Modèles Linéaires 2A
Période 1

Josephson Junior R.

July 9, 2024

Table des matières

La fonction de consommation

Comme vue en micro et macroéconomie on peut définir la fonction de consommation par :

$$\mathbf{C} = \mathbf{a_0} + \mathbf{a_1}\mathbf{Y}$$

où :

- $\mathbf{C}$: la consommation
- $\mathbf{a_0}$: consommation autonome incompressible
- $\mathbf{a_1}$: propension marginale à consommer
- $\mathbf{Y}$: le revenu

En modèle modèle linéaire on change juste d'appellation :

- $\mathbf{C}$: variable à expliquer ou endogène
- $\mathbf{Y}$: variable explicative ou exogène
- $\mathbf{a_0}$ et $\mathbf{a_1}$: coefficients ou paramètres du modèle

La fonction de production

Pour rappel en microéconomie on a parlé des fonctions type **Cobb-Douglas** qui modèle le niveau de production :

$$\mathbf{Y} = \mathbf{a} \times \mathbf{K}^{\alpha} \times \mathbf{L}^{\beta}$$

Par une transformation logarithmique on peut revenir à un modèle de regression simple défini comme suit :

$$\mathbf{y} = \mathbf{ln}(\mathbf{a}) + \alpha\ \mathbf{ln}(\mathbf{K}) + \beta\ \mathbf{ln}(\mathbf{L})$$

Plus simplement on pourra écrire :

$$\mathbf{y} = \beta_{\mathbf{0}} + \beta_{\mathbf{1}}\ \mathbf{K}^{*} + \beta_{\mathbf{2}}\ \mathbf{L}^{*}$$

NB : l'interprétation des coefficients du modèle diffère selon les **transformations** appliquées pour avoir une régression.

Modèle de régression simple

Un modèle de régression simple se définit comme suit :

$$\mathbf{y_i} = \beta_\mathbf{0} + \beta_\mathbf{1}\ \mathbf{x_i} + \varepsilon_\mathbf{i} \quad ; \quad \mathbf{i} \in [\mathbf{1}, \mathbf{n}]$$

Ce modèle doit vérifier des hypothèses :

- **H1** : $\mathbf{E}(\varepsilon_\mathbf{i}) = \mathbf{0}$
- **H2** : $\mathbf{V}(\varepsilon_\mathbf{i}) = \sigma_\varepsilon^\mathbf{2}$
- **H3** : $\mathbf{Cov}(\varepsilon_\mathbf{i}, \varepsilon_\mathbf{i'}) = \mathbf{0}$
- **H4** : $\mathbf{Cov}(\mathbf{x_i}, \varepsilon_\mathbf{i}) = \mathbf{0}$

Les problèmes liés aux non-respects de ces hypothèses seront traités ultérieurement en **économétrie**.

Formulation des estimateurs

L'estimateur des paramètres de régression est obtenu par la MCO (**M**oindres **C**arrés **O**rdinaires) définie comme suite :

$$\mathbf{Min}_{\beta_0,\beta_1} \sum_{\mathbf{i=1}}^{\mathbf{n}} \varepsilon_{\mathbf{i}}^{\mathbf{2}} = \mathbf{Min\ E} \quad \textbf{où} \quad \varepsilon_{\mathbf{i}} = \mathbf{y_i} - \beta_{\mathbf{0}} - \beta_{\mathbf{1}}\ \mathbf{x_i}$$

Par la suite on calcule :

$$\frac{\partial E}{\partial \beta_0} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1\ x_i) = 0 \quad (1)$$

$$\frac{\partial E}{\partial \beta_1} = -2 \sum_{i=1}^{n} x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1\ x_i) = 0 \quad (2)$$

Tout calcul fait on aura :

$$\hat{\beta}_1 = \frac{\sum(\mathbf{y_i} - \bar{\mathbf{y}})(\mathbf{x_i} - \bar{\mathbf{x}})}{\sum(\mathbf{x_i} - \bar{\mathbf{x}})^2} = \frac{\sum \mathbf{x_i y_i} - \mathbf{n\bar{x}\bar{y}}}{\sum \mathbf{x_i^2} - \mathbf{n\bar{x}^2}} = \frac{\mathbf{Cov(x, y)}}{\mathbf{V(x)}}$$

$$\hat{\beta}_0 = \bar{\mathbf{y}} - \hat{\beta}_1 \bar{\mathbf{x}}$$

Le modèle de prédiction s'écrit comme suit :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \ x_i$$

