Topic 6

Hypothesis Tests and Confidence Intervals

- Consider the multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} + u_i, \;\; i = 1, 2, ...n. \tag{1}$$

  Many hypotheses of interest may be formulated as a linear restriction on one or more of the regression coefficients in (1).
- For example, the linear restriction

$$\beta_1 = 0$$

  corresponds to the hypothesis that x_1 has no effect on y once we control for $x_2, x_3, ..., x_k$.
- In this unit, we restrict our attention to linear restrictions on the regression coefficients.

Hypothesis Tests of a Single Linear Restriction II

- The approach adopted in hypothesis testing is to specify a **null hypothesis** about one or more of the regression coefficients and use the observed data to determine whether or not there is enough evidence to reject the null hypothesis.
- For example, in the linear regression model

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935, \qquad (2)$$

the null hypothesis that education has no effect on wages once we control for IQ, corresponds to the restriction

$$\beta_1 = 0.$$

- Notice that we begin with a null hypothesis expressed in words and we translate it into a restriction on one or more of the regression coefficients.

- Frequently, the null hypothesis of interest implies a restriction on more than one regression coefficient.
- For example, the null hypothesis that an extra year of education has the same effect on average wages as a one point increase in IQ score implies that

  $$\beta_1 = \beta_2,$$

  or equivalently,

  $$\beta_1 - \beta_2 = 0,$$

  in (2).
- Note that a null hypothesis is always expressed as a restriction on one or more the βs, not the $\widehat{\beta}s$.

- It would **not** make sense to test the restriction

$$\widehat{\beta}_1 = 0$$

since once we have estimated the regression equation, we already know the value of $\widehat{\beta}_1$!!

- By convention, the notation H_0 is used to denote the null hypothesis. For example, in the regression model given by (2) above, the notation

$$H_0 : \beta_1 = 0$$

is used to denote the null hypothesis that education has no effect on wages once we control for IQ score.

- Once we have specified the null hypothesis in terms of a restriction on a subset of the regression coefficients, the next step in conducting a hypothesis test is to specify the **alternative hypothesis.**

Hypothesis Tests of a Single Linear Restriction V

1.1 The null and alternative hypotheses

- The alternative hypothesis is the hypothesis which we take to be true if the null hypothesis is false.
- By convention, the notation H_1 is used to denote the alternative hypothesis.
- While the null hypothesis is always expressed as an equality, for example,

$$H_0 : \beta_1 = 0,$$

the alternative hypothesis may be expressed as any one of the following inequalities:

$$H_1 : \beta_1 \neq 0, \tag{3a}$$

$$H_1 : \beta_1 > 0, \tag{3b}$$

$$H_1 : \beta_1 < 0. \tag{3c}$$

Hypothesis Tests of a Single Linear Restriction VI

- 3(a) is called a **two-sided alternative** hypothesis, while (3b) and (3c) are both called **one-sided alternatives**.
- Which form of the alternative hypothesis we adopt in a given hypothesis test is generally determined by appealing to economic/finance theory and/or common sense.
- For example, in the linear regression model

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935,$$

  we would test

$$H_0 : \beta_1 = 0$$

  against

$$H_1 : \beta_1 > 0,$$

  since common sense suggests that additional years of education either have no effect on wages or have a positive effect.

- The next essential ingredient required to conduct a hypothesis test is a **test statistic**.
- A test statistic is a random variable (or sometimes a random vector) whose pdf is known when the null hypothesis is true.
- The pdf of a test statistic when the null hypothesis is true is called the **null distribution** of the test statistic.
- In the linear regression model

$$y = X\beta + u \tag{4}$$

  test statistics for testing hypotheses about the regression coefficients are based on the null distribution of the OLS estimator $\widehat{\beta}$.
- However, in order to derive the null distribution of $\widehat{\beta}$ we need to make the following assumptions:

A1 The model is linear in the parameters. That is,

$$y = X\beta + u.$$

A2 **No perfect multicollinearity.** That is, no column of the X matrix is an exact linear function of any subset of the remaining columns. If this assumption is violated, the OLS estimator of β,

$$\widehat{\beta} = (X'X)^{-1}X'y,$$

is not defined.

A3

$$E(u|X) = \mathbf{0}.$$

A4

$$Var(y|X) = \sigma^2 I_n,$$

or equivalently,

$$Var(u|X) = \sigma^2 I_n.$$

A4 implies:

- Homoskedasticity,

$$Var(y_i|X) = \sigma^2, i = 1, 2, ..., n,$$

  or equivalently,

$$Var(u_i|X) = \sigma^2, i = 1, 2, ..., n.$$

- No serial correlation,

$$Cov(y_i, y_j|X) = 0, j \neq i,$$

or equivalently,

$$Cov(u_i, u_j|X) = 0, j \neq i.$$

A5

$$y|X \sim N(X\beta, \sigma^2 I_n),$$

or equivalently,

$$u|X \sim N(\mathbf{0}, \sigma^2 I_n).$$

- We have discussed assumptions A1-A4 in previous lectures.
- A5 states that:

- $$y_i|X \sim N(x_i'\beta, \sigma^2), i = 1, 2, ..., n,$$

  where

  $$x_i'\beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik},$$

  or equivalently,

  $$u_i|X \sim N(0, \sigma^2), i = 1, 2, ..., n.$$

- Assumptions A1-A5 are illustrated using Fig 1 below for the bivariate linear regression model

$$y_i = \beta_0 + \beta_1 x_i + u_i, i = 1, 2, ...n.$$

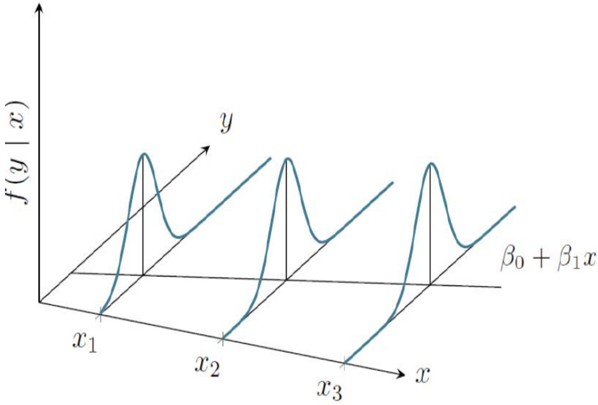


Figure: 1

Hypothesis Tests of a Single Linear Restriction VII

1.2 Test statistics

- If we assume that

$$y|X \sim N(X\beta, \sigma^2 I_n), \tag{5}$$

then in Figure 1

$$\begin{aligned} y_1|x_1 &\sim N(\beta_0 + \beta_1 x_1, \sigma^2), \\ y_2|x_2 &\sim N(\beta_0 + \beta_1 x_2, \sigma^2), \\ y_3|x_3 &\sim N(\beta_0 + \beta_1 x_3, \sigma^2). \end{aligned}$$

In addition, conditional on X, (y_1, y_2, y_3) are uncorrelated.

- The regression model defined by A1-A5 is called the **Classical Normal Linear Regression Model**.
- We can use assumptions A1-A5 to derive the pdf of $\widehat{\beta}$, conditional on X.
- We breakdown the derivation into a sequence of steps.

S1

- We saw in Topic 5 that

$$\begin{aligned}\widehat{\beta} &= (X'X)^{-1}X'y \\ &= (X'X)^{-1}X'y_1 + (X'X)^{-1}X'y_2 + ... + (X'X)^{-1}X'y_n. \quad (6)\end{aligned}$$

That is, $\widehat{\beta}$ can be expressed as a linear function of $(y_1, y_2, ..., y_n)$.

- Linear functions of normally distributed random variables are also normally distributed and

$$y|X \sim N(X\beta, \sigma^2 I_n), \quad \text{(A5)}$$

it follows that

$$\widehat{\beta}|X \sim N(.). \quad (7)$$

S2 Under A1-A3 it can be shown that

$$E(\widehat{\beta}|X) = \beta. \tag{8}$$

(See Appendix 1).

S3 Under A1-A4 it can be shown that

$$Var(\widehat{\beta}|X) = \sigma^2(X'X)^{-1}. \tag{9}$$

(See Appendix 2).

