Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2021 – 2022

Econométrie II

Contrôle continu n°1, mars 2022

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

Soit p suites indépendantes de v.a.i.i.d. $(W_i^{(j)})_{i\in\mathbf{N},\,1\le j\le p}$ telles que $\mathbb{E}[W_0^{(j)}]=m_j\in\mathbf{R}$ et $\mathrm{var}(W_0^{(j)})=\sigma_j^2>0$ pour $j=1,\cdots,p$. Soit $(Y_i)_{i\in\mathbf{N}}$ la suite de variables aléatoires définie par:

$$Y_i=\theta_0+\sum_{k=1}^p\theta_k\,W_i^{(k)}+\varepsilon_i\qquad\text{pour tout } i\in\mathbf{N}\text{, où:}\tag{1}$$

$\theta={}^t(\theta_0,\theta_1,\ldots,\theta_p)$ est un vecteur composé de $p+1$ réels inconnus et la suite $(\varepsilon_i)_{i\in\mathbf{N}}$ est une suite non observée de v.a.i.i.d. de loi gaussienne centrée de variance $\sigma^2>0$, indépendante de $(W_i^{(j)})_{i\in\mathbf{N},\,1\le j\le p}$.

1. On note $\varepsilon^{(n)}=(\varepsilon_i)_{1\le i\le n}$. Préciser la loi de $\varepsilon^{(n)}$. La loi de Y_i est-elle gaussienne?

2. Pour $i\in\mathbf{N}$ et $j=1,\cdots,p$, soit $Z_i^{(j)}=W_i^{(j)}-m_j$. Déterminer $\mathbb{E}[Z_i^{(j)}]$ et $\mathrm{var}(Z_i^{(j)})$.

3. Ecrire le modèle (1) sous la forme $Y_i=\beta_0+\sum_{k=1}^p\beta_k\,Z_i^{(k)}+\varepsilon_i$ pour tout $i\in\mathbf{N}$, en explicitant les β_k en fonction des θ_k.

4. Pour $n\ge p+1$, soit X_n la matrice de taille $(n,p+1)$ telle que la première colonne de X_n est constituée de 1, et pour $j=1,\cdots,p$, la colonne $j+1$ de X_n est constituée de $Z_1^{(j)},\cdots,Z_n^{(j)}$. Ecrire en détail $\frac{1}{n}\,{}^tX_n\,X_n$, et en déduire que $(\frac{1}{n}\,{}^tX_n\,X_n)_{n\in\mathbf{N}^*}$ tend p.s. vers une limite lorsque $n\to\infty$ et la préciser. En déduire que la suite de matrices $((\frac{1}{n}\,{}^tX_n\,X_n)^{-1})_{n\in\mathbf{N}^*}$ existe p.s. lorsque n est suffisamment grand et préciser sa limite Σ lorsque $n\to\infty$ en fonction des σ_j^2.

5. On a observé $Y^{(n)}=(Y_i)_{1\le i\le n}$ et X_n. Déterminer la loi de $Y^{(n)}$ sachant X_n. Pour n suffisamment grand, déterminer la loi sachant X_n de l'estimateur $\widehat{\beta}^{(n)}$ par moindres carrés de β. Montrer que
$$\sqrt{n}\,(\widehat{\beta}^{(n)}-\beta)\ \xrightarrow[n\to\infty]{\mathcal{L}}\ \mathcal{N}(0\,,\,\sigma^2\,\Sigma).\tag{2}$$

6. Lorsqu'elle existe on définit la matrice $H^{(n)}=(H_{ij}^{(n)})_{1\le i,j\le n}=X_n\,({}^tX_n\,X_n)^{-1}\,{}^tX_n$. Pour $i\in\{1,\ldots,n\}$, montrer que $H_{ii}^{(n)}={}^tz_i\,({}^tX_n\,X_n)^{-1}z_i$ où $z_i={}^t(1,Z_i^{(1)},\ldots,Z_i^{(p)})$. Avec $\lambda^{(n)}$ la plus grande valeur propre de $({}^tX_n\,X_n)^{-1}$, en déduire que $H_{ii}^{(n)}\le\lambda^{(n)}\,\|z_i\|^2$, puis qu'il existe un réel $C>0$ et une variable aléatoire N_0 tel que pour tout $n\ge N_0$, tel que pour tout $i\in\{1,\ldots,n\}$, $H_{ii}^{(n)}\le\frac{C}{n}\,\|z_i\|^2$ presque sûrement.

