Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2015 – 2016

Econométrie II

Contrôle continu n°2, avril 2016

Examen de 1h30. Tout document ou calculatrice est interdit.

1. **(Sur 20 points)** Pour n et p deux entiers tels que $n \geq p+1$, soit $Y = {}^t(Y_i)_{1 \leq i \leq n}$ défini par:

$$Y_i = \theta_0 + \theta_1 X_i^{(1)} + \cdots + \theta_p X_i^{(p)} + \varepsilon_i \qquad \text{pour tout } i = 1, \cdots, n, \tag{1}$$

avec:

- $(X_i^{(j)})_{1 \leq i \leq n,\, 1 \leq j \leq p}$ étant connues et telles que $X = \begin{pmatrix} 1 & X_1^{(1)} & \cdots & X_1^{(p)} \\ \vdots & \vdots & & \vdots \\ 1 & X_n^{(1)} & \cdots & X_n^{(p)} \end{pmatrix}$ soit une matrice de rang $p+1$;
- $\theta = {}^t(\theta_j)_{0 \leq j \leq p}$ un vecteur de nombres réels inconnus;
- $\varepsilon = {}^t(\varepsilon_i)_{1 \leq i \leq n}$ un vecteur d'erreur non observé tel que $\mathbb{E}(\varepsilon) = 0$ et $\mathrm{cov}(\varepsilon) = \Sigma$, où Σ est une matrice définie positive inconnue de valeurs propres $0 < \lambda_1 \leq \cdots \leq \lambda_n$ (inconnues).

(a) On effectue une régression par moindres carrés ordinaires pour obtenir un estimateur $\widehat{\theta}$. Montrer que $\widehat{\theta}$ est un estimateur sans biais de θ et $\mathrm{cov}(\widehat{\theta}) = \big({}^tX\,X\big)^{-1}\,{}^tX\,\Sigma\,X\,\big({}^tX\,X\big)^{-1}$.

(b) Pour une matrice carrée $A = (a_{ij})_{1 \leq i,j \leq m}$ où $m \in \mathbf{N}^*$, soit la norme matricielle définie par

$$\|A\|_2 = \sup_{U \in \mathbf{R}^m,\ U \neq 0} \frac{\|A\,U\|_2}{\|U\|_2} \quad \text{où} \quad \|Z\|_2^2 = {}^tZ\,Z \quad \text{pour tout } Z \in \mathbf{R}^m.$$

Montrer que si A est une matrice symétrique, alors $\|A\|_2 = \max\big(\big|\text{Valeurs propres de } A\big|\big)$.

(c) Montrer que ${}^tX\,X$ est une matrice définie positive et on notera $0 < \mu_1 \leq \cdots \leq \mu_{p+1}$ ses valeurs propres. Montrer que ${}^tV\,\Sigma\,V \leq \lambda_n \|V\|_2^2$ pour tout $V \in \mathbf{R}^n$, et en déduire que ${}^tW\,\mathrm{cov}(\widehat{\theta})\,W \leq \lambda_n \,\|X\,\big({}^tX\,X\big)^{-1}\,W\|_2^2$ pour tout $W \in \mathbf{R}^{p+1}$, puis que ${}^tW\,\mathrm{cov}(\widehat{\theta})\,W \leq \frac{\lambda_n}{\mu_1}\|W\|_2^2$. En déduire une condition suffisante de convergence de $\widehat{\theta}$ vers θ quand $n \to \infty$.

(d) On se place désormais dans le cadre d'hétéroscédasticité suivant: les (ε_i) sont indépendantes et $\mathrm{var}(\varepsilon_i) = \sigma_i^2$, avec $\underline{\sigma}^2 \leq \sigma_i^2 \leq \overline{\sigma}^2$ pour tout $i = 1, \ldots, n$ et $\frac{1}{n}\sum_{i=1}^n \sigma_i^2 \underset{n\to\infty}{\longrightarrow} \sigma^2$ où $0 < \underline{\sigma}^2 \leq \sigma^2 \leq \overline{\sigma}^2$ sont trois réels fixés inconnus. Montrer que $\widehat{\theta} \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} \theta$ si $\mu_1 \underset{n\to\infty}{\longrightarrow} +\infty$.

