Least Squares Asymptotics

- Convergence of Estimators: Review
- Least Squares Assumptions
- Least Squares Estimator
- Asymptotic Distribution
- Hypothesis Testing

Convergence of Estimators

- Let $\hat{\boldsymbol{\theta}}_n$ be an estimator of a parameter vector $\boldsymbol{\theta}$ based on a sample of size n. $\{\hat{\boldsymbol{\theta}}_n\}$ is a sequence of random variables.
- $\hat{\boldsymbol{\theta}}_n$ is **consistent** for $\boldsymbol{\theta}$ if

$$\underset{n\to\infty}{\mathrm{p}\lim}\,\hat{\boldsymbol{\theta}}_n = \boldsymbol{\theta} \quad \text{or} \quad \hat{\boldsymbol{\theta}}_n \to_p \boldsymbol{\theta}$$

Convergence of Estimators

- A consistent estimator $\hat{\boldsymbol{\theta}}_{\mathrm{n}}$ is **asymptotically normal** if $\sqrt{\mathrm{n}}(\hat{\boldsymbol{\theta}}_{\mathrm{n}} - \boldsymbol{\theta}) \to_{\mathrm{d}} \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$
  Such an estimator is called **√n-consistent**.
- The **asymptotic variance** $Asy\,\mathrm{var}(\hat{\boldsymbol{\theta}}_{n})$ is derived from the variance of the limiting distribution of $\sqrt{\mathrm{n}}(\hat{\boldsymbol{\theta}}_{\mathrm{n}} - \boldsymbol{\theta})$
- $Asy\,\mathrm{var}(\hat{\boldsymbol{\theta}}_{n}) = \boldsymbol{\Sigma} / n$

Convergence of Estimators

- Delta Method:
  $\sqrt{n}(\theta_n - \theta) \rightarrow_d N(\mathbf{0}, \Sigma)$
  $\Rightarrow$
  $\sqrt{n}[\mathbf{a}(\theta_n) - \mathbf{a}(\theta)] \rightarrow_d N(\mathbf{0}, \mathbf{A}(\theta)\Sigma\mathbf{A}(\theta)')$

  where $\mathbf{a}(.)$: $R^K \rightarrow R^r$ has continuous first derivatives with $\mathbf{A}(\theta)$ defined by

$$\mathbf{A}(\boldsymbol{\theta}) = \frac{\partial \mathbf{a}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$$

Least Squares Assumptions

- A1: *Linearity (Data Generating Process)*
  - $y_i = \mathbf{x}_i'\beta + \varepsilon_i$ (i=1,2,…,n)
  - The stochastic process that generated the finite sample $\{y_i, \mathbf{x}_i\}$ must be stationary and asymptotic independent:
    - $\lim_{n\to\infty} \sum_i \mathbf{x}_i\mathbf{x}_i'/n = E(\mathbf{x}_i\,\mathbf{x}_i')$
    - $\lim_{n\to\infty} \sum_i \mathbf{x}_i' y_i/n = E(\mathbf{x}_i' y_i)$
  - Finite 4th Moments for Regressors: $E[(x_{ik}x_{ij})^2]$ exists and is finite for all k,j (=1,2,…,K).

Least Squares Assumptions

- A2: *Exogeneity*
  - $E(\varepsilon_i|\mathbf{X}) = 0$ or $E(\varepsilon_i|\mathbf{x}_i) = 0$
  - $E(\varepsilon_i|\mathbf{X}) = 0 \Rightarrow E(\varepsilon_i|\mathbf{x}_i) = 0 \Rightarrow E(x_{ik}\varepsilon_i) = 0$.
- A2': *Weak Exogeneity*
  - $E(x_{ik}\varepsilon_i) = 0$ for all $i=1,2,\ldots,n$; $k=1,2,\ldots K$.
  - It is possible that $E(x_{jk}\varepsilon_i) \neq 0$ for some j,k and $j \neq i$.
- Define $\mathbf{g}_i = \mathbf{x}_i\varepsilon_i = [x_{i1},x_{i2},\ldots,x_{iK}]\varepsilon_i$, then
  $E(\mathbf{g}_i) = E(\mathbf{x}_i\varepsilon_i) = \mathbf{0}$
  What is $Var(\mathbf{g}_i) = E(\mathbf{g}_i\mathbf{g}_i') = E(\mathbf{x}_i\mathbf{x}_i'\varepsilon_i^2)$?

