ECON 480-3
LECTURE 3: BASIC INFERENCE & ENDOGENEITY

Ivan A. Canay
Northwestern University

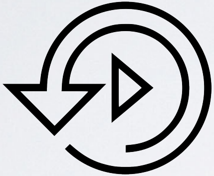
Past & Future

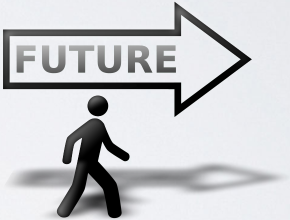
Last Class

- Solving and estimating sub-vectors of β
- Properties of LS
- Estimating $\mathbb{V}$

TODAY

- Basic Principles for Inference
- Linear Regression when $E[XU] \neq 0$





Inference

- Let (Y, X, U) be a random vector where Y and U take values in $\mathbf{R}$ and $X \in \mathbf{R}^{k+1}$. Assume further that the first component of X is a constant equal to one. Let $\beta \in \mathbf{R}^{k+1}$ be such that

$$Y = X'\beta + U \,.$$

- Assume (1) $E[XU] = 0$, (2) $E[XX'] < \infty$, (3) no perfect collinearity in X, and (4) $\text{Var}[XU] < \infty$.

- Under these assumptions, we established the asymptotic normality of the OLS estimator $\hat{\beta}_n$,

$$\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, \mathbb{V})$$

  with

$$\mathbb{V} = E[XX']^{-1} E[XX'U^2] E[XX']^{-1} \,.$$

- We also described a consistent estimator $\hat{\mathbb{V}}_n$ of the limiting variance $\mathbb{V}$.

- We now develop methods for inference under the assumption that (5) $E[XX'U^2]$ is non-singular.

Background

- Consider the following version of a testing problem. One observes an i.i.d. sample $W_i = (Y_i, X_i), i = 1, \ldots, n$, from the dist. $P \in \mathbf{P} = \{P_\beta : \beta \in \mathbf{R}^{k+1}\}$ and wishes to test

$$H_0 : \beta \in \mathbf{B}_0 \quad \text{versus} \quad H_1 : \beta \in \mathbf{B}_1$$

  where $\mathbf{B}_0$ and $\mathbf{B}_1$ form a partition of $\mathbf{R}^{k+1}$.

- In our context, β will be the coefficient in a linear regression.

- **Test function:** a function

$$\phi_n = \phi_n(W_1, \ldots, W_n)$$

  that returns the probability of rejecting the null hypothesis after observing $W_1, \ldots, W_n$.

- **non-randomized tests:** means that the function ϕ_n takes only two values:
  - it takes the value 1 for rejection
  - it takes the value 0 for non-rejection.

Common Case

Most often, ϕ_n is the indicator function of a certain test statistic $T_n = T_n(W_1, \ldots, W_n)$ being greater than some critical value $c_n(1-\alpha)$; i.e.:

$$\phi_n = I\{T_n > c_n(1-\alpha)\} \ .$$

- Examples of tests that take the above form as: Wald tests, quasi-likelihood ratio tests, and Lagrange multiplier tests.
- The critical value could be deterministic (e.g., the quantile of a normally distributed random variable) or could be a random variable itself (e.g., the bootstrap). We will cover both cases in class.
- The test is said to be (pointwise) asymptotically of level α (or consistent in levels) if,

$$\limsup_{n\to\infty} E_{P_\beta}[\phi_n] = \limsup_{n\to\infty} P_\beta\{\phi_n = 1\} \leqslant \alpha \ , \quad \forall \beta \in \mathbf{B}_0 \ .$$

Tests of a Single Linear Restriction

- Let r be a nonzero $(k+1)$-dimensional vector and c be a scalar. Consider testing
$$H_0 : r'\beta = c \text{ versus } H_1 : r'\beta \neq c \ .$$

- **Important case:** r selects the sth component of β,
$$H_0 : \beta_s = c \text{ versus } H_1 : \beta_s \neq c \ .$$

- The CMT implies: $\sqrt{n}(r'\hat{\beta}_n - r'\beta) \xrightarrow{d} N(0, r'\mathbb{V}r) \quad \text{as} \quad n \to \infty \ .$

- Since $\mathbb{V}$ is non-singular, $r'\mathbb{V}r > 0$. The CMT implies that $r'\hat{\mathbb{V}}_n r \xrightarrow{P} r'\mathbb{V}r$ as $n \to \infty$.

- A natural choice of test statistic for this problem is the **absolute value of the t-statistic**,
$$t_{\text{stat}} = \frac{\sqrt{n}(r'\hat{\beta}_n - c)}{\sqrt{r'\hat{\mathbb{V}}_n r}} \ ,$$
so that $T_n = |t_{\text{stat}}|$. When r selects the sth component of β, we get $r'\hat{\mathbb{V}}_n r = \hat{\mathbb{V}}_{n,[s,s]}$, i.e., the sth diagonal element of $\hat{\mathbb{V}}_n$.

