Université Paris I, Panthéon - Sorbonne

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

Econométrie II

Contrôle continu n°1, mars 2016

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

(Sur 26 points) On observe la température mensuelle T dans une ville pendant n années. On disposera donc des données $(T_k)_{1\leq k\leq 12n}$. On suppose que sur ces n années la température peut être modélisée par:

$$T_k = \theta_0 + \theta_1 \cos\big(k\,\frac{\pi}{6}\big) + \theta_2 \sin\big(k\,\frac{\pi}{6}\big) + \varepsilon_i \qquad \text{pour tout } k\in\{1,\dots,12\,n\}, \qquad \text{où:} \tag{1}$$

$\theta = {}^t(\theta_0,\theta_1,\theta_2)$ est un vecteur composé de 3 réels inconnus, $(\varepsilon_i)_i$ est une suite non observée de v.a.i.i.d. de loi centrée et de variance $\sigma^2_\varepsilon > 0$. On appellera respectivement X_1 et X_2 les variables constituées des $\cos(k\,\frac{\pi}{6})$ et des $\sin(k\,\frac{\pi}{6})$.

1. Montrer que pour $P\in\mathbb{N}^*$, $\sum_{k=1}^P \cos(2\pi\,k/P) = \sum_{k=1}^P \sin(2\pi\,k/P) = 0$, puis que $\sum_{k=1}^P \cos^2(2\pi\,k/P) = P/2$ et $\sum_{k=1}^P \cos(2\pi\,k/P)\sin(2\pi\,k/P) = 0$ **(2 pts)**.
2. On note $\widehat{\theta} = {}^t\big(\widehat{\theta}_0,\widehat{\theta}_1,\widehat{\theta}_2\big)$ l'estimateur de θ par moindres carrés. Donner l'expression des $\widehat{\theta}_i - \theta_i$ en fonction des ε_k (montrer notamment que $\widehat{\theta}_1 - \theta_1 = \frac{1}{6n}\sum_{k=1}^{12n}\cos\big(k\,\frac{\pi}{6}\big)\varepsilon_k\big)$ **(3 pts)**. En déduire, en fonction de $\widehat{\theta}$, la valeur prédite pour la température d'un mois de mars **(1 pt)**.
3. Démontrer que $\sqrt{n}(\widehat{\theta}-\theta)$ tend en loi vers une limite que l'on précisera lorsque $n\to\infty$ **(3 pts)**.
4. Ecrire l'estimateur $\widehat{\sigma}^2$ par moindres carrés et non biaisé de σ^2. Quel est son comportement asymptotique quand $n\to\infty$? **(1 pt)**
5. On veut tester que le signal est bien périodique dans le sens où les variables X_1 et/ou X_2 sont bien significatives. Après avoir poser les hypothèses du test, montrer que la statistique de test de Fisher associée à ce test s'écrit $\widehat{F} = \frac{3\,n}{\widehat{\sigma}^2}\big(\widehat{\theta}_1^2 + \widehat{\theta}_2^2\big)$ **(3 pts)**. On suppose que n est grand et que l'on a obtenu pour valeur de ce test 19.036. Que conclure? **(2 pts)**
6. On veut tester que la variable X_2 est bien significative. Ecrire explicitement en fonction des $\widehat{\theta}_i$ et $\widehat{\sigma}^2$ la statistique de test de Student relative à un tel test **(2 pts)**. On suppose que n est grand et que l'on a obtenu pour valeur de ce test -1.347. Que conclure? **(1 pt)**
7. On suppose que $\theta_0 = 4$, $\theta_1 = -10$ et $\theta_2 = -5$. On effectue une régression par moindres carrés sans la variable X_2 et on note $\widehat{\sigma}_2^2$ la variance estimée par un tel modèle. Montrer alors que $\widehat{\sigma}_2^2 = \frac{1}{12n-2}\Big(6\,n\,\theta_2^2 + \|P_A\,\varepsilon\|^2 + 2\,\theta_2{}^tX_2\,P_A\,\varepsilon\Big)$, en notant $\varepsilon = (\varepsilon_k)_{1\leq k\leq 12n}$ et $X_2 = \big(\sin\big(k\,\frac{\pi}{6}\big)\big)_{1\leq k\leq 12n}$ et P_A la projection orthogonale sur un sous-espace vectoriel A que l'on précisera **(3 pts)**. En déduire que $\widehat{\sigma}_2^2 \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 + \frac{1}{2}\,\theta_2^2$ **(3 pts)**. Utiliser ce résultat pour montrer que la probabilité que le critère BIC choisisse un modèle avec X_1 et X_2 plutôt qu'un modèle avec uniquement X_1 tend vers 1 quand $n\to\infty$ **(2 pts)**.

