Lecture 8: Hypothesis testing (Part 1)

Economics 326 — Methods of Empirical Research in Economics

Vadim Marmer, UBC

Hypothesis testing

$

$

- Hypothesis testing is one of a fundamental problems in statistics.
- A hypothesis is (usually) an **assertion** about the unknown population parameters such as β_1 in $Y_i = \beta_0 + \beta_1 X_i + U_i$.
- Using the data, the econometrician has to determine whether an assertion is **true** or **false**.
- Example: **Phillips curve**:

$$\text{Unemployment}_t = \beta_0 + \beta_1 \text{Inflation}_t + U_t.$$

In this example, we are interested in testing if $\beta_1 = 0$ (no Phillips curve) against $\beta_1 < 0$ (Phillips curve).

Assumptions

- β_1 is unknown, and we have to rely on its OLS estimator $\hat{\beta}_1$.
- We need to know the distribution of $\hat{\beta}_1$ or of its certain functions.
- We will assume that the assumptions of the **Normal Classical Linear Regression** model are satisfied:
  1. $Y_i = \beta_0 + \beta_1 X_i + U_i, \qquad i = 1, \ldots, n.$
  2. $\mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] = 0$ for all i's.
  3. $\mathrm{E}\left[U_i^2 \mid X_1, \ldots, X_n\right] = \sigma^2$ for all i's.
  4. $\mathrm{E}\left[U_i U_j \mid X_1, \ldots, X_n\right] = 0$ for all $i \neq j$.
  5. U's are jointly normally distributed conditional on X's.
- Recall that, in this case, conditionally on X's:

$$\hat{\beta}_1 \sim N\left(\beta_1, \mathrm{Var}\left(\hat{\beta}_1\right)\right), \text{ where } \mathrm{Var}\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}.$$

Null and alternative hypotheses

- Usually, we have **two competing hypotheses**, and we want to draw a conclusion, based on the data, as to which of the hypotheses is true.
- **Null hypothesis**, denoted as H_0: A hypothesis that is held to be true, unless the data provides a **sufficient** evidence against it.
- **Alternative hypothesis**, denoted as H_1: A hypothesis against which the null is tested. It is held to be true if the null is found false.

- Usually, the econometrician has to carry the **"burden of proof,"** and the case that he is interested in is stated as H_1.
- The econometrician has to prove that his assertion (H_1) is true by showing that the data rejects H_0.
- The two hypotheses must be **disjoint**: it should be the case that either H_0 is true or H_1 but never together simultaneously.

Decision rule

- The econometrician has to choose between H_0 and H_1.
- The **decision rule** that leads the econometrician to **reject or not to reject** H_0 is based on a **test statistic**, which is a **function of the data** $\{(Y_i, X_i) : i = 1, ... n\}$.
- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.

Errors

- There are two types of errors that the econometrician can make:

	Truth: H_0	Truth: H_1
Decision: H_0	✓	Type II error
Decision: H_1	Type I error	✓

- **Type I error** is the error of rejecting H_0 when H_0 is true.
- The probability of Type I error is denoted by α and called **significance level** or **size** of a test:

$$P(\text{Type I error}) = P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha.$$

- **Type II error** is the error of not rejecting H_0 when H_1 is true.
- **Power** of a test:

$$1 - P(\text{Type II error}) = 1 - P(\text{Accept } H_0 | H_0 \text{ is false}).$$

Errors

- The decision rule depends on a **test statistic** T.
- The real line is split into two regions: **acceptance region** and **rejection region** (critical region).
- When T is in the acceptance region, we accept H_0 (and risk making a Type II error).
- When T is in the rejection (critical) region, we reject H_0 (and risk making a Type I error).
- Unfortunately, the probabilities of Type I and II errors are inversely related. By decreasing the probability of Type I error α, one makes the critical region smaller, which increases the probability of the Type II error. Thus it is impossible to make both errors arbitrary small.
- By convention, α is chosen to be a small number, for example, $\alpha = 0.01, 0.05$, or 0.10. (This is in agreement with the econometrician carrying the burden of proof).

Steps

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.