Tests of A Single Linear Restriction

Critical value: A suitable choice of critical value for this test statistic is $z_{1-\frac{\alpha}{2}}$.

Comments

- The construction may be modified in a straightforward fashion for testing "one-sided" hypotheses, i.e.,

$$H_0 : r'\beta \leqslant c \text{ versus } H_1 : r'\beta > c\,.$$

- By using the duality between hypothesis testing and the construction of confidence regions, we may construct a confidence region of level α for each component β_s of β as

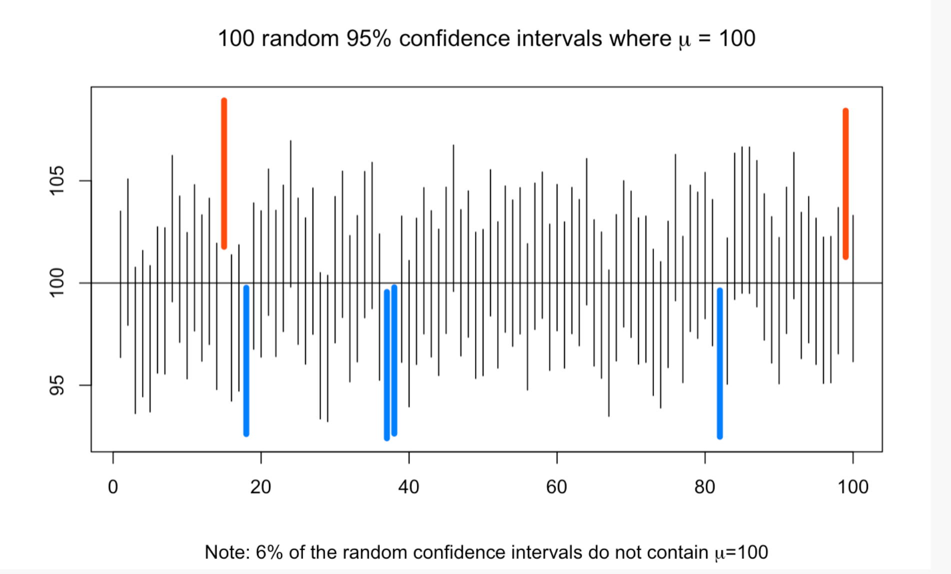
$$\begin{aligned} C_n &= \left\{ c \in \mathbf{R} : \left| \frac{\sqrt{n}(\hat{\beta}_{n,s} - c)}{\sqrt{\hat{\mathbb{V}}_{n,[s,s]}}} \right| \leqslant z_{1-\frac{\alpha}{2}} \right\} \\ &= \left\{ \hat{\beta}_{n,s} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\mathbb{V}}_{n,[s,s]}}{n}}, \hat{\beta}_{n,s} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\mathbb{V}}_{n,[s,s]}}{n}} \right\}\,. \end{aligned}$$

- This confidence region satisfies

$$P\{\beta_s \in C_n\} \to 1 - \alpha$$

as $n \to \infty$. It is straightforward to modify this to construct a confidence region of level α for $r'\beta$.

CI: Graphical Illustration



QUESTIONS?

Tests of Multiple Linear Restrictions

- Let R be a $p \times (k+1)$-dimensional matrix and c be a p-dimensional vector. Consider testing,

$$H_0 : R\beta = c \text{ versus } H_1 : R\beta \neq c \ .$$

- No redundant equations: the rows of R are linearly independent.
- The CMT implies that

$$\sqrt{n}(R\hat{\beta}_n - R\beta) \xrightarrow{d} N(0, R\mathbb{V}R') \quad \text{as} n \to \infty \ .$$

- Because $\mathbb{V}$ is assumed to be non-singular, $R\mathbb{V}R'$ is also non-singular.

Tests of Multiple Linear Restrictions

- From our earlier results, as $n \to \infty$

$$n(R\hat{\beta}_n - R\beta)'(R\hat{\mathbb{V}}_n R')^{-1}(R\hat{\beta}_n - R\beta) \xrightarrow{d} \chi_p^2$$

- **Test statistic:** A natural choice is $T_n = n(R\hat{\beta}_n - c)'(R\hat{\mathbb{V}}_n R')^{-1}(R\hat{\beta}_n - c)$

- **Critical value:** A suitable choice is $c_{p,1-\alpha}$ - the $1-\alpha$ quantile of χ_p^2.

- The test that rejects H_0 when $T_n > c_{p,1-\alpha}$ is consistent in levels.