Least Squares Assumptions

– We assume

- $\lim_{n\to\infty} \sum_i \mathbf{x}_i\varepsilon_i/n = E(\mathbf{g}_i)$
- $\lim_{n\to\infty} \sum_i \mathbf{x}_i\mathbf{x}_i'\varepsilon_i^2/n = E(\mathbf{g}_i\mathbf{g}_i')$
- $E(\mathbf{g}_i) = \mathbf{0}$ (A3)
- $Var(\mathbf{g}_i) = E(\mathbf{g}_i\mathbf{g}_i') = \mathbf{\Omega}$ is nonsingular

– CLT states that

$$\sqrt{n}\bar{\mathbf{g}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{g}_i \to_d N(\mathbf{0}, \mathbf{\Omega})$$

$$where\ Asy\ \mathrm{var}(\bar{\mathbf{g}}) = \mathbf{\Omega}/n$$

Least Squares Assumptions

- A3: *Full Rank*
  - $E(\mathbf{x}_i\mathbf{x}_i') = \boldsymbol{\Sigma}_{xx}$ is nonsingular.
  - Let $\mathbf{Q} = \sum_i \mathbf{x}_i\mathbf{x}_i'/n = \mathbf{X}'\mathbf{X}/n$
  - $\lim_{n\to\infty}\mathbf{Q} = E(\mathbf{x}_i\mathbf{x}_i') = \boldsymbol{\Sigma}_{xx}$ (A1)
  - Therefore **Q** is nonsingular (no multicolinearity in the limit).

Least Squares Assumptions

- A4: *Spherical Disturbances*
  - Homoscedasticity and No Autocorrelation
    - Var($\boldsymbol{\varepsilon}$|**X**) = E($\boldsymbol{\varepsilon\varepsilon}$'|**X**) = $\sigma^2\mathbf{I}_n$
    - Var($\mathbf{g}_i$) = E($\mathbf{x}_i\mathbf{x}_i'\varepsilon_i^2$) = $\sigma^2\mathbf{X}'\mathbf{X}$
    - CLT:

$$\sqrt{n}\bar{\mathbf{g}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{g}_i = \frac{1}{\sqrt{n}}\mathbf{X}'\boldsymbol{\varepsilon} \rightarrow_d N(\mathbf{0}, \sigma^2\mathbf{X}'\mathbf{X})$$

Least Squares Assumptions

- A4': Non-*Spherical Disturbances:*
  - Heteroscedasticity
    - Var($\boldsymbol{\varepsilon}$|**X**) = E($\boldsymbol{\varepsilon\varepsilon}$'|**X**) = **V**(**X**)
    - Var($\mathbf{g}_i$) = E($\mathbf{x}_i\mathbf{x}_i'\varepsilon_i^2$) = **X′VX**
    - CLT:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

$$\sqrt{n}\bar{\mathbf{g}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{g}_i = \frac{1}{\sqrt{n}}\mathbf{X}'\boldsymbol{\varepsilon} \rightarrow_d N(\mathbf{0}, \mathbf{X}'\mathbf{V}\mathbf{X})$$

Least Squares Estimator

- OLS
  - $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$
    $= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\boldsymbol{\varepsilon}/n)$
  - $\mathbf{b} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\boldsymbol{\varepsilon}/n)$
  - $(\mathbf{b} - \boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta})' = (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\boldsymbol{\varepsilon}/n)(\mathbf{X}'\boldsymbol{\varepsilon}/n)'\,(\mathbf{X}'\mathbf{X}/n)^{-1}$

Least Squares Estimator

- Consistency of **b**
  - $\mathbf{b} - \boldsymbol{\beta} = (\mathbf{X'X}/n)^{-1}(\mathbf{X'\varepsilon}/n)$
  - $\text{plim}(\mathbf{b} - \boldsymbol{\beta}) = \text{plim}[(\mathbf{X'X}/n)^{-1}(\mathbf{X'\varepsilon}/n)]$
    $= \text{plim}[(\mathbf{X'X}/n)^{-1}]\text{plim}[(\mathbf{X'\varepsilon}/n)]$
    $= [\text{plim}(\mathbf{X'X}/n)]^{-1}\text{plim}[(\mathbf{X'\varepsilon}/n)]$
    $= \mathbf{Q}^{-1}\,\mathbf{0}$

Least Squares Estimator

- Consistency of s^2
  - $s^2 = \mathbf{e'e}/(n\text{-}K) = \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}/(n\text{-}K)$
    $= [n/(n\text{-}K)](1/n)\ \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}$
  - $\text{plim}(s^2) = \text{plim}[(1/n)\ \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}]$
    $= \text{plim}[\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/n - \boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}/n)]$
    $= \text{plim}[\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/n]$
    $-\text{plim}[(\boldsymbol{\varepsilon}'\mathbf{X}/n)][\text{plim}(\mathbf{X}'\mathbf{X}/n)]^{-1}\text{plim}[(\mathbf{X}'\boldsymbol{\varepsilon}/n)]$
    $= \text{plim}[\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/n] - \mathbf{0}'\mathbf{Q}^{-1}\mathbf{0}$
    $= E(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) = \sigma^2$

Least Squares Estimator

- Asymptotic Distribution
  - $\mathbf{b} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\boldsymbol{\varepsilon}/n)$
  - $\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) = (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\boldsymbol{\varepsilon}/\sqrt{n}) \rightarrow_d \mathbf{Q}^{-1}N(\mathbf{0}, \boldsymbol{\Omega})$
  - This is because of A4 or A4'

$$\frac{1}{\sqrt{n}}\mathbf{X}'\boldsymbol{\varepsilon} \rightarrow_d N(\mathbf{0}, \boldsymbol{\Omega})$$

