TD 3

Exercice 1 :

1- Propriétés des estimateurs des MCO :

a- Les estimateurs des MCO des coefficients du modèle linéaire sont sans biais.

b- Ces estimateurs sont linéaires par rapport aux variables dépendantes (généralement notées y_i)

c- Ils sont de variance minimale dans l'ensemble des estimateurs linéaires sans biais.

2- Modèle linéaire simple : $y_t = ax_t + b + \varepsilon_t \;\; ; \;\; t = 1, 2, ..., T$

L'estimateur des MCO du paramètre a est :

$$\hat{a} = \frac{\displaystyle\sum_{t=1}^{T} x_t y_t - T\bar{x}\bar{y}}{\displaystyle\sum_{t=1}^{T} x_t^2 - T\bar{x}^2} \tag{1}$$

où $\bar{x}$ (resp. $\bar{y}$) est la moyenne empirique des x_t (resp. y_t) :

$$\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t \quad et \quad \bar{y}_t = \frac{1}{T}\sum_{t=1}^{T} x_t$$

Posons $S_{xy} = \displaystyle\sum_{t=1}^{T} x_t y_t - T\bar{x}\bar{y}$ et $S_{xx} = \displaystyle\sum_{t=1}^{T} x_t^2 - T\bar{x}^2$. On a alors, d'après ce qui précède:

$$\hat{a} = \frac{S_{xy}}{S_{xx}}$$

Notons $cov(x, y)$ la covariance **empirique** entre $(x_1, x_2, ..., x_T)$ et $(y_1, y_2, ..., y_T)$, et $V(x)$ la variance **empirique** de $(x_1, x_2, ..., x_T)$. On a :

$$cov(x, y) = \frac{1}{T}\sum_{t=1}^{T} x_t y_t - \bar{x}\bar{y} = \frac{1}{T} S_{xy}$$

$$V(x) = \frac{1}{T}\sum_{t=1}^{T} x_t^2 - \bar{x}^2 = \frac{1}{T} S_{xx}$$

ce qui nous fournit une seconde expression pour $\hat{a}$:

$$\hat{a} = \frac{cov(x, y)}{V(x)} \tag{2}$$

L'estimateur des MCO de b est :

$$\hat{b} = \bar{y} - \hat{a}\bar{x} \tag{3}$$

Les variances des estimateurs $\hat{a}$ et $\hat{b}$ sont données par :

$$V(\hat{a}) = \frac{\sigma^2}{\sum_{t=1}^{T}(x_t - \bar{x})^2} = \frac{\sigma^2}{\sum_{t=1}^{T}x_t^2 - T\bar{x}^2} \tag{4}$$

$$V(\hat{b}) = \frac{\sigma^2}{T}\frac{\sum_{t=1}^{T}x_t^2}{\sum_{t=1}^{T}(x_t - \bar{x})^2} = \frac{\sigma^2}{T}\frac{\sum_{t=1}^{T}x_t^2}{\sum_{t=1}^{T}x_t^2 - T\bar{x}^2} \tag{5}$$

où $\sigma^2 = V(\varepsilon_t)$.

Modèle linéaire multiple :

Posons $Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_T \end{pmatrix}$ $X = \begin{pmatrix} x_{11} & x_{21} & . & . & x_{p1} \\ x_{12} & & & & . \\ . & & & & . \\ . & & & & . \\ x_{1T} & . & . & . & x_{pT} \end{pmatrix}$ $a = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_p \end{pmatrix}$ et $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_T \end{pmatrix}$

Le modèle linéaire multiple peut s'écrire sous la forme :

$$y_t = \sum_{j=1}^{p} a_j x_{jt} + \varepsilon_t \quad ; \quad t = 1, 2, ..., T$$

ou sous la forme matricielle suivante :

$$Y = Xa + \varepsilon$$

L'estimateur des MCO du paramètre multi-dimensionnel a est :

$$\hat{a} = (X'X)^{-1} X'Y \tag{6}$$

et sa variance est :

$$V(\hat{a}) = \sigma^2 (X'X)^{-1} \tag{7}$$

où $\sigma^2 = V(\varepsilon_t)$.