3. Define a decision rule (critical region).
4. Perform the test using the data: given the data compute the test statistic and see if it falls into the critical region.

- The decision depends on the significance level α: larger values of α correspond to bigger critical regions (probability of Type I error is larger).
- It is **easier** to reject the null for larger values of α.
- p**-value**: Given the data, the **smallest** significance level at which the null can be rejected.

Two-sided tests

- For $Y_i = \beta_0 + \beta_1 X_i + U_i$, consider testing

$$H_0 : \beta_1 = \beta_{1,0},$$

against

$$H_1 : \beta_1 \neq \beta_{1,0}.$$

- β_1 is the true **unknown** value of the slope parameter.
- $\beta_{1,0}$ is a **known** number specified by the econometrician. (For example $\beta_{1,0}$ is zero if you want to test $\beta_1 = 0$).
- Such a test is called **two-sided** because the alternative hypothesis H_1 does not specify in which direction β_1 can deviate from the asserted value $\beta_{1,0}$.

Two-sided tests when σ^2 is known (infeasible test)

- Suppose for a moment that σ^2 is known.
- Consider the following **test statistic**:

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\mathrm{Var}\left(\hat{\beta}_1\right)}}, \text{ where } \mathrm{Var}\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}.$$

- Consider the following **decision rule** (test):

$$\text{Reject } H_0 : \beta_1 = \beta_{1,0} \text{ when } |T| > z_{1-\alpha/2},$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution (**critical value**).

Test validity and power

- We need to establish that:
  1. The test is **valid**, where the validity of a test means that it has **correct size** or $P\,(\text{Type I error}) = \alpha$:

  $$P\left(|T| > z_{1-\alpha/2} | \beta_1 = \beta_{1,0}\right) = \alpha.$$

  2. The test has **power**: when $\beta_1 \neq \beta_{1,0}$ (H_0 is false), the test rejects H_0 with probability that exceeds α:

  $$P\left(|T| > z_{1-\alpha/2} | \beta_1 \neq \beta_{1,0}\right) > \alpha.$$

- We want $P\left(|T| > z_{1-\alpha/2} | \beta_1 \neq \beta_{1,0}\right)$ to be as large as possible.
- Note that $P\left(|T| > z_{1-\alpha/2} | \beta_1 \neq \beta_{1,0}\right)$ depends on the true value β_1.

The distribution of T when σ^2 is known (infeasible test)

- Write

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}} = \frac{\hat{\beta}_1 - \beta_1 + \beta_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}}$$
$$= \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}} + \frac{\beta_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}}.$$

- Under our assumptions and conditionally on X's:

$$\hat{\beta}_1 \sim N\left(\beta_1, \text{Var}\left(\hat{\beta}_1\right)\right), \text{ or } \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}} \sim N(0,1).$$

- We have that conditionally on X's: $T \sim N\left(\frac{\beta_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}}, 1\right)$.

Size when σ^2 is known (infeasible test)

- We have that

$$T \sim N\left(\frac{\beta_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}}, 1\right).$$

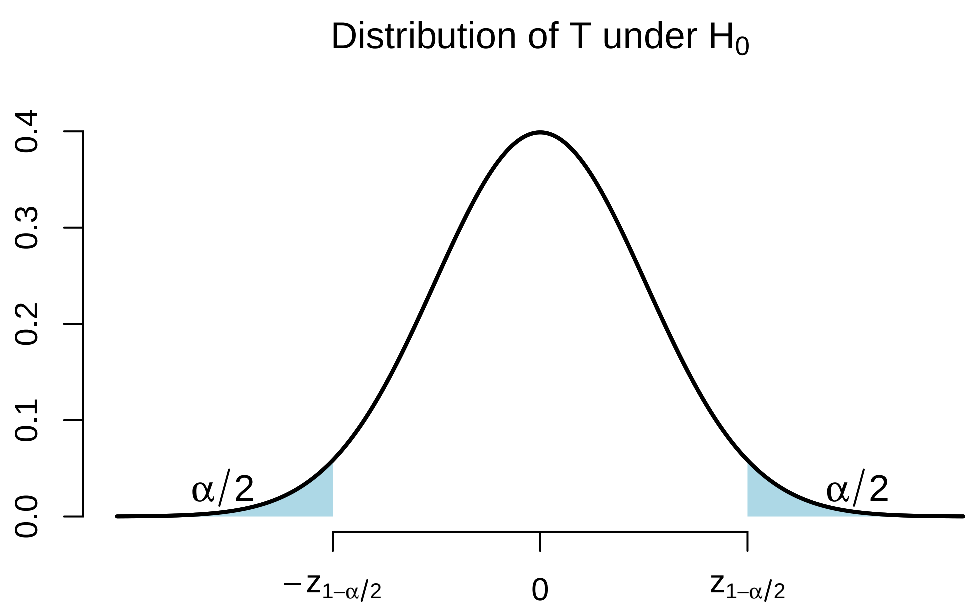
- When $H_0 : \beta_1 = \beta_{1,0}$ is true, $T \overset{H_0}{\sim} N(0,1)$.
- We reject H_0 when