S4 Combining S1, S2 and S3 we obtain

$$\widehat{\beta}|X \sim N[\beta, \sigma^2(X'X)^{-1}] \tag{10}$$

- The result in (10) implies that

$$\widehat{\beta}_j|X \sim N(\beta_j, Var(\widehat{\beta}_j)), j = 0, 1, 2, ..., k,$$

  where

$$Var(\widehat{\beta}_j) = \sigma^2 \alpha_{jj}, j = 0, 1, 2, ..., k,$$

  and α_{jj} denotes the jth diagonal element (the element in row j, column j) of the matrix $(X'X)^{-1}$.
- Note that since we observe X, we can compute $(X'X)^{-1}$ and α_{jj}, $j = 0, 1, 2, ..., k$.

- In summary, we have shown that when assumptions A1-A5 hold,

$$\widehat{\beta}|X \sim N[\beta, \sigma^2 (X'X)^{-1}], \tag{10}$$

where $\widehat{\beta}$ is the OLS estimator of β in the linear regression model

$$y = X\beta + u.$$

- The result in (10) is important because we can use it to derive test statistics for testing hypotheses about the regression coefficients.
- Since

$$\widehat{\beta}_j|X \sim N(\beta_j, \sigma^2 \alpha_{jj}),\ j = 0, 1, 2, ..., k,$$

it follows that

$$\frac{\widehat{\beta}_j - \beta_j}{sd(\widehat{\beta}_j)} = \frac{\widehat{\beta}_j - \beta_j}{\sigma\sqrt{\alpha_{jj}}} \sim N(0, 1), j = 0, 1, 2, ..., k. \tag{11}$$

- Of course, (11) cannot be used to test hypotheses about β_j because it depends on the unknown parameter σ.
- To obtain a valid test statistic we replace σ with the estimator

$$\widehat{\sigma} = \sqrt{\widehat{\sigma}^2} = \sqrt{\frac{SSR}{(n-k-1)}}, \tag{12}$$

where SSR is the sum of squared residuals we get when we estimate

$$y = X\beta + u$$

by OLS.

- Recall that in econometrics

$$se(\widehat{\beta}_j) = \widehat{sd(\widehat{\beta}_j)}.$$

- Replacing σ with $\widehat{\sigma}$ in

$$\frac{\widehat{\beta}_j - \beta_j}{sd(\widehat{\beta}_j)} = \frac{\widehat{\beta}_j - \beta_j}{\sigma\sqrt{\alpha_{jj}}} \sim N(0,1), j = 0, 1, 2, ..., k. \tag{11}$$

we obtain

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} = \frac{\widehat{\beta}_j - \beta_j}{\widehat{\sigma}\sqrt{\alpha_{jj}}} \sim t_{(n-k-1)}, \tag{13}$$

where the notation $t_{(n-k-1)}$ denotes a t distribution with $(n-k-1)$ degrees of freedom.

- Note that while

$$\frac{\widehat{\beta}_j - \beta_j}{\sigma\sqrt{\alpha_{jj}}} \sim N(0,1), \tag{11}$$

the test statistic

$$\frac{\widehat{\beta}_j - \beta_j}{\widehat{\sigma}\sqrt{\alpha_{jj}}} \sim t_{(n-k-1)}. \tag{13}$$

(The proof of (13) is beyond the scope of this unit).

- The t distribution has a very similar shape to that of the standard normal distribution. However, the former has fatter tails. The two distributions become more and more similar as the number of degrees of freedom increases and they converge as

$$(n-k-1) \to \infty.$$

- The standard normal distribution and a t distribution with 6 degrees of freedom are shown in Figure 2 below.

Hypothesis Tests of a Single Linear Restriction XVI

1.2 Test statistics

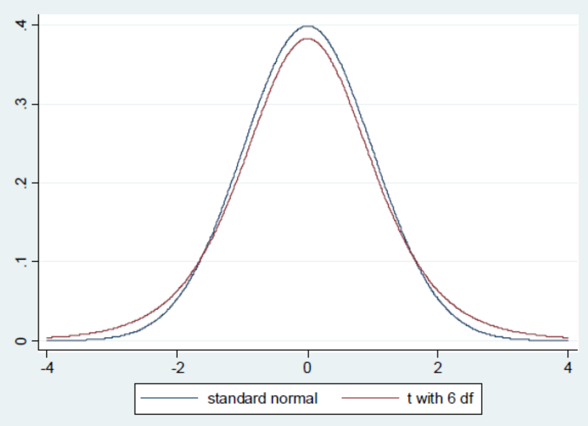


Figure: 2

Hypothesis Tests of a Single Linear Restriction XVII

- Since

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} \sim t_{(n-k-1)},$$

it follows that when the null hypothesis

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

is true

$$\frac{\widehat{\beta}_j - \beta_j^0}{se(\widehat{\beta}_j)} \sim t_{(n-k-1)}. \tag{14}$$

- The distribution $t_{(n-k-1)}$ is called the **null distribution** of our test statistic, $\frac{\widehat{\beta}_j - \beta_j^0}{se(\widehat{\beta}_j)}$, because it is the distribution that our test statistic possesses when the null hypothesis is true.

- For example, when

$$H_0 : \beta_j = 1$$

is true,

$$\frac{\widehat{\beta}_j - 1}{se(\widehat{\beta}_j)} \sim t_{(n-k-1)}.$$

- When a regression equation is estimated it is common practise to test the null hypotheses

$$H_0 : \beta_j = 0, j = 1, 2, ..., k.$$

Testing

$$H_0 : \beta_j = 0$$

is known as testing the **individual significance** of the regressor x_j since, if H_0 is true, x_j is not significant in predicting the value of the dependent variable y.

- When we estimate the linear regression equation

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935,$$

in Eviews we obtain the output shown in Figure 3 below.

Dependent Variable: WAGE
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-128.8899	92.18232	-1.398206	0.1624
EDUC	42.05762	6.549836	6.421171	0.0000
IQ	5.137958	0.955827	5.375403	0.0000

R-squared	0.133853	Mean dependent var	957.9455
Adjusted R-squared	0.131995	S.D. dependent var	404.3608
S.E. of regression	376.7300	Akaike info criterion	14.70414
Sum squared resid	1.32E+08	Schwarz criterion	14.71967

Figure: 3

- The numbers in column 4 are the sample values of the test statistics for testing the individual significance of the regressors.

- For example, the number 6.42 in column 4 of Figure 3 is the **sample value** of the test statistic for testing the null hypothesis

$$\beta_1 = 0.$$

- Notice that in this example

$$\frac{\widehat{\beta}_1}{se(\widehat{\beta}_1)} = \frac{42.06}{6.55} = 6.42.$$

- Carefully note that while the test statistic $\frac{\widehat{\beta}_j - \beta_j^0}{se(\widehat{\beta}_j)}$ is a random variable with pdf $t_{(n-k-1)}$, the sample value of the test statistic is not a random variable.
- In the above example, the sample value of the test statistic is the number 6.42, which is **not** a random variable.

- As the name suggests, a decision rule is a rule we use to determine whether or not to **reject** the null hypothesis of interest.
- The four possible outcomes that can arise when we conduct a hypothesis test are summarized in Figure 4 below.

	Truth in the population	
Decision	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Do not reject H_0	Correct	Type II error

Figure: 4

There are two ways to make an incorrect decision when conducting a hypothesis test:

  - "Reject the truth". This is called a **type 1 error**.
  - "Fail to reject a falsehood". This is called a **type 2 error**.
- When conducting a hypothesis test, the convention is to control the probability of a type 1 error by setting the **significance level** of the test, which we denote by α, at a small specified level, such as 0.01, 0.05 (the most commonly used level) or 0.10.
- In a given hypothesis test, the probability of committing a type 2 error is generally unknown, since it depends on the true value of the parameter of interest (which is unknown).
- Having determined α we next identify a **critical region** or **rejection region** for our test statistic.
- The critical region is chosen so that the probability of the test statistic falling in the critical region if the null hypothesis is true is α%.

- The precise construction of the critical region is determined by how the alternative hypothesis if formulated.
- A boundary of the critical region is marked by a **critical value.**
- Consider the case of testing at the 5% significance level

$$H_0 : \beta_1 = 0$$

against

$$H_1 : \beta_1 > 0$$

in the linear regression model

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935. \qquad (15)$$

This is an example of a **one-sided test**.