7. On note $\widehat{Y}^{(n)}=X_n\,\widehat{\beta}^{(n)}$, $\widehat{\varepsilon}^{(n)}=(\widehat{\varepsilon}_i^{(n)})_{1\le i\le n}=Y^{(n)}-\widehat{Y}^{(n)}$ et $\widehat{\sigma}_n^2=\frac{1}{n-p-1}\sum_{i=1}^n(\widehat{\varepsilon}_i^{(n)})^2$. Montrer que $\widehat{\sigma}_n^2\ \xrightarrow[n\to+\infty]{\mathcal{P}}\ \sigma^2$. En vous servant de la question 6., montrer que $\dfrac{\widehat{\varepsilon}_i^{(n)}}{\widehat{\sigma}_n}\ \xrightarrow[n\to\infty]{\mathcal{L}}\ \mathcal{N}(0,1)$ pour tout $1\le i\le n$.

8. Soit une suite de v.a.i.i.d. positives $(U_i)_{i\in\mathbf{N}}$ de loi de probabilité $\mathbb{P}_U$ telle que $\mathbb{E}[U_i]<\infty$. Montrer que pour tout $x>0$, $\displaystyle\int_{nx}^{\infty}t\,d\mathbb{P}_U(t)\ \xrightarrow[n\to\infty]{}\ 0$ puis que $n\,\mathbb{P}(U_0>nx)\ \xrightarrow[n\to\infty]{}\ 0$. Pour $M_n=\dfrac{1}{n}\max\limits_{1\le i\le n}U_i$, déduire de ce qui précède que $\mathbb{P}(M_n\ge x)\ \xrightarrow[n\to\infty]{}\ 0$ puis que $\max\limits_{1\le i\le n}|H_{ii}^{(n)}|\ \xrightarrow[n\to+\infty]{\mathcal{P}}\ 0$.

9. On peut même montrer (ne pas le faire!) que $\max\limits_{1\le i\le n}|H_{ii}^{(n)}|\ \xrightarrow[n\to\infty]{p.s.}\ 0$. Qu'en déduire pour (2)?

Proof. 1. D'après les hypothèses, comme les ε_i sont indépendantes et gaussiennes, alors le vecteur $\varepsilon_i^{(n)}$ est un vecteur gaussien. Et on a $\varepsilon_i^{(n)} \overset{\mathcal{L}}{\sim} \mathcal{N}\big(0\,,\,\sigma^2\,I_n\big)$.

Y_i n'est pas forcément une v.a. gaussienne. Par exemple, si $p=1$, $\theta_0=0$, $\theta_1=1$ et $Z_i^{(1)} \overset{\mathcal{L}}{\sim} \mathcal{B}(1/2)$, alors alors $\mathbb{P}(Y_i=0)=\mathbb{P}(Y_i=1)=1/4$ donc Y_i ne peut être gaussienne.

2. On a $\mathbb{E}[Z_i^{(j)}]=m_j-m_j=0$ et $\mathrm{var}(Z_i^{(j)})=\sigma_j^2$.
3. On a $Y_i=\theta_0+\sum_{k=1}^p\theta_k\,W_i^{(k)}+\varepsilon_i=\theta_0+\sum_{k=1}^p\theta_k\,\big(Z_i^{(k)}+m_k\big)+\varepsilon_i=\big(\theta_0+\sum_{k=1}^p m_k\big)+\sum_{k=1}^p\theta_k\,Z_i^{(k)}+\varepsilon_i$. D'où $\beta_0=\theta_0+\sum_{k=1}^p m_k$ et $\beta_k=\theta_k$ pour $1\le k\le p$.
4. On a $\frac{1}{n}\,{}^tX_n\,X_n$ qui est une matrice de taille $(p+1)\times(p+1)$ que l'on peut encore écrire sous la forme:

$$\begin{pmatrix} 1 & \frac{1}{n}\sum_{i=1}^n Z_i^{(1)} & \frac{1}{n}\sum_{i=1}^n Z_i^{(2)} & \cdots & \frac{1}{n}\sum_{i=1}^n Z_i^{(n)} \\ \frac{1}{n}\sum_{i=1}^n Z_i^{(1)} & \frac{1}{n}\sum_{i=1}^n (Z_i^{(1)})^2 & \frac{1}{n}\sum_{i=1}^n Z_i^{(1)}Z_i^{(2)} & \cdots & \frac{1}{n}\sum_{i=1}^n Z_i^{(1)}Z_i^{(p)} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n}\sum_{i=1}^n Z_i^{(p)} & \frac{1}{n}\sum_{i=1}^n Z_i^{(p)}Z_i^{(1)} & \frac{1}{n}\sum_{i=1}^n Z_i^{(p)}Z_i^{(2)} & \cdots & \frac{1}{n}\sum_{i=1}^n (Z_i^{(p)})^2 \end{pmatrix}.$$

En utilisant la loi forte des grands nombres pour chacune des moyennes empiriques, comme $\mathbb{E}[Z_0^{(j)}]=0$, on a $\frac{1}{n}\sum_{i=1}^n Z_i^{(j)} \xrightarrow[n\to\infty]{p.s.} 0$, comme $\mathbb{E}\big[(Z_0^{(j)})^2\big]=\sigma_j^2$, on a $\frac{1}{n}\sum_{i=1}^n (Z_i^{(j)})^2 \xrightarrow[n\to\infty]{p.s.} \sigma_j^2$ et comme $\mathbb{E}\big[(Z_0^{(j_1)}Z_0^{(j_2)})\big]=\mathrm{cov}(Z_0^{(j_1)},Z_0^{(j_2)})=0$ pour $j_1\neq j_2$, on a $\frac{1}{n}\sum_{i=1}^n Z_i^{(j_1)}Z_i^{(j_2)} \xrightarrow[n\to\infty]{p.s.} 0$. Donc la suite de matrices tend presque-sûrement vers la matrice:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_1^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_p^2 \end{pmatrix}.$$

Comme cette matrice est définie positive, et comme l'inversion d'une matrice est une fonction mesurable continue sur les matrices définies positives, on en déduit que la suite $(\frac{1}{n}\,{}^tX_n\,X_n)_n^{-1}$ converge vers la matrice Σ telle que:

$$\Sigma=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1/\sigma_1^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1/\sigma_p^2 \end{pmatrix}.$$

5. Comme n est suffisamment grand, on sait que la matrice $({}^tX_n\,X_n)^{-1}$ existe. On a alors $\widehat{\beta}_n=\beta+({}^tX_n\,X_n)^{-1\,t}X_n\,\varepsilon^{(n)}$. Sachant X_n, alors $\widehat{\beta}_n=\beta+({}^tX_n\,X_n)^{-1\,t}X_n\,\varepsilon^{(n)}$ a une loi gaussienne car $\varepsilon^{(n)}$ est un vecteur gaussien et $({}^tX_n\,X_n)^{-1\,t}X_n$ est une matrice "non aléatoire". Donc $\widehat{\beta}_n$ est un vecteur gaussien et $\widehat{\beta}_n \overset{\mathcal{L}}{\sim} \mathcal{N}\big(\beta\,,\,\sigma^2\,({}^tX_n\,X_n)^{-1}\big)$.

On a d'après ce qui précède $(\frac{1}{n}\,{}^tX_n\,X_n)^{1/2}\,\sqrt{n}\,\big(\widehat{\beta}_n-\beta\big) \overset{\mathcal{L}}{\sim} \mathcal{N}\big(\,,\,\sigma^2\,I_{p+1}\big)$. Mais $(\frac{1}{n}\,{}^tX_n\,X_n)^{1/2} \xrightarrow[n\to+\infty]{\mathcal{P}} \Sigma^{-1/2}$ d'après la question précédente. Donc d'après le Lemme de Slutsky, $\Sigma^{-1/2}\,\sqrt{n}\,\big(\widehat{\beta}_n-\beta\big) \overset{\mathcal{L}}{\sim} \mathcal{N}\big(\,,\,\sigma^2\,I_{p+1}\big)$, d'où le résultat.

