16.3 Asymptotic normality of LSE under homoscedasticity

10

+19

Repetition

(A0) $(Y_i, X_i^T)^T \overset{iid}{\sim} (Y, X^T)^T$, $\mathbb{E}(Y|X) = X^T\beta$

(A1) $\mathbb{E}|X_j X_\ell| < \infty$, $\mathbb{E}XX^T = W > 0$

elements of X

$V := W^{-1} = (\mathbb{E}(XX^T))^{-1}$

(A2 homoscedastic)

$$\sigma^2(X) := \mathrm{var}(Y|X) = \mathrm{var}(\varepsilon|X) = \sigma^2$$

$$\Rightarrow \mathrm{var}\,\varepsilon = \sigma^2$$

Theorem 16.4 Asymptotic normality of LSE in homoscedastic case

Assumptions (A0), (A1), (A2 homoscedastic),
additionally $\mathbb{E}|\varepsilon^2 X_j X_\ell| < \infty$ $\forall j, \ell$.

elements of a generic $X = (X_0, \dots, X_{k-1})^T$

Then

$$\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}_k(0_k, \sigma^2 V), \quad n \to \infty$$

$$\hat{\theta}_n = \ell^T\hat{\beta}_n, \quad \sqrt{n}(\hat{\theta}_n - \underbrace{\ell^T\beta}_{\theta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2 \ell^T V \ell), \quad n \to \infty$$

$$\hat{\xi}_n = L\hat{\beta}_n, \quad \sqrt{n}(\hat{\xi}_n - \underbrace{L\beta}_{\xi}) \xrightarrow{\mathcal{D}} \mathcal{N}_m(0_m, \sigma^2 L V L^T), \quad n \to \infty$$

Proof: It is sufficient to show $\underbrace{\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(\dots)}$.
The rest follows from Cramér-Wold & properties of a normal distribution.

will be shown jointly with heteroscedastic case

Usage of Theorem 16.4

$$\sqrt{n}(\hat{\beta}_n - \beta) \overset{\cdot}{\sim} N(0, \sigma^2 V)$$

$$(\hat{\beta}_n - \beta) \overset{\cdot}{\sim} N(0, \sigma^2 \tfrac{1}{n} V)$$

$$\hat{\beta}_n \overset{\cdot}{\sim} N(\beta, \sigma^2 \cdot \tfrac{1}{n} V)$$

If I knew σ^2 and matrix $V = (\mathbb{E} X X^T)^{-1}$, asymptotic confidence intervals (& tests) can be constructed for elements of $\beta = (\beta_0, \dots, \beta_{k-1})^T$:

$$\hat{\beta}_{n,j} \pm \sqrt{\sigma^2 \cdot \tfrac{v_{jj}}{n}}\; \underbrace{u(1-\tfrac{\alpha}{2})}_{(1-\frac{\alpha}{2}) \text{ quantile of } N(0,1)}$$

or for linear combinations $l^T \beta$:

$$l^T \hat{\beta}_n \pm \sqrt{\sigma^2 \cdot \tfrac{1}{n} l^T V l}\; u(1-\tfrac{\alpha}{2})$$

We don't know σ^2, V.

How about to replace them by their consistent estimators: $MSe_n \xrightarrow{P} \sigma^2$

$$V_n = (X_n^T X_n)^{-1} = \left(\sum_{i=1}^{n} X_i X_i^T\right)^{-1} \qquad \underbrace{n V_n \xrightarrow{P} V}_{V_n \approx \frac{1}{n} V}$$

$\to$ $(1-\alpha)\cdot 100\%$ asymptotic conf. interval for $l^T\beta$?

$$l^T \hat{\beta}_n \pm \sqrt{MSe_n\, l^T V_n l}\; \cdot u(1-\tfrac{\alpha}{2})$$