- By **duality**, we may construct a confidence region of level α for β as

$$C_n = \{c \in \mathbf{R}^{k+1} : n(\hat{\beta}_n - c)'\hat{\mathbb{V}}_n^{-1}(\hat{\beta}_n - c) \leqslant c_{k+1,1-\alpha}\} .$$

- This confidence region satisfies

$$P\{\beta \in C_n\} \to 1-\alpha \quad \text{as} \quad n \to \infty .$$

Tests of Nonlinear Restrictions

- Consider testing
$$H_0 : f(\beta) = 0 \text{ versus } H_1 : f(\beta) \neq 0 \,,$$
where $f : \mathbf{R}^{k+1} \to \mathbf{R}^p$, at level α.

- Assume that f is continuously differentiable at β and denote by $D_\beta f(\beta)$ the $p \times (k+1)$-dimensional matrix of partial derivatives of f evaluated at β.

- Assume that the rows of $D_\beta f(\beta)$ are linearly independent.

- The **Delta Method** implies that
$$\sqrt{n}(f(\hat{\beta}_n) - f(\beta)) \xrightarrow{d} N(0, D_\beta f(\beta) \mathbb{V} D_\beta f(\beta)') \quad \text{as} \quad n \to \infty \,.$$
The **CMT** implies that
$$D_\beta f(\hat{\beta}_n) \hat{\mathbb{V}}_n D_\beta f(\hat{\beta}_n)' \xrightarrow{P} D_\beta f(\beta) \mathbb{V} D_\beta f(\beta)' \quad \text{as} \quad n \to \infty \,.$$

- Straightforward to develop a test and/or a confidence region in this setting following steps as before.

QUESTIONS?

Linear Regression when $E[XU] \neq 0$

- Let (Y, X, U) be a random vector where Y and U take values in $\mathbf{R}$ and $X \in \mathbf{R}^{k+1}$. Assume further that $X = (X_0, X_1, \ldots, X_k)'$ with $X_0 = 1$ and let $\beta = (\beta_0, \beta_1, \ldots, \beta_k)' \in \mathbf{R}^{k+1}$ be such that
$$Y = X'\beta + U\,.$$

- We do not assume $E[XU] = 0$. Any X_j such that $E[X_jU] = 0$ is said to be *exogenous*; any X_j such that $E[X_jU] \neq 0$ is said to be *endogenous*. Normalizing β_0 if necessary, we view X_0 as exogenous.

- Note that it must be the case that we are interpreting this regression as a causal model. **WHY?**

Linear Regression when $E[XU] \neq 0$

Question: What about OLS in this setting?

Omitted Variables

- Suppose $k = 2$, so
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + U \, .$$
- We are interpreting this regression as a causal model and are willing to assume that $E[XU] = 0$ (i.e., $E[U] = E[X_1 U] = E[X_2 U] = 0$), but X_2 is unobserved. An example of a situation like this is when Y is wages, X_1 is education, and X_2 is ability.
- Given unobserved ability, we may rewrite this model as:

Omitted Variables

$$Y = \beta_0^* + \beta_1^* X_1 + U^* \quad \text{with} \quad \begin{cases} \beta_0^* &= \beta_0 + \beta_2 E[X_2] \\ \beta_1^* &= \beta_1 \\ U^* &= \beta_2 (X_2 - E[X_2]) + U \,. \end{cases}$$

Measurement Error

- Partition X into X_0 and X_1, where $X_0 = 1$ and X_1 takes values in $\mathbf{R}^k$. Partition β analogously.
$$Y = \beta_0 + X_1'\beta_1 + U \ .$$
- We are interpreting this regression as a causal model and are willing to assume that $E[XU] = 0$, but X_1 is *not* observed. Instead, $\hat{X}_1$ is observed, where
$$\hat{X}_1 = X_1 + V \ .$$
- Assume (a) $E[V] = 0$, (b) $\text{Cov}[X_1, V] = 0$, and (c) $\text{Cov}[U, V] = 0$.
- We may therefore rewrite this model as:

Measurement Error

$$Y = \beta_0^* + \hat{X}_1' \beta_1^* + U^* \quad \text{with} \quad \begin{cases} \beta_0^* &= \beta_0 \\ \beta_1^* &= \beta_1 \\ U^* &= -V' \beta_1 + U\,. \end{cases}$$

Simultaneity

- **Classical example**: supply and demand. Let Q^s be quantity supplied and Q^d be quantity demanded. As a function of (non-market clearing) price $\tilde{P}$, assume

$$\begin{aligned} Q^d &= \beta_0^d + \beta_1^d \tilde{P} + U^d \\ Q^s &= \beta_0^s + \beta_1^s \tilde{P} + U^s \ , \end{aligned}$$

  where $E[U^s] = E[U^d] = E[U^s U^d] = 0$. We observe (Q, P), where Q and P are such that the market clears, i.e., $Q^s = Q^d$. This implies:

QUESTIONS?