3- Dans le modèle linéaire, les variables dépendantes y_t sont des variables aléatoires (elles sont généralement obtenues par échantillonnage aléatoire dans une population). Les estimateurs des MCO sont des fonctions de ces variables, et par conséquent sont également des variables aléatoires. L'une des conséquences de ce caractère aléatoire est qu'il va falloir s'intéresser aux lois de ces estimateurs ou du moins à leur espérance et à leur variance. A titre d'exemple, dans le modèle linéaire simple, sous l'hypothèse supplémentaire de normalité des résidus ε_t, l'estimateur des MCO $\hat{a}$ suit la loi $N(a, V(\hat{a}))$ où l'expression de $V(\hat{a})$ est donnée par (4). Toujours sous l'hypothèse de normalité des résidus dans le modèle linéaire simple, l'estimateur des MCO de la variance σ^2 des résidus (dont on rappelle l'expression : $\hat{\sigma}^2 = \frac{1}{T-2}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$ avec $\hat{\varepsilon}_t = y_t - \hat{a}x_t + \hat{b}$) suit, à une constante multiplicative près que nous précisons c–après, une loi du $\chi_2(T-2)$:

$$(T-2)\,\frac{\hat{\sigma}^2}{\sigma^2} \rightsquigarrow \chi_2(T-2)$$

Dans le cadre du modèle linéaire multiple détaillé ci-dessus, l'estimateur $\hat{a}$ suit la loi normale multidimensionnelle $N_p(a, V(\hat{a}))$ où l'expression de $V(\hat{a})$ est donnée par (7), et l'estimateur de la variance σ^2 dont l'expression est $\hat{\sigma}^2 = \frac{1}{T-p}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$ où $\hat{\varepsilon}_t = y_t - \sum_{j=1}^{p}\hat{a}_j x_{jt}$, suit, à une constante multiplicative près, une loi du $\chi_2(T-p)$:

$$(T-p)\,\frac{\hat{\sigma}^2}{\sigma^2} \rightsquigarrow \chi_2(T-p)$$

Exercice 3 (Compléments : résolution des deux équations normales et démonstration de la formule des estimateurs MCO) :

On considère le modèle linéaire suivant :

$$C_t = a_1 Y_t + a_2 + \varepsilon_t$$

Dire que la fonction estimée $C = \hat{a}_1 Y + \hat{a}_2$ passe par le point moyen de l'échantillon $(\bar{Y}, \bar{C})$ revient à dire que:

$$\bar{C} = \hat{a}_1 \bar{Y} + \hat{a}_2$$

La procédure d'estimation par les moindres carrés linéaires consiste à minimiser par rapport au couple (a_1, a_2) la somme des résidus au carré:

$$\sum_{t=1}^{T}\varepsilon_t^2 = \sum_{t=1}^{T}\left(C_t - a_1 Y_t - a_2\right)^2$$

Le couple d'estimateurs des MCO $(\hat{a}_1, \hat{a}_2)$ est donc solution des deux conditions du premier ordre suivantes (qu'on appelle équations normales) :

$$\frac{\partial}{\partial a_1}\left(\sum_{t=1}^{T}\left(C_t - a_1 Y_t - a_2\right)^2\right) = 0 \tag{8}$$

$$\frac{\partial}{\partial a_2}\left(\sum_{t=1}^{T}(C_t - a_1 Y_t - a_2)^2\right) = 0 \tag{9}$$

On a : $\frac{\partial}{\partial a_1}\left((C_t - a_1 Y_t - a_2)^2\right) = -2Y_t\left(C_t - a_1 Y_t - a_2\right)$ donc

$$\frac{\partial}{\partial a_1}\left(\sum_{t=1}^{T}(C_t - a_1 Y_t - a_2)^2\right) = -2\sum_{t=1}^{T}\left(Y_t C_t - a_1 Y_t^2 - a_2 Y_t\right)$$

et $\frac{\partial}{\partial a_2}\left((C_t - a_1 Y_t - a_2)^2\right) = -2\left(C_t - a_1 Y_t - a_2\right)$ donc

$$\frac{\partial}{\partial a_2}\left(\sum_{t=1}^{T}(C_t - a_1 Y_t - a_2)^2\right) = -2\sum_{t=1}^{T}\left(C_t - a_1 Y_t - a_2\right)$$

