L3 2006/2007
Statistique et Économétrie
Interrogation 2, Éléments de correction

Exercice 1

Soit un échantillon d'hommes âgés de 26 à 54 ans[1]. On dispose de trois variables : *salaire* = taux de salaire horaire en euros, *age* = âge de l'individu, *etudes* = nombre d'années d'études. En régressant la variable *salaire* sur *age* et *etudes*, on obtient l'équation suivante :

$$\widehat{salaire} = -0,25 + 0,17 \cdot age + 0,55 \cdot etudes$$

1. Écrire le modèle économétrique qu'on a estimé.

$$salaire = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot etudes + \varepsilon$$

2. Rappeler l'interprétation du terme d'erreur.

   Le terme ε est une perturbation aléatoire. En effet, il « perturbe » une relation qui, autrement, est stable. La perturbation a plusieurs sources, par exemple, l'effet net (positif ou négatif) des facteurs omis, les erreurs de mesure des variables économiques, etc.

3. Interpréter la régression estimée.

   β_1 mesure l'impact de l'âge sur le taux de salaire horaire : un homme qui est plus âgé d'un an gagne 0,17 euros de plus par heure, toutes choses égales par ailleurs. β_2 mesure l'impact du nombre d'année d'études sur le taux de salaire horaire : une année d'études supplémentaire augmente le taux de salaire horaire de 0,55 euros, toutes choses égales par ailleurs.

Exercice 2

On dispose d'un échantillon $(X_1, \ldots, X_n)$ de taille n d'une variable suivant une loi normale de densité :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

1. Écrire la vraisemblance de cet échantillon.

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x_i-m)^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^{n} exp\left(-\frac{(x_i-m)^2}{2\sigma^2}\right)$$

2. Déterminez l'estimateur du maximum de vraisemblance de m.

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-m)^2}{2\sigma^2}$$

$$\frac{\partial \ln(L)}{\partial m} = \frac{\sum_{i=1}^{n} 2(x_i-m)}{2\sigma^2} = \frac{\sum_{i=1}^{n} x_i - nm}{\sigma^2} = 0 \Rightarrow \hat{m}_{MV} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}.$$

[1] Source : l'Enquête « Emploi » 2002 de l'INSEE.

3. Montrez que la méthode des moments donne le même estimateur.

   $E(X) = m$, puisque $X \rightsquigarrow N(m, \sigma^2)$. Donc, $m = E(X)$ le premier moment théorique, $\hat{m}_{MM} = \bar{x}$ le premier moment empirique.

4. Est-ce un estimateur sans biais ? convergent ? Expliquez.

   $E(\hat{m}) = E(\bar{x}) = E(\frac{\sum_{i=1}^n x_i}{n}) = \frac{\sum_{i=1}^n E(x_i)}{n} = \frac{\sum_{i=1}^n m}{n} = \frac{nm}{n} = m$ donc $\hat{m}$ est un estimateur sans biais de m.
   $V(\hat{m}) = V(\bar{x}) = V(\frac{\sum_{i=1}^n x_i}{n}) = \frac{\sum_{i=1}^n V(x_i)}{n^2} = \frac{\sum_{i=1}^n \sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$. Alors $\hat{m}$ est un estimateur sans biais et sa variance tend vers 0 quand n tend vers l'infini. Donc, $\hat{m}$ est un estimateur convergent de m.

Exercice 3

Soit X une variable aléatoire de loi normale d'espérance m et de variance σ^2. On dispose d'un échantillon de 16 tirages indépendants de X. La moyenne de cet échantillon est égale à 5, la variance empirique (l'estimateur sans biais) est égale à 9. Construire un intervalle de confiance à 80% pour m.

Donc, on dispose d'un échantillon $x_1, \ldots, x_{16}$ de $X \rightsquigarrow N(m, \sigma^2)$.
$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = 5,\ s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{16-1} = 9.$
$\bar{x} \rightsquigarrow N(m, \frac{\sigma^2}{16}) \Rightarrow \frac{\bar{x}-m}{\sqrt{\sigma^2/16}} \rightsquigarrow N(0,1).$

Mais la variance théorique σ^2 est inconnue donc on la remplace par son estimateur s^2.

$\frac{\bar{x}-m}{\sqrt{s^2/16}} \rightsquigarrow t_{16-1}.$

$\Pr(-1,341 < \frac{\bar{x}-m}{\sqrt{s^2/16}} < 1,341) = 0,80$

où 1,341 est la valeur critique d'une distribution t à 15 degrés de liberté.

$\Pr(\bar{x} - 1,341\frac{s}{4} < m < \bar{x} + 1,341\frac{s}{4}) = 0,80$

$\Pr(5 - 1,341\frac{3}{4} < m < 5 + 1,341\frac{3}{4}) = 0,80$

L'intervalle de confiance à 80% pour m est $]3,99; 6,01[$.

Exercice 4

Soit X une variable aléatoire d'espérance m et de variance σ^2. En utilisant l'inégalité de Tchebychev, trouver ε telle que la probabilité pour X de s'écarter de son espérance mathématique d'une valeur au plus égale à ε soit supérieure ou égale à 96%.

Il faut trouver ε telle que

$$\Pr\left(|X - E(X)| \leq \varepsilon\right) \geq 0,96.$$

L'inégalité de Tchebychev :

$$\Pr\left(|X - E(X)| > \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2},$$

alors

$$1 - \Pr\left(|X - E(X)| \leq \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2}$$

$$\Pr\left(|X - E(X)| \leq \varepsilon\right) \geq 1 - \frac{V(X)}{\varepsilon^2}$$

Donc

$$1 - \frac{V(X)}{\varepsilon^2} = 0,96 \Rightarrow \frac{V(X)}{\varepsilon^2} = 0,04 \Rightarrow \varepsilon = 5\sigma.$$

Barème :

Exercice 1 :
1. 1 point
2. 1 point
3. 1 point

Exercice 2 :
1. 1 point
2. 4 points
3. 2 points
4. 2 points

Exercice 3 : 4 points

Exercice 4 : 4 points