On introduit la notion de résidu d'estimation :

$$\hat{\varepsilon}_i = y_i - \hat{y}_i$$

- Les estimateurs sont-t-ils **sans biais** ?
  On commence par reécrire $\hat{\beta}_1$ en remarquant que :

$$(y_i - \bar{y}) = \beta_1(x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon}) \Rightarrow \hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon})}{\sum(x_i - \bar{x})^2}$$

$$\mathbf{E}\left(\hat{\beta}_1\right) = \mathbf{E}(\beta_1) + \frac{\sum x_i \ \mathbf{E}(\varepsilon_i) - n\bar{x} \ \mathbf{E}\left(\bar{\varepsilon}\right)}{\sum(x_i - \bar{x})^2} = \beta_1$$

$$E\left(\hat{\beta}_0\right) = E\left(\bar{y} - \hat{\beta}_1\bar{x}\right) = E\left(\beta_0 + \beta_1\bar{x} + \bar{\varepsilon} - \hat{\beta}_1\bar{x}\right) = \beta_0$$

  Donc les estimateurs obtenus par MCO sont **sans biais**.
- Les estimateurs sont-ils **convergents** ?

$$\mathbf{V}\left(\hat{\beta}_1\right) = \frac{\sigma_\varepsilon^2}{\sum(x_i - \bar{x})^2} \quad \textbf{et} \quad \mathbf{V}\left(\hat{\beta}_0\right) = \sigma_\varepsilon^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right)$$

En calculant la limite des variances des estimateurs on voit bien que **ça tend vers 0** donc les estimateurs obtenus par MCO sont **convergents**.

- Les estimateurs par MCO sont **BLUE** (**B**est **L**inear **U**nbiaised **E**stimator).
- Comment estimer la variance des erreurs ?
  Il faut partir de l'expression des résidus :

$$\hat{\varepsilon}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 \bar{x} = y_i - \bar{y} - \hat{\beta}_1 (x_i - \bar{x})$$

$$\Rightarrow \hat{\varepsilon}_i = \left(\beta_1 - \hat{\beta}_1\right)(x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon})$$

Après on fait $\sum \hat{\varepsilon}_i^2$ et on calcule sont espérance pour arriver à une estimation de la variance des erreurs :

$$\hat{\sigma}_\varepsilon^2 = \frac{\mathbf{1}}{\mathbf{n-2}} \sum_{\mathbf{i=1}}^{\mathbf{n}} \hat{\varepsilon}_{\mathbf{i}}^{\mathbf{2}}$$

Significativité individuelle

$$\begin{cases} H0 & : & \beta_1 = 0 \\ H1 & : & \beta_1 \neq 0 \end{cases}$$

Sous l'hypothèse de normalité des erreurs $\varepsilon_\mathbf{i} \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^\mathbf{2})$; on a la statistique pivotale suivante :

$$T = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim \mathcal{N}(0, 1)$$

On distingue deux cas :

- σ_ε^2 connue :

$$T = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim \mathcal{N}(0, 1)$$

- σ_ε^2 inconnue :

$$T = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim \mathcal{T}(n-2)$$

Sous H0 la statistique du test est défini par :

$$\mathbf{T} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}} \quad \textbf{où} \quad \hat{\sigma}_{\hat{\beta}_1} = \sqrt{\hat{\mathbf{V}}\left(\hat{\beta}_1\right)} = \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{\sum(\mathbf{x_i} - \bar{\mathbf{x}})^2}}$$

Sa loi dépend des cas cités précedemment.
Règle de décision : .

$$\textbf{si} \quad |\mathbf{T}| < \mathbf{t}_{\mathbf{n-2}}^{\alpha/2} \quad \textbf{on accepte H0}$$

Sous H0 l'intervalle de confiance est défini comme suit :

$$\mathbf{IC}_{\beta_1}^{\alpha} = \hat{\beta}_{\mathbf{1}} \quad \pm \quad \mathbf{t}_{\mathbf{n-2}}^{\mathbf{1-\alpha/2}} \times \hat{\sigma}_{\hat{\beta}_{\mathbf{1}}}$$

Règle de décision :

$$\textbf{si} \quad \mathbf{0} \in \mathbf{IC}_{\beta_1}^{\alpha} \qquad \textbf{on accepte H0}$$