Dire que $(\hat{a}_1, \hat{a}_2)$ est solution du système des deux équations normales (8) et (9) équivaut donc à dire que

$$\sum_{t=1}^{T}\left(Y_t C_t - \hat{a}_1 Y_t^2 - \hat{a}_2 Y_t\right) = 0 \tag{10}$$

$$\sum_{t=1}^{T}\left(C_t - \hat{a}_1 Y_t - \hat{a}_2\right) = 0 \tag{11}$$

L'équation (11) est équivalente à : $\sum_{t=1}^{T} C_t = \hat{a}_1 \sum_{t=1}^{T} Y_t + T\hat{a}_2$. En divisant par T les deux membres de cette équation, on trouve le résultat demandé, à savoir :

$$\bar{C} = \hat{a}_1 \bar{Y} + \hat{a}_2 \tag{12}$$

qui peut également s'écrire :

$$\hat{a}_2 = \bar{C} - \hat{a}_1 \bar{Y} \tag{13}$$

Quant à l'équation (10), on commence par la réécrire sous la forme:

$$\sum_{t=1}^{T} Y_t C_t - \hat{a}_1 \sum_{t=1}^{T} Y_t^2 - \hat{a}_2 \sum_{t=1}^{T} Y_t = 0$$

puis on remplace $\hat{a}_2$ par son expression en fonction de $\hat{a}_1$, à savoir $\bar{C} - \hat{a}_1 \bar{Y}$; on obtient alors :

$$\sum_{t=1}^{T} Y_t C_t - \hat{a}_1 \sum_{t=1}^{T} Y_t^2 - \left(\bar{C} - \hat{a}_1 \bar{Y}\right) \sum_{t=1}^{T} Y_t = 0$$

soit

$$\sum_{t=1}^{T} Y_t C_t - \bar{C} \sum_{t=1}^{T} Y_t - \hat{a}_1 \sum_{t=1}^{T} Y_t^2 + \hat{a}_1 \bar{Y} \sum_{t=1}^{T} Y_t = 0$$

En tenant compte du fait que $\sum_{t=1}^{T} Y_t = T\bar{Y}$, on obtient :

$$\sum_{t=1}^{T} Y_t C_t - T\bar{C}\bar{Y} - \hat{a}_1 \left(\sum_{t=1}^{T} Y_t^2 - T\bar{Y}^2 \right) = 0$$

ce qui donne l'expression (bien connue) de l'estimateur des MCO de a_1 :

$$\hat{a}_1 = \frac{\sum_{t=1}^{T} Y_t C_t - T\bar{C}\bar{Y}}{\sum_{t=1}^{T} Y_t^2 - T\bar{Y}^2} \tag{14}$$

La relation (13) permet alors de calculer $\hat{a}_2$ une fois $\hat{a}_1$ calculé.

Remarque : la notation générale la plus utilisée pour le modèle linéaire simple est la suivante :

$$y_i = ax_i + b + \varepsilon_i \qquad i = 1, ..., n$$

Dans le cas du modèle étudié $C_t = a_1 Y_t + a_2 + \varepsilon_t$ où $t = 1, ..., T$, les indices t jouent le rôle des indices i, T celui de n, les Y_t celui des x_i, les C_t celui des y_i, a_1 celui de a et a_2 celui de b. On vérifie immédiatement que les formules (13) et (14) correspondent bien aux expressions générales des estimateurs des MCO.

On nous demande également de montrer que la moyenne de la consommation estimée c'est-à-dire $\hat{a}_1\bar{Y} + \hat{a}_2$ est égale à la moyenne empirique $\bar{C}$ or c'est justement un des résultats démontrés précédemment (Cf égalité ((12)).