- We know that under A1-A5 above, if

$$H_0 : \beta_1 = 0$$

is true,

$$\frac{\widehat{\beta}_1}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{(935-3)} = t_{932}.$$

Because

$$H_1 : \beta_1 > 0,$$

we only regard large positive values as providing evidence against the null hypothesis.

- Our strategy is to reject the null hypothesis if the sample value of our test statistic is so large that the probability of observing it if the null hypotheses were true is $< 5\%$.

- Let c_α denote the critical value for our test statistic. c_α is defined by

$$P(t_{932} > c_\alpha) = 0.05.$$

The rejection region for the test is shown in Figure 5 below.

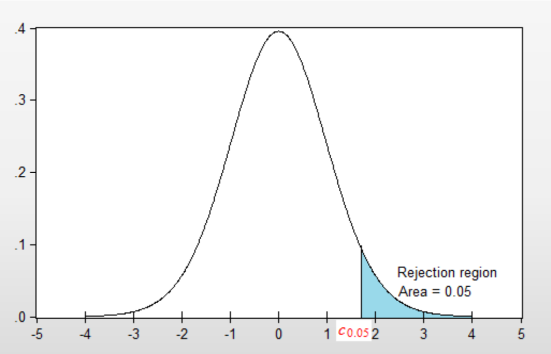


Figure 5

Hypothesis Tests of a Single Linear Restriction VII

1.3 Decision rules

- Since there is a 5% chance that the test statistic will fall in the rejection region if the null hypothesis is true, and we will reject the null if the test statistic falls in the rejection region, it follows that there is a 5% chance of committing a type 1 error (rejecting the truth).
- The 1%, 5% and 10% critical values for $t_{(n-k-1)}$ are given in Table G2 in the textbook for selected degrees of freedom.
- However, critical values for t_{932} are not provided, so we use the closest number of degrees of freedom which is less than 932.
- Adopting this rule, we use the 5% critical value for t_{120} as our t_{crit}. Therefore, from Table G2 in the textbook,

$$t_{crit} = 1.658.$$

- From Figure 3 below we see that the sample vale of our test statistic is 6.421.

Dependent Variable: WAGE
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-128.8899	92.18232	-1.398206	0.1624
EDUC	42.05762	6.549836	6.421171	0.0000
IQ	5.137958	0.955827	5.375403	0.0000

R-squared	0.133853	Mean dependent var	957.9455
Adjusted R-squared	0.131995	S.D. dependent var	404.3608
S.E. of regression	376.7300	Akaike info criterion	14.70414
Sum squared resid	1.32E+08	Schwarz criterion	14.71967

Figure: 3

- Since

$$6.421 > 1.658$$

we reject

$$H_0 : \beta_1 = 0$$

in favor

$$H_1 : \beta_1 > 0.$$

- More specifically, in the context of this example, we reject the null hypothesis that education has no effect on wages when we control for IQ score, in favor of the alternative hypothesis that it has a positive effect.

- The preceding test may be succinctly summarized as follows:

$$
\begin{aligned}
H_0 &: \quad \beta_1 = 0 \\
H_1 &: \quad \beta_1 > 0 \\
\text{Significance level} &: \quad \alpha = 0.05 \\
\text{Test statistic and null distribution} &: \quad \frac{\widehat{\beta}_1}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{932}. \\
t_{calc} &= \quad 6.421 \\
t_{crit} &= \quad 1.658 \\
\text{Decision rule} &: \quad reject\ H_0 \text{ if } t_{calc} > t_{crit} \\
\text{Decision} &: \quad Since\ 6.421 > 1.658, \ \text{reject } H_0
\end{aligned}
$$

- This is the template you should use to present your hypothesis tests in your assignments and in the final exam.

Hypothesis Tests of a Single Linear Restriction XI

1.3 Decision rules

- To test

$$
\begin{aligned}
H_0 &: \beta_1 = 0 \\
H_1 &: \beta_1 < 0
\end{aligned}
$$

  we need only slightly tweak the process described in the example above.

- Because

$$H_1 : \beta_1 < 0,$$

  we only regard large negative values as providing evidence against the null hypothesis.

- Our strategy is to reject the null hypothesis if the sample value of our test statistic is so (algebraically) small that the probability of observing it if the null hypotheses were true is $< 5\%$.

- The rejection region for this test is shown in Figure 6 below.

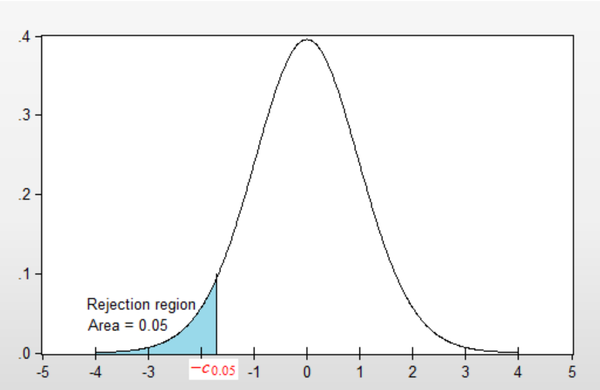


Figure 6

- The test is summarized below:

$$
\begin{aligned}
H_0 &: \quad \beta_1 = 0 \\
H_1 &: \quad \beta_1 < 0 \\
\text{Significance level} &: \quad \alpha = 0.05 \\
\text{Test statistic and null distribution} &: \quad \frac{\widehat{\beta}_1}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{932}. \\
t_{calc} &= \quad 6.421 \\
t_{crit} &= \quad -1.658 \\
\text{Decision rule} &: \quad reject\ H_0 \text{ if } t_{calc} < t_{crit} \\
\text{Decision} &: \quad Since\ 6.421 > -1.658,\ \text{dnr } H_0,
\end{aligned}
$$

where

$$\text{dnr} = \text{do not reject.}$$

- In this case, we fail to reject the null hypothesis that education has no effect on wages when we control for IQ score, against the alternative hypothesis that it has a negative effect.
- Carefully note that is this example, whether or not we reject the null hypothesis that education has no effect of wages when we control for IQ score, depends on how we formulate the alternative hypothesis.
- Of course, in this example common sense strongly suggests that

$$H_1 : \beta_1 > 0$$

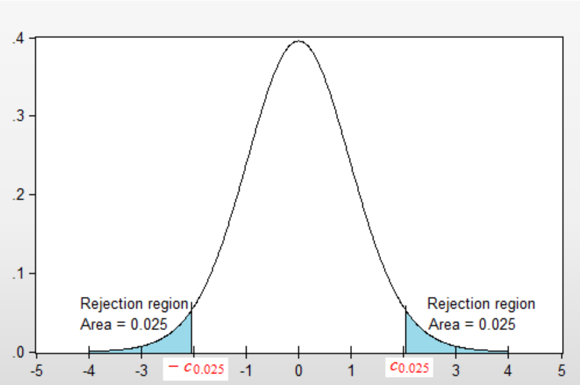
is a much more appropriate alternative hypothesis than

$$H_1 : \beta_1 < 0.$$

- To test

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 \neq 0 \end{aligned}$$

  we once again slightly tweak the process.
- Because of the form of the alternative hypothesis, we will now regard both large positive and large negative sample values of our test statistic as providing evidence against the null hypothesis.
- The rejection region for the test is shown in Figure 7 below.



.4
.3
.2
.1
.0
-5
-4
-3
$-c_{0.025}$
-1
0
1
$c_{0.025}$
3
4
5
Rejection region
Area = 0.025
Rejection region
Area = 0.025

Figure 7

- The key features of the two-sided test are:
  - c_α is chosen so that

$$P(t_{932} > c_\alpha) = \frac{\alpha}{2} = \frac{0.05}{2} = 0.025.$$

  - Our decision rule is to reject H_0 if

$$|t_{calc}| > t_{crit}.$$

  This rejection rule is equivalent to rejecting H_0 if

$$t_{crit} < t_{calc} < -t_{crit}.$$

- A two-sided test of the individual significance of the regressor educ In the wage equation

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935, \qquad (15)$$

is summarized below.

$$
\begin{aligned}
H_0 &: \quad \beta_1 = 0 \\
H_1 &: \quad \beta_1 \neq 0 \\
\text{Significance level} &: \quad \alpha = 0.05 \\
\text{Test statistic and null distribution} &: \quad \frac{\widehat{\beta}_1}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{932}. \\
t_{calc} &= \quad 6.421 \\
t_{crit} &= \quad 1.98 \\
\text{Decision rule} &: \quad reject\ H_0 \text{ if } \left|t_{calc}\right| > t_{crit} \\
\text{Decision} &: \quad Since\ 6.421 > 1.98, \text{ we reject } H_0
\end{aligned}
$$

- In this example, we reject the null hypothesis that, controlling for IQ score, education has no effect on wages in favour of the alternative hypothesis that education has an effect (either a positive or a negative effect)

Hypothesis Tests of a Single Linear Restriction I

1.4 p-values

- A very convenient alternative approach to deciding whether or not to reject a null hypothesis is to use a p-value rather than a critical value, if the p-value is available.
- When testing the

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 > 0 \end{aligned}$$

in the wage equation above,

$$t_{crit} = 1.658.$$

Therefore, any value of t_{932}, say t^*, such that

$$t^* > t_{crit} = 1.658$$

leads to rejection of H_0.