NB : Le calcul des estimateurs peut aussi se faire par MV (**M**aximum **V**raisemblance) et donc :

$$\hat{\beta}_1 \;\sim\; \mathcal{N}\left(\beta_1, V\left(\hat{\beta}_1\right)\right)$$

Equation d'analyse de la variance

La décomposition de la variance se fait par :

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}\hat{\varepsilon}_i^2$$

$$\mathbf{SCT} = \mathbf{SCE} + \mathbf{SCR}$$

On peut ainsi définir le coefficient de détermination : la proportion de la variabilité de Y expliqué par la variable X.

$$\mathbf{R}^2 = \frac{\mathbf{SCE}}{\mathbf{SCT}} = \mathbf{1} - \frac{\mathbf{SCR}}{\mathbf{SCT}}$$

On peut aussi reécrire l'estimateur de la variance des erreurs comme suit :

$$\hat{\sigma}_{\varepsilon}^{\mathbf{2}} = \frac{\mathbf{SCR}}{\mathbf{n - 2}}$$

Tableau d'analyse de la variance

Source de variation	Somme des carrés	ddl	Carré moyens
x	SCE	1	SCE/1
Résidu	SCR	n-2	SCR/n-2
Total	SCT	n-1	

Pourquoi avons nous besoin de ce tableau ?
En fait ce tableau nous permet d'avoir une autre statistique pour notre test de significativité. Dans le prochain nous allons plus entrer en détails.

$$\textbf{Sous H0}: \quad \mathbf{F} = \frac{\mathbf{SCE/1}}{\mathbf{SCR/n-2}} \sim \mathcal{F}(\mathbf{1}, \mathbf{n-2})$$

Cette statistique n'a pas réelement de nouveauté car c'est juste un rearrangement :

$$\mathbf{F} = \mathbf{T}^2 = \left(\frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}\right)^2 = \frac{\hat{\beta}_1^2 \sum(\mathbf{x_i} - \bar{\mathbf{x}})^2}{\hat{\sigma}_\varepsilon^2} = \frac{\mathbf{SCE/1}}{\mathbf{SCR/n-2}}$$

En fonction du coefficient de détermination on aura :

$$\mathbf{F} = \frac{\mathbf{R}^2}{(1 - \mathbf{R}^2)/(\mathbf{n} - 2)}$$

Règle de décision :

$$\textbf{si} \quad \mathbf{F} < \mathbf{F}^{\alpha/2}_{(1,\mathbf{n}-2)} \quad \textbf{on accepte H0}$$

NB : La partie prévision est juste une application des estimations dejà vu ; en outre certaines formules d'intervalle de confiance doivent être apprises .

Soit le modèle suivant :

$$\mathbf{Y_i} = \beta_\mathbf{0} + \beta_\mathbf{1}\mathbf{X_{1i}} + \beta_\mathbf{2}\mathbf{X_{2i}} + ... + \beta_\mathbf{k}\mathbf{X_{ki}} + \varepsilon_\mathbf{i}$$

Que l'on peut écrire sous forme matricielle :

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \Rightarrow \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Les dimensionnalités sont :

$$(\mathbf{n}, \mathbf{1}) = (\mathbf{n}, \mathbf{k} + \mathbf{1})(\mathbf{k} + \mathbf{1}, \mathbf{1}) + (\mathbf{n}, \mathbf{1})$$

Moindres Carrés Ordinaires

$$\textbf{Min } \varepsilon'\varepsilon \quad \textbf{où} \quad \varepsilon = \mathbf{Y} - \mathbf{X}\beta \Rightarrow \textbf{Min } (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$$

Pour trouver l'estimateur :

$$\frac{\partial E}{\partial \beta} = -2X'Y + 2X'X\hat{\beta} = 0 \Rightarrow \hat{\beta} = (\mathbf{X'X})^{-1}(\mathbf{X'Y})$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} n & \sum X_{1i} & \dots & \sum X_{ki} \\ \sum X_{1i} & \sum X_{1i}^2 & \dots & \sum X_{ki}X_{1i} \\ \vdots & \vdots & \vdots & \vdots \\ \sum X_{ki} & \sum X_{1i}X_{ki} & \dots & \sum X_{ki}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum Y_i \\ \sum Y_i X_{1i} \\ \vdots \\ \sum Y_i X_{ki} \end{pmatrix}$$

