Lecture 8: Hypothesis testing

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

April 5, 2026

Hypothesis testing

- Hypothesis testing is one of the 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).

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 **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 both 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** $\{(Y_i, X_i) : i = 1, \dots, n\}$.
- 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:

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

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

$$1 - P\left(\text{Type II error}\right) = 1 - P\left(\text{Do not reject } H_0 | H_0 \text{ is false}\right).$$

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 do not reject 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 arbitrarily 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.

Assumptions

- Recall: under the **Normal Classical Linear Regression** model and conditionally on $\mathbf{X}$:

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

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 test (σ^2 known)

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

  $$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)}},$$

  where $\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right) = \dfrac{\sigma^2}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}$.

- 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.

Distribution of T (σ^2 known)

- Write

  $$\begin{aligned} T &= \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)}} = \frac{\hat{\beta}_1 - \beta_1 + \beta_1 - \beta_{1,0}}{\sqrt{\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)}} \\ &= \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)}} + \frac{\beta_1 - \beta_{1,0}}{\sqrt{\operatorname{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)}}. \end{aligned}$$

- Under our assumptions and conditionally on $\mathbf{X}$:

$$\hat{\beta}_1 \sim N\left(\beta_1, \text{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)\right),$$

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

- We have that conditionally on $\mathbf{X}$:

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

Size of the test (σ^2 known)

- Conditionally on $\mathbf{X}$, we have that

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

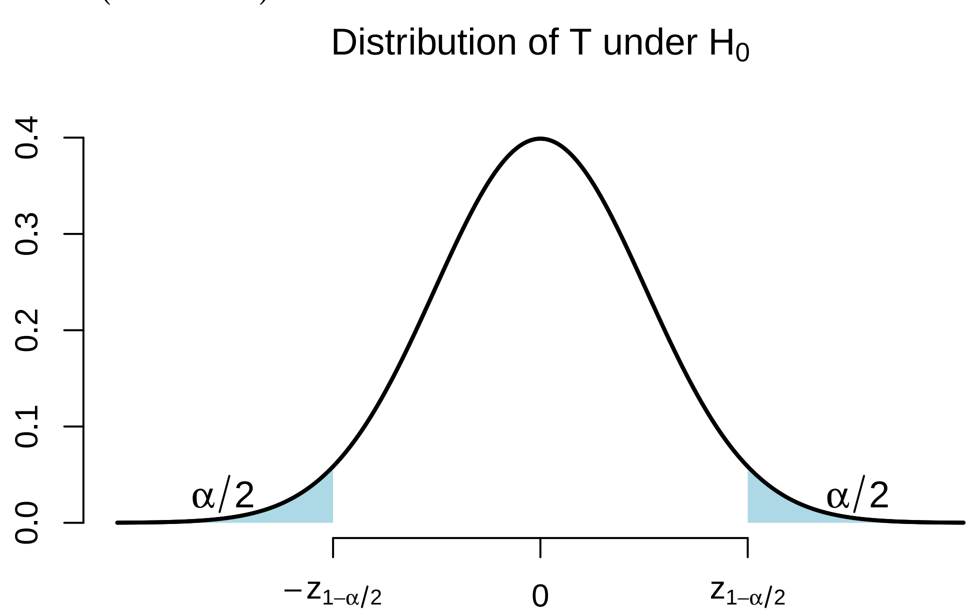
- When $H_0 : \beta_1 = \beta_{1,0}$ is true, $T \overset{H_0}{\sim} N(0,1)$ conditionally on $\mathbf{X}$.
- We reject H_0 when

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

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

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

Distribution of T (σ^2 known)