- Also, as we can see in Figure 5 below, if

$$t^* > t_{crit} = 1.658,$$

then

$$P(t_{932} > t^*) < 0.05.$$

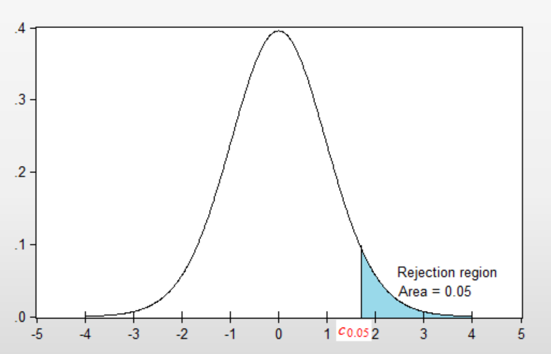


Figure 5

- Therefore, instead of adopting the decision rule to reject H_0 if

$$t_{calc} > t_{crit},$$

we could adopt the equivalent decision rule to reject H_0 if

$$P(t_{932} > t_{calc}) < 0.05. \tag{16}$$

- Define

$$p = P(t_{932} > t_{calc}).$$

Then our decision rule is reject

$$H_0 : \beta_1 = 0$$

in favor of

$$H_1 : \beta_1 > 0$$

if

$$p < 0.05.$$

- In the case of testing

  $$H_0 : \beta_1 = 0$$

  against

  $$H_1 : \beta_1 < 0,$$

  we reject H_0 if

  $$t_{calc} < t_{crit},$$

  or equivalently, if

  $$p < 0.05$$

  where now we define

  $$p = P(t_{932} < t_{calc}).$$

- In the case of testing

$$H_0 : \beta_1 = 0$$

against

$$H_1 : \beta_1 \neq 0,$$

we reject H_0 if

$$|t_{calc}| > t_{crit},$$

where now t_{crit} is chosen so that

$$P(t_{932} > t_{crit}) < \frac{\alpha}{2} = 0.025$$

- Therefore, we reject H_0 if

$$P(t_{932} > |t_{calc}|) < 0.025,$$

or equivalently, if

$$2P(t_{932} > |t_{calc}|) < 0.05.$$

- Define

$$p = 2P(t_{932} > |t_{calc}|). \tag{17}$$

Then in the two-sided test we reject H_0 if

$$p < 0.05.$$

- Carefully note that the **p-values reported by Eviews assume that the researcher wishes to conduct a two-sided test**. That is the p-values are based on (17).

- These p-values must be modified as follows when conducting a one-sided test at the 5% significance level.
    - Case 1:

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 > 0 \\ \alpha &= 0.05 \\ \text{reject } H_0 \text{ if } \mathrm{t}_{calc} &> 0 \text{ and } p < 0.10, \end{aligned}$$

or equivalently,

$$\text{reject } H_0 \text{ if } \mathrm{t}_{calc} > 0 \text{ and } \frac{p}{2} < 0.05,$$

where p is the p-value reported by Eviews.

- Case 2:

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 < 0 \\ \alpha &= 0.05 \\ \text{reject } H_0 \text{ if } \mathrm{t}_{calc} &< 0 \text{ and } p < 0.10, \end{aligned}$$

or equivalently,

$$\text{reject } H_0 \text{ if } \mathrm{t}_{calc} < 0 \text{ and } \frac{p}{2} < 0.05,$$

where p is the p-value reported by Eviews.

- When we estimate a linear regression in Eviews the p-values for testing

$$\begin{aligned} H_0 &: \beta_j = 0 \\ H_1 &: \beta_j \neq 0 \end{aligned}$$

  for each

$$j = 0, 1, 2, ..., k$$

  are automatically reported.

- For example, as we have seen, when we estimate the wage equation we obtain the output in Figure 3 below.

Hypothesis Tests of a Single Linear Restriction XI

Dependent Variable: WAGE
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-128.8899	92.18232	-1.398206	0.1624
EDUC	42.05762	6.549836	6.421171	0.0000
IQ	5.137958	0.955827	5.375403	0.0000

R-squared	0.133853	Mean dependent var	957.9455
Adjusted R-squared	0.131995	S.D. dependent var	404.3608
S.E. of regression	376.7300	Akaike info criterion	14.70414
Sum squared resid	1.32E+08	Schwarz criterion	14.71967

Figure: 3

Note:

- The p-values for a **two-sided test** of the individual significance of the regressors are reported in column 5.

- Based on the reported p-values, at the 5% significance level, we would reject the null hypothesis that educ is statistically insignificant in a two-sided test (p-value < 0.05), or in one-sided test against the alternative that

$$\beta_1 > 0 \ (p - value < 0.10).$$

- We would also reject the null hypothesis that IQ is statistically insignificant in a two-sided test, or in one-sided test against the alternative that

$$\beta_2 > 0 \ (p - value < 0.10).$$

- In addition to testing the individual significance of the regressors, there are other tests involving a single parameter that may be of interest.

Hypothesis Tests of a Single Linear Restriction XIII

1.4 p-values

- For example, suppose that having estimated the wage equation

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935,$$

we wish to test the null hypothesis that, controlling for IQ score, the increase in average weekly wages associated with an extra year of education is \$10, against the alternative hypothesis that it is greater than \$10.

- In terms of restrictions on the parameters of the wage equation, the null and alternative hypotheses are

$$\begin{aligned} H_0 &: \beta_1 = 10 \\ H_1 &: \beta_1 > 10. \end{aligned}$$

- Recall, when A1-A5 hold,

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} = \frac{\widehat{\beta}_j - \beta_j}{\widehat{\sigma}\sqrt{\alpha_{jj}}} \sim t_{(n-k-1)}, j = 0, 1, 2, ..., k.$$

- Therefore, when the null hypothesis is true,

$$\frac{\widehat{\beta}_1 - 10}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{932}.$$

- A test of

$$
\begin{aligned}
H_0 &: \beta_1 = 10 \\
H_1 &: \beta_1 > 10
\end{aligned}
$$

is summarized below:

$$
\begin{aligned}
H_0 &: \quad \beta_1 = 10 \\
H_1 &: \quad \beta_1 > 10 \\
\text{Significance level} &: \quad \alpha = 0.05 \\
\text{Under the null hypothesis} &: \quad \frac{\widehat{\beta}_1 - 10}{se(\widehat{\beta}_1)} \sim t_{(n-k-1)} = t_{932}. \\
t_{calc} &= \quad \frac{42.06 - 10}{6.55} = 4.89 \text{ (see Figure 3)} \\
t_{crit} &= \quad 1.66 \text{ (from table G2 in the textbook)} \\
\text{Decision rule} &: \quad \textit{reject } H_0 \text{ if } t_{calc} > t_{crit} \\
\text{Decision} &: \quad \textit{Since } 4.89 > 1.66, \text{ we reject } H_0
\end{aligned}
$$

- We conclude that, controlling for IQ score, the increase in average weekly wages associated with an extra year of education exceeds $10.