Proof.

1. Il est clair que $\sum_{k=1}^P e^{2i\pi\,k/P} = e^{2i\pi/P}(1-e^{2i\pi})(1-e^{2i\pi/P})^{-1} = 0$. D'où la nullité des deux premières sommes.
   En posant $\cos^2(2\pi\,k/P) = \frac{1}{2}\big(1+\cos(4\pi\,k/P)\big)$, on a donc $\sum_{k=1}^P \cos^2(2\pi\,k/P) = \frac{P}{2} + \frac{1}{2}\sum_{k=1}^P\cos(4\pi\,k/P)$. En utilisant encore l'exponentielle complexe, on montre facilement que la dernière somme est nulle.
   Enfin, $\cos(2\pi\,k/P)\sin(2\pi\,k/P) = \frac{1}{2}\sin(4\pi\,k/P)$, et comme précédemment, avec l'exponentielle complexe, $\sum_{k=1}^P \sin(4\pi\,k/P) = 0$.
2. On commence par écrire le modèle sous la forme matricielle $T = X\theta + \varepsilon$. Grâce à la question précédente, on montre alors que:

$$ {}^tX\,X = \begin{pmatrix} 12n & 0 & 0 \\ 0 & 6n & 0 \\ 0 & 0 & 6n \end{pmatrix} \quad\Longrightarrow\quad \big({}^tX\,X\big)^{-1} = \frac{1}{12n}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Comme $\widehat{\theta} = \theta + \left({}^tX\,X\right)^{-1}{}^tX\,\varepsilon$, on en déduit que

$$\left\{\begin{array}{l} \widehat{\theta}_0 - \theta_0 = \frac{1}{12n}\sum_{k=1}^{12n}\varepsilon_k \\ \widehat{\theta}_1 - \theta_1 = \frac{1}{6n}\sum_{k=1}^{12n}\cos(\pi\,k/6)\,\varepsilon_k \\ \widehat{\theta}_2 - \theta_2 = \frac{1}{6n}\sum_{k=1}^{12n}\sin(\pi\,k/6)\,\varepsilon_k \end{array}\right. .$$

Pour la prédiction d'un mois de mars, comme le signal non bruité est périodique, il suffit de choisir $k = 3$ et on obtient: $\widehat{\theta}_0 + \widehat{\theta}_2$.

3. Si on considère $H = X\left({}^tX\,X\right)^{-1}{}^tX$, un calcul rapide permet de montrer que cette matrice s'écrit $H = \frac{1}{12n}\big(1+2\cos(\pi\,i/6)\cos(\pi\,j/6)+2\sin(\pi\,i/6)\sin(\pi\,j/6)\big)\big)_{1\leq i,j\leq 12n}$. D'où $\max_{1\leq i\leq n}|H_{ii}| \leq \frac{5}{12n} \underset{n\to\infty}{\longrightarrow} 0$. Aussi $\widehat{\theta}$ est asymptotiquement gaussien. Comme $\mathbb{E}(\widehat{\theta}) = \theta$ et $\operatorname{cov}(\widehat{\theta}) = \sigma^2\left({}^tX\,X\right)^{-1}$, on obtient:

$$\sqrt{12\,n}\,(\widehat{\theta}-\theta) \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}\left(0\,,\,\sigma^2\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}\right).$$

4. On définit $\widehat{\sigma}^2 = \frac{1}{12n-3}\sum_{i=1}^{12n}(T_i-\widehat{T}_i)^2$, avec $\widehat{T}_i = \widehat{\theta}_0+\widehat{\theta}_1\cos(\pi\,i/6)+\widehat{\theta}_2\sin(\pi\,i/6)$. Un résultat du cours montre que sous les hypothèses vérifiées par les ε_i, alors $\widehat{\sigma}^2 \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} \sigma^2$.