Exercice 4 :

1) Cf. tableau distribué. Notons que la période 4 manque. On dispose donc de $T = 11$ périodes.

2) L'estimateur des MCO du paramètre a est donné par :

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

Or $\sum_{t=1}^{T} x_t y_t = 12,9913$; $T = 11$; $\bar{x} = 0,1759$; $\bar{y} = 6,8052$ et $\sum_{t=1}^{T} x_t^2 = 0,4329$ donc

$$\hat{a} = \frac{12,9913 - 11 \times 0,1759 \times 6,8052}{0,4329 - 11 \times 0,1759^2} \simeq -1,903$$

L'estimateur des MCO de b s'obtient à partir de celui de a de la façon suivante :

$$\hat{b} = \bar{y} - \hat{a}\bar{x}$$

donc

$$\hat{b} = 6,8052 + -1,903 \times 0,1759 = 7,140$$

Il n'est pas demandé de calculer des estimateurs des variances de $\hat{a}$ et $\hat{b}$. Néanmoins sans ces estimations nous n'avons aucune idée de la précision de ces estimateurs et en particulier nous ne pouvons pas statuer sur le caractère significatif ou pas des valeurs trouvées.

Rappelons que $V(\hat{a}) = \frac{\sigma^2}{\sum_{t=1}^{T} x_t^2 - T\bar{x}^2}$ et $V(\hat{b}) = \frac{\sigma^2}{T} \frac{\sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{T} x_t^2 - T\bar{x}^2}$.Or σ^2 est un paramètre inconnu du modèle qu'on peut estimer par

$$\hat{\sigma}^2 = \frac{1}{T-2}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 = \frac{1}{T-2}\sum_{t=1}^{T}\left(y_t - \hat{a}x_t - \hat{b}\right)^2$$

Par conséquent on peut estimer $V(\hat{a})$ par

$$\hat{V}(\hat{a}) = \frac{\hat{\sigma}^2}{\sum_{t=1}^{T} x_t^2 - T\bar{x}^2}$$

On trouve $\hat{\sigma}^2 = 0,00407$, ce qui implique que $\hat{V}(\hat{a}) = 0,044$ (l'écart-type estimé est donc $\sqrt{\hat{V}(\hat{a})} \simeq 0,21$).

On construit de façon similaire un estimateur $\hat{V}(\hat{b})$ de $V(\hat{b})$ en remplaçant σ^2 par $\hat{\sigma}^2$ dans l'expression de $V(\hat{b})$ et on calcule la valeur prise par $\hat{V}(\hat{b})$ en utilisant le fait que $\hat{\sigma}^2 = 0,00407$ (on ne fait pas ce calcul car seul le paramètre a nous intéresse dans la suite).

Interprétation du paramètre a : $y_t = ax_t + b + \varepsilon_t$ est un modèle économétrique "dérivé" du modèle théorique $y = ax + b$ c'est-à-dire $\ln q = a \ln p + b$. Il s'ensuit que $\frac{\partial \ln q}{\partial \ln p} = a$ or $-\frac{\partial \ln q}{\partial \ln p} = -\frac{\partial q}{\partial p}\frac{p}{q}$ n'est autre que l'élasticité de la demande par rapport au prix. Par conséquent $-a$ peut être interprété comme l'élasticité de la demande par rapport au prix et donc $-\hat{a}$ est une estimation de cette élasticité.

3) Les propriétaires cherchent à maximiser leur chiffre d'affaires pq (ce qui revient à maximiser le profit lorsque les coûts marginaux sont nuls). On a : $\frac{\partial(pq)}{\partial p} = q + p\frac{\partial q}{\partial p} = q(1 + \frac{\partial q}{\partial p}\frac{p}{q})$ or d'après la question précédente $a = -\frac{\partial \ln q}{\partial \ln p} = -\frac{\partial q}{\partial p}\frac{p}{q}$ donc $\frac{\partial(pq)}{\partial p} = q(1+a)$. A ce stade, on peut procéder de deux façons différentes :

La première (peu rigoureuse, mais ne nécessitant pas l'utilisation de tests statistiques) consiste à dire qu'on a une estimation du paramètre a, à savoir $\hat{a} \simeq -1,903$ et donc une estimation de $\frac{\partial(pq)}{\partial p} = q(1 + a)$ à savoir $-0.903q$. L'estimation de $\frac{\partial(pq)}{\partial p}$ étant négative, on peut raisonnablement penser que le chiffre d'affaires pq est une fonction décroissante du prix p (du moins dans la plage des prix considérée). Par conséquent, les propriétaires ont intérêt à diminuer le prix pour augmenter leur chiffre d'affaires.