- In applied econometrics we often wish to test hypotheses that involve several regression coefficients.
- Consider the linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, i = 1, 2, \ldots n. \tag{18}$$

- Suppose, for example, that we wish to test

$$\begin{aligned} H_0 &: \beta_1 = \beta_2, \\ H_1 &: \beta_1 > \beta_2, \end{aligned}$$

or equivalently,

$$\begin{aligned} H_0 &: \beta_1 - \beta_2 = 0, \\ H_1 &: \beta_1 - \beta_2 > 0. \end{aligned} \tag{19}$$

- Under A1-A5,

$$\frac{(\widehat{\beta}_1 - \widehat{\beta}_2) - (\beta_1 - \beta_2)}{se(\widehat{\beta}_1 - \widehat{\beta}_2)} \sim t_{(n-k-1)} = t_{(n-4)}.$$

- Using the properties of the variance operator discussed in Topic 2,

$$\widehat{Var}(\widehat{\beta}_1 - \widehat{\beta}_2) = \widehat{Var}(\widehat{\beta}_1) + \widehat{Var}(\widehat{\beta}_2) - 2\widehat{Cov}(\widehat{\beta}_1, \widehat{\beta}_2),$$

  implying that

$$se(\widehat{\beta}_1 - \widehat{\beta}_2) = \sqrt{\widehat{Var}(\widehat{\beta}_1) + \widehat{Var}(\widehat{\beta}_2) - 2\widehat{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)}. \qquad (20)$$

- The quantity $se(\widehat{\beta}_1 - \widehat{\beta}_2)$ is not automatically reported when we estimate

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, i = 1, 2, ...n, \qquad (18)$$

  in Eviews and it takes a lot of work to calculate it manually.
- We can avoid having to calculate (20) manually by using the following "trick".
- Define

$$\delta = \beta_1 - \beta_2 \qquad (21)$$

$\Rightarrow$

$$\beta_1 = \delta + \beta_2. \qquad (22)$$

- Substituting (22) into

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, i = 1, 2, ...n, \quad (18)$$

we obtain

$$\begin{aligned} y_i &= \beta_0 + (\delta + \beta_2) x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, \\ &= \beta_0 + \delta x_{i1} + \beta_2 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, \\ &= \beta_0 + \delta x_{i1} + \beta_2 (x_{i1} + x_{i2}) + \beta_3 x_{i3} + u_i. \quad (23) \end{aligned}$$

- Since

$$\delta = \beta_1 - \beta_2, \tag{21}$$

it follows that testing

$$\begin{aligned} H_0 &: \beta_1 - \beta_2 = 0 \\ H_1 &: \beta_1 - \beta_2 > 0 \end{aligned}$$

in

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i, i = 1, 2, ...n, \tag{18}$$

is **equivalent** to testing

$$\begin{aligned} H_0 &: \delta = 0 \\ H_1 &: \delta > 0 \end{aligned}$$

in

$$y_i = \beta_0 + \delta x_{i1} + \beta_2 (x_{i1} + x_{i2}) + \beta_3 x_{i3} + u_i. \tag{23}$$

- The great advantage of rewriting (18) in the form (23) is that when we estimate (23) by OLS Eviews automatically reports the t statistic for testing
$$H_0 : \delta = 0.$$
- Note that in order to estimate (23) we must first create the new regressor $x_{i1} + x_{i2\,i}$. However, this is easily done in Eviews.
- For example, consider the regression equation

$$Log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 \text{exper}_i + \beta_3 IQ_i + u_i, i = 1, 2, ...935. \tag{24}$$

- As we will see later in the unit, the coefficients β_1 and β_2 in (24) measures the % change in the average wage arising from an additional year of education and from an additional year of experience respectively.

- Suppose that we wish to test the null hypothesis that, controlling for IQ, the % change in the average wage arising from an additional year of education is the same as the % change arising from an extra year of experience, against the alternative hypothesis that the % change arising from an extra year of education is greater.
- In terms of restrictions on the regression coefficients in (24), this null and alternative hypothesis may be expressed as

$$\begin{aligned} H_0 &: \beta_1 - \beta_2 = 0 \\ H_1 &: \beta_1 - \beta_2 > 0. \end{aligned}$$

- As in the discussion above, define

$$\delta = \beta_1 - \beta_2$$

$\Rightarrow$

$$\beta_1 = \delta + \beta_2,$$

and rewrite

$$Log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 \text{exper}_i + \beta_3 IQ_i + u_i, \tag{24}$$

as

$$\begin{aligned} Log(wage_i) &= \beta_0 + (\delta + \beta_2) educ_i + \beta_2 \text{exper}_i + \beta_3 IQ_i + u_i \\ &= \beta_0 + \delta educ_i + \beta_2 (educ_i + \text{exper}_i) + \beta_3 IQ_i + u_i. \end{aligned} \tag{25}$$

Hypothesis Tests of a Single Linear Restriction IX

- Note that in going from (24) to (25) we have replaced the regressor exper with the regressor educ+exper.
- When we estimate (25) by OLS we obtain the output reported in Figure 8 below.

Dependent Variable: LOG(WAGE)
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.198085	0.121543	42.76759	0.0000
EDUC	0.037583	0.006717	5.594875	0.0000
EDUC+EXPER	0.019525	0.003244	6.018132	0.0000
IQ	0.005786	0.000980	5.905770	0.0000

R-squared	0.162244	Mean dependent var	6.779004
Adjusted R-squared	0.159545	S.D. dependent var	0.421144
S.E. of regression	0.386089	Akaike info criterion	0.938773
Sum squared resid	138.7795	Schwarz criterion	0.959481
Log likelihood	-434.8764	Hannan-Quinn criter.	0.946669
F-statistic	60.10079	Durbin-Watson stat	1.811646
Prob(F-statistic)	0.000000		

Figure: 8

- It follows from

$$Log(wage_i) = \beta_0 + \delta educ_i + \beta_2(educ_i + \text{exper}_i) + \beta_3 IQ_i + u_i. \quad (25)$$

that $\widehat{\delta}$ is the estimated coefficient associated with the regressor *educ*.

- From the output in Figure 8

$$\widehat{\delta} = 0.04$$

with a t statistic of 5.59 and a p-value of zero.

- Since
$$p = 0.0 < 0.10$$
we reject
$$H_0 : \delta = 0$$
in favour of
$$H_1 : \delta > 0,$$
or equivalently, we reject
$$H_0 : \beta_1 = \beta_2$$
in favour of
$$H_1 : \beta_1 > \beta_2.$$
- Therefore, we conclude that, controlling for IQ, an extra year of education leads to a greater % increase in the average wage than does an extra year of experience.

- When we estimated the wage equation

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935,$$

we found (see Figure 3 above) that

$$\widehat{\beta}_1 = 42.06. \quad (26)$$

- The number 42.06 is called a **point estimate** of β_1.
- However, if we had collected a different sample from the population of interest we would have obtained a different value for $\widehat{\beta}_1$.
- This observation raises the question of how much confidence we have in the accuracy of our estimate of β_1, given that it is based on single sample.

- A **confidence interval** for β_1 provides a way of making statements about β_1 which takes into account the inherent uncertainty about the accuracy of $\widehat{\beta}_1$ due to the fact that $\widehat{\beta}_1$ is a random variable whose value varies from sample to sample.
- Recall that when A1-A5 hold

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} = t_{(n-k-1)}. \tag{27}$$

- Let $c_{\alpha/2}$ be chosen such that

$$P(|t_{(n-k-1)}| > c_{\alpha/2}) = \alpha,$$

or equivalently,

$$P(|t_{(n-k-1)}| \leq c_{\alpha/2}) = 1 - \alpha.$$

That is, $c_{\alpha/2}$ is chosen such that

$$P[-c_{\alpha/2} \leq t_{(n-k-1)} \leq c_{\alpha/2}] = 1 - \alpha. \tag{28}$$

- Substituting (27) into (28) we obtain

$$P\left[-c_{\alpha/2} \leq \frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} \leq c_{\alpha/2}\right] = 1 - \alpha. \tag{29}$$

- Rearranging (29) we obtain

$$P\left[\widehat{\beta}_j - c_{\alpha/2} se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{\alpha/2} se(\widehat{\beta}_j)\right] = 1 - \alpha \tag{30}$$

- Equation (30) is called a $100(1-\alpha)\%$ confidence interval for β_j.

