,95
3- Quelle est l'estimation de σ l'écart-type de la régression ? cette estimation est-elle sans biais ? Justifier
4- Calculer le coefficient de détermination de la régression ainsi que le coefficient de corrélation linéaire entre y_t et z_t

Corrigé

1-

$\ln(PIB_t) = b + a\ln(CAE_t) + u_t$

· *Le modèle décrit la contribution de l'évolution des crédits alloués à l'économie dans la variation du PIB*

· β : *Le modèle étant logarithmique, donc β sera l'élasticité du PIB par rapport aux crédits alloués à l'économie*

2-

Tester la significativité de la variable (z) *à un niveau 95% revient au même de tester la significativité du paramètre a, à un même niveau de signification.*

Il s'agit donc de confronter $H_0 : a = 0$ *contre l'alternative* $H_1 : a \neq 0$

$T = \dfrac{\hat{a} - a}{\hat{\sigma}_{\hat{a}}} \sim \mathcal{T}(T-2)$, *la loi de Student de d.d.l* $(10-2)$

Règle de décision: on rejette H_0 *avec un risque de première espèce* α, *si* $|T_0| \geq t_{1-\frac{\alpha}{2}}(T-2)$

or $|T_0| = \left|\dfrac{\hat{a} - a_{sous\ H_0}}{\hat{\sigma}_{\hat{a}}}\right| = \left|\dfrac{1,24 - 0}{0,19}\right| = 6,526$ *et* $t_{0,975}(8) \cong 2$ *ainsi* $|T_0| \geq t_{0,975}(8)$

La variabilité des crédits alloués à l'économie contribue à l'explication du PIB

la variable z est significative avec un risque de première espèce $\alpha = 5\%$

3-

On démontre que: $E(SCR) = E\left(\sum_{t=1}^{T} \hat{u}_t^2\right) = (T-2)\,\sigma^2$, *il est donc biaisé de* σ^2 *en corrigeant ce biais, on obtient :* $\hat{\sigma}^2 = \dfrac{\sum_{t=1}^{T} \hat{u}_t^2}{T-2} = \dfrac{SCR}{T-2} = \dfrac{38,27}{8}$ *estimateur sans biais de* σ^2

$\boxed{\hat{\sigma}^2 = 4,78}$

4-

· $R^2 = 1 - \dfrac{SCR}{SCT} = 1 - \dfrac{38,27}{240} = 84,05\%$

$84,05\%$ *de la variation du PIB* (y_t) *est expliquée par la variabilité des crédits alloués l'économie* (z_t)

· $R^2 = r_{z,y}^2 \Leftrightarrow r_{z,y} = \pm\sqrt{0,8405} = \pm 0,917$, $or\ \hat{a} = 1,24 > 0$

D'où $\boxed{r_{z,y} = 0,917}$ $Il\ s'agit\ donc\ d'une\ corrélation\ linéaire\ est\ très\ forte$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXVIIIème PROMOTION(BANQUE)
JUILLET 2018

Exercice 4 : (5 points : un point par question)

ÉNONCÉ

$On\ propose\ de\ relier\ les\ évolutions\ temporelles\ du\ taux\ d'inflation\ \ y_t\ et\ le\ taux\ de$

$croissance\ de\ la\ masse\ monétaire\ x_t\ sous\ la\ forme$:

$y_t = ax_t + b + \varepsilon_t$, où ε_t $suit\ une\ loi\ normale\ centrées\ réduite$ ($V(\varepsilon_t) = 1$),

$avec\ \varepsilon_t\ indépendantes\ pour\ t = 1, 2, \dots, T$

$L'estimation\ du\ modèle\ par\ les\ moindres\ carrés\ a\ fourni\ les\ résultats\ suivants$:

$\hat{a} = 1,3\ et\ V(\hat{a}) = 0,04$

1)

i. $Interpréter$ $économiquement\ la\ relation\ et\ les\ résultats\ numériques\ précédents$

ii. $La\ variable\ masse\ monétaire\ est\text{-}elle\ statistiquement\ significative$? $Justifier\ votre\ réponse$

2)

i. $A\ partir\ des\ valeurs\ précédentes, en\ déduire\ les\ valeurs\ de$ $\sum_{t=1}^{T}(x_t - \overline{x})^2\ et\ de\ \sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y})$ où $\overline{x}\ et\ \overline{y}\ sont\ les\ moyennes\ empiriques\ de\ x\ et\ de\ y$

ii. $Calculer\ le\ coefficient\ de\ détermination\ sachant\ que\ \sum_{t=1}^{T}(y_t - \overline{y})^2 = 4$

iii. $Quelle\ interprétation$ $économique\ peut\text{-}on\ faire\ de\ ce\ dernier\ résultat$?

Corrigé

1)

i. $\textit{En négligeant l'effet des termes aléatoires, on obtient :}\ \Delta y_t = a\Delta x_t$

En conséquence, toutes choses égales par ailleurs (vitesse de circulation de la monnaie constante et niveau des transactions stable), une augmentation d'une unité de la variation absolue du taux de croissance de de la masse monétaire entraîne une hausse de la variation absolue du taux d'inflation de l'ordre de $1,3$

ii. *Tester la significativité de la masse monétaire, revient au même de tester l'hypothèse nulle* $H_0: \beta = 0$ *contre l'alternative* $H_1: \beta \neq 0$, *(sous un niveau de signification* $\alpha = 5\%$*)*

Les termes d'erreurs aléatoires $(\varepsilon_t)_{1\le t\le T}$ *sont de variance connue :* $\varepsilon_t \sim \mathcal{N}(0,1)$

Ainsi la statistique de décision sera : $U = \dfrac{\hat{a} - a}{\sigma_{\hat{a}}} \sim \mathcal{N}(0,1)$

Règle de décision : on rejette H_0 *avec un risque de première espèce* $\alpha = 5\%$, *si*

$|U_0| \geq \Phi^{-1}(0,975)$

$$or \begin{cases} |U_0| = \left|\dfrac{\hat{a} - 0}{\sigma_{\hat{a}}}\right| = \left|\dfrac{1,3}{0,2}\right| = 6,5 \\ \Phi^{-1}(0,975) = 1,96 \end{cases} \Rightarrow |U_0| \geq \Phi^{-1}(0,975)$$

D'où la variable, masse monétaire est statistiquement significative

2)

i.

$$\cdot\ \sigma_{\hat{a}}^2 = \frac{\sigma^2}{\sum_{t=1}^{T}(x_t - \overline{x})^2} \Leftrightarrow \sum_{t=1}^{T}(x_t - \overline{x})^2 = \frac{\sigma^2}{\sigma_{\hat{a}}^2} = \frac{1}{0,04} = 25$$

$$\cdot\ \hat{a} = \frac{\sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y})}{\sum_{t=1}^{T}(x_t - \overline{x})^2} \Leftrightarrow \sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y}) = \hat{a}\sum_{t=1}^{T}(x_t - \overline{x})^2 = 1,3 \times 25 = 32,5$$

ii. *Reformulons la question : Calculer le coefficient de détermination sachant que* $\sum_{t=1}^{T}(y_t - \overline{y})^2 = 44$

$$R^2 = \frac{SCE}{SCT} = \frac{\hat{a}\sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y})}{\sum_{t=1}^{T}(y_t - \overline{y})^2} = \frac{1,3 \times 32,5}{44} = 96,02\%$$

iii. $96,02\ \%$ *de la variation du taux d'inflation* (y_t) *est expliquée par la variation* *du taux de croissance de de la masse monétaire* (x_t)

Ce qui confirme la théorie quantitative de la monnaie : l'augmentation de la masse monétaire est la première cause d'inflation

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXVIème PROMOTION
Juillet 2006

Exercice 5 : *(12 points : 1,5+1.5+1.5+1.5+1+2+1.5+1.5)*

ÉNONCÉ

On considère 20 observations individuelles pour le logarithme népérien de consommation de tabac (Y), *le logarithme népérien du prix du tabac* (X) *et le logarithme népérien du revenu disponible* (Z). *On considère l'estimation de la relation suivante :*

$$Y_i = \alpha + \beta X_i + \varepsilon_i \, , i = 1, 2, \dots, 20 \qquad (1)$$

Où $\varepsilon_i \, , i = 1, 2, \dots, 20$ *sont des termes aléatoires identiquement et indépendamment distribués d'espérance mathématique zéro et de variance* σ^2, α *et* β *sont des paramètres à estimer.*

1) *Donner l'interprétation économique des deux paramètres* α *et* β.
2) *Donner l'expression des estimateurs par la méthode des moindres carrés ordinaire de* α *et* β, *notés respectivement* $\hat{\alpha}$ *et* $\hat{\beta}$
3) *Sachant que* $V(X) = 0,74$; $Cov(X, Y) = -0,925$; $\overline{X} = 0,2$ *et* $\overline{Y} = 4,85$, *calculer les valeurs numériques de* $\hat{\alpha}$ *et* $\hat{\beta}$. *Commenter les résultats obtenus.*
4) *En déduire la valeur du coefficient de détermination* (R^2), *pour* $V(Y) = 1,6$. *Commenter cette valeur .*
5) *L'ajout de la variable Z comme variable explicative de Y permet d'écrire le modèle (1) comme suit :* $Y_i = a + bX_i + cZ_i + \varepsilon_i \, , i = 1, 2, \dots, 20 \qquad (2)$, *où* ε_i *garde les mêmes hypothèses que précédemment.*
   a) *Quelle est la signification économique de c.*
   b) *Donner l'expression des estimateurs de b et de c, par la méthode des moindres carrés ordinaires.*
   c) *Sachant que* $Cov(X, Z) = 0,6$; $V(Z) = 2,02$ *et* $Cov(Y, Z) = -0,45$. *Calculer les*

estimateurs par la méthode des moindres carrés ordinaires de b et c.
Interpréter le résultat obtenu.

d) *Calculer la nouvelle valeur du coefficient de détermination. Commenter*

Corrigé

1) $\ln(cons\ tabac_i) = \alpha + \beta \ln(prix\ tabac_i) + \varepsilon_i$

· β : *Le modèle étant logarithmique, donc β sera l'élasticité prix de la demande de tabac*

· α : *La consommation correspondante à un prix unitaire du tabac serait égale à e^{α}*

2)

· $\widehat{\beta} = \dfrac{S_{X,Y}}{S_X^2}$ *et* · $\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$

3)

· $\widehat{\beta} = \dfrac{S_{X,Y}}{S_X^2} = \dfrac{-0,925}{0,74} = -1,25$

L'élasticité prix de la demande de tabac est estimée à $-1,25$

Ainsi une augmentation des prix de tabac (X) de 1% entraine une légère baisse de la consommation de l'ordre de $1,25\%$

· $\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} = 4,85 - (-1,25 \times 0,2) = 5,1$

La consommation correspondante à un prix unitaire du tabac serait égale à $e^{5,1} = 164,02\ UC$

4)

· $SCE_1 = \widehat{\beta}^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 = n\widehat{\beta}^2 S_x^2 = 20 \times (-1,25)^2 \times 0,74 = 23,125$

· $SCT = \sum_{i=1}^{n} (Y_i - \overline{Y})^2 = nS_Y^2 = 20 \times 1,6 = 32$

· $R_1^2 = \dfrac{SCE_1}{SCT} = \dfrac{23,125}{32} = 72,27\%$

$27,73\%$ de la consommation du tabac reste inexpliquée par le modèle (1), la part des prix du tabac dans l'explication du modèle est à $72,27\%$

5) $\ln(cons\ tabac_i) = a + b \ln(prix\ tabac_i) + c \ln(rev\ dispo_i) + \varepsilon_i$

a) c : *l'élasticité revenu de la demande de tabac*

b) $\widehat{\beta}^c = \begin{pmatrix} \widehat{b} \\ \widehat{c} \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c) = (X_c'X_c/n)^{-1}(X_c'Y_c/n)$

où $X'_cX_c = n\begin{pmatrix} S_X^2 & S_{X,Z} \\ S_{X,Z} & S_Z^2 \end{pmatrix}$ et $X'_cY_c = n\begin{pmatrix} S_{X,Y} \\ S_{Z,Y} \end{pmatrix}$, avec $(X'_cX_c)^{-1} = \dfrac{[Com(X^{c\prime}X^c)]'}{det(X^{c\prime}X^c)}$

c) $X'_cX_c = n\begin{pmatrix} S_X^2 & S_{X,Z} \\ S_{X,Z} & S_Z^2 \end{pmatrix} = \begin{pmatrix} 14,8 & 12 \\ 12 & 40,4 \end{pmatrix}$ et $X'_cY_c = n\begin{pmatrix} S_{X,Y} \\ S_{Z,Y} \end{pmatrix} = \begin{pmatrix} -18,5 \\ -9 \end{pmatrix}$

$$Com(X'_cX_c) = Com\begin{pmatrix} 14,8 & 12 \\ 12 & 40,4 \end{pmatrix} = \begin{pmatrix} 40,4 & -12 \\ -12 & 14,8 \end{pmatrix} \Rightarrow [Com(X'_cX_c)]' = Com(X'_cX_c) = \begin{pmatrix} 40,4 & -12 \\ -12 & 14,8 \end{pmatrix}$$

$$det(X'_cX_c) = \begin{vmatrix} 14,8 & 12 \\ 12 & 40,4 \end{vmatrix} = (14,8 \times 40,4) - 12^2 = 453,92$$

$$\hat{\beta}^c = \begin{pmatrix} \hat{b} \\ \hat{c} \end{pmatrix} = \frac{1}{453,92}\begin{pmatrix} 40,4 & -12 \\ -12 & 14,8 \end{pmatrix}\begin{pmatrix} -18,5 \\ -9 \end{pmatrix} = \frac{1}{453,92}\begin{pmatrix} -639,4 \\ 88,8 \end{pmatrix} \Rightarrow \boxed{\hat{\beta}^c = \begin{pmatrix} \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} -1,41 \\ 0,2 \end{pmatrix}}$$

· $\hat{b} = -1,41$

L'élasticité prix de la demande de tabac est estimée à $-1,41$

Ainsi une augmentation des prix de tabac (X) *de* 1% *entraine une légère baisse de la consommation de l'ordre de* $-1,41\%$

· $\hat{c} = 0,2$

L'élasticité revenu de la demande de tabac est estimée à $0,2$

Ainsi une augmentation des revenus disponibles (Z) *de* 1% *entraine une légère augmentation de la consommation de l'ordre de* $0,2\%$

d)

· $SCE_2 = \hat{\beta}^{c\prime}(X'_cX_c)\hat{\beta}^c = (-1,41 \quad 0,2)\begin{pmatrix} 14,8 & 12 \\ 12 & 40,4 \end{pmatrix}\begin{pmatrix} -1,41 \\ 0,2 \end{pmatrix} = (-18,47 \quad -8,84)\begin{pmatrix} -1,41 \\ 0,2 \end{pmatrix}$

$\Rightarrow \boxed{SCE_2 = 24,27}$

· $SCT = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = nS_Y^2 = 20 \times 1,6 = 32$

· $R_2^2 = \dfrac{SCE_2}{SCT} = \dfrac{24,27}{32} = 75,84\%$

$24,16\%$ *de la consommation du tabac reste inexpliquée par le modèle* (2) *, la part des prix du tabac et des revenus disponibles dans l'explication du modèle est de l'ordre de* $72,27\%$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXIème PROMOTION
Juillet 2011

Exercice 6 : (12 points : 1+1,5+1,5+1+1+1+1+1.5+1.5+1)

ÉNONCÉ

On considère la relation entre le logarithme de la productivité du travail (Y) *et le logarithme de l'intensité capitalistique* (X) *d'un pays donné pour une période de 32 années :*

$$Y_t = \alpha + \beta X_t + \varepsilon_t \,, t = 1, 2, \ldots, 32$$

Où ε_t, *sont des termes aléatoires identiquement et indépendamment distribués d'espérance mathématique zéro et de variance* σ^2, α *et* β *sont des paramètres à estimer.*

On dispose des statistiques suivantes : $\sum_{t=1}^{32} Y_t = 120$, $\sum_{t=1}^{32} X_t = 110$, $\sum_{t=1}^{32} Y_t^2 = 540$, $\sum_{t=1}^{32} X_t^2 = 510$

et $\sum_{t=1}^{32} X_t Y_t = 500$

1) *Donner l'interprétation économique de la relation définie précédemment ainsi que celle des paramètres* α *et* β
2) *Calculer les variances empiriques de* Y *et de* X *ainsi que leur covariance*
3) *Estimer par la méthode des moindres carrés ordinaires les paramètres* α *et* β, *que l'on note* $\hat{\alpha}$ *et* $\hat{\beta}$
4) *Donner la valeur numérique de l'estimateur sans biais de la variance des résidus, notée* $\hat{\sigma}^2$
5) *Calculer la variance estimée de* $\hat{\beta}$
6) *Sous l'hypothèse de la normalité des résidus, tester, au seuil de 5%, l'hypothèse* $H_0: \beta = 0,7$ *contre* $H_a: \beta \neq 0,7$. *Interpréter ce résultat*
7) *On introduit à présent la variable* (T), *comme variable explicative supplémentaire. La relation précédente devient :*

$$Y_t = \alpha + \beta X_t + \theta t + \varepsilon_t \,, t = 1, 2, \ldots, 32$$

   a) *Que représente le paramètre* θ ?
   b) *Sachant que :*

$$\sum_{t=1}^{32} t = 528\,, \sum_{t=1}^{32} t^2 = 11440, \sum_{t=1}^{32} tX_t = 420 \;\; et \;\; \sum_{t=1}^{32} tY_t = 450$$

i. *Calculer la variance empirique de* (T) *ainsi que les covariances entre* T *et les variables* Y *et* X

ii. *En déduire les estimateurs par moindres carrés ordinaires des paramètres* β *et* θ

iii. Commenter

Corrigé

Intensité capitalistique :

L'intensité capitalistique d'un secteur d'activité économique est définie par le rapport entre les immobilisations corporelles (valeur brute à la clôture de l'exercice) et les effectifs salariés moyens, ou bien par le rapport entre les immobilisations corporelles et la valeur ajoutée.

La productivité du travail :

La productivité du travail est le rapport entre la quantité ou la valeur ajoutée de la production et le nombre d'heures nécessaires pour la réaliser. Elle dépend de la capacité du personnel à produire une quantité, dite standard, de biens ou de services selon les normes ou les règles prédéfinies.