6. On a $H_i^{(n)}i={}^tJ_i\,H^{(n)}\,J_i={}^t\big({}^tX_n\,J_i\big)\,({}^tX_n\,X_n)^{-1}\big({}^tX_n\,J_i\big)$ avec $J_i={}^t\big(0,0,\cdots,0,1,0,\cdots,0\big)$. Mais ${}^tX_n\,J_i=z_i$, d'où le résultat.
Comme $({}^tX_n\,X_n)^{-1}$ est une matrice symétrique définie positive, on a $({}^tX_n\,X_n)^{-1}=Q\,D\,{}^tQ$, où Q est une matrice diagonale et D est une matrice diagonale composée des valeurs propres de $({}^tX_n\,X_n)^{-1}$. Ainsi $H_{ii}^{(n)}={}^t({}^tQz_i)\,D\,({}^tQz_i)={}^tz_i'\,D\,z_i'$ avec $z_i'={}^tQz_i$. Si $\lambda_1,\ldots,\lambda_{p+1}$ sont les valeurs propres de $({}^tX_n\,X_n)^{-1}$ et $z_i'={}^t(z_{i,1}',\ldots,z_{i,p+1}')$ alors ${}^tz_i'\,D\,z_i'=\sum_{k=1}^{p+1}\lambda_k\,(z_{i,k}')^2\le\lambda^{(n)}\sum_{k=1}^{p+1}(z_{i,k}')^2=\lambda^{(n)}\,\|z_i'\|^2$ car $\lambda^{(n)}$ est la plus grande valeur propre. Mais comme Q orthogonale, $\|z_i\|^2=\|z_i'\|^2$ d'où le résultat.
On a montrer que $(\frac{1}{n}\,{}^tX_n\,X_n)^{-1} \xrightarrow[n\to\infty]{p.s.} \Sigma$. Or l'application qui a une matrice associe ses valeurs propres est une application continue donc les valeurs propres de $(\frac{1}{n}\,{}^tX_n\,X_n)^{-1}$ convergent vers celles de Σ qui sont $(1,\sigma_1^{-2},\ldots,\sigma_1^{-2})$, d'où $n\,\lambda^{(n)} \xrightarrow[n\to\infty]{p.s.} \max(1,\sigma_1^{-2},\ldots,\sigma_1^{-2})$. En conséquence il existe une v.a. N_0 telle que $n\,\lambda^{(n)}\le 2\,\max(1,\sigma_1^{-2},\ldots,\sigma_1^{-2})$ p.s. pour tout $n\ge N_0$. D'où le résultat.
7. On a $\widehat{\sigma}_n^2=\frac{1}{n-p-1}\big(\|\varepsilon^{(n)}\|^2-\|H^{(n)}\varepsilon^{(n)}\|^2\big)$. Par la loi forte des grands nombres $\frac{1}{n-p-1}\,\|\varepsilon^{(n)}\|^2 \xrightarrow[n\to\infty]{p.s.} \sigma^2$. De plus $\mathbb{E}\big[\|H^{(n)}\varepsilon^{(n)}\|^2\big]=\mathbb{E}\big[{}^t\varepsilon^{(n)}\,H^{(n)}\,\varepsilon^{(n)}\big]$. En utilisant la trace, on en déduit que

$$\mathbb{E}\big[\|H^{(n)}\varepsilon^{(n)}\|^2=\mathbb{E}\big[\mathrm{Trace}\big({}^t\varepsilon^{(n)}\,H^{(n)}\,\varepsilon^{(n)}\big)\big]=\mathrm{Trace}\big(\mathbb{E}\big[H^{(n)}\big]\mathbb{E}\big[\varepsilon^{(n)}\,{}^t\varepsilon^{(n)}\big]\big)=\sigma^2\,\mathrm{Trace}\big(\mathbb{E}\big[H^{(n)}\big]\big)=\sigma^2\,(p+1).$$

