Estatística 2 - Lista 4

Luciano Fabio Busatto Venturin
09/08/21

1) (a) Let $\theta_0 = \underset{\theta \in \mathbb{R}^k}{\arg\min}\, E[(y - x'\theta)'(y - x'\theta)\tau(x)]$

The FOC of this problem gives us

$$0 = \frac{\partial}{\partial\theta} E[(y - x'\theta)^2 \tau(x)]\Big|_{\theta=\theta_0} = 2E[\tau(x)x(y - x'\theta_0)] \iff$$

$$E(\tau(x)xx')\theta_0 = E(\tau(x)xy) \Rightarrow \theta_0 = (E(\tau(x)xx'))^{-1} E(\tau(x)xy)$$

The moment estimator of θ_0 is then

$$\hat\theta = (1/n \sum_{i=1}^n \tau(x_i)x_i x_i')^{-1}(1/n \sum_{i=1}^n \tau(x_i)x_i y_i)$$

Indeed, the sample MTSE is

$$\hat T(\theta) = 1/n \sum_{i=1}^n (y_i - x_i'\theta)^2 \tau(x_i)$$

If we let $\tilde\theta = \underset{\theta \in \mathbb{R}^k}{\arg\min}\, 1/n \sum_{i=1}^n (y_i - x_i'\theta)^2 \tau(x_i)$

we have, by the FOC,

$$0 = \frac{2}{n}\sum_{i=1}^n x_i(y_i - x_i'\tilde\theta)\tau(x_i) \iff$$

$$(1/n \sum_{i=1}^n x_i x_i' \tau(x_i))\tilde\theta = (1/n \sum_{i=1}^n x_i y_i \tau(x_i)) \iff$$

$$\tilde\theta = (1/n \sum_{i=1}^n x_i x_i' \tau(x_i))^{-1}(1/n \sum_{i=1}^n x_i y_i \tau(x_i)) = \hat\theta \text{, i.e.,}$$

the moment estimator of the minimizer of the MTSE is the minimizer of the sample MTSE.

(b) Assuming $y_i = x_i'\theta + e_i \quad \forall i = 1, \dots, n$

$$\hat{\theta} = \left(1/n \sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \left(1/n \sum_{i=1}^n x_i x_i' \theta \tau(x_i) + 1/n \sum_{i=1}^n x_i e_i \tau(x_i)\right)$$
$$= \theta + \left(1/n \sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \left(1/n \sum_{i=1}^n x_i e_i \tau(x_i)\right)$$

Thus, $E(\hat{\theta} \mid X) = \theta + \left(1/n \sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \left(1/n \sum_{i=1}^n x_i \tau(x_i) E(e_i \mid X)\right)$

If $E(e_i \mid X) = 0 \quad \forall i = 1, \dots, n$, then

$$E(\hat{\theta} \mid X) = \theta \Rightarrow E(\hat{\theta}) = E(E(\hat{\theta} \mid X)) = \theta$$

(c) Under the condition above, we have that

$$Var(\hat{\theta} \mid X) = E((\hat{\theta} - \theta)(\hat{\theta} - \theta)' \mid X) =$$
$$= E\left(\left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \left(\sum_{i=1}^n x_i \tau(x_i) e_i\right) \left(\sum_{i=1}^n x_i' \tau(x_i) e_i\right) \left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \Big| X\right)$$
$$= \left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} E\left(\sum_{i=1}^n x_i x_i' (\tau(x_i))^2 e_i^2 + \sum_{i \neq j} x_i x_j' \tau(x_i) \tau(x_j) e_i e_j \mid X\right) \left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1}$$

$$= \left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1} \left(\sum_{i=1}^n x_i x_i' (\tau(x_i))^2 E(e_i^2 \mid X)\right) \left(\sum_{i=1}^n x_i x_i' \tau(x_i)\right)^{-1}$$