1)

· ***Le modèle décrit la contribution la contribution de l'évolution de l'intensité capitalistique dans la variation de la productivité du travail***

· β ***: Le modèle étant logarithmique, donc*** β ***sera l'élasticité de la productivité du travail par rapport à l'intensité capitalistique***

· α ***: Pour une intensité capitalistique réduite à l'unité, la productivité du travail serait égale à*** e^{α}

2)

· $\bar{X} = \dfrac{1}{32}\sum_{t=1}^{32} X_t = 3,438$ · $\bar{Y} = \dfrac{1}{32}\sum_{t=1}^{32} Y_t = 3,75$

· $S_X^2 = \dfrac{1}{32}\sum_{t=1}^{32} X_t^2 - \bar{X}^2 = 4,121 \Rightarrow S_X = 2,03$ · $S_Y^2 = \dfrac{1}{32}\sum_{t=1}^{32} Y_t^2 - \bar{Y}^2 = 2,813 \Rightarrow S_Y = 1,677$

$\cdot\ S_{X,Y} = \frac{1}{32}\sum_{t=1}^{32} X_tY_t - \overline{X}\overline{Y} = 2,734 \qquad \cdot\ r_{X,Y} = \frac{S_{X,Y}}{S_XS_Y} = 0,803$

3)

$\cdot\ \hat{\beta} = \frac{S_{X,Y}}{S_X^2} = \frac{2,734}{4,121} \Rightarrow \boxed{\hat{\beta} = 0,663}$

l'élasticité de la productivité du travail par rapport à l'intensité capitalistique est estimée à $0,663$.

Ainsi une variation de l'intensité capitalistique de 1% *entraine une variation de la productivité du travail de* $0,663\%$ *, l'élasticité étant inférieure à l'unité donc la variation de l'intensité capitalistique est moins proportionnelle que la variation de la productivité du travail .*

$\cdot\ \hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} = 3,75 - (0,663 \times 3,438) \Rightarrow \boxed{\hat{\alpha} = 1,47}$

4)

$\cdot\ SCE = T\hat{\beta}^2S_X^2 = 32 \times (0,663)^2 \times 4,121 = 57,967$

$\cdot\ SCT = TS_Y^2 = 32 \times 2,813 = 90,016$

$\cdot\ SCR = SCT - SCE = 32,049$

$\cdot\ \hat{\sigma}^2 = \frac{SCR}{T-2} = \frac{32,049}{30} = 1,068 \Rightarrow \boxed{\hat{\sigma}^2 = 1,068} \Rightarrow \boxed{\hat{\sigma} = 1,034}$

5)

$\hat{\sigma}_{\hat{\beta}}^2 = \frac{\hat{\sigma}^2}{TS_X^2} = \frac{1,068}{32 \times 4,121} \Rightarrow \boxed{\hat{\sigma}_{\hat{\beta}}^2 = 0,008} \Rightarrow \boxed{\hat{\sigma}_{\hat{\beta}} = 0,09}$

6)

On se propose de tester $H_0: \beta = 0,7$ *contre* $H_a: \beta \neq 0,7$ *, au risque de* 5%

$T = \frac{\hat{\beta} - \beta}{\hat{\sigma}_{\hat{\beta}}} \rightsquigarrow \mathcal{T}(32-2)$*, la loi de Student de d.d.l* (30)

Règle de décision : on rejette H_0 *avec un risque de première espèce* α*, si* $|T_0| \geq t_{1-\frac{\alpha}{2}}(30)$

$$or \begin{cases} |T_0| = \left|\frac{\hat{\beta} - \beta_0}{\hat{\sigma}_{\hat{\beta}}}\right| = \left|\frac{0,663 - 0,7}{0,09}\right| = 0,411 \\ t_{1-\frac{\alpha}{2}}(30) = t_{0,975}(30) = 2.0423 \end{cases} \Rightarrow |T_0| \ngeq t_{1-\frac{\alpha}{2}}(30)$$

On ne rejette pas l'hypothèse d'une élasticité $\varepsilon_{PT/IC}$ *égale à* $0,7$ *au risque de* 5%

7)

a)

Progrès technique :

Le progrès technique désigne l'ensemble des éléments mis en œuvre pour permettre l'amélioration du processus productif en accroissant la productivité. La variable temps (T) introduite est généralement utilisée comme variable proxy du progrès technique, pour substituer la variable utile (progrès technique) dans l'explication d'un modèle mais qui reste non observable ou non mesurable.

· θ : ***La variable temps (T) étant variable proxy du facteur progrès technique, difficile à mesurer. En effet une année de plus implique (100θ)% dans l'évolution de la productivité du travail***

b)

i.

$$\cdot\ \overline{T}=\frac{1}{32}\sum_{t=1}^{32}t=16,5 \qquad \cdot\ S_T^2=\frac{1}{32}\sum_{t=1}^{32}t^2-\overline{T}^2=85,25\Rightarrow S_X=9,233$$

$$\cdot\ S_{X,T}=\frac{1}{32}\sum_{t=1}^{32}tX_t-\overline{X}\overline{T}=-43,602 \qquad \cdot\ S_{T,Y}=\frac{1}{32}\sum_{t=1}^{32}tY_t-\overline{T}\overline{Y}=-47,813$$

ii.

$$\cdot\ Y_t=\alpha+\beta X_t+\theta t+\varepsilon_t\ ,t=1,2,\dots,32$$

$$\cdot\ \tilde{\beta}=\begin{pmatrix}\hat{\beta}\\\hat{\theta}\end{pmatrix}=(X_c'X_c)^{-1}(X_c'Y_c)=\left(T\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}\right)^{-1}\left(T\begin{pmatrix}S_{X,Y}\\ S_{T,Y}\end{pmatrix}\right)=\frac{1}{T}\times T\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}^{-1}\begin{pmatrix}S_{X,Y}\\ S_{T,Y}\end{pmatrix}$$

$$\tilde{\beta}=\begin{pmatrix}\hat{\beta}\\\hat{\theta}\end{pmatrix}=\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}^{-1}\begin{pmatrix}S_{X,Y}\\ S_{T,Y}\end{pmatrix}\ ,avec\ \begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}^{-1}=\frac{\left[Com\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}\right]'}{\begin{vmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{vmatrix}}$$

$$\cdot\ Com\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}=Com\begin{pmatrix}4,121 & -43,602\\ -43,602 & 85,25\end{pmatrix}=\begin{pmatrix}85,25 & 43,602\\ 43,602 & 4,121\end{pmatrix}$$

$$\Rightarrow\left[Com\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}\right]'=Com\begin{pmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{pmatrix}=\begin{pmatrix}85,25 & 43,602\\ 43,602 & 4,121\end{pmatrix}$$

$$\cdot\ \begin{vmatrix}S_X^2 & S_{X,T}\\ S_{X,T} & S_T^2\end{vmatrix}=\begin{vmatrix}4,121 & -43,602\\ -43,602 & 85,25\end{vmatrix}=-1549,819$$

· $\begin{pmatrix} S_X^2 & S_{X,T} \\ S_{X,T} & S_T^2 \end{pmatrix}^{-1} = -\dfrac{1}{1549,819}\begin{pmatrix} 85,25 & 43,602 \\ 43,602 & 4,121 \end{pmatrix}$

· $\widetilde{\beta} = \begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = \begin{pmatrix} S_X^2 & S_{X,T} \\ S_{X,T} & S_T^2 \end{pmatrix}^{-1} \begin{pmatrix} S_{X,Y} \\ S_{T,Y} \end{pmatrix} = -\dfrac{1}{1549,819}\begin{pmatrix} 85,25 & 43,602 \\ 43,602 & 4,121 \end{pmatrix}\begin{pmatrix} 2,734 \\ -47,813 \end{pmatrix}$

$\boxed{\widetilde{\beta} = \begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = \begin{pmatrix} 1,195 \\ 0,05 \end{pmatrix}}$

· $\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} - \widehat{\theta}\overline{T} = 3,75 - (1,195 \times 3,438) - (0,05 \times 16,5) = -1,183$

$\boxed{\widehat{\alpha} = -1,183}$

iii.

· $\widehat{\beta} = 1,195$

L'élasticité de la productivité du travail par rapport à l'intensité capitalistique est estimée à 1,195.

Ainsi une variation de l'intensité capitalistique de 1% entraine une variation de la productivité du travail de 1,195%, l'élasticité étant supérieure à l'unité donc la variation de l'intensité capitalistique est plus proportionnelle que la variation de la productivité du travail .

· $\widehat{\theta} = 0,05$

Une année de plus (progrès technique) implique 5%, en moyenne de croissance dans l'évolution de la productivité du travail

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXIVème PROMOTION
AOÛT 2014

Exercice 7 : *(12 points : 1 point par question et 2 points pour la dernière question)*

ÉNONCÉ

On considère un échantillon de $n = 100$ logements pour lesquels on observe le logarithme du prix (Y), le logarithme de la superficie en m^2 (X), et le nombre de pièces (Z).

On fournit les données suivantes :

$$\sum_{i=1}^{100} Y_i = 1179, \sum_{i=1}^{100} X_i = 733, \sum_{i=1}^{100} Z_i = 300, \sum_{i=1}^{100}(Y_i-\bar{Y})^2 = 24, \sum_{i=1}^{100}(X_i-\bar{X})^2 = 13,$$

$$\sum_{i=1}^{100}(Z_i-\bar{Z})^2 = 42, \sum_{i=1}^{100}(Y_i-\bar{Y})(X_i-\bar{X}) = 16, \sum_{i=1}^{100}(Y_i-\bar{Y})(Z_i-\bar{Z}) = 14, \sum_{i=1}^{100}(Z_i-\bar{Z})(X_i-\bar{X}) = 14$$

On adopte le modèle définit par :

$$Y_i = \alpha + \beta X_i + \varepsilon_i \, , i = 1, 2, \dots, 100 \qquad (1)$$

Avec ε_i : un terme d'erreur vérifiant les hypothèses suivantes :

$$E(\varepsilon_i) = 0 \; \forall\, i \, , V(\varepsilon_i) = \sigma^2 \; \forall\, i \, , Cov(\varepsilon_i, \varepsilon_j) = 0 \; \forall\, i \neq j \, , \; Cov(\varepsilon_i, X_i) = 0 \; \forall\, i$$

1) *Calculer les coefficients de corrélation linéaire entre Y et X d'une part et entre Y et Z d'autre part.*
2) *Donner une interprétation économique de ce modèle ainsi que du paramètre β.*
3) *Commenter économiquement les hypothèses $Cov(\varepsilon_i, \varepsilon_j) = 0$ pour $i \neq j$ et $Cov(\varepsilon_i, X_i) = 0$*
4) *Estimer par la méthode des moindres carrés ordinaires (MCO) les paramètres α et β, notés respectivement $\hat{\alpha}$ et $\hat{\beta}$. Commenter.*
5) *Calculer la somme des carrés de résidus, en déduire la variance estimée de $\hat{\beta}$.*
6) *Sous l'hypothèse de la normalité des résidus, tester la significativité de β au seuil de 5%.*

On rajoute à présent la variable nombre de pièces comme variable explicative, le modèle devient :

$$Y_i = \alpha + \beta X_i + \theta Z_i + \varepsilon_i \, , i = 1, 2, \dots, 100 \qquad (2)$$

7) *Reprendre l'estimation du modèle par la méthode des moindres carrés ordinaires. Commenter.*
8) *Ecrire l'équation d'analyse de la variance et calculer ses composantes.*
9) *En déduire le coefficient de détermination linéaire* (R^2). *Commenter.*
10) *Sous l'hypothèse de la normalité des résidus, tester au seuil d 5% la significativité globale de ce modèle.*
11) *On rajoute à présent une variable indicatrice pour le nouveau logement* $(N = 1)$ *s'il s'agit d'un nouveau logement et* $(N = 0)$ *pour un ancien logement. L'estimation du modèle fournit les résultats suivants :*

$$\widehat{Y}_i = \underset{(0{,}5)}{1{,}7} + \underset{(0{,}06)}{1{,}16}X_i - \underset{(0{,}025)}{0{,}08}Z_i + \underset{(0{,}07)}{0{,}21}N_i$$

Commenter ces résultats en termes d'ordre de grandeurs et de signes des coefficients sachant que les nombres entre parenthèses sont les écarts types estimés.

Indication : Pour une loi de Fisher $\mathcal{F}(2,97)$ *et une loi de Student* $\mathcal{T}(98)$, *nous avons :*

$$P(\mathcal{F}(2,97) > 3{,}09) = 0{,}05 \text{ et } P(|\mathcal{T}(98)| < 2{,}276) = 0{,}95$$

Corrigé

1)

· $$\rho_{Y,X} = \frac{\sum_{i=1}^{100}(Y_i - \overline{Y})(X_i - \overline{X})}{\sqrt{\sum_{i=1}^{100}(Y_i - \overline{Y})^2 \sum_{i=1}^{100}(X_i - \overline{X})^2}} = \frac{16}{\sqrt{24 \times 13}} \Rightarrow \boxed{\rho_{Y,X} = 0{,}91}$$

· $$\rho_{Y,Z} = \frac{\sum_{i=1}^{100}(Y_i - \overline{Y})(Z_i - \overline{Z})}{\sqrt{\sum_{i=1}^{100}(Y_i - \overline{Y})^2 \sum_{i=1}^{100}(Z_i - \overline{Z})^2}} = \frac{14}{\sqrt{24 \times 42}} \Rightarrow \boxed{\rho_{Y,Z} = 0{,}44}$$

2)

Le modèle étant logarithmique (variation des prix des logements en fonction de leurs superficies) donc β *sera l'élasticité des prix des logements* (Y) *par rapport à la superficie* (X)

3)

· $Cov(\varepsilon_i, \varepsilon_j) = 0$ *pour* $i \neq j$: *Indépendance des erreurs, autrement dit les erreurs relatives à 2 observations aléatoires différentes des prix des logements sont indépendantes.*

· $Cov(\varepsilon_i, X_i) = 0$: *Le terme d'erreur aléatoire* ε_i *est indépendant de la variable exogène non-aléatoire* X *(la superficie d'un logement).*

En d'autres termes l'omission ou l'ajout d'autres variables certaines (non-aléatoire) sera *non corrélé avec les perturbations aléatoires.*

En d'autres termes l'omission ou l'ajout d'autres variables certaines (non-aléatoire) sera non corrélé avec les perturbations aléatoires.

4)

$$\widehat{\beta} = \sum_{i=1}^{100}(Y_i - \overline{Y})(X_i - \overline{X}) \Big/ \sum_{i=1}^{100}(X_i - \overline{X})^2 = \frac{16}{13} \Rightarrow \boxed{\widehat{\beta} = 1,231}$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - \widehat{\beta}\sum_{i=1}^{n} X_i\right) = \frac{1}{100} \times \left(\sum_{i=1}^{100} Y_i - \widehat{\beta}\sum_{i=1}^{100} X_i\right) = \frac{1179 - \left(\frac{16}{13} \times 733\right)}{100} \Rightarrow \boxed{\widehat{\alpha} = 2,77}$$

L'élasticité des prix des logements (Y) par rapport à la superficie(X)est estimée à $1,231$.

Ainsi une augmentation de la superficie(X)de 1% entraine une augmentation des prix de $1,231\%$, l'élasticité étant supérieure à l'unité donc la variation du prix est plus que proportionnelle à la variation de la superficie.

5)

$$SCR = \sum_{i=1}^{100} \hat{\varepsilon}_i^2 = \sum_{i=1}^{100}(Y_i - \overline{Y})^2 - \widehat{\beta}^2 \sum_{i=1}^{100}(X_i - \overline{X})^2 = 24 - (1,231^2 \times 13) \Rightarrow \boxed{SCR = 4,3}$$

$$\widehat{\sigma}^2 = \frac{SCR}{100-2} = \frac{4,3}{98} = 0,044 \Rightarrow \boxed{\widehat{\sigma}^2 = 0,044} \Rightarrow \boxed{\widehat{\sigma} = 0,209}$$

$$\widehat{\sigma}_{\widehat{\beta}}^2 = \widehat{\sigma}^2 \Big/ \sum_{i=1}^{100}(X_i - \overline{X})^2 = 0,044/13 \Rightarrow \boxed{\widehat{\sigma}_{\widehat{\beta}}^2 = 0,0034} \Rightarrow \boxed{\widehat{\sigma}_{\widehat{\beta}} = 0,058}$$

6)

On se propose de tester $H_0 : \beta = 0$ contre $H_a : \beta \neq 0$, sous un niveau de signification $\alpha = 5\%$

$$T = \frac{\widehat{\beta} - \beta}{\widehat{\sigma}_{\widehat{\beta}}} \backsim \mathcal{T}(100-2), \text{la loi de Student de d.d.l } (98)$$

Règle de décision : on rejette H_0 avec un risque de première espèce α, si $|T_0| \geq t_{1-\frac{\alpha}{2}}(98)$

$$or \begin{cases} |T_0| = \left|\dfrac{\widehat{\beta} - 0}{\widehat{\sigma}_{\widehat{\beta}}}\right| = \left|\dfrac{1,231}{0,058}\right| = 21,224 \\ t_{1-\frac{\alpha}{2}}(98) = t_{0,975}(98) = 2,276 \end{cases} \Rightarrow |T_0| \geq t_{1-\frac{\alpha}{2}}(98)$$

D'où β est significativement non nul au risque de 5%

7)

$$Y_i = \alpha + \beta X_i + \theta Z_i + \varepsilon_i \,, i = 1,2, \dots, 100 \qquad (2)$$

$$\tilde{\beta} = \begin{pmatrix} \hat{\beta} \\ \hat{\theta} \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c)$$

$$\text{où } X_c'X_c = \begin{pmatrix} \sum_{i=1}^{n}(X_i-\bar{X})^2 & \sum_{i=1}^{n}(X_i-\bar{X})(Z_i-\bar{Z}) \\ \sum_{i=1}^{n}(Z_i-\bar{Z})(X_i-\bar{X}) & \sum_{i=1}^{n}(Z_i-\bar{Z})^2 \end{pmatrix} = \begin{pmatrix} 13 & 14 \\ 14 & 42 \end{pmatrix}$$

$$et\ X_c'Y_c = \begin{pmatrix} \sum_{i=1}^{n}(Y_i-\bar{Y})(X_i-\bar{X}) \\ \sum_{i=1}^{n}(Y_i-\bar{Y})(Z_i-\bar{Z}) \end{pmatrix} = \begin{pmatrix} 16 \\ 14 \end{pmatrix}$$

$$(X_c'X_c)^{-1} = \frac{[Com(X^{c\prime}X^c)]'}{det(X^{c\prime}X^c)}$$

$$Com(X_c'X_c) = Com\begin{pmatrix} 13 & 14 \\ 14 & 42 \end{pmatrix} = \begin{pmatrix} 42 & -14 \\ -14 & 13 \end{pmatrix} \Rightarrow [Com(X_c'X_c)]' = Com(X_c'X_c) = \begin{pmatrix} 42 & -14 \\ -14 & 13 \end{pmatrix}$$

$$det(X_c'X_c) = \begin{vmatrix} 13 & 14 \\ 14 & 42 \end{vmatrix} = (13 \times 42) - 14^2 = 350$$

$$\tilde{\beta} = \begin{pmatrix} \hat{\beta} \\ \hat{\theta} \end{pmatrix} = \frac{1}{350}\begin{pmatrix} 42 & -14 \\ -14 & 13 \end{pmatrix}\begin{pmatrix} 16 \\ 14 \end{pmatrix} = \frac{1}{350}\begin{pmatrix} 476 \\ -42 \end{pmatrix} \Rightarrow \boxed{\tilde{\beta} = \begin{pmatrix} \hat{\beta} \\ \hat{\theta} \end{pmatrix} = \begin{pmatrix} 1,36 \\ -0,12 \end{pmatrix}}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} - \hat{\theta}\bar{Z} = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - \hat{\beta}\sum_{i=1}^{n} X_i - \hat{\theta}\sum_{i=1}^{n} Z_i\right) = \frac{1179 - (1,36 \times 733) + (0,12 \times 300)}{100}$$

$$\Rightarrow \boxed{\hat{\alpha} = 2,18}$$

La semi -élasticité des prix des logements (Y) par rapport aux nombre de pièces (Z) étant estimée à $\hat{\theta} = -0,12$. En effet une augmentation d'une pièce implique une diminution du prix de l'ordre de 12%, chose qui est bizarre voir absurde de point de vue économique.