2 Confidence Intervals IV

- For example, if we choose

$$\alpha = 0.05 \Rightarrow 1 - \alpha = 0.95,$$

  then the equation

$$P\left[\widehat{\beta}_j - c_{0.025}se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{0.025}se(\widehat{\beta}_j)\right] = 0.95 \qquad (30a)$$

  is a 95% confidence interval for β_j.
- When we say that (30a) is a 95% confidence interval for β_j we mean that if we were to take an arbitrarily large number of samples of size n from the population of interest and compute (30a) for each sample, (30a) would contain β_j for 95% of the samples.

- We can express

$$P\left[\widehat{\beta}_j - c_{\alpha/2} se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{\alpha/2} se(\widehat{\beta}_j)\right] = 1 - \alpha \qquad (30)$$

  more compactly by stating that a $100(1-\alpha)\%$ confidence interval for β_j is given by

$$\widehat{\beta}_j \pm c_{\alpha/2} se(\widehat{\beta}_j), \qquad (31)$$

  where $c_{\alpha/2}$ is chosen such that

$$P(|t_{(n-k-1)}| > c_{\alpha/2}) = \alpha.$$

- By convention, $1-\alpha$ is chosen to be 0.90 or 0.95.

2 Confidence Intervals VI

- In the former case

$$P\left[\widehat{\beta}_j - c_{\alpha/2} se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{\alpha/2} se(\widehat{\beta}_j)\right] = 1 - \alpha \tag{30}$$

  is a 90% confidence interval for β_j and in the latter case it is a 95% confidence interval.

- Given $1 - \alpha$, the narrower the interval

$$P\left[\widehat{\beta}_j - c_{\alpha/2} se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{\alpha/2} se(\widehat{\beta}_j)\right] = 1 - \alpha, \tag{30}$$

  the more faith we have in the accuracy of $\widehat{\beta}_j$ as a point estimate of β_j.

- For example, we have more faith in the accuracy of $\widehat{\beta}_j$ as an estimate of β_j if
$$P\left[\widehat{\beta}_j - 2 \leq \beta_j \leq \widehat{\beta}_j + 2\right] = 1 - \alpha$$
than if
$$P\left[\widehat{\beta}_j - 6 \leq \beta_j \leq \widehat{\beta}_j + 6\right] = 1 - \alpha.$$
- Note that, given $\widehat{\beta}_j$, the confidence interval in (30) will be wider the larger is $se(\widehat{\beta}_j)$.
- Therefore, an imprecise estimate of β_j, in the sense that $se(\widehat{\beta}_j)$ is large, will lead to a wide $100(1-\alpha)\%$ confidence interval for β_j
- The interval in (30) is sometimes called an **interval estimate for** β_j, since it is an interval which is estimated to contain β_j with probability $100(1-\alpha)\%$.

2 Confidence Intervals VIII

- We can also use

$$P\left[\widehat{\beta}_j - c_{\alpha/2}se(\widehat{\beta}_j) \leq \beta_j \leq \widehat{\beta}_j + c_{\alpha/2}se(\widehat{\beta}_j)\right] = 1 - \alpha \qquad (30)$$

to test

$$\begin{aligned} H_0 &: \beta_j = \beta_j^0 \\ H_1 &: \beta_j \neq \beta_j^0. \end{aligned}$$

- If the null hypothesis is true then

$$P\left[\widehat{\beta}_j - c_{\alpha/2}se(\widehat{\beta}_j) \leq \beta_j^0 \leq \widehat{\beta}_j + c_{\alpha/2}se(\widehat{\beta}_j)\right] = 1 - \alpha,$$

or equivalently, if H_0 is true there is only a $100\alpha\%$ probability of β_j^0 lying outside this interval.

- Therefore, if

$$\widehat{\beta}_j \pm c_{\alpha/2} se(\widehat{\beta}_j) \tag{31}$$

does not contain β_j^0, we would reject

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

in favor of

$$H_1 : \beta_j \neq \beta_j^0.$$

- Let's use the output in Figure 3 below to construct a 95% confidence interval for β_1 in the wage regression equation

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + u_i, i = 1, 2, ...935. \tag{15}$$

Dependent Variable: WAGE
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-128.8899	92.18232	-1.398206	0.1624
EDUC	42.05762	6.549836	6.421171	0.0000
IQ	5.137958	0.955827	5.375403	0.0000

R-squared	0.133853	Mean dependent var	957.9455
Adjusted R-squared	0.131995	S.D. dependent var	404.3608
S.E. of regression	376.7300	Akaike info criterion	14.70414
Sum squared resid	1.32E+08	Schwarz criterion	14.71967

Figure: 3

- In this example $c_{\alpha/2}$ is chosen so that

$$P(|t_{(932)}| > c_{0.025}) = 0.05.$$

  Therefore, using Table G2 from the textbook,

$$c_{0.025} = 1.98.$$

- From the output in Figure 3

$$\widehat{\beta}_1 = 42.06, se(\widehat{\beta}_1) = 6.55.$$

2 Confidence Intervals XII

- Plugging these numbers into

$$\widehat{\beta}_j \pm c_{\alpha/2} se(\widehat{\beta}_j) \tag{31}$$

we obtain a 95% confidence interval for β_1 given by

$$\begin{aligned} 42.06 \pm 1.98(6.55) &= 42.06 \pm 12.97 \\ &= (29.09, 55.03). \end{aligned}$$

That is,

$$P(29.09 < \beta_1 < 55.03) = 0.95. \tag{32}$$

- Based on equation (32), we are 95% confident that, controlling for experience and IQ, an extra year of education would increase the average weekly wage by between \$29.09 and \$55.03.

2 Confidence Intervals XIII

- The confidence interval given by (32) is quite wide, so we don't have a lot of faith in the accuracy of our point estimate of β_1, which is \$42.06.
- This is not surprising since, from Figure 3,

$$\frac{se(\widehat{\beta}_1)}{\widehat{\beta}_1} = \frac{6.55}{42.06} = 0.16.$$

  Therefore, the standard error of $\widehat{\beta}_1$ is approximately 16% of the value of $\widehat{\beta}_1$.
- Since zero does not lie in the interval (29.09, 55.03) we would reject

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_1 &: \beta_1 \neq 0. \end{aligned}$$

- In the multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_k x_{ik} + u_i, i = 1, 2, ...n, \tag{33}$$

  many hypotheses of interest can be expressed as multiple linear restrictions on the regression coefficients.

- For example, it may be of interest to test

$$H_0 : \beta_1 = \beta_2, \beta_1 = -\beta_3 \text{ and } \beta_k = 0,$$

  against

$$H_1 : \beta_1 \neq \beta_2 \text{ and/or } \beta_1 \neq -\beta_3, \text{and/or } \beta_k \neq 0.$$

- Note that the null hypothesis above imposes **three linear restrictions** on the regression coefficients.

- Note also that the null hypothesis is false if any one of the postulated restrictions is false.
- Null hypotheses which impose multiple linear restrictions on the regression coefficients can be tested by performing an F test.
- When performing an F test of multiple linear restrictions on the coefficients of

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_k x_{ik} + u_i, i = 1, 2, ...n, \tag{33}$$

we distinguish between the **unrestricted model**, which is given by (33), and the **restricted model**, which is obtained by imposing on (33) the restrictions that we wish to test.

- Performing an F test of multiple linear restrictions

$$\beta_1 = \beta_2, \beta_1 = -\beta_3 \text{ and } \beta_k = 0,$$

involves the following steps:

S1 We impose the restrictions

$$\beta_1 = \beta_2, \beta_1 = -\beta_3 \text{ and } \beta_k = 0$$

on (33) and derive the restricted model. Imposing the restrictions we obtain

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} - \beta_1 x_{i3} + ... + \beta_{k-1} x_{ik-1} + u_i, i = 1, 2, ...n, \\ &= \beta_0 + \beta_1 (x_{i1} + x_{i2} - x_{i3}) + ... + \beta_{k-1} x_{ik-1} + u_i, i = 1, 2, ...n, \end{aligned} \tag{33a}$$

S2 Estimate the unrestricted model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_k x_{ik} + u_i, i = 1, 2, ...n,$$

by OLS and obtain the SSR, which we denote by SSR_{ur}.

S3 Estimate the restricted model

$$y_i = \beta_0 + \beta_1 (x_{i1} + x_{i2} - x_{i3}) + ... + \beta_{k-1} x_{ik-1} + u_i, i = 1, 2, ...n,$$

by OLS and obtain the SSR, which we denote by SSR_r.