② (a) Since $e_i | x_i \sim N(0, \sigma_i^2)$,

$y_i | x_i \sim N(x_i'\beta, \sigma_i^2)$. By independence,

$$f(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_{i=1}^n f(y_i | x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{ \frac{-1}{2\sigma_i^2} (y_i - x_i'\beta)^2 \right\}$$

$$= (2\pi)^{-\frac{n}{2}} \left(\prod_{i=1}^n \sigma_i^2 \right)^{-1/2} \exp\left\{ -\frac{1}{2} \sum_{i=1}^n \frac{(y_i - x_i'\beta)^2}{\sigma_i^2} \right\} \equiv L(\beta; \sigma_1^2, \dots, \sigma_n^2)$$

and $l(\beta; \sigma_1^2, \dots, \sigma_n^2) = \operatorname{Log} L(\beta; \sigma_1^2, \dots, \sigma_n^2) =$

$$= -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^n \log \sigma_i^2 - \frac{1}{2} \sum_{i=1}^n \frac{(y_i - x_i'\beta)^2}{\sigma_i^2}$$

Let $(\hat{\beta}_{mle}, \sigma^2_{1mle}, \dots, \sigma^2_{nmle}) = \underset{\beta \in \mathbb{R}^k, \sigma_1^2, \dots, \sigma_n^2 > 0}{\operatorname{argmin}} \; l(\beta; \sigma_1^2, \dots, \sigma_n^2)$

The FOCs give us

$$0 = \frac{\partial}{\partial \beta} l(\beta; \sigma_1^2, \dots, \sigma_n^2) \bigg|_{\beta = \hat{\beta}_{mle}} = \sum_{i=1}^n \frac{x_i (y_i - x_i'\hat{\beta}_{mle})}{\hat{\sigma}^2_{imle}} \quad \text{and}$$

$$0 = \frac{\partial}{\partial \sigma_i^2} l(\beta; \sigma_1^2, \dots, \sigma_n^2) \bigg|_{\sigma_i^2 = \sigma^2_{imle}} = \frac{-1}{2\hat{\sigma}^2_{imle}} + \frac{1}{2} \frac{(y_i - x_i'\hat{\beta}_{mle})^2}{\hat{\sigma}^4_{imle}} \quad \forall i.$$

The first equation gives us

$$\hat{\beta}_{mle} = \left(\sum_{i=1}^n \frac{x_i x_i'}{\hat{\sigma}^2_{imle}} \right)^{-1} \left(\sum_{i=1}^n \frac{x_i y_i}{\hat{\sigma}^2_{imle}} \right)^2$$

Moreover, for each i,

$$\hat{\sigma}^2_{imle} = (y_i - x_i'\hat{\beta}_{mle})^2$$

(b) Let $\Sigma = \operatorname{Var}(e | X) = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{pmatrix}$

Consider the model

$$\tilde{y} = \tilde{X}\beta + \tilde{e} \quad \text{where } \tilde{y} = \Sigma^{-1/2} y,\ \tilde{X} = \Sigma^{-1/2} X$$

and $\tilde{e} = \Sigma^{-1/2} e$. The GLS estimator is given by

$$\hat{\beta}_{GLS} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} = (X'\Sigma^{-1}X)^{-1} X\Sigma^{-1}y$$

Now, see that $X'\Sigma^{-1}X = (x_1 \cdots x_n)\begin{pmatrix} 1/\sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/\sigma_n^2 \end{pmatrix}\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$

$= \sum_{i=1}^{n} \frac{x_i x_i'}{\sigma_i^2}$ and $X\Sigma^{-1}y = (x_1 \cdots x_n)\begin{pmatrix} 1/\sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/\sigma_n^2 \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$= \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2}$, thus

$$\hat{\beta}_{GLS} = \left(\sum_{i=1}^{n} \frac{x_i x_i'}{\sigma_i^2}\right)^{-1}\left(\sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2}\right) = \hat{\beta}_{MLE}$$