8)

$$SCT = \sum_{i=1}^{n}(Y_i-\bar{Y})^2 \Rightarrow \boxed{SCT = 24}$$

$$SCE = \tilde{\beta}'(X'X)\tilde{\beta} = (1,36 \quad -0,12)\begin{pmatrix} 13 & 14 \\ 14 & 42 \end{pmatrix}\begin{pmatrix} 1,36 \\ -0,12 \end{pmatrix} = (16 \quad 14)\begin{pmatrix} 1,36 \\ -0,12 \end{pmatrix} \Rightarrow \boxed{SCE = 20,08}$$

$$SCR = SCT - SCE = 24 - 20,08 \Rightarrow \boxed{SCR = 3,92}$$

9)

$R^2 = \frac{SCE}{SCT} = \frac{20,08}{24} \Rightarrow \boxed{R^2 = 83,67\%} \Rightarrow 83,67\%$ *De la variabilité logarithme du prix* (Y) *est expliquée par le logarithme de la superficie* (X) *et le nombre de pièces* (Z).

10)

On se propose de confronter $\begin{cases} H_0: \breve{\beta} = 0 \\ \text{contre} \\ H_1: \breve{\beta} \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0: \begin{pmatrix} \beta \\ \theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \text{contre} \\ H_1: \begin{pmatrix} \beta \\ \theta \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$

sous un niveau de signification $\alpha = 5\%$

Stat de décision: $F = \frac{\widetilde{\beta}'(X_c'X_c)\widetilde{\beta}/2}{SCR/100-3} =\backsim \mathcal{F}(2,97)$

$$or \begin{cases} F_{obs} = \frac{SCE/2}{SCR/97} = \frac{20,08/2}{3,92/97} = 248,44 \\ \quad et \\ f_{95\%}(2,97) = 3,09 \end{cases} \Rightarrow F_{obs} \gg f_{95\%}(2,97)$$

Donc le modèle est globalement significatif au risque de 5%

11)

- *Le nouveau modèle étant* : $Y_i = \alpha + \beta X_i + \theta Z_i + \lambda N_i + \varepsilon_i \,, i = 1,2,\ldots,100 \qquad (3)$

L'estimation du modèle fournit les résultats suivants :

$$\hat{Y}_i = \underset{(0,5)}{1,7} + \underset{(0,06)}{1,16}X_i - \underset{(0,025)}{0,08}Z_i + \underset{(0,07)}{0,21}N_i$$

En calculant le coefficient de détermination partielle associé à la variable indicatrice N_i *introduite au modèle, on trouve* :

$$r^2_{Y,N_i} = \frac{t^2_{\hat{\lambda}}}{t^2_{\hat{\lambda}} + (n-k_2)} \text{ où } t^2_{\hat{\lambda}} = \left(\frac{\hat{\lambda}}{\hat{\sigma}_{\hat{\lambda}}}\right)^2 = \left(\frac{0,21}{0,07}\right)^2 = 9 \text{ et } k_2 = 4 \Rightarrow \boxed{r^2_{Y,N_i} = 8,57\%}$$

En éliminant l'effet de la superficie des logements, ainsi que le nombre de pièces, on constate que l'impact de cette variable muette introduite n'est pas négligeable et il est de l'ordre de $8,57\%$. *Ce qui va systématiquement réduire l'effet des autres variables.*

On pourra aussi identifier cette réduction de l'effet des variables (X) *et* (Z) *par une simple remarque apportée à travers les paramètres du nouveau modèle* :

$$\begin{cases} \hat{\beta}_{(2)} = 1,36 \\ \hat{\beta}_{(3)} = 1,16 \end{cases} \Rightarrow \hat{\beta}_{(3)} < \hat{\beta}_{(2)}$$

$$et \begin{cases} \widehat{\theta}_{(2)} = -0,12 \\ \widehat{\theta}_{(3)} = -0,08 \text{ (plus proche de 0 donc d'une significativité inférieure à celle du modèle(2))} \end{cases}$$

· Y *étant le logarithme du prix donc un nouveau logement coûte* ,

$\left((e^{\hat{\lambda}} \times 100) - 100\right)\% = \left((e^{0,21} \times 100) - 100\right)\% = 23,37\%$ *plus cher qu'un ancien logement, toute autre chose égale par ailleurs c.-à-d. en fixant la superficie et le nombre de pièces.*

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXVIIème PROMOTION
JUILLET 2007

Exercice 8 : *(7 points : 1+1+1+1+1+1+1)*

ÉNONCÉ

On considère la régression qui relie le rendement mensuel d'un titre donné (y) *et le rendement mensuel de l'indice* (x) *de la bourse des valeurs mobilières d'un pays donné et ce pour la période courant janvier 2005 à mars 2007 (soit 27 mois) :*

$y_t = \alpha + \beta x_t + \varepsilon_t \, , t = 1, 2, \dots, 27.$

Où ε_t *sont des termes aléatoires identiquement et indépendamment distribués d'espérance mathématique zéro et de variance* σ^2 , α *et* β *sont des paramètres à estimer .*

On dispose des statistiques suivantes :

$$\sum_{t=1}^{27} y_t = 5,3 \, , \sum_{t=1}^{27} x_t = 22 \, , \sum_{t=1}^{27} y_t^2 = 485 \, , \sum_{t=1}^{27} x_t^2 = 112 \, , \sum_{t=1}^{27} x_t y_t = 135$$

1) *Commenter brièvement cette relation.*
2) *Calculer la variance de* y*, la variance de* x *et la covariance entre* y *et* x *.*
3) *En déduire les valeurs numériques des estimateurs des paramètres* α *et* β *obtenus par la méthode des moindres carrés ordinaires*
4) *Donner la valeur numérique de l'estimateur sans biais de la variance des résidus, noté* $\hat{\sigma}^2$
5) *La valeur du rendement de l'indice du pays pour le mois d'avril 2007 est :* $x_{28} = 3$ *.*
   a) *Calculer la valeur de la prévision du rendement de ce titre pour le mois*

d'avril $2007.$

b) *Calculer la variance estimée de cette prévision.*

c) *Sous l'hypothèse de la normalité des résidus, construire un intervalle de prévision au seuil de 95%. On rappelle que pour une loi de Student à 25 degré de liberté, nous avons :* $P(|S| < 2,06) = 0,95$

Corrigé

1) *Il s'agit d'une relation linéaire entre le rendement du titre et le rendement global du marché. Cette relation est dérivée du modèle d'évaluation des actifs financiers (MEDAF)*

2)

$$\cdot\ S_y^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^2 - \bar{y}^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T} y_t\right)^2 = \frac{1}{T^2}\left[T\sum_{t=1}^{T} y_t^2 - \left(\sum_{t=1}^{T} y_t\right)^2\right]$$

$$S_y^2 = \frac{1}{27^2}\times[(27\times 485) - (5,3)^2] \Rightarrow \boxed{S_y^2 = 17,92}$$

$$\cdot\ S_x^2 = \frac{1}{T}\sum_{t=1}^{T} x_t^2 - \bar{x}^2 = \frac{1}{T}\sum_{t=1}^{T} x_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T} x_t\right)^2 = \frac{1}{T^2}\left[T\sum_{t=1}^{T} x_t^2 - \left(\sum_{t=1}^{T} x_t\right)^2\right]$$

$$S_x^2 = \frac{1}{27^2}\times[(27\times 112) - (22)^2] \Rightarrow \boxed{S_x^2 = 3,48}$$

$$\cdot\ S_{x,y} = \frac{1}{T}\sum_{t=1}^{T} x_t y_t - \overline{xy} = \frac{1}{T}\sum_{t=1}^{T} x_t y_t - \left(\frac{1}{T}\sum_{t=1}^{T} x_t\right)\left(\frac{1}{T}\sum_{t=1}^{T} y_t\right) = \frac{1}{T^2}\left[T\sum_{t=1}^{T} x_t y_t - \left(\sum_{t=1}^{T} x_t\right)\left(\sum_{t=1}^{T} y_t\right)\right]$$

$$S_{x,y} = \frac{1}{27^2}[(27\times 135) - (22\times 5,3)] \Rightarrow \boxed{S_{x,y} = 4,84}$$

3)

$$\cdot\ \hat{\beta} = \frac{S_{X,Y}}{S_X^2} = \frac{4,84}{3,48} \Rightarrow \boxed{\hat{\beta} = 1,39}$$

$$\cdot\ \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} = \frac{1}{T}\left[\sum_{t=1}^{T} y_t - \hat{\beta}\sum_{t=1}^{T} x_t\right] = \frac{1}{27}[5,3\ - (1,39\times 22)] \Rightarrow \boxed{\hat{\alpha} = -0,94}$$

4)

$$\cdot\ SCT = TS_y^2 = 27\times 17,92 = 483,84$$

$$\cdot\ SCE = T\hat{\beta}^2 S_x^2 = 27\times 1,39^2\times 3,48 = 181,54$$

$$\cdot\ SCR = SCT - SCE = 483,84 - 181,54 = 302,3$$

$$\cdot\ \hat{\sigma}^2 = \frac{SCR}{T-2} = \frac{302,3}{27-2} \Rightarrow \boxed{\hat{\sigma}^2 = 12,09 = 3,48^2}$$

5) $x_{28} = 3$

a) $\boldsymbol{Supposons\ que\ le\ modèle\ reste\ inchangé\ ainsi\ que\ ses\ hypothèses\ MCO, nous\ aurons:}\ y_{28} = \alpha + \beta x_{28} + \varepsilon_{28}$

$\hat{y}_{28} = \hat{\alpha} + \hat{\beta} x_{28}$ $\boldsymbol{Sera\ un\ estimateur\ sans\ biais\ de}\ y_{28}$

$\hat{y}_{28} = -0,94 + (1,39 \times 3) \Rightarrow \boxed{\hat{y}_{28} = 3,23}$

b)

$\cdot\ \bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t = \frac{22}{27} = 0,81$

$\cdot\ \hat{V}(\hat{y}_{28}) = \hat{\sigma}^2\left(\frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right) = \hat{\sigma}^2\left(\frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{TS_x^2}\right) = \frac{12,09}{27} \times \left(1 + \frac{(3-0,81)^2}{3,48}\right)$

$\Rightarrow \boxed{\hat{V}(\hat{y}_{28}) = 1,06 = 1,03^2}$

c)

$\cdot\ \hat{\varepsilon}_{28} = y_{28} - \hat{y}_{28}$

$\cdot\ \hat{V}(\hat{\varepsilon}_{28}) = \hat{V}(y_{28} - \hat{y}_{28}) = \hat{V}(y_{28}) + \hat{V}(\hat{y}_{28}) = \hat{\sigma}^2 + \hat{\sigma}^2\left(\frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right) = \hat{\sigma}^2\left(1 + \frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right)$

$\cdot\ \hat{\varepsilon}_{28} \backsim \mathcal{N}\big(0, V(\hat{\varepsilon}_{28})\big) \Rightarrow \frac{\hat{\varepsilon}_{28}}{\sqrt{\hat{V}(\hat{\varepsilon}_{28})}} = \frac{y_{28} - \hat{y}_{28}}{\hat{\sigma}\sqrt{1 + \frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}}} \backsim \mathcal{T}(T-2)$

$\boldsymbol{D'où\ l'intervalle\ de\ prévision:}\ IP_{1-\alpha}(y_{28}) = \left[\hat{y}_{28} \pm t_{1-\frac{\alpha}{2}}(T-2)\hat{\sigma}\sqrt{1 + \frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}}\right]$

$\cdot\ \begin{cases} P(|S| < 2,06) = 0,95 \\ S \backsim \mathcal{T}(25) \end{cases} \Rightarrow t_{1-\frac{\alpha}{2}}(T-2) = t_{0,975}(25) = 2,06$

$\cdot\ \hat{V}(\hat{\varepsilon}_{28}) = \hat{\sigma}^2 + \hat{\sigma}^2\left(\frac{1}{T} + \frac{(x_{28}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right) = \hat{\sigma}^2 + \hat{V}(\hat{y}_{28}) = 12,09 + 1,06 = 13,15 \Rightarrow \hat{\sigma}_{\hat{\varepsilon}_{28}} = 3,63$

$IP_{95\%}(y_{28}) = [3,23 \pm (2,06 \times 3,63)] \Rightarrow \boxed{IP_{95\%}(y_{28}) = [-4,25, 10,71\,]}$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXVIIIème PROMOTION
JUILLET 2008

Exercice 9 : *(12 points : 1 point par question)*

ÉNONCÉ

On considère la relation entre le logarithme du prix unitaire de l'or (noté y) et le logarithme du taux de change euro-dollar (noté x) observés dans le temps :

$y_t = \alpha + \beta x_t + \varepsilon_t$, $t = 1, 2, \dots, 360$.

Avec ε_t un terme d'erreur vérifiant toutes les hypothèses de la méthode des moindres carrés ordinaires (MCO) ; α et β des paramètres à estimer.

1) *Rappeler les hypothèses classiques du terme d'erreur permettant les moindres carrés ordinaires.*
2) *Commenter la relation définie précédemment en précisant l'interprétation de β.*

Pour : $\sum_{t=1}^{360} y_t^2 = 14800$, $\sum_{t=1}^{360} x_t^2 = 60$, $\sum_{t=1}^{360} x_t y_t = 800$, $\sum_{t=1}^{360} y_t = 2300$, $\sum_{t=1}^{360} x_t = 120$

3) *Calculer la variance de x et y ainsi que la covariance entre x et y.*
4) *Calculer les valeurs des estimateurs obtenus par MCO des paramètres α et β, notés respectivement $\hat{\alpha}$ et $\hat{\beta}$. Commenter.*
5)
   i. *Exprimer la variance expliquée par la régression en fonction de $\hat{\beta}$ et de la covariance entre x et y*
   ii. *En déduire l'équation d'analyse de la variance (connue aussi sous la relation d'équation de la variance)*
6) *Tester la significativité de l'estimateur $\hat{\beta}$ au seuil de 5%.*

On admet qu'une loi de Student est comprise entre -2 et 2 avec une probabilité de 95%.

7) *Pour $x_{361} = 0,6$. Construire un intervalle de prévision pour y_{361} au niveau de confiance 95%*
8) *En introduisant le logarithme du prix unitaire du pétrole (noé z), comme variable explicative supplémentaire, le modèle devient :*
   $y_t = \alpha + \beta x_t + \theta z_t + \varepsilon_t$, $t = 1, 2, \dots, 360$.

Avec $Var(z) = 0,054$, $Cov(y, z) = 0,070$ *et* $Cov(x, z) = 0,013$

   a) *Estimer par MCO les paramètres β et θ, notés respectivement $\tilde{\beta}$ et $\tilde{\theta}$.*

b) *Calculer l'estimateur de la matrice de variance-covariance du vecteur* $\begin{pmatrix}\tilde{\beta}\\ \tilde{\theta}\end{pmatrix}$

c) *Tester la significativité de* θ *au seuil de* 5%

d) *Commenter.*

Corrigé

1)

- $\mathcal{H}_1$(*Hypothèse sur la partie systématique*):

Hypothèses sur Y et X :

X et Y Sont deux grandeurs numériques mesurées sans erreur. X Est une donnée exogène dans le modèle. Elle est supposée non aléatoire (ou non stochastique). Y Est aléatoire par l'intermédiaire $\mathcal{E}$ *c-à-d. la seule erreur qu'on a sur Y provient des insuffisances de X à expliquer ses valeurs dans le modèle.*

- $\mathcal{H}_2$ (*Hypothèse sur la partie stochastique*):

Hypothèses sur les termes aléatoires ε_i :

Les ε_i *sont i.i.d (indépendants et identiquement distribués).*

- $\mathcal{H}_3$(*Hypothèse sur la partie stochastique*): $E(\varepsilon_i) = 0\ \forall\ i = 1, 2, \dots, n$

En moyenne les erreurs s'annulent c-à-d. le modèle est bien spécifié.

- $\mathcal{H}_4$(*Hypothèse sur la partie stochastique*): $Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_\varepsilon^2\ \forall\ i = 1, 2, \dots, n$

La variance de l'erreur est constante et ne dépend pas de l'observation. C'est l'hypothèse d'homoscédasticité des erreurs.

Il s'agit d'assumer que les variables explicatives omises dans le modèle influent toutes pratiquement de façon constante sur la variable expliquée.