Is this true? YES! $\to$ next page

16.3.1 Asymptotic validity of the classical inference under homoscedast. but non-normality

We know (for long), for $l = (l_0, \dots, l_{k-1})^T \in \mathbb{R}^k$

With homoscedasticity & NORMALITY, for each n:

$$T_n = \frac{l^T \hat{\beta}_n - l^T \beta}{\sqrt{MSe_{,n}\, l^T (X_n^T X_n)^{-1} l}} \sim t_{n-k}$$

→ $(1-\alpha)100\%$ conf. interval for $l^T\beta$ is

$$l^T \hat{\beta}_n \pm \sqrt{MSe_{,n}\, l^T \underbrace{(X_n^T X_n)}_{V_n}{}^{-1} l} \cdot t_{n-k}\left(1 - \frac{\alpha}{2}\right)$$

→ test of $H_0: l^T\beta = \theta^0$, for $\theta^0 \in \mathbb{R}$

Is the above inference valid even without normality, at least asymptotically? YES ↓

$$T_n = \frac{l^T \hat{\beta}_n - l^T \beta}{\sqrt{MSe_{,n}\, l^T (X_n^T X_n)^{-1} l}} = \underbrace{\frac{\sqrt{n}\,(l^T \hat{\beta}_n - l^T \beta)}{\sqrt{\sigma^2 l^T V l}}}_{\xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)} \cdot \underbrace{\sqrt{\frac{\sigma^2 l^T V l}{n\, MSe_{,n}\, l^T (X_n^T X_n)^{-1} l}}}_{\xrightarrow[n\to\infty]{P} 1}$$

Theorem 16.4 ; Lemma 16.1 + Theor. 16.3

Cramér–Слуцкий ⟹

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad n \to \infty$$

= Consequence of Theorem 16.4.

Hence $\forall \theta^0 \in \mathbb{R}$, $\theta^0 = l^T \beta^0$

$$P\left(|T_n| < u(1-\tfrac{\alpha}{2});\ \theta = \theta^0\right) \to 1-\alpha, \quad n \to \infty$$

$$P\left(\left|\frac{l^T\hat{\beta}_n - l^T\beta}{\sqrt{MSe_n\, l^T(X_n^T X_n)l}}\right| < u(1-\tfrac{\alpha}{2});\ \theta = \theta^0\right) \to 1-\alpha$$

$$P\Big(\underbrace{\big(l^T\hat{\beta}_n \pm \sqrt{MSe_n\, l^T(X_n^T X_n)l} \cdot u(1-\tfrac{\alpha}{2})\big)}_{\text{interval } I_n^W \text{ on slide } \underline{13}} \ni \theta^0;\ \theta = \theta^0\Big) \to 1-\alpha$$

It also holds that $T_\nu \sim t_\nu \Rightarrow T_\nu \xrightarrow{\mathcal{D}} N(0,1)$, $\nu \to \infty$.

Hence $t_{n-k}(1-\frac{\alpha}{2}) \longrightarrow u(1-\frac{\alpha}{2})$, $n \to \infty$.

$$\Rightarrow P\Big(\underbrace{\big(l^T\hat{\beta}_n \pm \sqrt{MSe_n\, l^T(X_n^T X_n)l} \cdot t_{n-k}(1-\tfrac{\alpha}{2})\big)}_{\text{interval } I_n^t \text{ on slide } \underline{13}} \ni \theta^0;\ \theta = \theta^0\Big) \to 1-\alpha$$

Standard inference used under assumption of normality is <u>asymptotically</u> (for $n \to \infty$) valid even <u>without</u> normality

(assumptions of theorem 16.4 are needed)

= reasonable behavior of

(a) error terms

(b) covariates

Analogous steps in case of a vector parameter $\xi = L\beta$, $L_{m \times k}$ matrix with linear indep. rows, $m \le k$

We know (for long)

with homoscedasticity & NORMALITY, for each n

$$Q_n = \frac{1}{m}(L\hat{\beta}_n - L\beta)^T \{ MSe_{,n} L (X_n^T X_n)^{-1} L^T \}^{-1} (L\hat{\beta}_n - L\beta) \sim F_{m, n-k}$$