5. Considérons ${}^tC = \begin{pmatrix}0&1&0\\0&0&1\end{pmatrix}$. Alors ${}^tC\,\theta = \begin{pmatrix}\theta_1\\\theta_2\end{pmatrix}$. On veut donc tester H_0: ${}^tC\,\theta = 0$ contre H_1: ${}^tC\,\theta \neq 0$.

   La statistique de Fisher $\widehat{F}$ pour un tel test s'écrit:

$$\begin{aligned}
\widehat{F} &= \frac{1}{2}\,{}^t\widehat{\theta}\,C\big(\widehat{\sigma}^2\,{}^tC\left({}^tX\,X\right)^{-1}C\big)^{-1}\,{}^tC\,\widehat{\theta} \\
&= \frac{1}{2\,\widehat{\sigma}^2}\,(\widehat{\theta}_1\,,\,\widehat{\theta}_2)\left(\begin{pmatrix}0&1&0\\0&0&1\end{pmatrix}\frac{1}{12n}\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}\begin{pmatrix}0&0\\1&0\\0&1\end{pmatrix}\right)^{-1}\begin{pmatrix}\widehat{\theta}_1\\\widehat{\theta}_2\end{pmatrix} \\
&= \frac{1}{2\,\widehat{\sigma}^2}\,(\widehat{\theta}_1\,,\,\widehat{\theta}_2)\left(\frac{1}{12n}\begin{pmatrix}2&0\\0&2\end{pmatrix}\right)^{-1}\begin{pmatrix}\widehat{\theta}_1\\\widehat{\theta}_2\end{pmatrix} \\
&= \frac{3\,n}{\widehat{\sigma}^2}\,(\widehat{\theta}_1\,,\,\widehat{\theta}_2)\begin{pmatrix}\widehat{\theta}_1\\\widehat{\theta}_2\end{pmatrix} = \frac{3\,n}{\widehat{\sigma}^2}\,(\widehat{\theta}_1^2+\widehat{\theta}_2^2).
\end{aligned}$$

   On sait que pour $n\to\infty$, $\widehat{F} \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \frac{1}{2}\,\chi^2(2)$ car le test porte sur 2 paramètres et $\max_{1\leq i\leq n}|H_{ii}| \underset{n\to\infty}{\longrightarrow} 0$. Or on sait que pour (Z_1,Z_2) un couple de variables gaussiennes centrées réduites indépendantes, $\mathbb{P}(Z_1^2+Z^2>4) \leq \mathbb{P}(Z_1>2)+\mathbb{P}(Z_2>2) \leq 0.05$. On en déduit que pour $n\to\infty$, on a $\mathbb{P}(\widehat{F}>8) \simeq \mathbb{P}(\chi^2(2)>4) \leq 0.05$. Ainsi avec $\widehat{F} \simeq 19.036$, alors l'hypothèse H_1 est choisi: le signal est bien périodique.

6. Le test proposé est: H_0: $\theta_2 = 0$ contre H_1: $\theta_2 \neq 0$, et on remarque que $\theta_2 = {}^tD\,\theta$ avec $D = {}^t\big(0,0,1\big)$. On choisit alors comme statistique de test la statistique de Student:

$$\widehat{T} = \frac{{}^tD\,\widehat{\theta}}{\left(\widehat{\sigma}^2\,{}^tD\,\left({}^tX\,X\right)^{-1}D\right)^{1/2}} = \frac{\sqrt{6\,n}}{\widehat{\sigma}}\,\widehat{\theta}_2,$$

   en utilisant les mêmes calculs que pour la statistique de Fisher.

   On sait que pour $n\to\infty$, $\widehat{T} \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(0\,,1)$ car $\max_{1\leq i\leq n}|H_{ii}| \underset{n\to\infty}{\longrightarrow} 0$. Or $\mathbb{P}(|Z|\geq 1.96) \simeq 0.05$ avec Z de loi $\mathcal{N}(0\,,1)$, donc si $\widehat{T} \simeq -1.346$, alors on choisit H_0: la variable X_2 n'est pas significative.