- $\mathcal{H}_5$(*Hypothèse d'indépendance entre la partie systématique et la partie aléatoire*):

L'erreur est indépendante de la variable exogène c-à-d. $Cov(\varepsilon_i, x_i) = 0\ \forall\ i = 1, 2, \dots, n$

Cette hypothèse signifie que l'erreur et les variables explicatives ont une influence séparée sur la variable endogène

- $\mathcal{H}_6$(*Hypothèse sur la partie stochastique*):

Indépendance des erreurs :

Les erreurs relatives à 2 observations différentessont indépendantes c-à-d.

BEN AHMED MOHSEN Téléphone: (+216) 97 619191 / 54 619191
omega.center.cp@gmail.com https://web.facebook.com/OMEGACENTER2014

$Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0 \ pour \ \ i \neq j$.

On parle de l'hypothèse de la non autocorrélation des erreurs

- $\mathcal{H}_7$(*Hypothèse sur la partie stochastique*): $\varepsilon_i \rightsquigarrow \mathcal{N}(0, \sigma_\varepsilon^2)$.

L'hypothèse de normalité des erreurs est un élément clé pour l'inférence statistique.

Elle est donc nécessaire pour mener les tests.

- $\mathcal{H}_8$(*Hypothèse sur la partie systématique*) :

x_i *prend au moins deux valeurs différentes*

- $\mathcal{H}_9$(*Hypothèse sur la partie systématique*):

$n > 2$, *c-à-d. le nombre d'observations n doit être supérieur au nombre des paramètres à estimer.*

2) *Il s'agit d'une spécification Log-Log* :

$$ln\left(\underbrace{prix\ unitaire\ de\ l'or}_{PUO}\right) = \alpha + \beta ln\left(\underbrace{taux\ de\ change\ euro\text{-}dollar}_{TCED}\right) + \varepsilon_t$$

Notons η *l'élasticité du prix unitaire de l'or* (PUO) *par rapport aux taux de change euro-dollar* $(TCED) \Rightarrow \eta = \dfrac{d(PUO)/PUO}{d(TCED)/TCED}$

$$Or \begin{cases} d\ln(PUO)/dPUO = 1/PUO \\ d\ln(TCED)/dTCED = 1/TCED \end{cases} \Rightarrow \begin{cases} dy = d\ln(PUO) = dPUO/PUO \\ dx = d\ln(TCED) = dTCED/TCED \end{cases}$$

et comme on a : $\beta = \dfrac{dy}{dx} = \dfrac{d(PUO)/PUO}{d(TCED)/TCED}$, *donc* $\beta = \eta$, *l'élasticité du prix unitaire de l'or* (PUO) *par rapport aux taux de change euro-dollar* $(TCED)$.

Ainsi une augmentation du taux de change euro-dollar de 1% *entraine une augmentation de* $\beta\%$ *des prix unitaire de l'or.*

3)

$$Var(x) = \frac{1}{360}\sum_{t=1}^{360} x_t^2 - \left(\frac{1}{360}\sum_{t=1}^{360} x_t\right)^2 = \frac{60}{360} - \left(\frac{120}{360}\right)^2 \Rightarrow \boxed{Var(x) = 0{,}056}$$

$$Cov(x,y) = \frac{1}{360}\sum_{t=1}^{360} x_t y_t - \left[\left(\frac{1}{360}\right)^2 \sum_{t=1}^{360} x_t \sum_{t=1}^{360} y_t\right] = \frac{800}{360} - \left(\frac{120 \times 2300}{360^2}\right) \Rightarrow \boxed{Cov(x,y) = 0{,}093}$$

Adresse : 2, Rue Ibn Arafa, av de l'environnement Manouba Centre, Manouba

4)

$$\hat{\beta} = \frac{\sum_{t=1}^{360}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{360}(x_t - \bar{x})^2} = \frac{Cov(x,y)}{Var(x)} = \frac{0,093}{0,056} \Rightarrow \boxed{\hat{\beta} = 1,67}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = \frac{1}{360} \times \left(\sum_{t=1}^{360} y_t - \hat{\beta}\sum_{t=1}^{360} x_t\right) = \frac{2300 - (1,67 \times 120)}{360} \Rightarrow \boxed{\hat{\alpha} = 5,83}$$

***Une augmentation du taux de change euro-dollar de** 1% **entraine une augmentation de** 1,67 % **des prix unitaire de l'or. Autrement dit l'élasticité du prix unitaire de l'or par rapport aux taux de change euro-dollar est de l'ordre de** 1,67*

5)

i.

$$SCE = \sum_{t=1}^{360}(\hat{y}_t - \bar{\hat{y}})^2 = \sum_{t=1}^{360}(\hat{y}_t - \bar{y})^2 = \hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2 = \hat{\beta} \times \hat{\beta}\sum_{t=1}^{360}(x_t - \bar{x})^2$$

$$= \hat{\beta} \times \left(\frac{\sum_{t=1}^{360}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{360}(x_t - \bar{x})^2}\right)\sum_{t=1}^{360}(x_t - \bar{x})^2 = \hat{\beta}\sum_{t=1}^{360}(x_t - \bar{x})(y_t - \bar{y})$$

$$SCE = 360\hat{\beta}Cov(x,y)$$

$$\boxed{SCE = 55,912}$$

ii.

$$On\ a\ \hat{\varepsilon}_t = y_t - \hat{y}_t = y_t - \hat{\alpha} - \hat{\beta}x_t = y_t - (\bar{y} - \hat{\beta}\bar{x}) - \hat{\beta}x_t = (y_t - \bar{y}) - \hat{\beta}(x_t - \bar{x})$$

***Par la suite** :*

$$SCR = \sum_{t=1}^{360}\hat{\varepsilon}_t^2 = \sum_{t=1}^{360}\left[(y_t - \bar{y}) - \hat{\beta}(x_t - \bar{x})\right]^2 = \sum_{t=1}^{360}(y_t - \bar{y})^2 - 2\hat{\beta}\underbrace{\sum_{t=1}^{360}(x_t - \bar{x})(y_t - \bar{y})}_{\hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2} + \hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2$$

$$= \sum_{t=1}^{360}(y_t - \bar{y})^2 - 2\hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2 + \hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2 = \sum_{t=1}^{360}(y_t - \bar{y})^2 - \underbrace{\hat{\beta}^2\sum_{t=1}^{360}(x_t - \bar{x})^2}_{\sum_{t=1}^{360}(\hat{y}_t - \bar{y})^2}$$

$$SCR = \underbrace{\sum_{t=1}^{360}(y_t - \bar{y})^2}_{SCT} - \underbrace{\sum_{t=1}^{360}(\hat{y}_t - \bar{y})^2}_{SCE}$$

***D'où l'équation d'analyse de la variance** :*

$$\underbrace{\sum_{t=1}^{360}(y_t-\bar{y})^2}_{SCT}=\underbrace{\sum_{t=1}^{360}(\hat{y}_t-\bar{y})^2}_{SCE}+\underbrace{\sum_{t=1}^{360}\hat{\varepsilon}_t^2}_{SCR}\quad ou\ encore\ SCT=SCE+SCR$$

$$SCT=\sum_{t=1}^{360}(y_t-\bar{y})^2=\sum_{t=1}^{360}y_t^2-360\left(\frac{1}{360}\sum_{t=1}^{360}y_t\right)^2=\sum_{t=1}^{360}y_t^2-\frac{1}{360}\left(\sum_{t=1}^{360}y_t\right)^2=14800-\frac{2300^2}{360}$$

$$\boxed{SCT=105,556}$$

$$SCR=SCT+SCE\Rightarrow\boxed{SCR=49,644}$$

6)

Calculons d'abord $\hat{\sigma}^2$ *et* $\hat{\sigma}^2_{\hat{\beta}}$:

$$\hat{\sigma}^2=\frac{SCR}{360-2}=\frac{49,644}{358}\Rightarrow\boxed{\hat{\sigma}^2=0,139}\Rightarrow\boxed{\hat{\sigma}=0,372}$$

$$et\ \hat{\sigma}^2_{\hat{\beta}}=\hat{\sigma}^2\Big/\sum_{t=1}^{360}(x_t-\bar{x})^2=0,139/20\Rightarrow\boxed{\hat{\sigma}^2_{\hat{\beta}}=0,007}\Rightarrow\boxed{\hat{\sigma}_{\hat{\beta}}=0,083}$$

On se propose maintenant de tester :

$H_0: \beta=0$ *contre* $H_1: \beta\neq 0$, *sous un niveau de signification* $\alpha=5\%$

$$T=\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}}\leadsto \mathcal{T}(360-2),\ la\ loi\ de\ Student\ de\ d.d.l\ (358)$$

Règle de décision : on rejette H_0 *avec un risque de première espèce* α, *si* $|T_0|\geq t_{1-\frac{\alpha}{2}}(358)$

$$or\ \begin{cases}|T_0|=\left|\dfrac{\hat{\beta}-0}{\hat{\sigma}_{\hat{\beta}}}\right|=\left|\dfrac{1,67}{0,083}\right|=20,12\\ t_{1-\frac{\alpha}{2}}(n)\in[-2,2]\end{cases}\Rightarrow |T_0|\geq t_{1-\frac{\alpha}{2}}(358)$$

D'où β *est significativement non nul au risque de* 5%

7)

Sous l'hypothèse que le modèle reste inchangé, nous aurons : $y_{361}=\alpha+\beta x_{361}+\varepsilon_{361}$

et $\hat{y}_{361}=\hat{\alpha}+\hat{\beta}x_{361}$ *Sera un estimateur sans biais de* y_{361}

$$\hat{y}_{361}=5,83+(1,67\times 0,6)\Rightarrow\boxed{\hat{y}_{361}=6,832}$$

$$L'intervalle\ de\ prévision : IC_{95\%}(y_{361})=\left[\hat{y}_{361}\pm t_{0,975}(358)\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_{361}-\bar{x})^2}{\sum_{t=1}^{360}(x_t-\bar{x})^2}}\right]$$

$$Avec : t_{0,975}(358) = \Phi^{-1}(0,975) = 1,96\ , \hat{\sigma} = 0,372, \overline{X} = 0,333\ et\ \sum_{t=1}^{360}(x_t - \overline{x})^2 = 20$$

$$IC_{95\%}(y_{361}) = \left[6,832 \pm 1,96 \times 0,372\sqrt{1 + \frac{1}{360} + \frac{(0,6-0,333)^2}{20}}\right]$$

$$\boxed{IC_{95\%}(y_{361}) = [6,832 \pm 0,731] = [6,101\ ;\ 7,763]}$$

8) $y_t = \alpha + \beta x_t + \theta z_t + \varepsilon_t\ , t = 1, 2, \dots, 360$

$$Notons\ \breve{\beta} = \begin{pmatrix} \beta \\ \theta \end{pmatrix}, X_c = \begin{pmatrix} (x_1 - \overline{x}) & (z_1 - \overline{z}) \\ \vdots & \vdots \\ (x_{360} - \overline{x}) & (z_{360} - \overline{z}) \end{pmatrix}, Y_c = \begin{pmatrix} (y_1 - \overline{y}) \\ \vdots \\ (y_{360} - \overline{y}) \end{pmatrix}\ et\ \mathcal{E}_c = \begin{pmatrix} (\varepsilon_1 - \overline{\varepsilon}) \\ \vdots \\ (\varepsilon_{360} - \overline{\varepsilon}) \end{pmatrix}$$

$$donc\ l'ecriture\ matricielle\ des\ données\ centrées\ sera : (M_c) :\ Y_c = X_c\widetilde{\beta} + \widehat{\mathcal{E}}$$

a)

$$\widetilde{\beta} = \begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c) = \left(\frac{X_c'X_c}{n}\right)^{-1}\left(\frac{X_c'Y_c}{n}\right) et\ \widehat{\alpha} = \overline{y} - (\widehat{\beta}\overline{x} + \widehat{\theta}\overline{z})$$

$$X_c'X_c = 360\begin{pmatrix} Var(x) & Cov(x,z) \\ Cov(x,z) & Var(z) \end{pmatrix} = \begin{pmatrix} 20 & 4,68 \\ 4,68 & 19,44 \end{pmatrix}$$

$$Com(X_c'X_c) = Com\begin{pmatrix} 20 & 4,68 \\ 4,68 & 19,44 \end{pmatrix} = \begin{pmatrix} 19,44 & -4,68 \\ -4,68 & 20 \end{pmatrix}$$

$$\Rightarrow [Com(X_c'X_c)]' = Com(X_c'X_c) = \begin{pmatrix} 19,44 & -4,68 \\ -4,68 & 20 \end{pmatrix}$$

$$et\ det(X_c'X_c) = \begin{vmatrix} 20 & 4,68 \\ 4,68 & 19,44 \end{vmatrix} = (20 \times 19,44) - 4,68^2 = 366,9$$

$$(X_c'X_c)^{-1} = \frac{[Com(X_c'X_c)]'}{det(X_c'X_c)} = \frac{1}{366,9}\begin{pmatrix} 19,44 & -4,68 \\ -4,68 & 20 \end{pmatrix}$$

$$X_c'Y_c = 360\begin{pmatrix} Cov(y,x) \\ Cov(y,z) \end{pmatrix} = \begin{pmatrix} 33,33 \\ 25,2 \end{pmatrix}$$

$$\widetilde{\beta} = \begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c) = \frac{1}{366,9}\begin{pmatrix} 19,44 & -4,68 \\ -4,68 & 20 \end{pmatrix}\begin{pmatrix} 33,33 \\ 25,2 \end{pmatrix}$$

$$\boxed{\begin{pmatrix} \widehat{\beta} \\ \widehat{\theta} \end{pmatrix} = \begin{pmatrix} 1,44 \\ 0,95 \end{pmatrix}}$$

b)

$$SCT = \sum_{t=1}^{360}(y_t - \overline{y}) = Y_c'Y_c \Rightarrow \boxed{SCT = 105,556}$$

$$SCE = \sum_{t=1}^{360}\left(\hat{y}_t - \bar{\hat{y}}\right)^2 = \sum_{t=1}^{360}(\hat{y}_t - \bar{y})^2 = \hat{Y}_c'\hat{Y}_c = \tilde{\beta}'X_c'Y_c = \tilde{\beta}'(X_c'X_c)\tilde{\beta}$$

$$= (1,44 \quad 0,95)\begin{pmatrix} 20 & 4,68 \\ 4,68 & 19,44 \end{pmatrix}\begin{pmatrix} 1,44 \\ 0,95 \end{pmatrix} = (32,246 \quad 25,21)\begin{pmatrix} 1,44 \\ 0,95 \end{pmatrix} \Rightarrow \boxed{SCE = 70,38}$$

$$SCR = SCT - SCE \Rightarrow \boxed{SCR = 35,176} \Rightarrow \hat{\sigma}^2 = \frac{SCR}{360-3} \Rightarrow \boxed{\hat{\sigma}^2 = 0,197} \Rightarrow \boxed{\hat{\sigma} = 0,444}$$

$$\hat{\sigma}^2_{\tilde{\beta}} = \hat{\sigma}^2(X_c'X_c)^{-1} = \frac{0,197}{366,9}\begin{pmatrix} 19,44 & -4,68 \\ -4,68 & 20 \end{pmatrix} \Rightarrow \boxed{\hat{\sigma}^2_{\tilde{\beta}} = \begin{pmatrix} 0,010 & -0,0025 \\ -0,0025 & 0,0107 \end{pmatrix}}$$

$$or\ \hat{\sigma}^2_{\tilde{\beta}} = \begin{pmatrix} \hat{\sigma}^2_{\hat{\beta}} & \widehat{Cov}(\hat{\beta},\hat{\theta}) \\ \widehat{Cov}(\hat{\beta},\hat{\theta}) & \hat{\sigma}^2_{\hat{\theta}} \end{pmatrix} \Rightarrow \boxed{\begin{cases} \hat{\sigma}^2_{\hat{\beta}} = 0,010 \Rightarrow \hat{\sigma}_{\tilde{\beta}} = 0,1 \\ \hat{\sigma}^2_{\hat{\theta}} = 0,0107 \Rightarrow \hat{\sigma}_{\tilde{\theta}} = 0,103 \\ \widehat{Cov}(\hat{\beta},\hat{\theta}) = -0,0025 \end{cases}}$$

c) *On se propose maintenant de tester :*

$H_0: \theta = 0$ *contre* $H_1: \theta \neq 0$, *sous un niveau de signification* $\alpha = 5\%$

$$T = \frac{\hat{\theta} - \theta}{\hat{\sigma}_{\hat{\theta}}} \rightsquigarrow \mathcal{T}(360-3), \textit{la loi de Student de d.d.l } (357)$$

Règle de décision : on rejette H_0 *avec un risque de première espèce* α, *si* $|T_0| \geq t_{1-\frac{\alpha}{2}}(357)$

$$or \begin{cases} |T_0| = \left|\dfrac{\hat{\theta}}{\hat{\sigma}_{\tilde{\theta}}}\right| = \left|\dfrac{0,95}{0,103}\right| = 9,22 \\ t_{1-\frac{\alpha}{2}}(n) \in [-2,2] \end{cases} \Rightarrow |T_0| \geq t_{0,975}(357)$$

D'où θ *est significativement non nul au risque de* 5%

d) *L'élasticité du prix unitaire de l'or par rapport aux prix unitaire du pétrole est significativement non nulle.*

L'influence de la variable prix unitaire du pétrole sur prix unitaire du pétrole n'est pas nulle et il est de l'ordre de $0,95\%$.

BEN AHMED MOHSEN Téléphone: (+216) 97 619191 / 54 619191
omega.center.cp@gmail.com https://web.facebook.com/OMEGACENTER2014

Exercice 10 : *(10 points :1 point par question)*

ÉNONCÉ

Sur une période de dix ans, on dispose des informations sur l'épargne (E_t), mesurée en millions de dinars et le taux d'intérêt (R_t) mesuré en pourcentage.

Ces informations sont résumées comme suit :

$$\sum_{t=1}^{10} E_t = 435\ ;\ \sum_{t=1}^{10} R_t = 43,5\ ;\ \sum_{t=1}^{10} E_t^2 = 20775\ ;\ \sum_{t=1}^{10} R_t^2 = 215,25 \text{ et } \sum_{t=1}^{360} R_t E_t = 2106,5$$

On se propose d'estimer une relation linéaire définie par :

$$E_t = \alpha + \beta R_t + \varepsilon_t\ ;\ t = 1, 2, \ldots, 10 \qquad (1)$$

Avec α et β des paramètres à estimer ; ε_t, $t = 1, 2, \ldots, 10$ sont des termes aléatoires identiquement et indépendamment distribués d'espérance mathématique zéro et de variance σ^2. On suppose que ces termes suivent la loi normale.