$$|T| > z_{1-\alpha/2} \Leftrightarrow T > z_{1-\alpha/2} \text{ or } T < -z_{1-\alpha/2}.$$

- Let $Z \sim N(0,1)$.

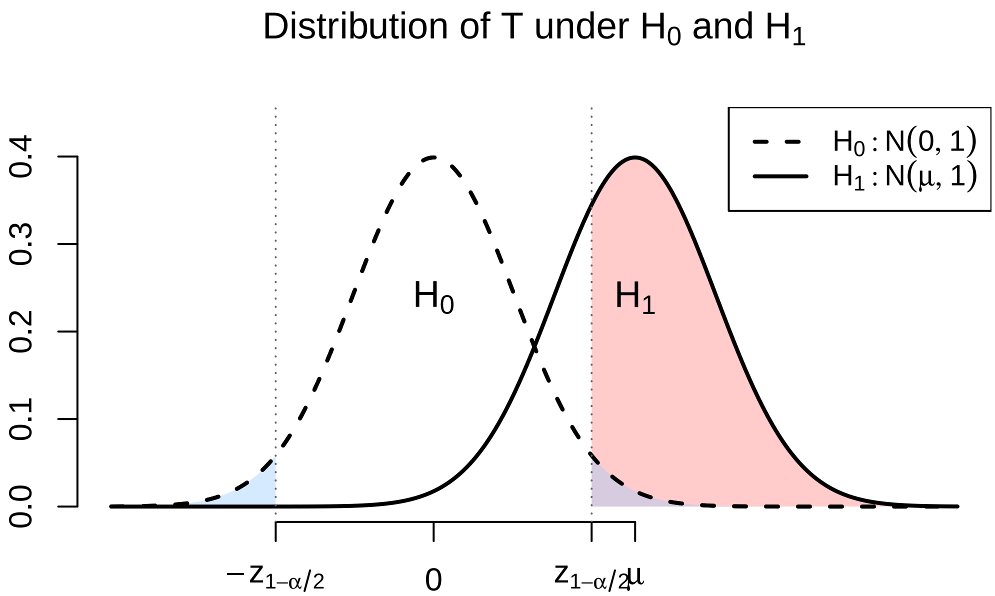
$$P(\text{Reject } H_0 | H_0 \text{ is true}) = P\left(Z > z_{1-\alpha/2}\right) + P\left(Z < -z_{1-\alpha/2}\right)$$
$$= \alpha/2 + \alpha/2 = \alpha$$

The distribution of T when σ^2 is known (infeasible test)



Power of the test when σ^2 is known (infeasible test)

- Under H_1, $\beta_1 - \beta_{1,0} \neq 0$ and the distribution of T is **not** centered zero: $T \sim N\left(\frac{\beta_1 - \beta_{1,0}}{\sqrt{\text{Var}\left(\hat{\beta}_1\right)}}, 1\right)$.
- When $\beta_1 - \beta_{1,0} > 0$:



- Rejection probability exceeds α under H_1: power increases with the distance from H_0 ($|\beta_{1,0} - \beta_1|$) and decreases with Var $\left(\hat{\beta}_1\right)$.