7. Notons Y la matrice constituée des 2 premières colonnes de X (donc comprenant les valeurs relatives à l'intercept et à X_1) et $Q = Y\left({}^tY\,Y\right)^{-1}{}^tY$, la matrice de la projection orthogonale sur le sous-espace vectoriel engendré par Y dans $\mathbf{R}^{12n}$. Alors

$$\widehat{\sigma}_2^2 = \frac{1}{12n-2}\,{}^t(T-Q\,T)\,(T-Q\,T) = \frac{1}{12n-2}\,{}^t(\theta_2X_2+(I_{12n}-Q)\,\varepsilon)\,(\theta_2X_2+(I_{12n}-Q)\,\varepsilon)$$

   car $QX_2 = 0$ puisque ${}^tY\,X_2 = 0$ et donc $(Q\,T)k = \theta_0+\theta_1\cos(\pi k/6)+(Q\varepsilon)_k$. Si on appelle $P_A = I_{12n}-Q$, P_A est la projection orthogonale sur le sous-espace orthogonal à celui engendré par Y. Ainsi:

$$\widehat{\sigma}_2^2 = \frac{1}{12n-2}\,{}^t(\theta_2X_2+P_A\,\varepsilon)\,(\theta_2X_2+P_A\,\varepsilon) = \frac{1}{12n-2}\big(\|\theta_2X_2\|^2+\|P_A\,\varepsilon\|^2+2\,\theta_2\,{}^tX_2\,P_A\,\varepsilon\big),$$

   d'où le résultat.

   Il est facile de montrer que $\|P_A\,\varepsilon\|^2 = \|\varepsilon\|^2-\|Q\,\varepsilon\|^2$. Or d'après la loi forte des grands nombres, $\frac{1}{12n-2}\,\|\varepsilon\|^2 \underset{n\to\infty}{\xrightarrow{p.s.}} \sigma^2$. De plus $\mathbb{E}\big(\|Q\,\varepsilon\|^2\big) = 2\,\sigma^2$ car $\dim(Y) = 2$. Donc comme $\mathbb{E}\big(\frac{1}{12n-2}\,\|Q\,\varepsilon\|^2\big) \underset{n\to\infty}{\longrightarrow} 0$ et $\|Q\,\varepsilon\|^2 \geq 0$, l'Inégalité de Markov implique que $\frac{1}{12n-2}\,\|Q\,\varepsilon\|^2 \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} 0$. Par suite $\frac{1}{12n-2}\,\|P_A\,\varepsilon\|^2 \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} \sigma^2$.

   Par ailleurs, $\frac{1}{12n-2}\,{}^tX_2\,P_A\,\varepsilon \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} 0$ en utilisant l'Inégalité de Bienaymé-Tchebichev et le fait que $\operatorname{var}({}^tX_2\,P_A\,\varepsilon) = \mathbb{E}\big({}^tX_2\,P_A\,\varepsilon\,{}^t\varepsilon P_AX_2\big) = \sigma^2\,\mathbb{E}\big({}^tX_2\,P_A\,X_2\big) \leq \|P_AX_2\|^2 \leq \|X_2\|^2 = 6\,n$. Par conséquent $\widehat{\sigma}_2^2 \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} \frac{1}{2}\,\theta_2^2+\sigma^2$.

   On calcule $\mathbb{P}(BIC(X_1) < BIC(X_1,X_2)) = \mathbb{P}\big(12n\,\ln(\widehat{\sigma}_2^2) \leq 12n\,\ln(\widehat{\sigma}^2)+\ln(12n)\big) = \mathbb{P}\Big(\frac{\widehat{\sigma}_2^2}{\widehat{\sigma}^2} \leq \exp\big(\frac{\ln(12n)}{12n}\big)\Big)$. Comme d'après le Lemme de Slutsky, $\frac{\widehat{\sigma}_2^2}{\widehat{\sigma}^2} \underset{n\to+\infty}{\xrightarrow{\mathcal{P}}} 1+\frac{1}{2}\,\frac{\theta_2^2}{\sigma^2}$ et $\exp\big(\frac{\ln(12n)}{12n}\big) \underset{n\to\infty}{\longrightarrow} 1$, on montre ainsi que $\mathbb{P}(BIC(X_1) < BIC(X_1,X_2)) \underset{n\to\infty}{\longrightarrow} 0$.

□