1)
   a) *Interpréter les coefficients α et β. Quels sont les signes attendus de ces paramètres ?*
   b) *Quelle interprétation économique peut-on donner au terme d'erreur ε_t ?*
2) *Estimer les paramètres α et β par la méthode des moindres carrés ordinaires.*
3) *Calculer la valeur numérique de l'estimateur sans biais de la variance des termes d'erreurs.*
4)
   a) *Calculer le coefficient de corrélation linéaire simple entre les deux variables E et R.*
   b) *En déduire le coefficient de détermination.*
5) *Tester la significativité de l'effet de la variable R_t sur E_t, au seuil de 5%.*
6) *Pour l'année 11, le taux d'intérêt observé est de 4. Construire un intervalle de prévision pour l'épargne de l'année correspondante, au niveau de 95%.*
7)
   a) *En admettant que le vrai modèle est celui défini par l'équation (1), montrer*

que l'estimation du modèle sans le terme constant (α) conduit à un estimateur biaisé de β.

b) *En déduire la variance de ce dernier estimateur.*

Corrigé

1) $Epargne_t = \alpha + \beta Taux\ d'intérêt_t + \varepsilon_t$

a)

· β : *En négligeant l'effet des termes aléatoires, on obtient :* $\Delta E_t = a\Delta R_t$

En conséquence, toutes choses égales par ailleurs, une augmentation d'une unité de la variation absolue du taux d'intérêt entraîne une hausse de la variation absolue de de l'épargne de β unités $(\beta > 0)$, puisque l'épargne est une fonction croissante du taux d'intérêt

· α : *Le montant d'épargne pour un taux d'intérêt nul (L'épargne minimum)*

b)

· ε_t : *Synthétise l'influence sur l'épargne de toutes les autres variables oubliées (l'évolution du revenu disponible, les prix des biens et services qui sont consommés, l'inflation, … etc.) mais aussi des erreurs éventuelles de spécification de la forme fonctionnelle dans le modèle spécifié par l'économiste.*

2)

· $\bar{R} = \frac{1}{T}\sum_{t=1}^{T} R_t = \frac{43,5}{10} = 4,35$ · $\bar{E} = \frac{1}{T}\sum_{t=1}^{T} E_t = \frac{435}{10} = 43,5$

· $\sum_{t=1}^{T}(R_t - \bar{R})^2 = \sum_{t=1}^{T} R_t^2 - T\bar{R}^2 = 215,25 - (10 \times 4,35^2) = 26,03$

· $\sum_{t=1}^{T}(E_t - \bar{E})^2 = \sum_{t=1}^{T} E_t^2 - T\bar{E}^2 = 20775 - (10 \times 43,5^2) = 1852,5$

· $\sum_{t=1}^{T}(R_t - \bar{R})(E_t - \bar{E}) = \sum_{t=1}^{T} R_t E_t - T\bar{R}\bar{E} = 2106,5 - (10 \times 4,35 \times 43,5) = 214,25$

· $\hat{\beta} = \frac{\sum_{t=1}^{T}(R_t - \bar{R})(E_t - \bar{E})}{\sum_{t=1}^{T}(R_t - \bar{R})^2} = \frac{214,25}{26,03} \Rightarrow \boxed{\hat{\beta} = 8,23}$

· $\hat{\alpha} = \bar{E} - \hat{\beta}\bar{R} = 43,5 - (8,23 \times 4,35) \Rightarrow \boxed{\hat{\alpha} = 7,7}$

3)

$$\cdot\ SCT = \sum_{t=1}^{T}(E_t - \bar{E})^2 = 1852,5 \qquad \cdot\ SCE = \hat{\beta}^2 \sum_{t=1}^{T}(R_t - \bar{R})^2 = 1763,09$$

$$\cdot\ SCR = SCT - SCE = 89,41$$

$$\cdot\ \hat{\sigma}^2 = \frac{SCR}{T-2} = \frac{89,41}{8} \Rightarrow \boxed{\hat{\sigma}^2 = 11,18 = 3,34^2}$$

4)

a)

$$\cdot\ \rho_{R,E} = \frac{\sum_{t=1}^{T}(R_t - \bar{R})(E_t - \bar{E})}{\sqrt{\sum_{t=1}^{T}(R_t - \bar{R})^2 \sum_{t=1}^{T}(E_t - \bar{E})^2}} = \frac{214,25}{\sqrt{26,03 \times 1852,5}} \Rightarrow \boxed{\rho_{R,E} = 0,98}$$

b)

$$\cdot\ R^2 = \rho_{R,E}^2 \Rightarrow \boxed{R^2 = 95,19\%}$$

5) *On se propose maintenant de tester :* $H_0 : \beta = 0$ *contre* $H_1 : \beta \neq 0$ *, sous un niveau de signification* $\alpha = 5\%$

$$\cdot\ T = \frac{\hat{\beta} - \beta}{\hat{\sigma}_{\hat{\beta}}} \backsim \mathcal{T}(10-2), \textit{la loi de Student de d.d.l } (8)$$

$$\cdot\ \hat{\sigma}_{\hat{\beta}}^2 = \hat{\sigma}^2 \Big/ \sum_{t=1}^{T}(R_t - \bar{R})^2 = 11,18/26,03 \Rightarrow \boxed{\hat{\sigma}_{\hat{\beta}}^2 = 0,43} \Rightarrow \boxed{\hat{\sigma}_{\hat{\beta}} = 0,66}$$

Règle de décision : on rejette H_0 *avec un risque de première espèce* α*, si* $|T_0| \geq t_{1-\frac{\alpha}{2}}(8)$

$$or \begin{cases} |T_0| = \left|\dfrac{\hat{\beta} - 0}{\hat{\sigma}_{\hat{\beta}}}\right| = \left|\dfrac{8,23}{0,66}\right| = 12,47 \\ t_{1-\frac{\alpha}{2}}(8) = t_{0,975}(8) = 2,306 \end{cases} \Rightarrow |T_0| \geq t_{1-\frac{\alpha}{2}}(8)$$

D'où la variable taux d'intérêt (R_t) *est statistiquement significative au risque de* 5%

6) *Sous l'hypothèse que le modèle reste inchangé, nous aurons :* $E_{11} = \alpha + \beta R_{11} + \varepsilon_{11}$

et $\hat{E}_{11} = \hat{\alpha} + \hat{\beta} R_{11}$ *Sera un estimateur sans biais de* E_{11}

$$\hat{E}_{11} = 7,7 + (8,23 \times 4) \Rightarrow \boxed{\hat{E}_{11} = 40,62}$$

$$\textit{L'intervalle de prévision} : IC_{95\%}(E_{11}) = \left[\hat{E}_{11} \pm t_{0,975}(8)\hat{\sigma}\sqrt{1 + \frac{1}{T} + \frac{(R_{11} - \bar{R})^2}{\sum_{t=1}^{T}(R_t - \bar{R})^2}}\right]$$

$$Avec : t_{0,975}(8) = 2,306\ , \hat{\sigma} = 3,34, \bar{R} = 4,35\ et\ \sum_{t=1}^{T}(R_t - \bar{R})^2 = 26,03$$

$$IC_{95\%}(E_{11}) = \left[40,62 \pm \left(2,306 \times 3,34 \times \sqrt{1 + \frac{1}{10} + \frac{(4-4,35)^2}{26,03}}\right)\right]$$

$$\boxed{IC_{95\%}(E_{11}) = [40,62 \pm 8,1] = [32,52\ ;\ 48,72]}$$

7)

a) *La méthode des moindres carrés ordinaire consiste à minimiser la somme des carrés des résidus* : $\min\limits_{\tilde{\beta}} \sum_{t=1}^{T} \tilde{\varepsilon}_t^2 = \min\limits_{\tilde{\beta}} \underbrace{\sum_{t=1}^{T}(E_t - \tilde{\beta}R_t)^2}_{\varphi(\tilde{\beta})}$

$$\cdot\ \varphi'(\tilde{\beta}) = \frac{d\sum_{t=1}^{T}(E_t - \tilde{\beta}R_t)^2}{d\tilde{\beta}} = \sum_{t=1}^{T}\frac{d(E_t - \tilde{\beta}R_t)^2}{d\tilde{\beta}} = \sum_{t=1}^{T} -2R_t(E_t - \tilde{\beta}R_t) = -2\sum_{t=1}^{T} R_t(E_t - \tilde{\beta}R_t)$$

$$\cdot\ \varphi''(\tilde{\beta}) = \frac{d\left[-2\sum_{t=1}^{T} R_t(E_t - \tilde{\beta}R_t)\right]}{d\tilde{\beta}} = -2\sum_{t=1}^{T}\frac{dR_t(E_t - \tilde{\beta}R_t)}{d\tilde{\beta}} = 2\sum_{t=1}^{T} R_t^2$$

• ***Condition nécessaire*** :

$$\varphi'(\tilde{\beta}) = 0 \Leftrightarrow -2\sum_{t=1}^{T} R_t(E_t - \tilde{\beta}R_t) = 0 \Leftrightarrow \sum_{t=1}^{T} R_tE_t - \tilde{\beta}\sum_{t=1}^{T} R_t^2 = 0 \Leftrightarrow \tilde{\beta} = \sum_{t=1}^{T} R_tE_t \Big/ \sum_{t=1}^{T} R_t^2$$

• ***Condition suffisante*** :

$$\varphi''(\tilde{\beta}) = 2\sum_{t=1}^{T} R_t^2 > 0$$

$$D'où\ \boxed{\tilde{\beta} = \frac{\sum_{t=1}^{T} R_tE_t}{\sum_{t=1}^{T} R_t^2}\ , minimise\ \sum_{t=1}^{T}\tilde{\varepsilon}_t^2}$$

$$\cdot\ \tilde{\beta} = \frac{\sum_{t=1}^{T} R_tE_t}{\sum_{t=1}^{T} R_t^2} = \frac{\sum_{t=1}^{T} R_t(\alpha + \beta R_t + \varepsilon_t)}{\sum_{t=1}^{T} R_t^2} = \frac{\alpha\sum_{t=1}^{T} R_t + \beta\sum_{t=1}^{T} R_t^2 + \sum_{t=1}^{T} R_t\varepsilon_t}{\sum_{t=1}^{T} R_t^2}$$

$$\tilde{\beta} = \beta + \frac{\alpha\sum_{t=1}^{T} R_t + \sum_{t=1}^{T} R_t\varepsilon_t}{\sum_{t=1}^{T} R_t^2} = \beta + \frac{\alpha\sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} + \frac{\sum_{t=1}^{T} R_t\varepsilon_t}{\sum_{t=1}^{T} R_t^2}$$

$$\cdot\ E(\tilde{\beta}) = E\left(\beta + \frac{\alpha\sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} + \frac{\sum_{t=1}^{T} R_t\varepsilon_t}{\sum_{t=1}^{T} R_t^2}\right) = E(\beta) + E\left(\frac{\alpha\sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2}\right) + E\left(\frac{\sum_{t=1}^{T} R_t\varepsilon_t}{\sum_{t=1}^{T} R_t^2}\right)$$

$$= \beta + \frac{\alpha \sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} + \frac{1}{\sum_{t=1}^{T} R_t^2} \sum_{t=1}^{T} E(R_t \varepsilon_t) = \frac{\alpha \sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} + \beta + \underbrace{\frac{1}{\sum_{t=1}^{T} R_t^2} \sum_{t=1}^{T} R_t \underbrace{E(\varepsilon_t)}_{0}}_{0}$$

$$E(\tilde{\beta}) = \beta + \frac{\alpha \sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} \Rightarrow E(\tilde{\beta}) \neq \beta. \boxed{D'\text{où } \tilde{\beta} \text{ estimateur biaisé de } \beta}$$

b)

$$\cdot\ \tilde{\beta} - E(\tilde{\beta}) = \left(\beta + \frac{\alpha \sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2} + \frac{\sum_{t=1}^{T} R_t \varepsilon_t}{\sum_{t=1}^{T} R_t^2}\right) - \left(\beta + \frac{\alpha \sum_{t=1}^{T} R_t}{\sum_{t=1}^{T} R_t^2}\right) = \frac{\sum_{t=1}^{T} R_t \varepsilon_t}{\sum_{t=1}^{T} R_t^2}$$

$$\cdot\ Var(\tilde{\beta}) = E\left[\left(\tilde{\beta} - E(\tilde{\beta})\right)^2\right] = E\left[\left(\frac{\sum_{t=1}^{T} R_t \varepsilon_t}{\sum_{t=1}^{T} R_t^2}\right)^2\right] = \frac{1}{\left(\sum_{t=1}^{T} R_t^2\right)^2} E\left[\left(\sum_{t=1}^{T} R_t \varepsilon_t\right)^2\right]$$

$$\left(\sum_{i=m}^{n} x_i\right)^2 = \sum_{i=m}^{n} x_i^2 + 2 \sum_{m \le i < j \le n} x_i x_j = \sum_{i=m}^{n} x_i^2 + 2 \sum_{i=m}^{n-1} \sum_{j=i+1}^{n} x_i x_j$$

$$On\ a: \left(\sum_{t=1}^{T} R_t \varepsilon_t\right)^2 = \sum_{t=1}^{T} x_i^2 + 2 \sum_{1 \le t < s \le T} (R_t \varepsilon_t)(R_s \varepsilon_s) = \sum_{t=1}^{T} (R_t \varepsilon_t)^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} (R_t \varepsilon_t)(R_s \varepsilon_s)$$

$$Par\ la\ suite:\ Var(\tilde{\beta}) = \frac{1}{\left(\sum_{t=1}^{T} R_t^2\right)^2} E\left[\sum_{t=1}^{T} (R_t \varepsilon_t)^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} (R_t \varepsilon_t)(R_s \varepsilon_s)\right]$$

$$= \frac{1}{\left(\sum_{t=1}^{T} R_t^2\right)^2} \left[\sum_{t=1}^{T} E[(R_t \varepsilon_t)^2] + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} E[(R_t \varepsilon_t)(R_s \varepsilon_s)]\right]$$

$$= \frac{1}{\left(\sum_{t=1}^{T} R_t^2\right)^2} \left[\sum_{t=1}^{T} R_t^2 \underbrace{E[\varepsilon_t^2]}_{\sigma^2} + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} R_t R_s \underbrace{E[\varepsilon_t \varepsilon_s]}_{0}\right] = \frac{\sigma^2 \sum_{t=1}^{T} R_t^2}{\left(\sum_{t=1}^{T} R_t^2\right)^2}$$

$$D'\text{où} \boxed{Var(\tilde{\beta}) = \frac{\sigma^2}{\sum_{t=1}^{T} R_t^2}}$$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XLème PROMOTION (BANQUE)
Octobre 2020

Exercice 11 : *(10 points :1 point par question)*

ÉNONCÉ

Pour un ensemble de n entreprises, on relie leurs profits y_i *(mesurés en logarithme* : $y_i = \ln(profit)$) *à leur niveau de vente* x_i *(mesurés en logarithme* : $x_i = \ln(vente)$) *selon l'équation* :

$$y_i = ax_i + b + u_i$$

Avec u_i *des termes d'erreurs indépendants tels que* $E(u_i) = 0$ *et* $V(u_i) = \sigma^2$

1. *Interpréter économiquement cette relation en précisant les signes attendus des paramètres a et b*
2. *Quelle signification économique revêt l'hypothèse* σ^2 *constante (valeur indépendante de i) ? Quelle hypothèse alternative peut-on envisager ?*
3. *Les observations relatives à* $n = 10$ *entreprises ont fourni les résultats suivants* :

$$\sum_{i=1}^{10} x_i = 40\ ;\ \sum_{i=1}^{10} y_i = 50\ ;\ \sum_{i=1}^{10} x_i^2 = 170\ ;\ \sum_{i=1}^{10} y_i^2 = 262\ ;\ \sum_{i=1}^{10} x_i y_i = 207$$

3-1 *Calculer* :

$$\sum_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})\ ;\ \sum_{i=1}^{10}(x_i-\overline{x})^2\ et\ \sum_{i=1}^{10}(y_i-\overline{y})^2$$

$$\text{où } \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ et\ \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

3-2 *Calculer le coefficient de corrélation linéaire entre les variables x et y*

3-3 *Calculer les estimations de a et de b par les moindres carrés ordinaires*

3-4 *Dresser le tableau analyse de la variance associé à cette régression*

3-5 *Calculer la variance estimée du paramètre a*

3-6 *La variable X est-elle significative ? Justifier votre réponse*

3-7 *Calculer le coefficient de détermination associé à ce modèle. Commenter ce résultat.*

3-8 *Que se passe-t-il au niveau de l'estimation du paramètre a si on oubli de mettre la constante b dans le modèle étudié ?*

Corrigé

1. *Le modèle décrit la variabilité des profits en fonction de la variabilité des niveaux de vente. Intuitivement la relation doit être croissante, en effet, toutes choses égales par ailleurs, il est normal que le profit augmente quand niveau de vente augmente, ce qui explique la positivité de la pente a*

· *a : Le modèle étant logarithmique, donc a sera l'élasticité du profit par rapport au niveau de vente*

· *b : Pour un niveau de vente réduit à l'unité, le profit serait égale à* e^b

2.

· *L'hypothèse d'homoscédasticité suppose une variance des termes d'erreur, constante et finie, il s'agit d'assumer que les variables explicatives omises dans le modèle influent toutes pratiquement de façon constante sur la variable expliquée.*
L'homoscédasticité nous épargne par conséquence de toute variation de de la variance de la variable que l'on souhaite expliquer et /ou prédire et qui est dépendante et aléatoire par le biais des termes d'erreurs.
· *Les profits des grandes entreprises seront sujets à des variations plus importantes que ceux des petites. L'hypothèse d'homoscédasticité ne semble donc pas convenir dans ce cas. L'hétéroscédasticité est une situation rencontrée fréquemment dans les données, c'est l'autre alternative qui sera confrontée à l'hypothèse d'homoscédasticité, Il existe toute une batterie de tests permettant de détecter l'hétéroscédasticité (test de White, test ARCH, … etc.)*

3.

