Lecture 5

Testing Economic Hypotheses

Ananya Iyengar

January 2026

St. Stephen's College

This corresponds to Chapter 5 of Studenmund (2017).

An econometric hypothesis is a *testable* implication that emerges from the combination of a probabilistic mathematical model and economic variables. We have learnt until now how to estimate β (i.e. quantify the relationship between two or more economic variables) and have established some desirable properties of the estimated $\hat{\beta}$ under certain conditions.

What we will not think about is the *size* of these estimates. For example, consider regressing $Wages$ on D, which represents participation in a training programme. The estimated coefficient on D i.e.$\hat{\beta}_1$ gives the average marginal impact of participation on wages. Suppose this estimated coefficient takes value 1.3. Is this a big effect? Is it small? Suppose this coefficient is 0.05. Is this a sizeable effect?

Here, we differentiate between the concepts of *numerical* and *statistical* differences. The numbers 0 and 0.05 are evidently numerically different, are they statistically different? If the true population parameter is 0 (i.e. there is no relationship between the variables in the population), *what is the probability of drawing a sample that produces an effect of 0.05*?

1 Some terminology

1. **Expected effect:** The range of values of the true population coefficient that we expect to hold true. In the above example, we may expect that the effect of the training programme on wages in ≥ 0. This number need not always be 0.

2. **Null hypothesis:** The range of values for which we observe a *null* or no expected effect. The null hypothesis is the baseline. In the above example, it may take form of the statement: *The training programme has no effect i.e.* $\beta_1 = 0$. Formally, this null hypothesis is denoted by $H_0 : \beta_1 = 0$. Two important points:
   - The Null Hypothesis need not always expect a coefficient of 0. It is possible that our baseline is an effect of 1 and we want to examine if the true β is different from 1.
   - The Null Hypothesis need not be an equality hypothesis. We can write $H_0 : \beta_1 < 0$!
3. **Alternate hypothesis:** The range of values that we expect the coefficient to take or the effect that we expect to see is encapsulated by the alternate hypothesis. This is denoted by H_a and can take the form $\beta_1 \neq 0$. Depending on the formulation of H_a, using the above example:
   - If I expect that the true population coefficient is positive, $H_a : \beta_1 > 0$. If I expect that the true population coefficient is non-positive, $H_a : \beta_1 \leq 0$. These are called *one-sided tests.*
   - If I expect that the true population coefficient is non-zero, but I think it can take both positive and negative values, $H_a : \beta_1 \neq 0$. This is called a *two-sided test.*
4. Note that hypotheses are *always* formulated in terms of the population coefficients (which we do not observe). We want to test if our expectations of the true population parameters are fulfilled in the data, and if so, by how much.
5. Formally, we want to either *reject the null hypothesis* (i.e. in the above example, we would infer from the data that the true effect is indeed non-zero) or *not reject the null* (i.e. in the above example, we would infer from the data that the true effect is zero). We never use the phrase “accept the null".
6. **Distribution of $\hat{\beta}$ under** H_0: The sampling distribution of the BLUE $\hat{\beta}$, centered at the mean as defined by the null hypothesis. For example, if $H_0 : \beta_1 = 0$, then *under the null, the sampling distribution of* $\hat{\beta}_1$ is normally distributed around mean 0.
7. **Distribution of $\hat{\beta}$ under** H_A: Same as above, but the sampling distribution is not considered centered around the mean as defined by the alternate hypothesis.
8. **Decision Rule:** A rule that tells us when to reject and when not to reject the null, using the sampling distributions of $\hat{\beta}$ under H_0 and H_a.

9. What is this decision rule? It is based on the *observed* value of the estimated $\hat{\beta}$. Specifically, we try and define a range of values for $\hat{\beta}$ such that in this range, we reject the null hypothesis. The boundary point of this region is called the **critical value**.

2 Decision Rules

Let us consider the above example where we are trying to estimate the impact of a training programme on wages. Assume that the correct regression specification is:

$$wage = \beta_0 + \beta_1 training + \epsilon$$

Here, we want to test if the programme had *any* impact on wages or not. Therefore, the null hypothesis or baseline considers the case where the true parameter is 0 (i.e. the programme is not associated with a change in wages) and the alternate hypothesis considers the case where the true parameter is positive (i.e. the programme raised wages). Formally, this very rudimentary test is given by:

$H_0 : \beta_1 = 0$

$H_a : \beta_1 > 0$

Decision Rules take the form: Reject H_0 if $\hat{\beta} > k$, where k is the critical value. Suppose k is taken to be 3. Then,

- If we observe $\hat{\beta} = 2$, we assume that this estimate has been computed from a sample drawn from a population that has mean 0 (i.e. under the null). Therefore, although 0 and 2 are numerically different, they are statistically similar.
- Similarly, if we observe $\hat{\beta} = 4$, we assume that this estimate has been computed from a sample drawn from a population that has mean 5 (i.e. under the alternate hypothesis). Therefore, we reject the null. Here, 0 and 4 are statistically different.