$\rightarrow$ $(1-\alpha)100\%$ conf. ellipsoid for $\xi = L\beta$ is

$$\left\{ \xi : \frac{1}{m}(L\hat{\beta}_n - \xi)^T \{ MSe_{,n} L (X_n^T X_n)^{-1} L^T \}^{-1} (L\hat{\beta}_n - \xi) < F_{m, n-k}(1-\alpha) \right\}$$

$\rightarrow$ test of $H_0 : L\beta = \xi^0$, $\xi^0 \in \mathbb{R}^m$

Is the above inference valid even without normality, at least asymptotically?

YES

$$m Q_n = \underbrace{\sqrt{n}(L\hat{\beta}_n - L\beta)}_{\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2 L V L^T)}{}^T \underbrace{\{ MSe_{,n} L\, n (X_n^T X_n)^{-1} L^T \}}_{\xrightarrow{P} (\sigma^2 L V L^T)^{-1}}{}^{-1} \underbrace{(L\hat{\beta}_n - L\beta)\sqrt{n}}_{\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2 L V L^T)}$$

Cramér - Слуцкий $\Longrightarrow$

$$m . Q_n \xrightarrow{\mathcal{D}} \chi^2_m, \quad n \to \infty$$

$\equiv$ Consequence of Theorem 16.4

Hence $\forall\, \xi^0 \in \mathbb{R}^m$, $\xi^0 = L\beta$ $\boxed{15}$

$$P\left((L\hat{\beta} - L\beta)^T \{MSe_n\, L\, (X_n^T X_n)^{-1} L^T\}^{-1} (L\hat{\beta} - L\beta) < \chi^2_m(1-\alpha);\ L\beta = \xi^0\right) \xrightarrow[n\to\infty]{} 1-\alpha$$

→ defines a confidence ellipsoid

$$K_n^{\chi} = \left\{\xi \in \mathbb{R}^m : (\xi - L\hat{\beta})^T \{MSe_n\, L\, (X_n^T X_n)^{-1} L^T\}^{-1} (\xi - L\hat{\beta}) < \chi^2_m(1-\alpha)\right\}$$

with asymptotic coverage of $1-\alpha$

$\boxed{16}$

It also holds that $Q_\nu \sim F_{m,\nu} \Rightarrow mQ_\nu \xrightarrow[\nu\to\infty]{\mathcal{D}} \chi^2_m$.

Hence $m \cdot F_{m,n-k}(1-\alpha) \xrightarrow[n\to\infty]{} \chi^2_m(1-\alpha)$

$$\Rightarrow P\left(K_n^F \ni \xi^0;\ L\beta = \xi^0\right) \xrightarrow[n\to\infty]{} 1-\alpha,$$

where

$$\underbrace{K_n^F = \left\{\xi \in \mathbb{R}^m : (\xi - L\hat{\beta})^T \{MSe_n\, L\, (X_n^T X_n)^{-1} L^T\}^{-1} (\xi - L\hat{\beta}) < m \cdot F_{m,n-k}(1-\alpha)\right\}}_{\text{confidence } \cancel{\text{el}} \text{ ellipsoid being used under assumption of normality}}$$

Overall summary

Even without normality (still with homoscedasticity) standard inference based on statistics

$$T_n = \frac{l^T\hat{\beta}_n - l^T\beta}{\sqrt{MSe_{,n}\; l^T(X_n^T X_n)^{-1} l}} \quad \sim t_{n-k}$$

and $$Q_m = \frac{1}{m}(L\hat{\beta}_n - L\beta)^T \{MSe_{,n} L(X_n^T X_n)^{-1} L^T\}^{-1} (L\hat{\beta}_n - L\beta) \quad \sim F_{m,n-k}$$

is asymptotically valid (under assumptions of Theorem 16.4 ($\approx$ CLT)).