3-1 · $n = 10$

· $$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{n}x_iy_i - n\bar{x}\bar{y} = \sum_{i=1}^{n}x_iy_i - n\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)$$

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{n}x_iy_i - \frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right) = 207 - \frac{40\times 50}{10}$$

$$\boxed{\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y}) = 7}$$

· $\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}x_i^2-n\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2=170-\frac{40^2}{10}$

$$\boxed{\sum_{i=1}^{10}(x_i-\bar{x})^2=10}$$

· $\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-n\bar{y}^2=\sum_{i=1}^{n}y_i^2-n\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)^2=\sum_{i=1}^{n}y_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2=262-\frac{50^2}{10}$

$$\boxed{\sum_{i=1}^{10}(y_i-\bar{y})^2=12}$$

3-2

$$r_{y,x}=\frac{S_{x,y}}{S_xS_y}\frac{\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{10}(x_i-\bar{x})^2\sum_{i=1}^{10}(y_i-\bar{y})^2}}=\frac{7}{\sqrt{10\times 12}}\Rightarrow\boxed{r_{y,x}=0,639}$$

3-3

· $\begin{cases}\hat{b}=\bar{y}-\hat{a}\bar{x}=\frac{1}{n}\left[\left(\sum_{i=1}^{n}y_i\right)-\hat{a}\left(\sum_{i=1}^{n}x_i\right)\right]=\frac{1}{10}[50-(0,7\times 40)]=2,2\\ \hat{a}=\frac{S_{x,y}}{S_x^2}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{7}{10}=0,7\end{cases}$

$$\boxed{\hat{a}=0,7\ et\ \hat{b}=2,2}$$

· ***L'élasticité du profit par rapport au niveau de vente est estimée à 0,7.***

Ainsi une augmentation du niveau de vente de 1% entraine une augmentation du profit de 0,7%, l'élasticité étant inférieure à l'unité donc la variation du profit est moins proportionnelle que la variation du niveau de vente.

· ***Pour un niveau de vente réduit à l'unité, le profit serait égale à*** $e^{2,2}\cong 9,025$

3-4

· $SCE=\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2=\hat{a}^2\sum_{i=1}^{n}(x_i-\bar{x})^2=(0,7)^2\times 10\Longrightarrow\boxed{SCE=4,9}$

· $SCT=\sum_{i=1}^{n}(y_i-\bar{y})^2\Longrightarrow\boxed{SCT=12}$

$\cdot\ SCR = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = SCT - SCE = 12 - 4,9 \Longrightarrow \boxed{SCR = 7,1}$

Source de la variabilité	Somme des carrés	Degré de liberté	Carrés moyens
Régression	$SCE = 4,9$	1	$\frac{SCE}{1} = 4,9$
Résidus	$SCR = 7,1$	$n-2=8$	$\hat{\sigma}^2 = \frac{SCR}{n-2} = 0,888$
Total	$SCT = 12$	$n-1=9$	$S^2 = \frac{SCT}{n-1} = 1,333$

3-5

$$\hat{\sigma}^2_{\hat{a}} = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{0,888}{10} \Rightarrow \boxed{\hat{\sigma}^2_{\hat{a}} = 0,089} \Rightarrow \boxed{\hat{\sigma}_{\hat{a}} = 0,298}$$

3-6

L'hypothèse de normalité des erreurs est un élément clé pour l'inférence statistique.

Elle est donc nécessaire pour mener les tests, des intervalles de confiance et de prévision

En effet, on doit imposer la normalité des termes d'erreurs !

$(u_i)_{1\le i\le n}$ ***termes d'erreurs i.i.d de la loi*** $\mathcal{N}(0,\sigma^2)$

Tester la significativité de la variable x (niveau de vente) revient au même de tester la significativité statistique du paramètre a .

En effet , on se propose de confronter l'hypothèse nulle $H_0 : a = 0$ ***, contre l'hypothèse alternative*** $H_a : a \neq 0$ ***, sous un risque de première espèce*** $\alpha = 5\%$

$\cdot$ ***Statistique de décision :*** $\frac{\hat{a} - a}{\hat{\sigma}_{\hat{a}}} \rightsquigarrow \mathcal{T}(n-2)$

$\cdot$ ***Règle de décision : on rejette*** H_0 ***si on observe*** $|T_0| > t_{1-\frac{\alpha}{2}}(n-2)$ ***, où*** $T_0 = \frac{\hat{a} - 0}{\hat{\sigma}_{\hat{a}}}$

$or\ |T_0| = \left|\frac{\hat{a}}{\hat{\sigma}_{\hat{a}}}\right| = \left|\frac{0,7}{0,298}\right| = 2,345\ et\ t_{1-\frac{\alpha}{2}}(n-2) = t_{0,975}(8) = 2.306$

$d'où\ |T_0| > t_{1-\frac{\alpha}{2}}(n-2)$ ***et a est significativement non nul au risque de*** 5%

En d'autres termes $\boxed{\text{la variable } x \text{ (niveau de vente) est significative avec un risque de première espèce } \alpha = 5\%}$

3-7

$$\cdot\ R^2 = \frac{SCE}{SCT} = r^2_{x,y} \Rightarrow \boxed{R^2 = 40,83\%}$$

La qualité de l'ajustement est moyenne, en effet 59,17% du profit reste inexpliqué par le modèle.

3-8

On va démontrer qu'en oubliant la constante b, on obtiendra un estimateur biaisé de a :

Régression sans constante : $(M_{sc})\ y_i = \alpha x_i + u_i \Rightarrow (\widehat{M}_{sc})\ \widehat{y}_i = \widehat{\alpha} x_i$

La MCO consiste à minimiser la somme des carrés des résidus :

$$\min_{\widehat{\alpha}} \sum_{i=1}^{n} \widehat{u}_i^2 = \min_{\widehat{\alpha}} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 = \min_{\widehat{\alpha}} \underbrace{\sum_{i=1}^{n} (y_i - \widehat{\alpha} x_i)^2}_{\varphi(\widehat{\alpha})}$$

$$\cdot\ \varphi'(\widehat{\alpha}) = \frac{d\sum_{i=1}^{n}(y_i - \widehat{\alpha}x_i)^2}{d\widehat{\alpha}} = \sum_{i=1}^{n} \frac{d(y_i - \widehat{\alpha}x_i)^2}{d\widehat{\alpha}} = \sum_{i=1}^{n} -2x_i(y_i - \widehat{\alpha}x_i) = -2\sum_{i=1}^{n} x_i(y_i - \widehat{\alpha}x_i)$$

$$\cdot\ \varphi''(\widehat{\alpha}) = \frac{d[-2\sum_{i=1}^{n} x_i(y_i - \widehat{\alpha}x_i)]}{d\widehat{\alpha}} = -2\sum_{t=1}^{T} \frac{dx_i(y_i - \widehat{\alpha}x_i)}{d\widehat{\alpha}} = 2\sum_{i=1}^{n} x_i^2$$

- *Condition nécessaire :*

$$\varphi'(\widehat{\alpha}) = 0 \Leftrightarrow -2\sum_{i=1}^{n} x_i(y_i - \widehat{\alpha}x_i) = 0 \Leftrightarrow \sum_{i=1}^{n} x_i y_i - \widehat{\alpha}\sum_{i=1}^{n} x_i^2 = 0 \Leftrightarrow \widehat{\alpha} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

- *Condition suffisante :* $\varphi''(\widehat{\alpha}) = 2\sum_{i=1}^{n} x_i^2 > 0$

$$D'où\ \boxed{\widehat{\alpha} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\ , minimise \sum_{i=1}^{n} \widehat{u}_i^2}$$

$$\cdot\ \widehat{\alpha} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i(ax_i + b + u_i)}{\sum_{i=1}^{n} x_i^2} = \frac{a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}$$

$$\widehat{\alpha} = a + \frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}$$

$$\cdot\ E(\widehat{\alpha}) = E\left(a + \frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = E(a) + E\left(\frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}\right) + E\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right)$$

$$= a + \frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} + \frac{1}{\sum_{i=1}^{n} x_i^2}\sum_{i=1}^{n} E(x_i u_i) = a + \frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} + \underbrace{\frac{1}{\sum_{i=1}^{n} x_i^2}\sum_{i=1}^{n} x_i \underbrace{E(u_i)}_{0}}_{0}$$

$$E(\widehat{\alpha}) = a + \frac{b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} \Rightarrow E(\widehat{\alpha}) \neq a.\ \boxed{D'où\ \widehat{\alpha}\ estimateur\ biaisé\ de\ a}$$

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXVIème PROMOTION (BANQUE)
AOÛT 2016

Exercice 12 : *(10 points : 1,5+1,5+1,5+1,5+2+2)*

ÉNONCÉ

On considère la régression entre le chiffre d'affaires y et le niveau d'investissement x observés sur un ensemble de $n = 20$ entreprises :

$$y_i = ax_i + b + \epsilon_i \text{ , pour } i = 1, 2, \dots, 20$$

où ϵ_i sont des termes aléatoires identiquement et indépendamment distribués selon une loi normale d'espérance mathématique nulle et de variance σ^2, a et b sont des paramètres à estimer. On dispose des statistiques suivantes : $\sum_{i=1}^{20} y_i = 198$; $\sum_{i=1}^{20} x_i = 210$

et $\sum_{i=1}^{20} (x_i - \overline{x})^2 = 665$; *où* $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ *est la moyenne empirique de x.*

Par ailleurs, $\widehat{a}$ l'estimation de a par la méthode des moindres carrés ordinaires et de sa sa variance $V(\widehat{a})$ sont égal à : $\widehat{a} = 0,8$ *et* $\widehat{V(\widehat{a})} = (0,01)^2$

1)
Rappeler les expressions mathématiques de $\widehat{a}$ et de $\widehat{b}$, les estimations de a et b par la méthode des moindres carrés ordinaires. En déduire la valeur de $\widehat{b}$

2)
Prouver que $\widehat{a}$ peut s'écrire sous la forme linéaire, $\widehat{a} = \sum_{i=1}^{n} \alpha_i y_i$ *avec α_i des pondérations vérifant :* $\sum_{i=1}^{n} \alpha_i = 0$ *et* $\sum_{i=1}^{n} \alpha_i x_i = 1$

3)

$En\ déduire\ que\ la\ variance\ de\ \hat{a}\ est\ V(\hat{a}) = \dfrac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$

4)

$Tester\ la\ significativité\ de\ la\ variable\ x\ à\ un\ niveau\ 95\%\ de\ significarivité\ .\ Interpréter\ ce\ résultat\ .\ On\ rappelle\ que\ pour\ S\ une\ variable\ de\ Student\ ,\ la\ probabilité\ P[-2 < S < 2]\ est\ approximativement\ égale\ à\ 0,95$

5)

$Déterminer\ la\ valeur\ de\ \sum_{i=1}^{20}(x_i - \overline{x})(y_i - \overline{y})\ ainsi\ que\ l'estimation\ de\ \sigma\ ,\ l'écarttype\ de\ \epsilon_i$

6)

$En\ déduire\ la\ somme\ des\ carrés\ des\ résidus\ \sum_{i=1}^{20}\hat{\epsilon}_i^2\ ainsi\ que\ le\ coefficient\ de\ détermination\ de\ la\ régression.$

<u>Corrigé</u>

1)

$Sous\ les\ hypothèses\ de\ la\ régression\ classique\ ,\ la\ méthode\ MCO\ ,\ qui\ consiste\ à\ minimiser\ la\ somme\ des\ carrés\ des\ résidus : \min_{\hat{b},\hat{a}}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \min_{\hat{b},\hat{a}}\sum_{i=1}^{n}\left(y_i - \hat{b} - \hat{a}x_i\right)^2\ ,fournira\ les\ estimations\ suivantes :$

$$\boxed{\begin{cases}\hat{b} = \overline{y} - \hat{a}\overline{x} = \dfrac{1}{n}\left[\left(\displaystyle\sum_{i=1}^{n} y_i\right) - \hat{a}\left(\displaystyle\sum_{i=1}^{n} x_i\right)\right] \\ \hat{a} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\end{cases}}$$

$$\begin{cases}\hat{b} = \overline{y} - \hat{a}\overline{x} = \dfrac{1}{n}\left[\left(\displaystyle\sum_{i=1}^{n} y_i\right) - \hat{a}\left(\displaystyle\sum_{i=1}^{n} x_i\right)\right] = \dfrac{1}{20}[\,198 - (0,8 \times 210)] = 1,5 \\ \hat{a} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0,8\end{cases} \Rightarrow \boxed{\begin{cases}\hat{b} = 1,5 \\ \hat{a} = 0,8\end{cases}}$$

2)

$On\ a : \ \hat{a} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\ ,notons\ S_x^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \dfrac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) - \overline{x}^2$

$ainsi\ ,nS_x^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 = \left(\sum_{i=1}^{n} x_i^2\right) - n\overline{x}^2$

$$donc\ , \hat{a} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n}[(x_i-\bar{x})y_i-(x_i-\bar{x})\bar{y}]}{nS_x^2} = \frac{[\sum_{i=1}^{n}(x_i-\bar{x})y_i]-[\sum_{i=1}^{n}(x_i-\bar{x})\bar{y}]}{nS_x^2}$$

$$= \left[\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{nS_x^2}\right)y_i\right] - \left[\frac{\bar{y}}{nS_x^2}\sum_{i=1}^{n}(x_i-\bar{x})\right] = \left[\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{nS_x^2}\right)y_i\right] - \frac{\bar{y}}{nS_x^2}\left[\left(\sum_{i=1}^{n}x_i\right)-\left(\sum_{i=1}^{n}\bar{x}\right)\right]$$

$$= \left[\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{nS_x^2}\right)y_i\right] - \frac{\bar{y}}{nS_x^2}\left[\left(\sum_{i=1}^{n}x_i\right)-n\bar{x}\right] = \left[\sum_{i=1}^{n}\underbrace{\left(\frac{x_i-\bar{x}}{nS_x^2}\right)}_{\alpha_i}y_i\right] - \frac{n\bar{y}}{nS_x^2}\underbrace{\left[\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)-\bar{x}\right]}_{0}$$

$$\boxed{\hat{a} = \sum_{i=1}^{n}\alpha_i y_i\ , \text{où}\ \alpha_i = \frac{x_i-\bar{x}}{nS_x^2} = \frac{x_i-\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \qquad (1)}$$

$$Vérifions\ maintenant\ que\ \sum_{i=1}^{n}\alpha_i = 0\ et\ \sum_{i=1}^{n}\alpha_i x_i = 1 :$$

$$\cdot\ \sum_{i=1}^{n}\alpha_i = \sum_{i=1}^{n}\frac{x_i-\bar{x}}{nS_x^2} = \frac{1}{nS_x^2}\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})}_{0}\quad d'où\ \boxed{\sum_{i=1}^{n}\alpha_i = 0 \qquad (2)}$$

$$\cdot\ \sum_{i=1}^{n}\alpha_i x_i = \sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{nS_x^2}\right)x_i = \frac{1}{nS_x^2}\sum_{i=1}^{n}(x_i-\bar{x})x_i = \frac{1}{nS_x^2}\left[\sum_{i=1}^{n}x_i^2 - \bar{x}\underbrace{\sum_{i=1}^{n}x_i}_{n\bar{x}}\right] = \frac{1}{nS_x^2}\left[\sum_{i=1}^{n}x_i^2 - n\bar{x}^2\right]$$

$$= \frac{n}{nS_x^2}\underbrace{\left[\frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2\right]}_{S_x^2} = \frac{S_x^2}{S_x^2} = 1$$

$$D'où\ \boxed{\sum_{i=1}^{n}\alpha_i x_i = 1 \qquad (3)}$$

$\color{red}{Conclusion}$: $\hat{a}\ peut\ s'écrire\ sous\ la\ forme\ linéaire, \hat{a} = \sum_{i=1}^{n}\alpha_i y_i$

$$avec\ \left(\alpha_i = \frac{x_i-\bar{x}}{nS_x^2} = \frac{x_i-\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)\ des\ pondérations\ vérifant :\ \sum_{i=1}^{n}\alpha_i = 0\ et\ \sum_{i=1}^{n}\alpha_i x_i = 1$$

3)

$$V(\hat{a}) = V\left(\sum_{i=1}^{n}\alpha_i y_i\right) = V\left[\sum_{i=1}^{n}\alpha_i(ax_i+b+\epsilon_i)\right] = V\left[a\underbrace{\left(\sum_{i=1}^{n}\alpha_i x_i\right)}_{1} + b\underbrace{\left(\sum_{i=1}^{n}\alpha_i\right)}_{0} + \left(\sum_{i=1}^{n}\alpha_i\epsilon_i\right)\right]$$

$$V(\hat{a}) = V\left[a + \underbrace{\left(\sum_{i=1}^{n} \alpha_i \epsilon_i\right)}_{\zeta_n}\right]$$

a étant non stochastique , de même pour les $(\alpha_i)_{1\le i\le n}$ *puisque* $\alpha_i = \frac{x_i - \overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$ *, donc elle ne dépend qur des variables exogènes non stochastiques.*

Par la suite $\zeta_n = \sum_{i=1}^{n} \alpha_i \epsilon_i$ *, est une combinaison linéaire d'une suite de v.a.i.i.d* $(\epsilon_i)_{1\le i\le n}$ *de la loi* $\mathcal{N}(0,\sigma^2)$ *, ce qui donne une suite de v.a.i.i.d* $(\alpha_i\epsilon_i)_{1\le i\le n}$ *de la loi* $\mathcal{N}\big(E(\alpha_i\epsilon_i), V(\alpha_i\epsilon_i)\big)$

On aura , donc , $V(\hat{a}) = V(a + \zeta_n) = V(\zeta_n) = V\left(\sum_{i=1}^{n} \alpha_i\epsilon_i\right) = \sum_{i=1}^{n} V(\alpha_i\epsilon_i) = \sum_{i=1}^{n} \alpha_i^2 \underbrace{V(\epsilon_i)}_{\sigma^2} = \sigma^2 \sum_{i=1}^{n} \alpha_i^2$

$$= \sigma^2 \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{nS_x^2}\right)^2 = \frac{\sigma^2}{n^2 S_x^4} \underbrace{\sum_{i=1}^{n}(x_i - \overline{x})^2}_{nS_x^2} = \frac{nS_x^2\sigma^2}{n^2S_x^4} = \frac{\sigma^2}{nS_x^2}$$

D'où $\boxed{V(\hat{a}) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}$

4)

Tester la significativité de la variable x (le niveau d'investissement) cela revient au même de tester la significativité statistique du paramètre a .