  - $\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) \rightarrow_d N(\mathbf{0}, \mathbf{Q}^{-1}\boldsymbol{\Omega}\mathbf{Q}^{-1})$

Least Squares Estimator

- Asymptotic Distribution
  - How to estimate $\mathbf{\Omega}$?
    - If A4 (homoscedasticity), $\mathbf{\Omega} = \sigma^2\mathbf{Q}$
      - $\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) \rightarrow_d N(\mathbf{0}, \sigma^2\mathbf{Q}^{-1})$
      - Approximately, $\mathbf{b} \rightarrow_a N(\boldsymbol{\beta}, (\sigma^2/n)\mathbf{Q}^{-1})$
      - Consistent estimator of $(\sigma^2/n)\mathbf{Q}^{-1}$ is $(s^2/n)\mathbf{Q}^{-1} = s^2(\mathbf{X}'\mathbf{X})^{-1}$

Least Squares Estimator

- Asymptotic Distribution
  - How to estimate $\mathbf{\Omega}$?
    - If A4' (heteroscedasticity),
      - $\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) \to_d N(\mathbf{0}, \mathbf{Q}^{-1}\mathbf{\Omega}\mathbf{Q}^{-1})$, where $\mathbf{\Omega} = \mathbf{X}'\mathbf{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathbf{X} = \mathbf{X}'\mathbf{V}\mathbf{X}$
      - Asy var($\mathbf{b}$) = $(1/n)\mathbf{Q}^{-1}\mathbf{\Omega}\mathbf{Q}^{-1}$
      - White Estimator of $\mathbf{\Omega}$:

$$\hat{\mathbf{\Omega}} = \frac{1}{n}\mathbf{X}'\hat{\mathbf{V}}\mathbf{X} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' e_i^2$$

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

$$\hat{\mathbf{V}} = \begin{bmatrix} e_1^2 & 0 & \cdots & 0 \\ 0 & e_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & e_n^2 \end{bmatrix}$$

$$e_i = y_i - \mathbf{x}_i'\mathbf{b}$$

Least Squares Estimator

- Asymptotic Distribution
  - White Estimator of Asy var(**b**)
    - Est(Asy var(**b**)) = $(1/n)\mathbf{Q}^{-1}[\sum_i \mathbf{x}_i\mathbf{x}_i' e_i^2/n]\mathbf{Q}^{-1}$
      $= (\mathbf{X}'\mathbf{X})^{-1}[\sum_i \mathbf{x}_i\mathbf{x}_i' e_i^2/n]\,(\mathbf{X}'\mathbf{X})^{-1}$
    - White Estimator is Consistent: Heteroscedasticity Consistent Variance-Covariance Matrix

Least Squares Estimator

- Asymptotic Results
  - How "fast" does **b** → **β**?
    - Under A4, Asy Var(**b**) = $(\sigma^2/n)\mathbf{Q}^{-1}$ is O(1/n)
      - √n**b** has variance of O(1)
      - Convergence is at the rate of 1/√n
  - Asymptotic distribution of **b** does not depend on normality assumption of **ε**.
  - Slutsky theorem and the delta method apply to function of **b**.

Hypothesis Testing

$$Under\ H_0 : \beta_k = \bar{\beta}_k$$

$$\sqrt{n}(b_k - \bar{\beta}_k) \rightarrow_d N(0, Asy\,\text{var}(b_k))$$

$$Est(Asy\,\text{var}(b_k)) \rightarrow_p Asy\,\text{var}(b_k)$$

Therefore,

$$t_k = \frac{\sqrt{n}(b_k - \bar{\beta}_k)}{\sqrt{Est(Asy\,\text{var}(b_k))}} = \frac{b_k - \bar{\beta}_k}{SE^*(b_k)} \rightarrow_d N(0,1)$$

$$where\ SE^*(b_k) = \sqrt{\frac{1}{n} Est(Asy\,\text{var}(b_k))}$$

Hypothesis Testing

- Under the null hypothesis H_0: $\mathbf{Rb} = \mathbf{q}$, where $\mathbf{R}$ is J by K matrix of full row rank and J is the number of restrictions (the dimension of $\mathbf{q}$),
  $W = n(\mathbf{Rb-q})'\{\mathbf{R}[\text{Est(Asy var}(\mathbf{b}))]\mathbf{R}'\}^{-1}(\mathbf{Rb-q})$
  $= n(\mathbf{Rb-q})'\{\mathbf{R}[ns^2(\mathbf{X'X})^{-1}]\mathbf{R}'\}^{-1}(\mathbf{Rb-q})$
  $= (\mathbf{Rb-q})'\{\mathbf{R}(\mathbf{X'X})^{-1}]\mathbf{R}'\}^{-1}(\mathbf{Rb-q})/s^2$
  $= JF$
  $= (SSR_r - SSR_{ur})/s^2 \sim \chi^2(J)$