(3) We have that $P(|\hat{\theta}_n - \theta_0| > \varepsilon) =$

$$P((\hat{\theta}_n - \theta_0)^2 > \varepsilon^2) \leq \frac{E(\hat{\theta}_n - \theta_0)^2}{\varepsilon^2} =$$

$$= \frac{E_{\theta_0}(\hat{\theta}_n - E_{\theta_0}(\hat{\theta}_n) + E_{\theta_0}(\hat{\theta}_n) - \theta_0)^2}{\varepsilon^2} =$$

$$= \frac{E_{\theta_0}(\hat{\theta}_n - E_{\theta_0}(\hat{\theta}_n))^2 + 2E_{\theta_0}(\hat{\theta}_n - E_{\theta_0}(\hat{\theta}_n)) E_{\theta_0}(E(\hat{\theta}_n) - \theta_0) + E_{\theta_0}(E_{\theta_0}(\hat{\theta}_n) - \theta_0)^2}{\varepsilon}$$

$$= \frac{Var(\hat{\theta}_n) + [Bias(\hat{\theta}_n)]^2}{\varepsilon} \xrightarrow{n} 0 \quad \forall \varepsilon > 0,$$

where the first inequality follows by the Markov inequality.

Thus, $\hat{\theta}_n \xrightarrow{P} \theta_0$.

(4) (a.6.7) First, let $\hat\theta = \log \hat\mu = \frac{1}{n}\sum_{i=1}^{n} \log X_i$

By the Weak Law of Large numbers,
$\hat\theta \xrightarrow{P} E(\log X_i) := \theta$

Now, see that $\hat\mu = \exp(\hat\theta)$. Since $\exp(\cdot)$ is a continuous function, by the Continuous Mapping Theorem
$\hat\mu = \exp(\hat\theta) \longrightarrow \mu = \exp(\theta) = \exp(E(\log X))$.

(b. 6.13) (a) First, by the Taylor expansion of $\hat\beta = f(\hat\mu) = \hat\mu^2$ around μ, we have

$\hat\mu^2 = \mu^2 + 2\bar\mu(\hat\mu - \mu)$ for some $\bar\mu$ between $\hat\mu$ and μ. Thus,

$\sqrt{n}(\hat\mu^2 - \mu^2) = (2\bar\mu)\sqrt{n}(\mu - \mu_0)$. Since $\hat\mu \xrightarrow{P} \mu$, $\bar\mu \xrightarrow{P} \mu$. By the Slutsky Theorem,

$$\sqrt{n}(\hat\mu^2 - \mu^2) \longrightarrow (2\mu)N(0, v^2) = N(0, (2\mu)^2 v^2)$$

(b) If $\mu = 0$, then $\sqrt{n}(\hat\mu^2 - \mu^2) = \sqrt{n}\,\hat\mu^2 = \frac{(\sqrt{n}\hat\mu)^2}{\sqrt{n}}$. Now, since $\sqrt{n}\hat\mu \longrightarrow N(0, v^2)$,

by the Slutsky Theorem, $\sqrt{n}(\hat\mu^2 - \mu^2) \xrightarrow{P} 0$ thus $\sqrt{n}(\hat\mu^2 - \mu^2) \xrightarrow{d} 0$. In the asymptotic distribution of item (a), the asymptotic variance of $\sqrt{n}(\hat\mu^2 - \mu^2)$ is $(2 \cdot 0)^2 v^2 = 0$, thus it converges to a degenerated distribution.

(c) $n\hat\mu^2 = (\sqrt{n}\hat\mu)$. Since $\sqrt{n}\hat\mu \xrightarrow{d} N(0, v^2)$, $\left(\frac{\sqrt{n}\hat\mu}{v}\right) \xrightarrow{d} N(0, 1)$. By the CMT,

$\frac{n\hat{\mu}^2}{v^2} \xrightarrow{d} \chi^2_1$, thus $n\hat{\mu}^2 \xrightarrow{d} v^2\chi^2_1$

(d) The difference shows us that when $\mu = 0$, the rate of convergence of $\hat{\mu}^2$ is higher and the delta method fails in providing the correct asymptotic distribution.