(e) On considère l'estimateur de la variance par MCO $\widehat{\sigma}^2 = \frac{1}{n-(p+1)}\,{}^t(Y - X\,\widehat{\theta})\,(Y - X\,\widehat{\theta})$. Montrer que $\mathbb{E}(\widehat{\sigma}^2) = \frac{1}{n-(p+1)}\big(\mathrm{Tr}(\Sigma) - \mathrm{Tr}(H\,\Sigma)\big)$ où vous exprimerez $H = (H_{ij})_{1 \leq i,j \leq n}$ en fonction X. Montrer que pour tout $i = 1, \ldots, n$, $H_{ii} = \sum_{k=1}^n H_{ik}^2 \geq 0$, puis que $\underline{\sigma}^2(p+1) \leq \mathrm{Tr}(H\,\Sigma) \leq \overline{\sigma}^2(p+1)$. En déduire que $\mathbb{E}(\widehat{\sigma}^2) \underset{n\to\infty}{\longrightarrow} \sigma^2$.

(f) Montrer que $\frac{1}{n-(p+1)}\sum_{i=1}^n \varepsilon_i^2 - \widehat{\sigma}^2 \geq 0$, et $\frac{1}{n-(p+1)}\sum_{i=1}^n \varepsilon_i^2 \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} \sigma^2$. En déduire que $\widehat{\sigma}^2 \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} \sigma^2$.

(g) On va montrer que les tests de Student peuvent être valables dans un tel cadre. On suppose que $p = 1$, n pair, et $X_i^{(1)} = 1$, $\sigma_i^2 = \overline{\sigma}^2$ pour $i = 1, \ldots, n/2$, $X_i^{(1)} = 0$, $\sigma_i^2 = \underline{\sigma}^2$ pour $i = n/2+1, \ldots, n$, avec $\underline{\sigma}^2 < \overline{\sigma}^2$. Donner l'expression exacte de $\widehat{\theta}$, de $\mathrm{cov}(\widehat{\theta})$, de σ^2. Rappeler l'expression matricielle du test de Student $\widehat{T}_1$ relatif à la significativité de $X^{(1)}$. En déduire que $\widehat{T}_1 = \sqrt{n}\,\frac{\widehat{\theta}_1}{2\,\widehat{\sigma}}$, puis montrer que $\widehat{T}_1 \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0\,,\,1)$.

2. **(Sur 8 points)** Exercice de TP utilisant le logiciel R

   (a) On a tapé les commandes suivantes:

```
n=100; Y=0; i=1:100;
Z1=i; Z2=5*rnorm(n,0)
Z3=runif(n,0,20); Z4=4*cos(2*pi*i/12)
epsilon=2*rnorm(n,0)^2
Y=10-3*Z3+4*sqrt(i+3)+5*epsilon
reg1=lm(Y~Z1+Z2+Z3+Z4)
summary(reg1)
```

   On obtient les résultats:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 30.52028    3.52702   8.653 1.24e-13 ***
Z1           0.30467    0.04661   6.537 3.12e-09 ***
Z2           0.13032    0.28060   0.464    0.643
Z3          -2.84563    0.22257 -12.785  < 2e-16 ***
Z4          -0.79840    0.48118  -1.659    0.100

Residual standard error: 13.44 on 95 degrees of freedom
Multiple R-squared:  0.6933,    Adjusted R-squared:  0.6804
F-statistic: 53.68 on 4 and 95 DF,  p-value: < 2.2e-16
```

   Questions 1: Quelle est la loi de `epsilon`*? Ecrire Y sous la forme $Y = X\,\theta + \varepsilon$ avec X dépendant des Z_i et $\mathbb{E}(\varepsilon) = 0$. Que pensez-vous des résultats de cette régression? Pouvait-on s'y attendre?*

   (b) On tape ensuite les commandes:

```
library(MASS)
ZZ=as.data.frame(cbind(Z1,Z2,Z3,Z4));
y.lm=lm(Y~.,data=ZZ);
y.bic=stepAIC(y.lm,k=log(n));
```

   Voici les résultats:

```
Start:  AIC=537.54
Y ~ Z1 + Z2 + Z3 + Z4

       Df Sum of Sq   RSS    AIC
- Z2    1      39.0 17198 533.16
- Z4    1     497.3 17656 535.79
<none>              17159 537.54
- Z1    1    7718.8 24878 570.08
- Z3    1   29524.2 46683 633.02

Step:  AIC=533.16
Y ~ Z1 + Z3 + Z4

       Df Sum of Sq   RSS    AIC
- Z4    1     515.2 17713 531.50
<none>              17198 533.16
- Z1    1    7710.1 24908 565.59
- Z3    1   31283.9 48482 632.19

Step:  AIC=531.5
Y ~ Z1 + Z3

       Df Sum of Sq   RSS    AIC
<none>              17713 531.50
- Z1    1    7551.5 25264 562.41
- Z3    1   30880.0 48593 627.82
```

   Questions 2: Qu'a-t-on fait avec ces commandes et que peut-on en conclure? Pouvait-on s'y attendre?

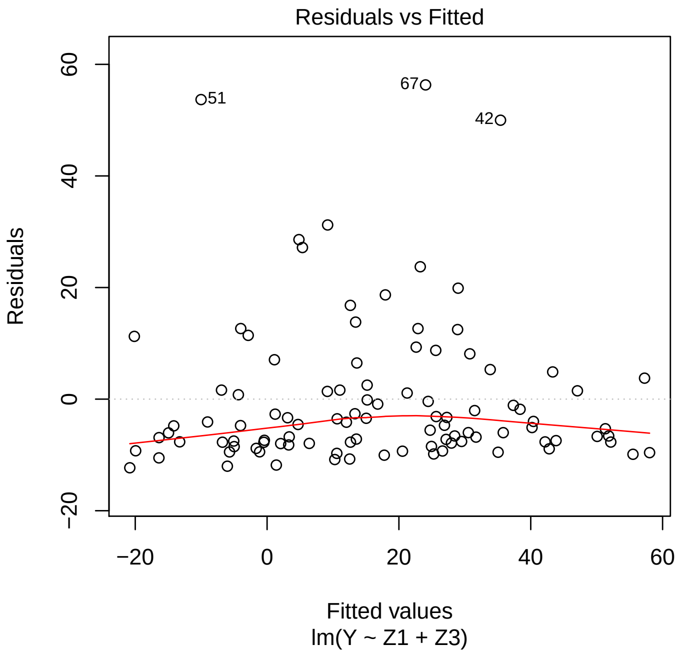
   (c) De nouvelles commandes sont tapées:

```
summary(y.bic)
plot(y.bic)
```

Voici les résultats et un des graphes obtenus:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 30.57434    3.52651   8.670 9.93e-14 ***
Z1           0.30105    0.04681   6.431 4.79e-09 ***
Z3          -2.84153    0.21851 -13.004  < 2e-16 ***

Residual standard error: 13.51 on 97 degrees of freedom
Multiple R-squared:  0.6834,    Adjusted R-squared:  0.6768
F-statistic: 104.7 on 2 and 97 DF,  p-value: < 2.2e-16
```



Questions 3: Qu'a-t-on fait avec ces commandes? Que signifie précisément la valeur numérique 30.57434 *et pouvait-on s'y attendre? Dire également en termes mathématiques ce que sont* 0.30105, -2.84153, 0.21851, -13.004 *et* $4.79e-09$. *Au final que pensez-vous des résultats obtenus?*

(d) On a enfin tapé les commandes:

```
Q=matrix(0,50,50)
for (a in c(1:50))
{ for (b in c(1:50))
{ Z1a=Z1^(a/10); Z3b=Z3^(b/10)
reg=lm(Y~Z1a+Z3b)
Q[a,b]=sum(reg$res^2) }}
which(Q==min(Q),arr.ind = TRUE)
```

Voici le résultat:

```
  row col
[1,]   6  10
```

Questions 4: Qu'a-t-on fait avec ces commandes et qu'a-t-on obtenu? Pouvait-on s'attendre au résultat obtenu? Si vous deviez prédire Y_i *pour* $i = 101$, *quelles commandes* `R` *utiliseriez vous?*