Cas particuliers : Données centrées

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} V(x_1) & Cov(x_1, x_2) & \dots & Cov(x_1, x_k \\ Cov(x_2, x_1) & V(x_1) & \dots & Cov(x_2, x_k) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_k, x_1) & cov(x_k, x_2) & \dots & V(x_k) \end{pmatrix}^{-1} \begin{pmatrix} Cov(y, x_1) \\ Cov(y, x_2) \\ \vdots \\ Cov(y, x_k) \end{pmatrix}$$

Avec :

$$\hat{\beta}_\mathbf{0} = \bar{\mathbf{Y}} - \hat{\beta}_\mathbf{1}\bar{\mathbf{X}}_\mathbf{1} - \hat{\beta}_\mathbf{2}\bar{\mathbf{X}}_\mathbf{2} - \dots - \hat{\beta}_\mathbf{k}\bar{\mathbf{X}}_\mathbf{k}$$

De plus on rajoute une hypothèse structurelle :

- **H5 :** La matrice $(X'X)$ est de plein rang.

Les mêmes propriétes que celles mentionnées dans le chapitre précèdent; juste un petit changement dans l'écriture.

- Variance des estimateurs :

$$\mathbf{V}(\hat{\beta}) = \sigma_{\varepsilon}^{\mathbf{2}}(\mathbf{X'X})^{-\mathbf{1}}$$

- Estimateur de la variance des erreurs :

$$\hat{\sigma}_{\varepsilon}^{\mathbf{2}} = \frac{\mathbf{SCR}}{\mathbf{n - k - 1}}$$

Dans certaines ouvrages on préfère directement noter : $\mathbf{K} = \mathbf{k} + \mathbf{1}$ donc la formule précédente devient :

$$\hat{\sigma}_{\varepsilon}^{\mathbf{2}} = \frac{\mathbf{SCR}}{\mathbf{n - K}}$$

En écriture matricielle on a :

$$\mathbf{Y'Y} = \hat{\beta}'\mathbf{X'X}\hat{\beta} + \hat{\varepsilon}'\hat{\varepsilon} \Rightarrow \mathbf{SCT} = \mathbf{SCE} + \mathbf{SCR}$$

Lorsque le degré de liberté est faible càd le nombre d'observation n'est pas assez grand il est préférable d'utiliser le R^2 ajusté :

$$\mathbf{R^2_{adj}} = \mathbf{1} - \frac{\mathbf{n-1}}{\mathbf{n-K}}(\mathbf{1} - \mathbf{R^2})$$

Source de variation	Somme des carrés	ddl	Carré moyens
X	SCE	k	SCE/k
Résidu	SCR	n-k-1	SCR/n-k-1
Total	SCT	n-1	

Statistique : $\mathbf{F} = \dfrac{\mathbf{SCE/k}}{\mathbf{SCR/n-k-1}} \sim \mathcal{F}(\mathbf{k}, \mathbf{n-k-1})$

Significativité individuelle

$$\begin{cases} H0 & : & \beta_k = 0 \\ H1 & : & \beta_k \neq 0 \end{cases}$$

Sous H0 la statistique de test est définie par :

$$T = \frac{\hat{\beta}_k}{\hat{\sigma}_{\hat{\beta}_k}} \sim \mathcal{T}(n-K)$$

Règle de décision :

$$\mathbf{si} \quad |\mathbf{T}| < \mathbf{t}_{\mathbf{n-K}}^{\alpha/\mathbf{2}} \quad \textbf{on accepte H0}$$

Intervalle de confiance :

$$\mathbf{IC}_{\beta_\mathbf{k}}^{\alpha} = \hat{\beta}_\mathbf{k} \quad \pm \quad \mathbf{t}_{\mathbf{n-K}}^{\mathbf{1}-\alpha/\mathbf{2}} \times \hat{\sigma}_{\hat{\beta}_\mathbf{k}}$$

Significativité globale

$$\begin{cases} H0 & : \quad \beta_1 = \beta_2 = \ldots = \beta_k = 0 \\ H1 & : \quad \beta_j \neq 0 \;\; \exists\, j \in [1, k] \end{cases}$$

Sous H0 la statistique pour le test est définie par :

$$\mathbf{F} = \frac{\mathbf{SCE/k}}{\mathbf{SCR/n-K}} \;\sim\; \mathcal{F}(\mathbf{k}, \mathbf{n}-\mathbf{K})$$

Règle de décision :

$$\textbf{si} \;\; \mathbf{F} < \mathbf{F}^{\alpha/2}_{(\mathbf{k},\mathbf{n}-\mathbf{K})} \;\; \textbf{on accepte H0}$$