S4 Form the test statistic

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q},$$

where:

- q denotes the number of restrictions imposed under the null hypothesis**.**
- k denotes the number of explanatory variables (excluding the constant) **in the unrestricted model**.

- When A1-A5 hold, and the restrictions are valid (the null hypothesis is true), it can be shown that

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} \sim F(q, n-k-1). \tag{34}$$

(The proof of this result is beyond the scope of this unit).

- Since in this example we are imposing 3 restrictions,

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{3} \sim F(3, n-k-1).$$

S5 We reject the null hypothesis that the restrictions are true if

$$F_{calc} > F_{crit},$$

where F_{calc} denotes the sample value of the test statistic and F_{crit} is chosen such that

$$P[F(q, n-k-1) > F_{crit}] = \alpha,$$

where α denotes the significance level of the test.

- The pdf for F(3,60) is shown in Figure 9 below.

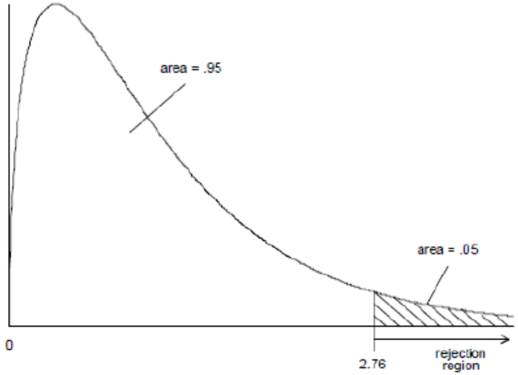


Figure: 9

From Table G3b of the textbook, at the 5% significance

$$F_{crit} = 2.76.$$

We reject H_0 if

$$F_{calc} > 2.76.$$

- Note that an F test of multiple linear restrictions is always a **one-tail test**. That is, the rejection region is always located in the right tail of the distribution.
- Whenever the linear regression equation

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} ..... + \beta_k x_{ik} + u_i, i = 1, 2, ...n, \quad (33)$$

is estimated by OLS, it is standard practise to test the null hypothesis that the explanatory variables are **jointly insignificant**.

- This is an important hypothesis to test, since if it cannot be rejected it means that the chosen explanatory variables have no ability to explain, or predict, changes in the dependent variable, in which case the regression model is useless.
- In terms of restrictions on the regression coefficients, the null hypothesis that the explanatory variables are jointly insignificant may be expressed as

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = ... = \beta_k = 0 \\ H_1 &: \beta_j \neq 0, \text{ for at least one } j = 1, 2, ..., k. \end{aligned} \tag{35}$$

- Notice that this null hypothesis imposes k restrictions on the regression coefficients.
- If (35) is true, then none of the explanatory variables help to predict the dependent variable.

- When we are unable to reject (35) we say that the regressors are **jointly insignificant.**
- For this reason, the test is often called **a test of the joint significance of the regressors**.
- It is important to note that an F test can only be used to test against two-sided alternatives.
- For example, we cannot use an F test to test

$$
\begin{aligned}
H_0 &: \beta_1 = \beta_2 = ... = \beta_k = 0 \\
H_1 &: \beta_j > 0, \text{ for at least one } j = 1, 2, ..., k.
\end{aligned}
$$

- The reason for this is that the unrestricted model is not well defined when the alternative hypothesis is

$$H_1 : \beta_j > 0, \text{ for at least one } j = 1, 2, ..., k.$$

- When we estimate a linear regression model in Eviews the sample value of the test statistic for testing the joint significance of the regressors, together with a p-value, are automatically reported.
- For example, consider the linear regression equation

$$Log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 \exp er + \beta_3 IQ_i + u_i, i = 1, 2, ...935. \tag{36}$$

In terms of restrictions on the regression coefficients, the null hypothesis that the regressors educ, exper and IQ are jointly insignificant is

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0. \tag{37}$$

- Imposing (37) on (36) we obtain the restricted model

$$Log(wage_i) = \beta_0 + u_i, i = 1, 2, ...935. \tag{38}$$

- When (37) holds,

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} = \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(935-4)}{3} \sim F(3, 931)$$

where

$$\begin{aligned} SSR_r &= SSR \text{ from estimating (38)}, \\ SSR_{ur} &= SSR \text{ from estimating (36)}. \end{aligned}$$

- If

$$F_{calc} > F_{crit}$$

we reject

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

in favor of

$$H_1 : \beta_j \neq 0 \text{ for at least one } j = 1, 2, 3.$$

- However, in the case of testing the joint significance of the regressors, we don't have to compute F_{calc} and F_{crit} manually, because Eviews automatically reports F_{calc} and an associated p-value whenever we estimate a linear regression model.
- When we estimate

$$Log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 \exp er + \beta_3 IQ_i + u_i, i = 1, 2, ...935 \tag{36}$$

we obtain the output reported in Figure 10 below.

Dependent Variable: LOG(WAGE)
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.198085	0.121543	42.76759	0.0000
EDUC	0.057108	0.007348	7.771960	0.0000
EXPER	0.019525	0.003244	6.018132	0.0000
IQ	0.005786	0.000980	5.905770	0.0000

R-squared	0.162244	Mean dependent var	6.779004
Adjusted R-squared	0.159545	S.D. dependent var	0.421144
S.E. of regression	0.386089	Akaike info criterion	0.938773
Sum squared resid	138.7795	Schwarz criterion	0.959481
Log likelihood	-434.8764	Hannan-Quinn criter.	0.946669
F-statistic	60.10079	Durbin-Watson stat	1.811646
Prob(F-statistic)	0.000000		

Figure: 10

- The sample value of the F statistic for testing the joint significance of the regressors is 60.10 and the associated p-value is zero.
- Therefore, at any significance level, we would reject the null hypothesis that the regressors educ, exper and IQ are jointly insignificant, in favor of the alternative hypothesis that at least one of the regressors is significant, since

$$p - value = 0.00 < \alpha$$

for any

$$\alpha > 0.$$

- In the following example, published in the Journal of Financial Economics, Vol. 58, pp. 261-300, 2000, a researcher wished to investigate whether or not political instability has any impact on economic growth once we control for a variety of economic variables.

- The researcher collected data on 65 countries and used the variables rev_coups and assassinations as measures of political instability.
- The unrestricted model is

$$gr_i = \beta_0 + \beta_1 rgdp60 + \beta_2 ys_i + \beta_3 ts_i + \beta_4 rc_i + \beta_5 ass_i + u_i, i = 1, 2, ...65. \tag{39}$$

- The null and alternative hypotheses are

$$\begin{aligned} H_o &: \beta_4 = \beta_5 = 0, \\ H_1 &: \beta_4 \neq 0 \text{ and/or } \beta_5 \neq 0. \end{aligned}$$

- The restricted model is

$$gr_i = \beta_0 + \beta_1 rgdp60 + \beta_2 ys_i + \beta_3 ts_i + u_i, i = 1, 2, ...65. \tag{40}$$

- The output from estimating the unrestricted model (39) is reported in Figure 11 below.

Dependent Variable: GROWTH
Method: Least Squares
Sample: 1 65
Included observations: 65

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.4898	0.6896	0.7102	0.4804
RGDP60/1000	-0.4693	0.1482	-3.1667	0.0024
YEARSSCHOOL	0.5748	0.1393	4.1256	0.0001
TRADESHARE	1.5617	0.7579	2.0604	0.0438
REV_COUPS	-2.1575	1.1103	-1.9432	0.0568
ASSASINATIONS	0.3541	0.4774	0.7417	0.4612

R-squared	0.3589	Mean dependent var	1.9427
Adjusted R-squared	0.3045	S.D. dependent var	1.8971
S.E. of regression	1.5821	Akaike info criterion	3.8431
Sum squared resid	147.6766	Schwarz criterion	4.0438
Log likelihood	-118.9017	Hannan-Quinn criter.	3.9223
F-statistic	6.6052	Durbin-Watson stat	2.1260
Prob(F-statistic)	0.0001		

Figure: 11

- The output from estimating the restricted model, (40), is reported in Figure 12 below.

