Econometrics

Chapter 5 Multiple Regression: Inference

In Choi

Sogang University

Distribution of the OLS estimator

- Assume, conditional on $\{x_{1t}, ..., x_{kt}\}_{t=1}^{n}$,

$$u_t \sim iidN(0, \sigma^2).$$

Then,

$$\hat{\beta}_j \sim N(\beta_j, Var(\hat{\beta}_j))$$

and

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{Var(\hat{\beta}_j)}} \sim N(0, 1).$$

The t-test

- Used for hypothesis tests on a single coefficient.
- Consider the null hypothesis

$$H_0 : \beta_j = \beta_j^0.$$

In many applications, we set $\beta_j^0 = 0$.

The t-test

- The t-test is defined as

$$\frac{\hat{\beta}_j - \beta_j^0}{\sqrt{\widehat{Var}(\hat{\beta}_j)}},$$

  where $\widehat{Var}(\hat{\beta}_j) = \frac{s^2}{\sum_{t=1}^{n}(x_{jt} - \bar{x}_j)^2(1 - R_j^2)}$ $(s^2 = \frac{1}{n-k-1}\sum_{t=1}^{n}\hat{u}_t^2)$ denotes the estimator of $Var(\hat{\beta}_j)$.
- The t-test has Student's t-distribution with $n - k - 1$ (# of sample size - # of regressors) degrees of freedom.

The t-test

- The alternatives can be either one-sided or two-sided. One-sided alternative hypotheses are

$$H_1 : \beta_j > \beta_j^0 \tag{1}$$

and

$$H_1 : \beta_j < \beta_j^0. \tag{2}$$

Two-sided hypothesis is

$$H_1 : \beta_j \neq \beta_j^0 \tag{3}$$

- For the alternative (1), we reject the null at the significance level α when the t-test is greater than $t_\alpha(n-k-1)$ where $t_\alpha(n-k-1)$ is the critical value taken from the right tail of the distribution.

The t-test

- For the alternative (3), we reject the null at the significance level α when the absolute value of the t-test is greater than $t_{\alpha/2}(n-k-1)$.
- The p-value for the one-sided alternative (1) is computed by

$$P(t(n-k-1) > t)$$

where t is the realized value of the t-test. The p-value for the two-sided alternative is computed by

$$P(|t(n-k-1)| > t).$$

Confidence interval

- We deduce from the distribution of the -t-test

$$
\begin{aligned}
& P\left(\hat{\beta}_j - t_{\alpha/2}(n-k-1)SE(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + t_{\alpha/2}(n-k-1)SE(\hat{
\right. \\
= \; & 1 - \alpha
\end{aligned}
$$

The $100(1-\alpha)\%$ confidence interval for β_j is

$$\left[\hat{\beta}_j - t_{\alpha/2}(n-k-1)SE(\hat{\beta}_j), \hat{\beta}_j + t_{\alpha/2}(n-k-1)SE(\hat{\beta}_j)\right].$$

F-test

- Consider the null hypothesis

$$H_0 : \beta_1 = ... = \beta_k = 0.$$

  The $F-$test for this null is defined as

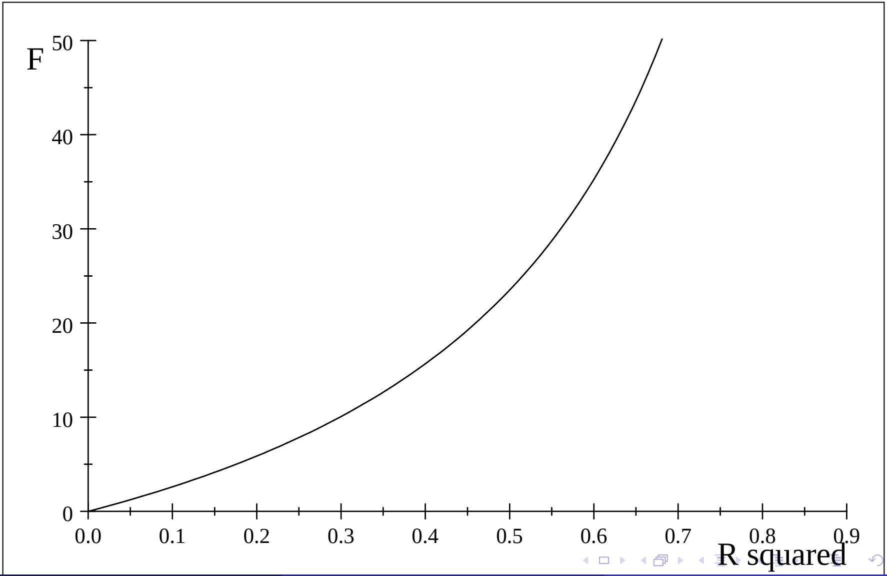
$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)}.$$

  The null distribution of F is $F(k, n-k-1)$.

- $F(a,b) = \frac{\chi^2(a)/a}{\chi^2(b)/b}$, where the numerator and denominator are independent.
- kF follows $\chi^2(k)$ approximately when n is large.

F-test

Plot F for $k = 2$ and $n = 50$.



F-test

- Consider the null hypothesis

$$H_0 : \beta_{k-q+1} = ... = \beta_k = 0.$$

Estimate the restricted linear regression model

$$y_t = \beta_0 + \beta_1 x_{1t} + ... + \beta_{k-q} x_{k-q,t} + u_t$$

and let the sum of squared residuals RSS_r.

- The sum of squared residuals from regressing y on $x_1, ..., x_k$ is denoted as RSS_{ur} (unrestricted sum of squared residuals).

F-test

- The $F-$test for this null is defined as

$$F = \frac{(RSS_r - RSS_{ur})\,/\,q}{RSS_{ur}\,/\,(n-k-1)}.$$

  The null distribution of F is $F(q, n-k-1)$.
- Note that $RSS_r \geq RSS_{ur}$.

F-test

- Consider the null hypothesis

$$H_0 : R\beta = r$$

where the $J \times (k+1)$ matrix R has full row rank J. The F-test for this null is defined as

$$F = \left(R\hat{\beta} - r\right)' \left[s^2 R \left(X'X\right)^{-1} R'\right]^{-1} \left(R\hat{\beta} - r\right) / J$$

The null distribution of F is $F\left(J, n-k-1\right)$.

F-test

- If
$$H_0 : \beta_1 = ... = \beta_k = 0,$$
the null hypothesis can be written as

$$R = \begin{bmatrix} 0 & 1 & \cdots & & 0 \\ & 0 & 1 & & \\ & & & \ddots & \\ 0 & \cdots & & 0 & 1 \end{bmatrix}_{k \times (k+1)}, \quad r = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{k \times 1}, \quad R\beta = r.$$