Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2018 – 2019

Econométrie II

Contrôle continu n°1, mars 2019

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

Soit $(Y_i)_{1\le i\le n}$ une famille de variables aléatoires telle que:

$$Y_i = \mu_i^* + \varepsilon_i \qquad \text{pour tout } i \in \{1,\dots,n\}, \tag{1}$$

où $(\varepsilon_i)_{i\in\mathbf{N}}$ est une suite de variables aléatoires identiquement distribuées indépendantes centrées et gaussiennes telles que $\mathrm{var}(\varepsilon_1)=\sigma^2<\infty$ et $\mu^*=(\mu_i^*)_{1\le i\le n}\in\mathbf{R}^n$, vecteur fixé, appartenant à F un sous-espace vectoriel de $\mathbf{R}^n$. Dans toute la suite on notera $\|\cdot\|$ la norme sur $\mathbf{R}^n$ telle que pour tout $Z\in\mathbf{R}^n$, $\|Z\|^2={}^tZ\,Z$ (${}^tA$ désignant la transposée de la matrice A). On définit les estimateurs respectifs de μ^* et σ^2 par

$$\widehat{\mu}=\mathrm{Argmin}_{\mu\in F}\|Y-\mu\| \qquad \text{et} \qquad \widehat{\sigma}_n^2=\frac{1}{n}\,\|Y-\widehat{\mu}\|^2.$$

1. On note $Y=(Y_i)_{1\le i\le n}$. Ecrire le modèle (1) sous une forme vectorielle. Calculer $\mathbb{E}(Y)$ et $\mathrm{cov}(Y)$ (matrice de variance-covariance de Y).

2. Soit $X_0\in\mathbf{R}^n$ tel que ${}^tX_0=(1,1,\dots,1)$. On suppose dans cette question que $F=\mathrm{Vect}(X_0)=[X_0]$, sous-espace engendré par X_0. Montrer alors que $\widehat{\mu}=\overline{Y}_n\,X_0$, où $\overline{Y}_n=\frac{1}{n}\sum_{i=1}^n Y_i$. Calculer le risque quadratique de $\widehat{\mu}$, c'est-à-dire $\mathbb{E}(\|\widehat{\mu}-\mu^*\|^2)$. Calculer $\mathbb{E}(\widehat{\sigma}_n^2)$ et montrer que $\widehat{\sigma}_n^2 \xrightarrow[n\to+\infty]{p.s.} \sigma^2$. Montrer que $(\widehat{\mu},\widehat{\sigma}_n^2)$ est l'estimateur par maximum de vraisemblance de (μ^*,σ^2). Est-ce un estimateur efficace?

3. Soit $X_1\in\mathbf{R}^n$, avec (X_0,X_1) non colinéaires. On suppose maintenant que $F=\mathrm{Vect}(X_0,X_1)=[X_0,X_1]$. Rappeler quelle est la loi de $\|\widehat{\mu}-\mu^*\|^2$, en déduire le risque quadratique de $\widehat{\mu}$. Montrer que $\widehat{\sigma}_n^2 \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$.

4. On suppose que la variable X_1 n'est en réalité pas explicative, c'est-à-dire que $F=[X_0]$ mais que $\widehat{\mu}$ a la même expression que dans la question précédente (donc construit avec X_0 et X_1) car on a cru que X_1 était explicative. Déterminer alors la loi de $\|Y-\widehat{\mu}\|^2$ puis le risque quadratique de $\widehat{\mu}$. Comparer ce risque avec celui obtenu dans le cas où l'on sait à l'avance que seule X_0 est explicative.

5. On veut tester $H_0:\ F=\mathrm{Vect}(X_0)$ contre $H_1:\ F=\mathrm{Vect}(X_0,X_1)$. Pour cela on considère la statistique
$$\widehat{S}=\frac{\|Y-P_{[X_0]}Y\|^2-\|Y-P_{[X_0,X_1]}Y\|^2}{\frac{1}{n-2}\,\|Y-P_{[X_0,X_1]}Y\|^2},$$
où P_E désigne la matrice de projection orthogonale sur le sev E. Montrer que $\widehat{S}$ est une variable positive, puis que $\widehat{S}$ suit une loi de Fisher $F(1,n-2)$ sous H_0.

6. On se place maintenant sous l'hypothèse H_1 avec précisément $\mu^*=a_0^*X_0+a_1^*X_1$, où $a_1^*\neq 0$. Montrer alors que $\mathbb{E}(\|Y-P_{[X_0]}Y\|^2-\|Y-P_{[X_0,X_1]}Y\|^2)=n(a_1^*)^2\sigma_{1,n}^2+\sigma^2$ avec $\sigma_{1,n}^2=\frac{1}{n}\,{}^tX_1X_1-(\frac{1}{n}{}^tX_0X_1)^2$. On suppose que $\sigma_{1,n}^2 \xrightarrow[n\to\infty]{} \sigma_1^2$ avec $\sigma_1^2>0$. En déduire alors que sous H_1, $\widehat{S} \xrightarrow[n\to+\infty]{\mathcal{P}} +\infty$. Montrer que ce sera notamment le cas si X_1 est la réalisation de n vaiid de variance σ_1^2, indépendantes de ε. Et que se passe-t-il si $X_1={}^t(1,2,3,\dots,n)$?