On en déduit que $\frac{1}{n-p-1}\,\mathbb{E}\big[\|H^{(n)}\varepsilon^{(n)}\|^2\big] \xrightarrow[n\to\infty]{} 0$ et avec l'Inégalité de Markov, $\frac{1}{n-p-1}\,\|H^{(n)}\varepsilon^{(n)}\|^2 \xrightarrow[n\to+\infty]{\mathcal{P}} 0$. Avec le Lemme Slutsky, $\widehat{\sigma}_n^2 \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$.
On a $\widehat{\varepsilon}^{(n)}=\big(I_n-H^{(n)}\big)\,\varepsilon^{(n)}$, ce qui implique que, sachant X_n, la loi de $\widehat{\varepsilon}^{(n)}$ est gaussienne. Et $\mathbb{E}[\widehat{\varepsilon}^{(n)}\mid X_n]=0$ et $\mathrm{cov}(\widehat{\varepsilon}^{(n)}\mid X_n])=\mathbb{E}[\widehat{\varepsilon}^{(n)\,t}\widehat{\varepsilon}^{(n)}\mid X_n]=\sigma^2\,\big(I_n-H^{(n)}\big)$. D'où $\widehat{\varepsilon}_i^{(n)} \overset{\mathcal{L}}{\sim} \mathcal{N}\big(0\,,\,\sigma^2\,(1-H_{ii}^{(n)})\big)$ sachant X_n, ou encore $\frac{\widehat{\varepsilon}_i^{(n)}}{\sigma\,(1-H_{ii}^{(n)})^{1/2}} \overset{\mathcal{L}}{\sim} \mathcal{N}\big(0\,,\,1\big)$ sachant X_n. Mais comme $H_{ii}^{(n)} \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ et $\widehat{\sigma}_n^2 \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$, on en déduit avec le Lemme de Sltusky que $\frac{\widehat{\varepsilon}_i^{(n)}}{\widehat{\sigma}_n} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$ sachant X_n, donc également sans savoir X_n puisque la limite ne dépend pas X_n.

8. On a $\mathbb{E}[U_0] < \infty$ donc $\mathbb{E}[U_0] = \displaystyle\int_0^\infty t\, d\mathbb{P}_U(t) < \infty$. Par définition d'une intégrale généralisée, et comme $nx \underset{n\to\infty}{\longrightarrow} +\infty$ on en déduit que $\displaystyle\int_{nx}^\infty t\, d\mathbb{P}_U(t) \underset{n\to\infty}{\longrightarrow} 0$.
   On a $\displaystyle\int_{nx}^\infty t\, d\mathbb{P}_U(t) \geq nx \int_{nx}^\infty d\mathbb{P}_U(t) = nx\, \mathbb{P}(U_0 \geq nx)$. D'où $nx\, \mathbb{P}(U_0 \geq nx) \underset{n\to\infty}{\longrightarrow} 0$.
   Pour $x > 0$, on a $\mathbb{P}(M_n \leq x) = \prod_{i=1}^n \mathbb{P}(U_i \leq nx) = \big(1 - \mathbb{P}(U_i > nx)\big)^n = \exp\Big(n \ln\big(1 - \mathbb{P}(U_i > nx)\big)\Big)$. En utilisant un DL d'ordre 1 de $\ln(1+u)$, on a $n \ln\big(1 - \mathbb{P}(U_i > nx)\big) \sim -n\, \mathbb{P}(U_i > nx) \underset{n\to\infty}{\longrightarrow} 0$. D'où $\mathbb{P}(M_n \leq x) \underset{n\to\infty}{\longrightarrow} 1$ impliquant $M_n \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0$.
   Si on pose $U_i = C\, \|z_i\|^2$, alors d'après la question 6., $\displaystyle\max_{1\leq i\leq n} \left|H_{ii}^{(n)}\right| \leq M_n$. Mais on a $\|z_i\|^2$ suite de v.a.i.i.d. telle que $\mathbb{E}[\|z_1\|^2] = 1 + \sigma_1^2 + \cdots + \sigma_p^2 < \infty$, et $\displaystyle\max_{1\leq i\leq n} \left|H_{ii}^{(n)}\right| \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0$.

9. En ayant montré la convergence presque sûre, on en déduit que la normalité asymptotique de $\widehat{\beta}^{(n)}$ est conservée même quand la loi de (ε_i) n'est pas gaussienne.

$\square$