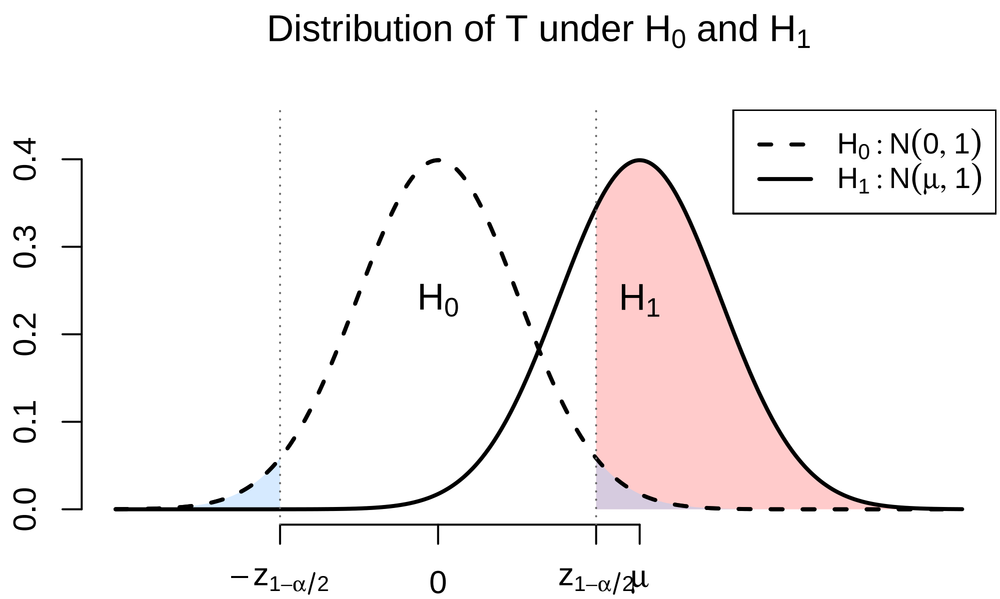


Power (σ^2 known)

- Under H_1, $\beta_1 - \beta_{1,0} \neq 0$ and, conditionally on $\mathbf{X}$, the distribution of T is **not** centered at zero:

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

- 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 $\mathrm{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right)$.

The two-sided t-test

- We are testing $H_0 : \beta_1 = \beta_{1,0}$ against $H_1 : \beta_1 \neq \beta_{1,0}$.
- When σ^2 is **unknown**, we replace it with $s^2 = \frac{1}{n-2}\sum_{i=1}^n \hat{U}_i^2$.
- Recall the **standard error** of $\hat{\beta}_1$:

$$\mathrm{se}\left(\hat{\beta}_1\right) = \sqrt{\widehat{\mathrm{Var}}\left(\hat{\beta}_1\right)} = \sqrt{\frac{s^2}{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}}.$$

- The **t-statistic**:

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\mathrm{se}\left(\hat{\beta}_1\right)}.$$

- We also replace the standard normal critical values $z_{1-\alpha/2}$ with the t_{n-2} critical values $t_{n-2,1-\alpha/2}$. However, **for large** n, $t_{n-2,1-\alpha/2} \approx z_{1-\alpha/2}$.
- The two-sided **t-test**:

Reject H_0 when $|T| > t_{n-2,1-\alpha/2}$.

The two-sided p-value

- The decision to reject or not reject H_0 depends on the critical value $t_{n-2,1-\alpha/2}$.
- If $\alpha_1 > \alpha_2$ then $t_{n-2,1-\alpha_1/2} < t_{n-2,1-\alpha_2/2}$.
- Thus, it is easier to reject H_0 with the significance level α_1 since it corresponds to a smaller acceptance region.
- p-value is the **smallest** significance level α for which we can reject H_0.

The two-sided p-value

- In order to find the p-value:
  1. Compute T.
  2. The p-value $= 2 \cdot P(t_{n-2} \leq -|T|)$.
  3. In R: `2 * pt(-abs(T), df = n - 2)`.
- Note that for all $\alpha > p$-value,

  $$|T| = t_{n-2,1-(p\text{-value})/2} > t_{n-2,1-\alpha/2}$$

  and we will reject H_0.
- For all $\alpha \leq p$-value,

  $$|T| = t_{n-2,1-(p\text{-value})/2} \leq t_{n-2,1-\alpha/2}$$

  and we will not reject H_0.