3 Joint Hypothesis Tests XX

3.1 The F test

Dependent Variable: GROWTH
Method: Least Squares
Sample: 1 65
Included observations: 65

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.1427	0.5495	-0.2597	0.7960
RGDP60/1000	-0.3909	0.1446	-2.7038	0.0089
YEARSSCHOOL	0.5667	0.1412	4.0123	0.0002
TRADESHARE	1.8424	0.7168	2.5702	0.0126

R-squared	0.3177	Mean dependent var	1.9427
Adjusted R-squared	0.2841	S.D. dependent var	1.8971
S.E. of regression	1.6051	Akaike info criterion	3.8439
Sum squared resid	157.1668	Schwarz criterion	3.9777
Log likelihood	-120.9259	Hannan-Quinn criter.	3.8967
F-statistic	9.4667	Durbin-Watson stat	2.0314
Prob(F-statistic)	0.0000		

Figure: 12

- Using the output reported in Figure 11 and Figure 12, we test the null hypothesis that political instability has no impact on economic growth as follows:

$$
\begin{aligned}
H_0 &: \quad \beta_4 = \beta_5 = 0,\ \ H_1 : \beta_4 \text{ and/or } \beta_5 \neq 0. \\
\text{Significance level} &: \quad \alpha = 0.05 \\
\text{Test stat} &: \quad \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} \\
\text{Under the null} &: \quad \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} \sim F(q, n-k-1) \\
\text{Under the null} &: \quad \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(65-6)}{2} \sim F(2, 59) \\
F_{calc} &= \quad \frac{(157.1668 - 147.6766)}{147.6766} \frac{59}{2} = 1.896 \\
F_{crit} &= \quad 3.23 \\
\text{Decision rule} &: \quad reject\ H_0 \text{ if } F_{calc}\ > Ft_{crit} \\
\text{Decision} &: \quad Since\ 1.896 < 3.23,\ \text{ we fail to reject } H_0.
\end{aligned}
$$

- There is insufficient evidence in the sample to reject the proposition that political instability has no effect on economic growth once we control for initial rgdp, average years of schooling and trade share.
- An obvious question to ask is "why should we conduct a joint hypothesis test rather than two t tests of individual significance"?
- Based on the p-values reported in Figure 11, if we had conducted two t tests of individual significance, we would have rejected

$$\begin{aligned} H_0 &: \beta_4 = 0 \\ H_1 &: \beta_4 < 0, \end{aligned}$$

but failed to reject

$$\begin{aligned} H_0 &: \beta_5 = 0 \\ H_1 &: \beta_5 < 0. \end{aligned}$$

- There are two reasons why a joint hypothesis test is preferred to multiple t tests when testing multiple restrictions.
  - Recall that when we conduct a hypothesis test we wish to set the probability of a type 1 error (rejecting the truth) equal to α. However, when we conduct two separate t tests, each t test has a significance level of α, so the two t tests together will have a significance level greater than α. We can avoid this problem by conducting a single joint hypothesis test.
  - When the regressors are highly correlated in our sample, we say that the regression model suffers from **near multicollinearity**. When this occurs, the reliability of t tests of individual significance is adversely affected. We will study the topics of near and exact multicollinearity later in the unit.

- In the case in which **the unrestricted and restricted models have the same dependent variable**, it can be shown that

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} = \frac{(R^2_{ur} - R^2_r)}{(1 - R^2_{ur})} \frac{(n-k-1)}{q}$$

where

R^2_{ur} is the R^2 we obtain when we estimate the unrestricted model,
R^2_r is the R^2 we obtain when we estimate the restricted model.

- In this case we have the option of using the test statistic

$$\frac{(R^2_{ur} - R^2_r)}{(1 - R^2_{ur})} \frac{(n-k-1)}{q} \sim F(q, n-k-1). \qquad (39)$$

- The R-squared form of the F test is particularly convenient for testing the joint significance of the regressors, since the restricted model is

$$y_i = \beta_0 + u_i,$$

which has an

$$R^2 = 0.$$

- In this case (39) reduces to

$$\frac{R^2_{ur}}{(1 - R^2_{ur})} \frac{(n - k - 1)}{q} \sim F(q, n - k - 1). \tag{40}$$

- Therefore, to conduct the F test of joint significance using (40) we need only estimate the unrestricted model, since we only need R^2_{ur} to compute the test statistic.

- In the previous example, using the R^2 reported in Figure 11, the sample value of the test statistic given by (40) is

$$\begin{aligned} \frac{R^2_{ur}}{(1 - R^2_{ur})} \frac{(n-k-1)}{q} &= \frac{0.3589}{(1 - 0.3589)} \frac{(59)}{5} \\ &= 6.60 \end{aligned}$$

which is the same as the sample value of the F statistic reported in Figure 11.

- The t tests and F tests which we have discussed above are **exact tests** in the sense that they are valid whatever the sample size. In deriving these tests we did not rely on arguments that are valid only as the sample size becomes arbitrarily large. (These are called asymptotic arguments).
- However, the assumptions required for exact hypothesis tests to be valid are very strong.
- The results that

$$\frac{\widehat{\beta}_j - \beta_j}{se(\widehat{\beta}_j)} \sim t_{(n-k-1)}$$

and

$$\frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \frac{(n-k-1)}{q} \sim F(q, n-k-1),$$

are ultimately derived from the result that

$$\frac{\widehat{\beta}_j - \beta_j}{sd(\widehat{\beta}_j)} = \frac{\widehat{\beta}_j - \beta_j}{\sigma\sqrt{\alpha_{jj}}} \sim N(0,1), j = 1, 2, ..., k. \tag{41}$$

- The derivation of (41) requires that assumptions A1-A5 (which we previously discussed) hold.
- In many cases one or more of these assumptions fail to hold.
- As we have previously noted:

A3

$$E(u|X) = 0,$$

rarely holds for time series data.

A4
$$Var(y|X) = \sigma^2 I_n,$$
is unrealistic for many data sets.

A5 In addition,
$$y|X \sim N(.,.),$$
is also often problematic.

- If any of A1-A5 are violated, then the t and F tests discussed above are no longer reliable.
- Later in the unit we will see how the t and F tests discussed above need to be modified when we have reason to believe that A4 or A5 fail to hold.

4 Appendix 1: Derivation of the conditional mean of the OLS estimator I

Since

$$\widehat{\beta} = (X'X)^{-1}X'y,$$

it follows that

$$\begin{aligned} E(\widehat{\beta}|X) &= E[(X'X)^{-1}X'y|X] \\ &= (X'X)^{-1}X'E(y|X) \\ &= (X'X)^{-1}X'(X\beta) \\ &= (X'X)^{-1}(X'X)\beta \\ &= I_k\beta \\ &= \beta. \end{aligned}$$

5 Appendix 2: Derivation of the conditional variance of the OLS estimator I

In order to derive $Var(\widehat{\beta}|X)$ we need to appeal to a result in matrix algebra which states that, for a random vector z and a conformable constant matrix A,

$$Var(Az|A) = AVar(z|A)A'. \tag{A2.1}$$

Let

$$\begin{aligned} \widehat{\beta} &= (X'X)^{-1}X'y \\ &= Ay, \end{aligned} \tag{A2.2}$$

where

$$\underset{(k+1xn)}{A} = (X'X)^{-1}X'. \tag{A2.3}$$

5 Appendix 2: Derivation of the conditional variance of the OLS estimator II

It follows from A2.2 that

$$\begin{aligned} Var(\widehat{\beta}|X) &= Var(Ay) \\ &= AVar(y|X)A' \text{ (using A2.1)} \end{aligned} \quad \text{(A2.4)}$$

Under A4,

$$Var(y|X) = \sigma^2 I_n. \quad \text{(A2.5)}$$

Substituting A2.5 into A2.4 we obtain

$$\begin{aligned} Var(\widehat{\beta}|X) &= A[\sigma^2 I_n]A' \\ &= \sigma^2 A I_n A' \\ &= \sigma^2 AA'. \end{aligned} \quad \text{(A2.6)}$$

Finally, substituting

$$\underset{(k+1xn)}{A} = (X'X)^{-1}X' \quad \text{(A2.3)}$$

5 Appendix 2: Derivation of the conditional variance of the OLS estimator III

into (A2.6) we obtain

$$
\begin{aligned}
Var(\widehat{\beta}|X) &= \sigma^2(X'X)^{-1}X'[(X'X)^{-1}X']' \\
&= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}I_{k+1} \\
&= \sigma^2(X'X)^{-1}.
\end{aligned}
$$