En effet , on se propose de confronter l'hypothèse nulle $H_0 : a = 0$ *, contre l'hypothèse alternative* $H_a : a \neq 0$ *, sous un risque de première espèce* $\alpha = 5\%$

· *Statistique de décision :* $\frac{\hat{a} - a}{\hat{\sigma}_{\hat{a}}} \rightsquigarrow \mathcal{T}(n-2)$

· *Règle de décision : on rejette* H_0 *si on observe* $|T_0| > t_{1-\frac{\alpha}{2}}(n-2)$ *, où* $T_0 = \frac{\hat{a} - 0}{\hat{\sigma}_{\hat{a}}}$

or $|T_0| = \left|\frac{\hat{a}}{\hat{\sigma}_{\hat{a}}}\right| = \left|\frac{0,8}{0,01}\right| = 80$ *et* $t_{1-\frac{\alpha}{2}}(n-2) = t_{0,975}(20-2) = t_{0,975}(18) \cong 2$

d'où $|T_0| > t_{1-\frac{\alpha}{2}}(n-2)$ *et a est significativement non nul au risque de 5%*

En d'autres termes $\boxed{\text{la variable } x \text{ (le niveau d'investissement) est significative avec un risque de première espèce } \alpha = 5\%}$

5)

· $\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y}) = \widehat{a}\sum_{i=1}^{20}(x_i-\overline{x})^2$ *provient du fait que* $\widehat{a} = \frac{\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{20}(x_i-\overline{x})^2}$

ainsi $\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y}) = 0,8\times 665$

$$\boxed{\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y}) = 532}$$

· *On a* $\widehat{V(\widehat{a})} = \frac{\widehat{\sigma}^2}{\sum_{i=1}^{20}(x_i-\overline{x})^2}$ *est un estimateur sans biais et convergent de* $V(\widehat{a})$

il en résulte l'expression de $\widehat{\sigma}^2$ *l'estimateurnon biaisé et convergent de la variance des erreurs* $\sigma^2 = V(\epsilon_i)$: $\widehat{\sigma}^2 = \widehat{V}(\epsilon_i) = \widehat{V(\widehat{a})}\sum_{i=1}^{20}(x_i-\overline{x})^2 = (0,01)^2\times 665 = 0,0665$

d'où l'estimation de σ *l'écart-type des* ϵ_i : $\boxed{\widehat{\sigma} = 0,2579}$

6)

· *On a* : $\widehat{\sigma}^2 = \frac{\sum_{i=1}^{20}\widehat{\epsilon}_i^2}{n-2} = \frac{SCR}{n-2} \Rightarrow SCR = (n-2)\widehat{\sigma}^2 = 18\times 0,0665$

$$\boxed{SCR = \sum_{i=1}^{20}\widehat{\epsilon}_i^2 = 1,197}$$

· *On a la somme des carrés expliqués* :

$$SCE = \sum_{i=1}^{20}\left(\widehat{y}_i-\overline{\widehat{y}}\right)^2 = \widehat{a}^2\sum_{i=1}^{20}(x_i-\overline{x})^2 = (0,8)^2\times 665 = 425,6$$

or l'équation d'analyse de la variance, nous donne : $SCT = SCE + SCR$ *, où* SCT *est la somme des carrés totaux* :

$$SCT = \sum_{i=1}^{20}(y_i-\overline{y})^2 = \widehat{a}^2\sum_{i=1}^{20}(x_i-\overline{x})^2 + \sum_{i=1}^{20}\widehat{\epsilon}_i^2 = 425,6 + 1,197 = 426,797$$

et comme on a le coefficient de détermination de la régression , $R^2 = \frac{SCE}{SCT}$ *, donc , on obtient* : $R^2 = \frac{425,6}{426,797}$

$\boxed{R^2 = 99,72\%}$

En effet il s'agit d'un meilleur ajustement. Le niveau d'investissement x contribue à une explication des chiffres d'affaires de l'ordre de 99,72% , ce qui nous permet de deviner avec précision les valeurs de y

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXVème PROMOTION
AOÛT 2015

Exercice 13 : (12 points : 1+2+3+2+2+2)

ÉNONCÉ

On considère les données relatives à un échantillon de $n = 103$ individus pour lesquels on observe le logarithme du revenu (Y), le nombre d'années de formation (X), et une variable indicatrice qui vaut 1 si l'individu est occupé et 0 si non (W)

On considère le modèle suivant :

$$y_i = \beta_1 x_i + \beta_2 w_i + \varepsilon_i$$

Où ε_i est un terme d'erreur vérifiant les hypothèses de la méthode des moindres carrés ordinaires. La variance de ε_i est notée σ^2. Les variables minuscules correspondent aux données centrées :

$x_i = X_i - \overline{X}$ où $\overline{X} = \dfrac{\sum_i X_i}{n}$; *de même pour w_i et y_i*

1) *Que représentes les paramètres β_1 et β_2 en termes d'analyse économique ?*
2) *Déterminer $\widehat{\beta}$, l'estimateur par la méthode des moindres carrés ordinaires du vecteur des paramètres $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$. Calculer l'espérance mathématique et la variance de $\widehat{\beta}$*
3) *Calculer la valeur de l'estimateur de β, sachant que :*

$$\sum_{i=1}^{103} x_i y_i = 14 \, , \sum_{i=1}^{103} w_i y_i = 2 \, , \sum_{i=1}^{103} y_i^2 = 1$$

$$\sum_{i=1}^{103} x_i^2 = 577\,, \sum_{i=1}^{103} w_i x_i = 18\,, \sum_{i=1}^{103} w_i^2 = 20$$

Commenter .

4) *Dresser le tableau d'analyse de la variance et calculer le coefficient de détermination linéaire du modèle.*
5) *On suppose que la variable âge de l'individu (A), une variable proxy de l'expérience professionnelle, a été omise du modèle. Evaluer le biais d'omission de cette variable si le vrai modèle était :*

$$y_i = \theta_1 x_i + \theta_2 w_i + \theta_3 a_i + \epsilon_i$$

6) *L'estimation de ce nouveau modèle a donné le résultat suivant :*

$$\hat{\theta} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_3 \end{pmatrix} = \begin{pmatrix} 0,03 \\ 0,16 \\ 0,01 \end{pmatrix}$$

Commenter .

NB : Arrondir les calculs à deux chiffres après la virgule

Corrigé

1)

- β_1: *mesure La contribution de la durée (en années) de formation à l'amélioration des revenus, ainsi une année de plus de formation (toutes choses égales par ailleurs) entraine une augmentation de* $100\beta_1$ *unités des revenus.* β_1 *mesure aussi la semi-élasticité des revenus par rapport aux nombres d'annéesde formation*
- β_2: *si ce paramètre est jugé statistiquement significatif, on en conclurait que le caractère (individu est occupé) est un facteur discriminant*

2)

- *Il s'agit d'une régression avec des données centrées, à cet effet* $\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c)$

$$\text{Où } X_c'X_c = n\begin{pmatrix} S_X^2 & S_{X,W} \\ S_{X,W} & S_W^2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n}(X_i-\bar{X})^2 & \sum_{i=1}^{n}(X_i-\bar{X})(W_i-\bar{W}) \\ \sum_{i=1}^{n}(X_i-\bar{X})(W_i-\bar{W}) & \sum_{i=1}^{n}(W_i-\bar{W})^2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} w_i x_i \\ \sum_{i=1}^{n} w_i x_i & \sum_{i=1}^{n} w_i^2 \end{pmatrix}$$

$$et\ X_c'Y_c = n\begin{pmatrix} S_{X,Y} \\ S_{Z,Y} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y}) \\ \sum_{i=1}^{n}(W_i-\overline{W})(Y_i-\overline{Y}) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} w_i y_i \end{pmatrix}$$

· *Sous les hypothèses, la MCO fournit un estimateur sans biais de β qui n'est autre que $\widehat{\beta}$*

ainsi $E(\widehat{\beta}) = \beta$ *ou encore* $\boxed{E\left(\begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix}\right) = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}}$

· *Elle fournit aussi un vecteur estimé de variance minimale (qui converge vers β):*

$$\boxed{Var(\widehat{\beta}) = \Omega_{\widehat{\beta}} = \sigma^2_{\widehat{\beta}} = \sigma^2 (X_c'Y_c)^{-1} = \begin{pmatrix} \sigma^2_{\widehat{\beta}_1} & Cov(\widehat{\beta}_1,\widehat{\beta}_2) \\ Cov(\widehat{\beta}_1,\widehat{\beta}_2) & \sigma^2_{\widehat{\beta}_2} \end{pmatrix}}$$

3)

$$X_c'X_c = \begin{pmatrix} \sum_{i=1}^{103} x_i^2 & \sum_{i=1}^{103} w_i x_i \\ \sum_{i=1}^{103} w_i x_i & \sum_{i=1}^{103} w_i^2 \end{pmatrix} = \begin{pmatrix} 577 & 18 \\ 18 & 20 \end{pmatrix}$$

$$|X_c'X_c| = 11216\ ;\ Com(X_c'X_c) = \begin{pmatrix} 20 & -18 \\ -18 & 577 \end{pmatrix}$$

$$\Rightarrow (X_c'X_c)^{-1} = \frac{1}{|X_c'X_c|}[Com(X_c'X_c)]' = \frac{1}{11216}\begin{pmatrix} 20 & -18 \\ -18 & 577 \end{pmatrix}$$

$$X_c'Y_c = \begin{pmatrix} \sum_{i=1}^{103} x_i y_i \\ \sum_{i=1}^{103} w_i y_i \end{pmatrix} = \begin{pmatrix} 14 \\ 2 \end{pmatrix}$$

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = (X_c'X_c)^{-1}(X_c'Y_c) = \frac{1}{11216}\begin{pmatrix} 20 & -18 \\ -18 & 577 \end{pmatrix}\begin{pmatrix} 14 \\ 2 \end{pmatrix} = \frac{1}{11216}\begin{pmatrix} 244 \\ 902 \end{pmatrix} = \begin{pmatrix} 0,02 \\ 0,08 \end{pmatrix}$$

$$\boxed{\widehat{\beta}_1 = 0,02\ et\ \widehat{\beta}_2 = 0,08}$$

· $\widehat{\beta}_1$: *Une année de plus de formation fait progresser les revenus de 2%*

· $\widehat{\beta}_2$: *Un individu est occupé voit ses revenus augmenter de* $\left((e^{\widehat{\beta}_2} \times 100) - 100\right)\%$

$$= \left((e^{0,08} \times 100) - 100\right)\% = 8,33\%$$

4) $SCE = \hat{\beta}'(X_c'X_c)\hat{\beta} = \hat{\beta}'X_c'Y_c = \frac{1}{11216}(244 \quad 902)\binom{14}{2} \Rightarrow \boxed{SCE = 0,47}$

$$SCT = \sum_{i=1}^{103}(Y_i - \bar{Y})^2 = \sum_{i=1}^{103} y_i^2 = 1 \Rightarrow \boxed{SCT = 1}$$

$SCR = SCT - SCE \Rightarrow \boxed{SCR = 0,53}$

Tableau d'analyse de variance

Source de la variabilité	Somme des carrés	Degré de liberté	Carrés moyens
Régression	$SCE = 0,47$	$k = 3$	$SCE/3 = 0,16$
Résidus	$SCR = 0,53$	$n - k = 103 - 3 = 100$	$\hat{\sigma}^2 = SCR/n - k = 0,053$
Total	$SCT = 1$	$n = 103$	$S^2 = SCT/n = 0,01$

$R^2 = \frac{SCE}{SCT} = \frac{0,47}{1} \Rightarrow \boxed{R^2 = 47\%}$

La qualité d'ajustement linéaire est faible.

5)

Le recours à une variable proxy (âge de l'individu: A) qui nous permet d'approcher l'expérience professionnelle indispensable pour l'explication du modèle et dont on éprouve des difficultés à mesurer.

Cette variable proxy nous permet de corriger ce biais de spécification néanmoins elle reste corrélée avec les autres variables explicatives ce qui va causer des biais d'estimations, des intervalles de confiance , tests d'hypothèse et prévisions éventuellement erronées .

- $Biais(\hat{\theta}_1) = \begin{cases} E(\hat{\theta}_1) - \theta_1 = \theta_3 \frac{S_{X,A}}{S_X^2} \text{ , Si A et X sont corrélés} \\ 0 \text{ , Si non} \end{cases}$

	$S_{X,A} > 0$	$S_{X,A} < 0$
$\theta_3 > 0$	*Biais positif*	*Biais négatif*
$\theta_3 < 0$	*Biais négatif*	*Biais positif*

- $Biais(\hat{\theta}_2) = \begin{cases} E(\hat{\theta}_2) - \theta_2 = \theta_3 \frac{S_{W,A}}{S_W^2} \text{ , Si A et W sont corrélés} \\ 0 \text{ , Si non} \end{cases}$

	$S_{W,A} > 0$	$S_{W,A} < 0$
$\theta_3 > 0$	*Biais positif*	*Biais négatif*
$\theta_3 < 0$	*Biais négatif*	*Biais positif*

6)

- $S_{X,A} > 0$ *Puisque l'âge de l'individu (A) et le nombre d'années de formation (X) sont* *en général positivement corrélés , comme on a* $\widehat{\theta}_3 > 0$ *donc il s'agit d'un biais positif*

et le modèle était sous-dimensionné.

Cependant ce biais négligeable puisque l'effet partiel de A sur Y est négligeable

$\widehat{\theta}_3 = 0,01 \cong 0$

· *En supposant valable, l'estimation du vrai modèle , on aura :*

- $\widehat{\theta}_1$: *Une année de plus de formation fait progresser les revenus de* 3%
- $\widehat{\theta}_2$: *Un individu est occupé voit ses revenus augmenter de* $\left((e^{\widehat{\theta}_2} \times 100) - 100\right)\%$
  $= \left((e^{0,16} \times 100) - 100\right)\% = 17,35\%$
- $\widehat{\theta}_3$: *les revenus augmentent avec l'âge de l'individu (l'expérience professionnelle)*

voir 1% *pour une année de plus*

Institut de Financement du Développement du Maghreb Arabe
CONCOURS DE RECRUTEMENT DE LA XXXIXème PROMOTION (ASSURANCE)
Juin 2022

Exercice 14 : (8 points : 1,5+1,5+1,5+1,5+2)

ÉNONCÉ

Considérons un modèle, appelé le vrai modèle (VM) ayant $K = 2$ *variables explicatives centrées sans constante :*

$$y_i = a_1 x_{1i} + a_2 x_{2i} + u_i \,.$$

Avec u_i *des termes d'erreur ayant les propriétés classiques indépendantes d'espérances nulles et de variances* σ^2, *pour* $i = 1, 2, \dots, n$..

L'erreur de l'analyste (l'économètre) a conduit à omettre (oublier) la variable x_2 *du vrai modèle, ce qui donne un modèle erroné, appelé le faux modèle (FM) défini par :*

$$y_i = \alpha x_{1i} + v_i$$

1- *Déterminer* $\widehat{\alpha}$ *l'estimateur par les MCO dans le faux modèle*
2- *Calculer l'espérance mathématique de* $\widehat{\alpha}$ *dans le vrai modèle*

3- *En déduire que cette estimation est biaisée. Dans quel cas ce biais est-il-nul ? Interpréter ce résultat.*
4- *Calculer la variance du coefficient de la variable x_1, dans le faux modèle.*
5- *Comparer cette variance alors à la vraie valeur de la variance du coefficient de x_1 dans le modèle à deux variables*

Corrigé

▪ *Initialement le vrai modèle non centré (VMNC) était :* $\mathcal{y}_i = a_0 + a_1\mathcal{x}_{1i} + a_2\mathcal{x}_{2i} + \epsilon_i$

L'écriture matricielle sera : $\mathcal{Y} = \mathcal{X}\mathcal{B} + \mathcal{E}$, *où* $\mathcal{Y} = \begin{pmatrix} \mathcal{y}_1 \\ \mathcal{y}_2 \\ \vdots \\ \mathcal{y}_n \end{pmatrix}$, $\mathcal{X} = \begin{pmatrix} 1 & \mathcal{x}_{11} & \mathcal{x}_{21} \\ 1 & \mathcal{x}_{12} & \mathcal{x}_{22} \\ \vdots & \vdots & \vdots \\ 1 & \mathcal{x}_{1n} & \mathcal{x}_{2n} \end{pmatrix}$, $\mathcal{B} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ *et*

$$\mathcal{E} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

En sommant sur les i, on obtient :

$$\sum_{i=1}^{n} \mathcal{y}_i = \sum_{i=1}^{n} a_0 + a_1 \sum_{i=1}^{n} \mathcal{x}_{1i} + a_2 \sum_{i=1}^{n} \mathcal{x}_{2i} + \sum_{i=1}^{n} \epsilon_i \Longrightarrow n\bar{\mathcal{y}} = na_0 + na_1\bar{\mathcal{x}}_1 + na_2\bar{\mathcal{x}}_2 + n\bar{\epsilon}$$

$$\Longrightarrow (\overline{VMNC}): \bar{\mathcal{y}} = a_0 + a_1\bar{\mathcal{x}}_1 + a_2\bar{\mathcal{x}}_2 + \bar{\epsilon}$$

En soustrayant $(\overline{VMNC})$ *à (VMNC), on a finalement :*

$$\underbrace{(\mathcal{y}_i - \bar{\mathcal{y}})}_{y_i} = a_1 \underbrace{(\mathcal{x}_{1i} - \bar{\mathcal{x}}_1)}_{x_{1i}} + a_2 \underbrace{(\mathcal{x}_{2i} - \bar{\mathcal{x}}_2)}_{x_{2i}} + \underbrace{(\epsilon_i - \bar{\epsilon})}_{u_i}$$

D'où le vrai modèle centré sans constante (VM): $y_i = a_1 x_{1i} + a_2 x_{2i} + u_i$

L'écriture matricielle sera : $Y = X\beta + U$, *où* $Y = \begin{pmatrix} \mathcal{y}_1 - \bar{\mathcal{y}} \\ \mathcal{y}_2 - \bar{\mathcal{y}} \\ \vdots \\ \mathcal{y}_n - \bar{\mathcal{y}} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$,

$$X = \begin{pmatrix} (\mathcal{x}_{11} - \bar{\mathcal{x}}_1) & (\mathcal{x}_{21} - \bar{\mathcal{x}}_2) \\ (\mathcal{x}_{12} - \bar{\mathcal{x}}_1) & (\mathcal{x}_{22} - \bar{\mathcal{x}}_2) \\ \vdots & \vdots \\ (\mathcal{x}_{1n} - \bar{\mathcal{x}}_1) & (\mathcal{x}_{2n} - \bar{\mathcal{x}}_2) \end{pmatrix} = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{pmatrix}, \beta = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \text{ et } U = \begin{pmatrix} \epsilon_1 - \bar{\epsilon} \\ \epsilon_2 - \bar{\epsilon} \\ \vdots \\ \epsilon_n - \bar{\epsilon} \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