Example of p-value calculation

- Suppose a regression with 19 observations produced the following output:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.18197    0.25094   40.58   <2e-16 ***
x           -0.67253    0.58049   -1.16    0.263
```

- Here, $\hat{\beta}_1 = -0.6725$, $\beta_{1,0} = 0$, and the `t value` column gives $t = -0.6725/0.5804 = -1.16$.
- Thus, $|T| = 1.16$ and df=17.
- The two-sided p-value:

```r
2 * pt(-abs(-1.16), df = 17)
```

```
[1] 0.2620816
```

- The p-value is large, so we cannot reject H_0 at conventional significance levels.

Computing in R

- We compute critical values and p-values using R.
- To compute standard normal quantiles use `qnorm( )`, where τ is a number between 0 and 1:

```r
# z critical value for a two-sided 5% test
qnorm(1 - 0.05 / 2)
```

```
[1] 1.959964
```

- For t critical values use `qt( , df)`, where df is the number of degrees of freedom and τ is the left-tail probability:

```r
# t critical value for a two-sided 5% test with 62 df
qt(1 - 0.05 / 2, df = 62)
```

```
[1] 1.998972
```

Computing in R

- To compute two-sided normal p-values use `2 * pnorm(-abs(T))`:

```
# Two-sided normal p-value for T = 1.96
2 * pnorm(-abs(1.96))
```

```
[1] 0.04999579
```

- To compute two-sided t-distribution p-values, use `2 * pt(-abs(T), df)`:

```
# Two-sided t p-value for T = 1.96 with 62 df
2 * pt(-abs(1.96), df = 62)
```

```
[1] 0.05449415
```

Example

- Data: `rental` from the `wooldridge` R package. 64 US cities in 1990.
  - `rent`: average monthly rent ($)
  - `avginc`: per capita income ($)
- Model: $\text{Rent}_i = \beta_0 + \beta_1 \text{AvgInc}_i + U_i$.
- Regression output:

```
library(wooldridge)
data("rental")
rental90 <- subset(rental, y90 == 1)
reg <- lm(rent ~ avginc, data = rental90)
summary(reg)$coefficients
```

```
                Estimate   Std. Error  t value     Pr(>|t|)
(Intercept) 148.77643972 32.097874358 4.635087 1.885260e-05
avginc        0.01158001  0.001308365 8.850748 1.340614e-12
```

- R reports the t-statistics and the p-value for $H_0 : \beta_1 = 0$.
- To test H_0 whether the coefficient of AvgInc is zero: $T = 0.01158/0.0013084 = 8.85$.
- The p-value is extremely close to zero:

```
2 * pt(-abs(8.85), df = 62)
```

```
[1] 1.344588e-12
```

  So for all reasonable significance levels α, we reject H_0 that the coefficient of AvgInc is zero.

- AvgInc is a **statistically significant** regressor.

Example (continued)

- Consider now testing H_0 that the coefficient of AvgInc is 0.009 against the alternative that it is different from 0.009.
- $T = (0.01158 - 0.009)/0.0013084 \approx 1.97$.

```
T_stat <- (0.01158 - 0.009) / 0.0013084
T_stat
```

```
[1] 1.971874
```

- At 5% significance level, $t_{62,0.975} \approx 1.999 > T$ and we do not reject H_0.

```
qt(0.975, df = 62)
```

```
[1] 1.998972
```

- At 10% significance level, $t_{62,0.95} \approx 1.67 < T$ and we reject H_0.

```r
qt(0.95, df = 62)
```

```
[1] 1.669804
```

- The two-sided p-value:

```r
2 * pt(-abs(T_stat), df = 62)
```

```
[1] 0.05308963
```

The p-value is ≈ 0.053.

- For $\alpha \leq 0.053$ we will not reject H_0 and for $\alpha > 0.053$ we will reject H_0.

Confidence intervals and hypothesis testing

- There is a one-to-one correspondence between confidence intervals and hypothesis testing.
- We cannot reject $H_0 : \beta_1 = \beta_{1,0}$ against a two-sided alternative if $|T| \leq t_{n-2,1-\alpha/2}$, i.e., if and only if:

$$
\begin{aligned}
& -t_{n-2,1-\alpha/2} \leq \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} \leq t_{n-2,1-\alpha/2} \\
& \iff \\
& \hat{\beta}_1 - t_{n-2,1-\alpha/2} \times \text{se}\left(\hat{\beta}_1\right) \\
& \qquad \leq \beta_{1,0} \leq \hat{\beta}_1 + t_{n-2,1-\alpha/2} \times \text{se}\left(\hat{\beta}_1\right) \\
& \iff \\
& \beta_{1,0} \in CI_{1-\alpha}.
\end{aligned}
$$

- Thus, for any $\beta_{1,0} \in CI_{1-\alpha}$, we **cannot reject** $H_0 : \beta_1 = \beta_{1,0}$ against $H_1 : \beta_1 \neq \beta_{1,0}$ at significance level α.

Example

- The 95% confidence interval for the coefficient of AvgInc:

```r
confint(reg, "avginc", level = 0.95)
```

```
             2.5 %     97.5 %
avginc 0.008964625 0.01419539
```

- A significance level 5% test of $H_0 : \beta_1 = \beta_{1,0}$ against $H_1 : \beta_1 \neq \beta_{1,0}$ will not reject H_0 if $\beta_{1,0}$ is in the 95% confidence interval.

One-sided tests

- Consider testing $H_0 : \beta_1 \leq \beta_{1,0}$ against $H_1 : \beta_1 > \beta_{1,0}$.
- It is reasonable to reject H_0 when $\hat{\beta}_1 - \beta_{1,0}$ is large and positive or when

$$
T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} > c_{1-\alpha}
$$

where $c_{1-\alpha}$ is a positive constant.

- The null hypothesis H_0 is **composite**. The probability of rejection under H_0 depends on β_1.
- We pick the critical value $c_{1-\alpha}$ so that

$$
P\left(\frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} > c_{1-\alpha} | \beta_1 \leq \beta_{1,0}\right) \leq \alpha
$$

for all $\beta_1 \leq \beta_{1,0}$.

One-sided tests

- For **all** $\beta_1 \leq \beta_{1,0}$,

$$\frac{\beta_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} \leq 0,$$

and

$$\begin{aligned}
&P\left(\frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} > c_{1-\alpha} | \beta_1 \leq \beta_{1,0}\right) \\
&= P\left(\frac{\hat{\beta}_1 - \beta_1}{\text{se}\left(\hat{\beta}_1\right)} + \frac{\beta_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} > c_{1-\alpha} | \beta_1 \leq \beta_{1,0}\right) \\
&\leq P\left(\frac{\hat{\beta}_1 - \beta_1}{\text{se}\left(\hat{\beta}_1\right)} > c_{1-\alpha} | \beta_1 \leq \beta_{1,0}\right) \\
&= \alpha \text{ if } c_{1-\alpha} = t_{n-2,1-\alpha}.
\end{aligned}$$

One-sided tests

- For size α test, we reject $H_0 : \beta_1 \leq \beta_{1,0}$ against $H_1 : \beta_1 > \beta_{1,0}$ when

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} > t_{n-2,1-\alpha},$$

where $t_{n-2,1-\alpha}$ is the critical value corresponding to the t-distribution with $n-2$ degrees of freedom.

  - Note that we use $1-\alpha$ and not $1-\alpha/2$ for choosing critical values in the case of one-sided testing.

- For size α test, we reject $H_0 : \beta_1 \geq \beta_{1,0}$ against $H_1 : \beta_1 < \beta_{1,0}$ when

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}\left(\hat{\beta}_1\right)} < -t_{n-2,1-\alpha}.$$

One-sided tests

- One-sided p-values for $H_0 : \beta_1 \leq \beta_{1,0}$ against $H_1 : \beta_1 > \beta_{1,0}$:
  1. Compute T.
  2. The p-value $= P(t_{n-2} \geq T) = 1 - P(t_{n-2} \leq T)$.
  3. In R: `1 - pt(T, df = n - 2)`.