

Lecture 6-2: OLS Inference

Goal: Develop statistical inference methods for multiple regression coefficients.

- Testing hypotheses about individual coefficients
 - t -values in multiple regression
 - Test procedure
 - Confidence intervals
- Testing linear hypotheses: simple case
 - Linear combinations of coefficients
 - Variance of linear combinations
- Testing joint significance
 - F tests for joint hypotheses
 - χ^2 and F distributions
 - Restricted vs unrestricted models
 - Ex1. Omnibus tests
 - Ex2. Equality of coefficients tests
 - Relationship between t and F tests
- Testing linear hypotheses: general case
 - Limitations of simple F tests
 - General restriction matrix form
 - Wald statistic
 - Large and small sample distributions

Worksheet

1. (*A linear combination* $(\beta_{PS} - 2\beta_S)$ *and its standard error*) Return to the Chile model with dummies S (secondary) and PS (post-secondary), with primary schooling as the baseline. The estimated model is

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_{fem} FEM + \hat{\beta}_{age} AGE + \hat{\beta}_S S + \hat{\beta}_{PS} PS.$$

We wish to test whether the 1-2-3 coding of education fits the data:

$$H_0 : \beta_{PS} - 2\beta_S = 0.$$

- (a) Derive $\text{Var}(\hat{\beta}_{PS} - 2\hat{\beta}_S)$
(b) Let $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1} = [a_{rs}]$ and $\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \mathbf{A}$. Explain that

$$\widehat{\text{Var}}(\hat{\beta}_{PS} - 2\hat{\beta}_S) = \hat{\sigma}^2 (a_{PS,PS} + 4a_{S,S} - 4a_{S,PS})$$

is an unbiased estimator of $\text{Var}(\hat{\beta}_{PS} - 2\hat{\beta}_S \mid \mathbf{X})$ under Assumptions I–V.

2. Consider the multiple regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + u,$$

and suppose we wish to test the joint null hypothesis

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \quad \text{versus} \quad H_1 : \text{at least one } \beta_j \neq 0.$$

- (a) Explain in words what this hypothesis is asking about the explanatory power of the regressors.
- (b) Describe step-by-step how to conduct the F test for this hypothesis, including:
 - i. the unrestricted and restricted models,
 - ii. how to compute the F statistic from the two sum of squared residuals, and
 - iii. how to determine whether to reject H_0 at significance level α .
- (c) Interpret the meaning of rejecting versus not rejecting H_0 in terms of the overall explanatory power of $X_1, \dots, X_k$.

3. Consider the multiple regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_k X_k + u,$$

and suppose we want to test whether the first three regressors have the same effect:

$$H_0 : \beta_1 = \beta_2 = \beta_3 \quad \text{versus} \quad H_1 : \text{at least one differs.}$$

- (a) Describe the step-by-step F test procedure, including:
 - i. how to construct the restricted model implied by H_0 ,
 - ii. how to compute the F statistic using the two sums of squared residuals, and
 - iii. the reference distribution used for inference.
- (b) Provide an intuitive interpretation of what rejecting or failing to reject H_0 implies about the relationship between X_1 , X_2 , and X_3 .

4. (*General procedure for testing linear hypotheses*) Consider the linear regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u,$$

with coefficient vector $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2 \ \beta_3]'$. We wish to test linear hypotheses of the form $R\boldsymbol{\beta} = r$. Construct the corresponding R and r in the following two scenarios.

- (a) Test the hypothesis that all three slope coefficients equal 3:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 3.$$

Write this null hypothesis in matrix form $R\boldsymbol{\beta} = r$ by explicitly constructing R and r .

- (b) Test the composite hypothesis

$$H_0 : \beta_1 = 2\beta_2 = 0.5\beta_3 + 1.$$