Lecture 14 - Economics 113

Professor Spearot

- Agenda
  1. P-values
  2. Comparing different parameters
  3. Prediction Intervals
- Homework #1 is posted
- Exam #3 on Friday

P-values

A different approach

- How likely is it that I falsely reject the null?
- P-value:

The probability that the null hypothesis is falsely rejected

- In the crime and enrollment example

$$H_0 : \beta_{Enroll} = 1,\ H_A : \beta_{Enroll} \neq 1$$

- T-statistic

$$t_{stat} = \frac{1.27 - 1}{0.11} = 2.45$$

- The p-value is the value of the following

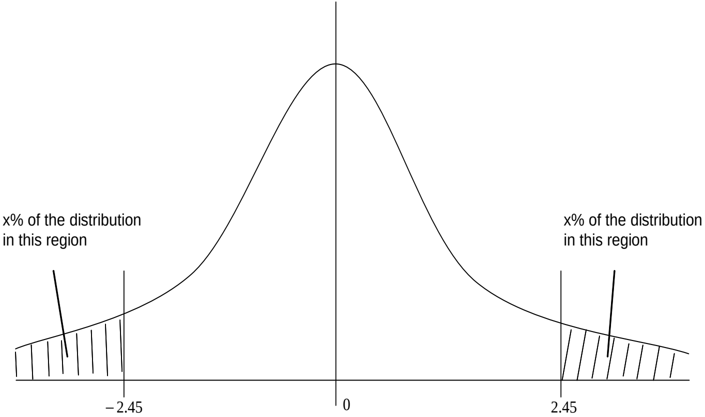
$$pvalue = \Pr(|T| > 2.45)$$

- The p-value is the probability that I randomly draw a value from the t-distribution that is larger (in absolute terms) than the estimated T-statistic

P-values

An illustration

- P-value: The smallest significance level such that the null would be rejected.



- The p-value for this is the probability of being in the shaded region.

P-values

A different approach

- To find the p-value
  1. Calculate the t-statistic
  2. Find the closest match to your t-statistic on the normal table
  3. The p-value is the significance level that generates this t-statistic.
- In the crime and enrollment example
  1. $t_{stat} = 2.45$
  2. $\Pr(|T| > 2.45)$?

$$
\begin{aligned}
\Pr(|T| &> 2.45) = \Pr(T > 2.45 \bigcup T < -2.45) \\
&= \Pr(T > 2.45 \bigcup T < -2.45) \\
&= \Pr(T > 2.45) + \Pr(T < -2.45) = 2 * \Pr(T > 2.45) \\
&= 2 * (1 - 0.9929) = 0.0142
\end{aligned}
$$

- Thus, we will falsely reject the null less than 1.42% of the time.

Linear combinations of parameters

Returns to schooling

- Suppose that we estimated

$$\begin{aligned} \widehat{\log(Wage)} &= \underset{(0.021)}{1.472} + \underset{(0.0068)}{0.0667}\, JC + \underset{(0.0023)}{0.0769}\, Univ + \underset{(0.0002)}{0.0049}\, exper \\ obs &= 6763,\ R^2 = 0.222 \end{aligned}$$

  1. *JC* is years at a junior college
  2. *Univ* is years at a university
  3. *exper* is experience.

- Clearly, $\widehat{\beta}_1 < \widehat{\beta}_2$.
- What does this mean?
- Is this significant?
- How do we test $H_0 : \beta_1 = \beta_2$
- $\beta_1 = \beta_2$ is the same as $\beta_1 - \beta_2 = 0$

Linear combinations of parameters

Returns to schooling

- Define $\theta = \beta_1 - \beta_2$.
- $H_0 : \theta = 0$.
- $H_A : \theta \neq 0$.
- How do we integrate θ into our regression?
  - Solve for $\beta_1 = \theta + \beta_2$
  - Substitute for β_1

$$
\begin{aligned}
\log(Wage) &= \beta_0 + \beta_1 JC + \beta_2 Univ + \beta_3 exper + u \\
\log(Wage) &= \beta_0 + (\theta + \beta_2)\, JC + \beta_2 Univ + \beta_3 exper + u
\end{aligned}
$$

  - Collect parameters

$$
\log(Wage) = \beta_0 + \theta JC + \beta_2 \left(Univ + JC\right) + \beta_3 exper + u
$$

- $Univ + JC$ is total schooling
- θ the effect of a JC year, controlling for total schooling.

Linear combinations of parameters

Returns to schooling

- Modified Model:

$$\begin{aligned}\widehat{\log(Wage)} &= \underset{(0.021)}{1.472} - \underset{(0.0069)}{0.0102}\,JC + \underset{(0.0023)}{0.0769}\,(Univ + JC) + \underset{(0.0002)}{0.0049}\,exper \\ obs &= 6763,\ R^2 = 0.222\end{aligned}$$

- Can we reject $H_0 : \theta = 0$ in favor of $H_A : \theta \neq 0$?
- 95% Confidence interval?
- Use the formula

$$\begin{aligned}-0.0102 - 1.96 * 0.0069 &< \theta < -0.0102 + 1.96 * 0.0069 \\ -0.023724 &< \theta < 0.003324\end{aligned}$$

- Interpret?

Predictions

Introduction

- Suppose you start with the equation:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$$

- You may wish to predict y for various individuals which are not in the sample.
- Solving for predictions is easy

$\longrightarrow$ Plug $x_1 = c_1, ..., x_k = c_k$ into equation
$\longrightarrow$ Produce prediction, $\widehat{\theta}$

- Also might want Standard Errors. Why?
- The prediction may be precise or imprecise - need standard errors to figure this out.

Predictions

Getting standard errors

- Prediction:

$$\theta = \beta_0 + \beta_1 c_1 + \cdots + \beta_k c_k$$

- Solve for β_0:

$$\beta_0 = \theta - \beta_1 c_1 - \cdots - \beta_k c_k$$

- Plug into estimating equation

$$y = (\theta - \beta_1 c_1 - \cdots - \beta_k c_k) + \beta_1 x_1 + \cdots + \beta_k x_k + u$$

- Simplify:

$$y = \theta + \beta_1 (x_1 - c_1) + \cdots + \beta_k (x_k - c_k) + u$$

- Estimate $\rightarrow$ gives us prediction θ and standard error.

Predictions

Earnings Example

- Predicted earnings of a person with
  - 10 years of education
  - 2 years of experience
  - 1 year on tenure
- Want prediction and standard error.
- Estimate:

$$wage = \theta + \beta_1 \,(educ - 10) + \beta_2(\textit{exper-2}) + \beta_3(tenure - 1) + u$$

- Confidence interval for θ?