▪ *Initialement le Faux modèle non centré (FMNC) était :* $\mathcal{y}_i = c + \alpha\mathcal{x}_{1i} + \varepsilon_i$

En sommant sur les i, on obtient :

$$\sum_{i=1}^{n} \mathcal{y}_i = \sum_{i=1}^{n} c + \alpha \sum_{i=1}^{n} \mathcal{x}_{1i} + \sum_{i=1}^{n} \varepsilon_i \Longrightarrow n\bar{\mathcal{y}} = nc + n\alpha\bar{\mathcal{x}}_1 + n\bar{\varepsilon}$$

$$\Longrightarrow (\overline{FMNC}): \bar{\mathcal{y}} = c + \alpha\bar{\mathcal{x}}_1 + \bar{\varepsilon}$$

En soustrayant $(\overline{FMNC})$ *à (FMNC), on a finalement :*

$$\underbrace{(y_i-\bar{y})}_{y_i}=\alpha\underbrace{(x_{1i}-\bar{x}_1)}_{x_{1i}}+\underbrace{(\varepsilon_i-\bar{\varepsilon})}_{v_i}$$

D'où le faux modèle centré sans constante (FM): $y_i=\alpha x_{1i}+v_i$

1-

(FM): $y_i=\alpha x_{1i}+v_i=\hat{\alpha}x_{1i}+\hat{v}_i$

L'estimateur MCO de α *minimise la somme des carrées résiduelles*

$$SCR_{FM}=\sum_{i=1}^{n}\hat{v}_i^2=\sum_{i=1}^{n}(y_i-\hat{\alpha}x_{1i})^2$$

$$\boxed{\hat{\alpha}=\frac{\sum_{i=1}^{n}x_{1i}y_i}{\sum_{i=1}^{n}x_{1i}^2}}$$

2-

- *Dans le* (VM) : $\hat{\alpha}=\dfrac{\sum_{i=1}^{n}x_{1i}(a_1x_{1i}+a_2x_{2i}+u_i)}{\sum_{i=1}^{n}x_{1i}^2}=\dfrac{a_1\sum_{i=1}^{n}x_{1i}^2+a_2\sum_{i=1}^{n}x_{1i}x_{2i}+\sum_{i=1}^{n}x_{1i}u_i}{\sum_{i=1}^{n}x_{1i}^2}$

$$\hat{\alpha}=\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}u_i\right]+a_1+\frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}$$

- $E(\hat{\alpha})=E\left[\left[\left(\dfrac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}u_i\right]+a_1+\dfrac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}\right]$

$$=\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)E\left(\sum_{i=1}^{n}x_{1i}u_i\right)+a_1+\frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}$$

$$=\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}E(x_{1i}u_i)\right]+a_1+\frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}$$

$$=\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}\underbrace{E(u_i)}_{0}\right]+a_1+\frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}$$

$$\boxed{E(\hat{\alpha})=a_1+\frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}}$$

3-

- $E(\hat{\alpha})=a_1+\dfrac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}\neq a_1$

$\boxed{\hat{\alpha}\ \textit{est biaisée de}\ a_1}$

- $\hat{\alpha}$ *est sans biais de* $a_1 \Leftrightarrow Biais(\hat{\alpha})=0\Leftrightarrow E(\hat{\alpha})-a_1=0\Leftrightarrow\dfrac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}=0$

$$\Leftrightarrow\sum_{i=1}^{n}x_{1i}x_{2i}=0$$

$Or\ rappelons\ que\ x_1\ et\ x_2\ sont\ des\ variables\ centrées\ donc\ avoir \sum_{i=1}^{n} x_{1i}x_{2i} = 0$

$revient\ au\ même\ d'avoir \sum_{i=1}^{n}(x_{1i}-\overline{x}_1)(x_{2i}-\overline{x}_2) = 0\ ou\ encore\ S_{x_1,x_2} = 0$

$(avec : x_{1i} = x_{1i} - \overline{x}_1\ et\ x_{2i} = x_{2i} - \overline{x}_2)$

$L'absence\ de\ corrélations\ entre\ les\ deux\ variables\ explicatives\ x_1\ et\ x_2\ entraine\ un\ biais\ nul$

$On\ rappelle\ l'expression\ du\ biais\ d'omission :$

$$\text{Biais}(\widehat{\alpha}) = E(\widehat{\alpha}) - a_1 = \frac{a_2 S_{x_1,x_2}}{S_{x_1}^2} = \frac{a_2\sum_{i=1}^{n}(x_{1i}-\overline{x}_1)(x_{2i}-\overline{x}_2)}{\sum_{i=1}^{n}(x_{1i}-\overline{x}_1)^2} = \frac{a_2\sum_{i=1}^{n}x_{1i}x_{2i}}{\sum_{i=1}^{n}x_{1i}^2}$$

4-

$Calculons\ la\ variance\ (\widehat{\alpha}) dans\ (FM): y_i = \alpha x_{1i} + v_i$

- $Dans\ le\ (FM) :\ \widehat{\alpha} = \frac{\sum_{i=1}^{n}x_{1i}y_i}{\sum_{i=1}^{n}x_{1i}^2} = \frac{\sum_{i=1}^{n}x_{1i}(\alpha x_{1i}+v_i)}{\sum_{i=1}^{n}x_{1i}^2} = \frac{\alpha\sum_{i=1}^{n}x_{1i}^2 + \sum_{i=1}^{n}x_{1i}v_i}{\sum_{i=1}^{n}x_{1i}^2}$

$$\widehat{\alpha} = \left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}v_i\right] + \alpha$$

- $Pour\ le\ (FM)\ on\ vérifie\ bien\ que\ \widehat{\alpha}\ est\ sans\ biais\ de\ \alpha :$

- $E_{FM}(\widehat{\alpha}) = E\left[\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}v_i\right] + \alpha\right] = \alpha + \left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)E\left(\sum_{i=1}^{n}x_{1i}v_i\right)$

$$E_{FM}(\widehat{\alpha}) = \alpha + \frac{1}{\sum_{i=1}^{n}x_{1i}^2}\sum_{i=1}^{n}E(x_{1i}v_i) = \alpha + \frac{1}{\sum_{i=1}^{n}x_{1i}^2}\sum_{i=1}^{n}x_{1i}\underbrace{E(v_i)}_{0}$$

$$\boxed{E_{FM}(\widehat{\alpha}) = \alpha}$$

- $V_{FM}(\widehat{\alpha}) = V\left[\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}v_i\right] + \alpha\right] = V\left[\left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)\sum_{i=1}^{n}x_{1i}v_i\right] = \left(\frac{1}{\sum_{i=1}^{n}x_{1i}^2}\right)^2 V\left(\sum_{i=1}^{n}x_{1i}v_i\right)$

$$= \frac{1}{\left(\sum_{i=1}^{n}x_{1i}^2\right)^2}\left[\sum_{i=1}^{n}V(x_{1i}v_i) + 2\sum_{1\le i<j\le n}Cov(x_{1i}v_i, x_{1j}v_j)\right]$$

$$= \frac{1}{\left(\sum_{i=1}^{n}x_{1i}^2\right)^2}\left[\sum_{i=1}^{n}x_{1i}^2\underbrace{V(v_i)}_{\sigma_v^2} + 2\sum_{1\le i<j\le n}x_{1i}x_{1j}\underbrace{Cov(v_i, v_j)}_{0}\right] = \frac{\sigma_v^2\sum_{i=1}^{n}x_{1i}^2}{\left(\sum_{i=1}^{n}x_{1i}^2\right)^2}$$

$$\boxed{V_{FM}(\widehat{\alpha}) = \sigma_{\widehat{\alpha}}^2 = \frac{\sigma_v^2}{\sum_{i=1}^{n}x_{1i}^2} = \frac{\sigma_v^2}{nS_{x_1}^2}}$$

5-

- $(VM): y_i = a_1x_{1i} + a_2x_{2i} + u_i \Rightarrow$ *L'écriture matricielle sera* : $Y = X\beta + U$

  - $$X'X = \begin{pmatrix} \sum_{i=1}^{n} x_{1i}^2 & \sum_{i=1}^{n} x_{1i}x_{2i} \\ \sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{2i}^2 \end{pmatrix} \Rightarrow \begin{cases} Com(X'X) = \begin{pmatrix} \sum_{i=1}^{n} x_{2i}^2 & -\sum_{i=1}^{n} x_{1i}x_{2i} \\ -\sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{1i}^2 \end{pmatrix} \\ |X'X| = \left(\sum_{i=1}^{n} x_{1i}^2\right)\left(\sum_{i=1}^{n} x_{2i}^2\right) - \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)^2 \end{cases}$$

$$(X'X)^{-1} = \frac{1}{|X'X|}[Com(X'X)]' = \frac{1}{(\sum_{i=1}^{n} x_{1i}^2)(\sum_{i=1}^{n} x_{2i}^2) - (\sum_{i=1}^{n} x_{1i}x_{2i})^2}\begin{pmatrix} \sum_{i=1}^{n} x_{2i}^2 & -\sum_{i=1}^{n} x_{1i}x_{2i} \\ -\sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{1i}^2 \end{pmatrix}$$

  - $$Or \begin{cases} \sigma_{\hat\beta}^2 = \sigma_u^2(X'X)^{-1} = \dfrac{\sigma_u^2}{(\sum_{i=1}^{n} x_{1i}^2)(\sum_{i=1}^{n} x_{2i}^2) - (\sum_{i=1}^{n} x_{1i}x_{2i})^2}\begin{pmatrix} \sum_{i=1}^{n} x_{2i}^2 & -\sum_{i=1}^{n} x_{1i}x_{2i} \\ -\sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{1i}^2 \end{pmatrix} \\ \sigma_{\hat\beta}^2 = \begin{pmatrix} \sigma_{\hat a_1}^2 & Cov(\hat a_1, \hat a_2) \\ Cov(\hat a_1, \hat a_2) & \sigma_{\hat a_2}^2 \end{pmatrix} \end{cases}$$

Par identification, on obtient :

$$V_{VM}(\hat a_1) = \sigma_{\hat a_1}^2 = \frac{\sigma_u^2 \sum_{i=1}^{n} x_{2i}^2}{(\sum_{i=1}^{n} x_{1i}^2)(\sum_{i=1}^{n} x_{2i}^2) - (\sum_{i=1}^{n} x_{1i}x_{2i})^2} = \frac{n\sigma_u^2 S_{x_2}^2}{(nS_{x_1}^2)(nS_{x_2}^2) - (nS_{x_1,x_2})^2}$$

$$V_{VM}(\hat a_1) = \sigma_{\hat a_1}^2 = \frac{n\sigma_u^2 S_{x_2}^2}{n^2(S_{x_1}^2 S_{x_2}^2 - S_{x_1,x_2}^2)} = \frac{\sigma_u^2 S_{x_2}^2}{n(S_{x_1}^2 S_{x_2}^2 - S_{x_1,x_2}^2)} = \frac{\sigma_u^2 S_{x_2}^2}{nS_{x_1}^2 S_{x_2}^2}\left(\frac{1}{1 - S_{x_1,x_2}^2/S_{x_1}^2 S_{x_2}^2}\right)$$

$$V_{VM}(\hat a_1) = \sigma_{\hat a_1}^2 = \frac{\sigma_u^2}{nS_{x_1}^2}\left(\frac{1}{1 - (S_{x_1,x_2}/S_{x_1}S_{x_2})^2}\right) = \frac{\sigma_u^2}{nS_{x_1}^2}\left(\frac{1}{1 - (S_{x_1,x_2}/S_{x_1}S_{x_2})^2}\right) = \frac{\sigma_u^2}{nS_{x_1}^2}\left(\frac{1}{1 - r_{x_1,x_2}^2}\right)$$

$$\boxed{V_{VM}(\hat a_1) = \sigma_{\hat a_1}^2 = \frac{\sigma_u^2}{nS_{x_1}^2(1 - r_{x_1,x_2}^2)}}$$

La variance de l'erreur est constante et ne dépend pas de l'observation. C'est l'hypothèse d'homoscédasticité des erreurs.

Il s'agit d'assumer que les variables explicatives omises dans le modèle influent toutes pratiquement de façon constante sur la variable expliquée.

Ce qui conduit à: $\sigma_v^2 = \sigma_u^2 = \sigma^2$

$$Par\ la\ suite : \begin{cases} V_{FM}(\widehat{\alpha}) = \sigma^2_{\widehat{\alpha}} = \dfrac{\sigma^2}{\sum_{i=1}^n x_{1i}^2} = \dfrac{\sigma^2}{nS^2_{x_1}} \\ V_{VM}(\widehat{a}_1) = \sigma^2_{\widehat{a}_1} = \dfrac{\sigma^2}{nS^2_{x_1}\left(1-r^2_{x_1,x_2}\right)} = \dfrac{1}{1-r^2_{x_1,x_2}}\left(\dfrac{\sigma^2}{nS^2_{x_1}}\right) = \dfrac{V_{FM}(\widehat{\alpha})}{1-r^2_{x_1,x_2}} \end{cases}$$

$$et\ \frac{V_{FM}(\widehat{\alpha})}{V_{VM}(\widehat{a}_1)} = \frac{V_{FM}(\widehat{\alpha})}{V_{FM}(\widehat{\alpha})/1-r^2_{x_1,x_2}} = V_{FM}(\widehat{\alpha})\left(\frac{1-r^2_{x_1,x_2}}{V_{FM}(\widehat{\alpha})}\right) = 1-r^2_{x_1,x_2} \le 1$$

$D'où\ \boxed{V_{FM}(\widehat{\alpha}) \le V_{VM}(\widehat{a}_1)}$

En conclusion l'omission d'une variable explicative importante conduit à un biais d'omission et l'estimateur MCO du coefficient de la variable x_1 dans le (FM) est de variance minimale par rapport à sa variance dans le (VM)
On démontre aussi qu'elle reste minimale même en la remplaçant dans le (VM),

autrement dit : $\boxed{V_{VM}(\widehat{\alpha}) \le V_{VM}(\widehat{a}_1)}$

Au fur et à mesure que le coefficient de corrélation entre les deux variables explicatives x_1 et x_2 devient plus élevé $\left(r_{x_1,x_2} \nearrow \Longrightarrow r^2_{x_1,x_2} \nearrow \Longrightarrow \left(1-r^2_{x_1,x_2}\right) \searrow\right)$ la variance du coefficient de la variable x_1 dans le (VM) va augmenter.

- ***Seule l'absence de corrélation entre x_1 et $x_2\left(r_{x_1,x_2}=0\right)$ qui conduit à l'égalité :***

$V_{FM}(\widehat{\alpha}) = V_{VM}(\widehat{a}_1)$

- ***Dans le (VM) :*** $\widehat{\alpha} = \left[\left(\dfrac{1}{\sum_{i=1}^n x_{1i}^2}\right)\sum_{i=1}^n x_{1i}u_i\right] + a_1 + \dfrac{a_2\sum_{i=1}^n x_{1i}x_{2i}}{\sum_{i=1}^n x_{1i}^2}$

- $E_{VM}(\widehat{\alpha}) = a_1 + \dfrac{a_2\sum_{i=1}^n x_{1i}x_{2i}}{\sum_{i=1}^n x_{1i}^2} \neq a_1$

- $V_{VM}(\widehat{\alpha}) = V\left[\left(\left(\dfrac{1}{\sum_{i=1}^n x_{1i}^2}\right)\sum_{i=1}^n x_{1i}u_i\right) + \underbrace{a_1 + \dfrac{a_2\sum_{i=1}^n x_{1i}x_{2i}}{\sum_{i=1}^n x_{1i}^2}}_{non-aléatoire}\right] = V\left[\left(\dfrac{1}{nS^2_{x_1}}\right)\sum_{i=1}^n x_{1i}u_i\right]$

$$V_{VM}(\widehat{\alpha}) = \left(\frac{1}{nS^2_{x_1}}\right)^2 V\left(\sum_{i=1}^n x_{1i}u_i\right) = \frac{1}{n^2S^4_{x_1}}\left[\sum_{i=1}^n V(x_{1i}u_i) + 2\sum_{1\le i<j\le n} Cov(x_{1i}u_i\,, x_{1j}u_j)\right]$$

$$V_{VM}(\widehat{\alpha}) = \frac{1}{n^2S^4_{x_1}}\left[\sum_{i=1}^n x_{1i}^2\underbrace{V(u_i)}_{\sigma^2} + 2\sum_{1\le i<j\le n} x_{1i}x_{1j}\underbrace{Cov(u_i\,,u_j)}_{0}\right] = \frac{\sigma^2\sum_{i=1}^n x_{1i}^2}{n^2S^4_{x_1}} = \frac{n\sigma^2S^2_{x_1}}{n^2S^4_{x_1}}$$

$$\boxed{V_{VM}(\widehat{\alpha}) = \frac{\sigma^2}{nS^2_{x_1}}}$$

$$\begin{cases} V_{VM}(\hat{\alpha}) = \sigma^2_{\hat{\alpha}} = \dfrac{\sigma^2}{\sum_{i=1}^{n} x^2_{1i}} = \dfrac{\sigma^2}{nS^2_{x_1}} \\ V_{VM}(\hat{a}_1) = \sigma^2_{\hat{a}_1} = \dfrac{V_{FM}(\hat{\alpha})}{1 - r^2_{x_1,x_2}} \end{cases}$$

$$\frac{V_{FM}(\hat{\alpha})}{V_{VM}(\hat{a}_1)} = \frac{V_{FM}(\hat{\alpha})}{V_{FM}(\hat{\alpha})/1 - r^2_{x_1,x_2}} = V_{FM}(\hat{\alpha})\left(\frac{1 - r^2_{x_1,x_2}}{V_{FM}(\hat{\alpha})}\right) = 1 - r^2_{x_1,x_2} \le 1 \Longrightarrow \boxed{V_{VM}(\hat{\alpha}) \le V_{VM}(\hat{a}_1)}$$

BEN AHMED MOHSEN Téléphone: (+216) 97 619191 / 54 619191
omega.center.cp@gmail.com https://web.facebook.com/OMEGACENTER2014

Axe ⑧ : Table des matières

Adresse : 2, Rue Ibn Arafa, av de l'environnement Manouba Centre, Manouba