3 Errors

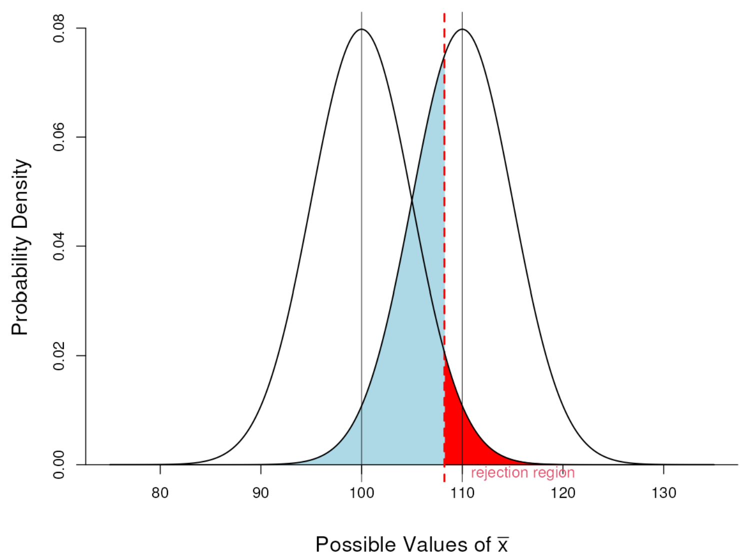
Observed test statistics are products of random draws from the population; and therefore we must beware of errors in inference. Table 1 outlines these errors:

Table 1: Types of Errors in Hypothesis Testing

	H_0 is true	H_a is true
Reject H_0	Type I Error	✓
Don't Reject H_0	✓	Type II Error

We use α to denote the probability of Type I error and β to denote the probability of Type II error. There is a trade-off between the two, which is depicted in Figure 1 (produced using the University of Albany's shiny app; access here).

Figure 1: Type I and Type II Error



P-Value: The probability of observing a value at least as adverse as the observed value *under the null hypothesis*.

4 Test Statistics

A **test statistic** is a real valued function T_n from data to $\mathbb{R}$, where the decision rule is as follows:

1. Do not reject H_0 if $T_n \leq c$
2. Reject H_0 if $T_n > c$

We use the t-statistic: $t = \frac{\hat{\beta - \beta}}{se(\hat{\beta})}$. When we use a 0 null, the t-statistic is reduced to $t = \frac{\hat{\beta}}{se(\hat{\beta})}$.

The distribution of the t-statistic is a bell-shaped curve, much like the normal distribution, but has fatter tails and is used for inference when the true population parameters are not known and we have a finite sample size. When the sample size $n \to \infty$, the t-distribution approximates the standard normal distribution. The t-statistic has $N-k-1$ degrees of freedom, where N is the sample size and k is the number of regressors (therefore, $k+1$ coefficients are estimated).

Critical values for the t-statistic can be found here.

Note that:

1. The t-test does not test theoretical validity.
2. The t-test does not test importance.
3. The t-test is not intended for a sample size approximating the population ($se(\hat{\beta}) \to 0$).

5 Confidence Intervals

Consider the two-sided t-test with α significance level. We want to find an interval s.t. the probability that the population parameter β *is contained* in this interval is $1-\alpha$ (the confidence level). This is equivalent to not rejecting the null at a $(1-\alpha)100\%$ significance level when β is the true parameter.

$$(1-\alpha)100\%\ CI = \hat{\beta} \pm t_{\frac{\alpha}{2}} se(\hat{\beta})$$

The Confidence Interview can be used to as a decision rule in the two-tailed hypothesis test.

6 F-Test

The F-Test allows us to test joint hypotheses on multiple coefficients. Therefore, H_0 is an equation of constraints. The alternate hypothesis H_A is the unconstrained model.

The F-Test Statistic is given by:

$$F = \frac{(RSS_R - RSS_{UR})/M}{RSS_{UR}/N-K-1}$$

Here,

1. RSS_R: Residual Sum of Squares from the restricted regression.

2. RSS_{UR}: Residual Sum of Squares from the unrestricted regression.

3. M: The number of linear constraints.

4. $N - K - 1$: Degrees of freedom of the unrestricted regression.

Reject H_0 if $F_{obs} < F_c$, where F_c is the critical value. These values can be read from the F-table, which can be found here. The F-statistic has df on the basis of the numerator and denominator i.e. M & $N - K - 1$.

The F-Test allows us to *test* the goodness-of-fit of a regression.

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

H_a : At least one $\beta_j \neq 0$

$F = \frac{ESS/K}{RSS/(N-K-1)}$

Here, the numerator is ESS since $RSS_R = TSS$.

Other applications of the F-test include testing for seasonality.

Weekly Assignment: Solve Questions 1, 5, and 6.