La seconde (beaucoup plus rigoureuse mais nécessitant le recours à la théorie des tests qui n'a pas encoré été vue en cours) consiste à tester l'hypothèse $a < -1$ (contre l'hypothèse $a \geq 1$). Un test au seuil de 5% conduit à accepter cette hypothèse (vérifiez-le dès que la question des tests sera traitée en cours). On peut donc raisonnablement affirmer que $\frac{\partial(pq)}{\partial p} = q(1 + a) < 0$ ce qui implique que le chiffre d'affaires pq est une fonction décroissante du prix p . Par conséquent, les propriétaires ont intérêt à diminuer le prix pour augmenter leur chiffre d'affaires.

Texte :

1) Considérons une spécification en double-log du type :

$$\log Q_m = b_0 + b_1 \log P_m + b_2 \log E + b_3 \log S + e \tag{15}$$

où Q_m est la quantité consommée de marijuana, P_m est le prix de la marijuana, E représente les dépenses mensuelles moyennes (cette variable sert ici comme mesure approximative du revenu), S représente une mesure de la dispersion des dépenses, et e est le terme d'erreur. Dans ce cas, les élasticités de la consommation par rapport au revenu et au prix se lisent directement sur l'équation de régression. En effet, il est aisé de voir que $\frac{\partial \log Q_m}{\partial \log P_m} = b_1$ et $\frac{\partial \log Q_m}{\partial \log E} = b_2$. Par conséquent b_1 est l'élasticité-prix de la consommation de marijuana et b_2 est l'élasticité-revenu de la consommation de marijuana.

2) La différence essentielle entre des estimations sur données d'enquête (données récoltées en demandant à des individus quelles quantités ils consommeraient si le prix de la marijuana prenait certaines valeurs, ce qui permet de déterminer leur fonction de demande hypothétique) et des données de marché (données sur les prix de marché et les quantités effectivement consommées) provient du fait que le comportement des individus lorsqu'on leur demande ce qu'ils feraient dans une situation hypothétique n'est pas nécessairement le même que leur comportement si la situation en question venait à se présenter réellement.

3) Les résultats de l'estimation des élasticités de la consommation par rapport au revenu et au prix peuvent être sous-estimés ou surestimés pour plusieurs raisons. D'abord, il se peut qu'il y ait des erreurs de mesure, notamment sur la variable E qui est une mesure indirecte du revenu. Ensuite, il se peut qu'il y ait des variables omises dans la régression. Ceci ne pose pas de problème en soi puisque le terme d'erreur e est censé capturer, entre autres, ces variables. Le problème vient du fait que ces variables omises peuvent être corrélées avec les variables explicatives P_m, E et S, et peuvent ainsi induire une corrélation entre le terme d'erreur e et les variables explicatives (dans ce

cas, les estimateurs MCO de la régression (15) ne seraient plus sans biais) Considérons par exemple la probabilité d'incarcération suite à la consommation de marijuana. Si pour des raisons exogènes, cette probabilité venait à baisser, la demande pourrait augmenter entraînant ainsi une augmentation du prix. La probabilité d'incarcération est donc une variable omise probablement corrélée à la variable explcative P_m (autre exemple de variable omise : le prix d'autres drogues concurrentes qui pourraient représenter des substituts à la marijuana). Enfin, les résultats peuvent être également biaisés à cause du caractère statique du modèle qui ne prend pas en compte les dépendances dynamiques dans le comportement de consommation de marijuana, notamment les effets d'habitude et d'addiction.