University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

3. Hypotheses Testing

Prof. Dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Ljubljana, October 2025

Motivation

The basic task of econometrics is to confront the existing economic theories with the economic reality by using the methods of statistical inference.

P. Samuelson, *Econometrica*, 1954 (22)

Statistical Inference

Null hypothesis → **Alternative hypothesis**

$$H_0 : \beta_j = \beta_0 \quad ; \quad H_1 : \beta_j \neq \beta_0$$

$$H_0 : \beta_j \leq \beta_0 \quad ; \quad H_1 : \beta_j > \beta_0$$

$$H_0 : \beta_j \geq \beta_0 \quad ; \quad H_1 : \beta_j < \beta_0$$

Simple hypothesis ↔ **Composite hypothesis**

One-tailed test ↔ **Two-tailed test**

Statistical Inference

Goal: REJECT H_0

What to put into the null hypothesis?

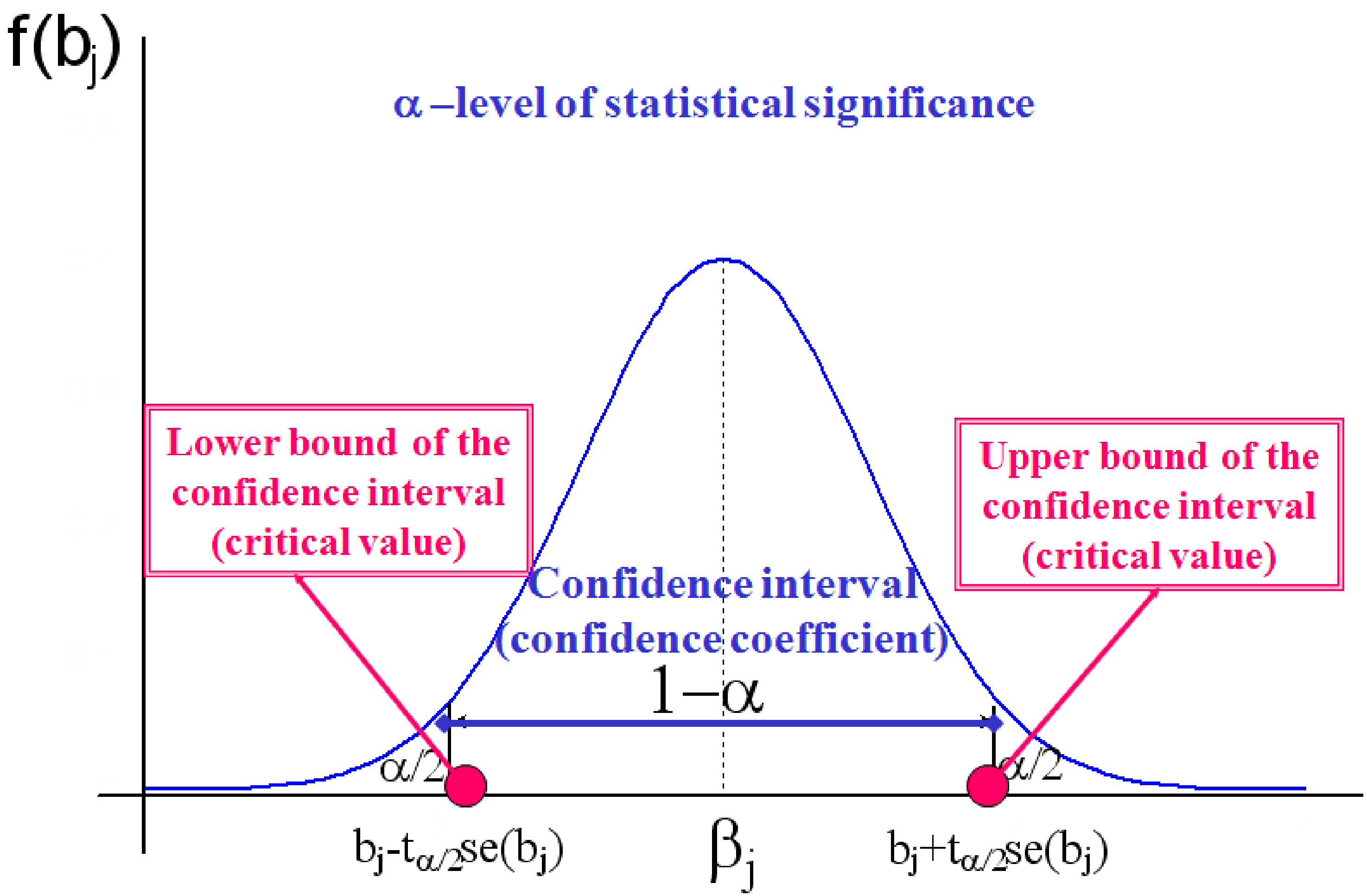
H_o is not rejected! → **H_o is rejected!**

Computer software includes (direct) testing of the following hypothesis:

$$H_0 : \beta_j = \beta_0 \quad ; \quad H_1 : \beta_j \neq \beta_0$$

Simple hypothesis! **Two-tailed test!**

Defining the Confidence Intervals



Hypotheses about Individual Values of β_j

Essential elements:

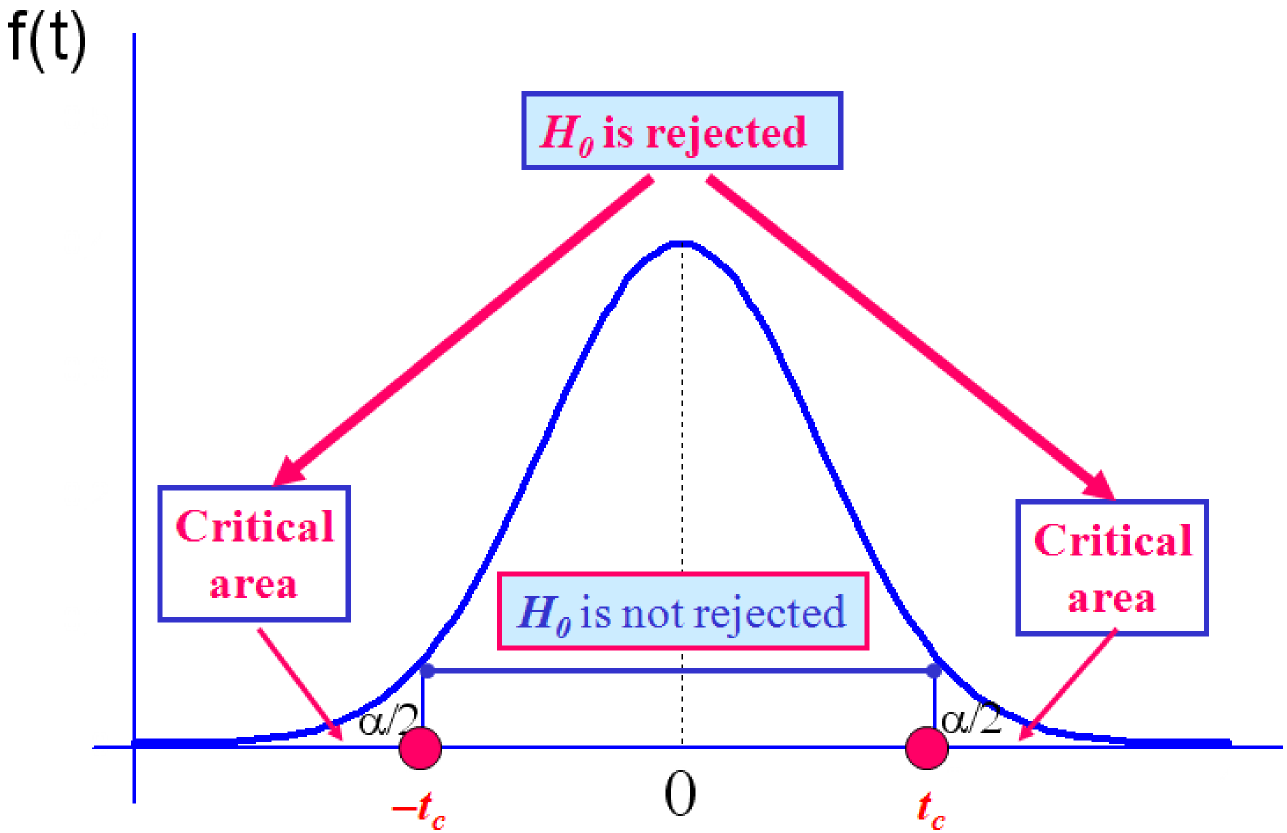
① test statistic

② sample distribution of the test statistic taking into account H_o

$$t = \frac{b_j - \beta_0}{\sqrt{\mathrm{var}(b_j)}} = \frac{b_j - \beta_0}{se(b_j)} \underset{H_0}{\sim} t_{(n-k)}$$

$$\Pr(-t_{\alpha/2} \leq \frac{b_j - \beta_0}{se(b_j)} \leq t_{\alpha/2}) = 1 - \alpha$$

Hypotheses about Individual Values of β_j



If the value of t–statistic is in the critical area,
H_0 is rejected at the level of statistical significance α!

Hypotheses about Individual Values of β_j

Hypothesis testing scheme about the values of regression coefficients based on *t*–statistic

Hypothesis type	H_0	H_1	H_0 is rejected, if
Two-tailed test	$\beta_j = \beta_{j0}$	$\beta_j \neq \beta_{j0}$	$\lvert t \rvert \geq t_{\alpha/2}$
One-tailed test	$\beta_j \leq \beta_{j0}$	$\beta_j > \beta_{j0}$	$t \geq t_{\alpha}$
One-tailed test	$\beta_j \geq \beta_{j0}$	$\beta_j < \beta_{j0}$	$t \leq -t_{\alpha}$

Type I Error ! ⟷ **Type II Error !**

Rule of thumb:

***H_0* is rejected at $\alpha = 0.05$**

- **$t = 2$ or above; two-tailed test**
- **$t = 1.65$ or above; one-tailed test**

Exact Level of Statistical Significance

***p*–value** or **exact level of statistical significance** is the *smallest* statistical significance level that enables *rejection of the null hypothesis* based on a *given data sample*.

E.g. $p = \Pr\left(|t| \geq t_{\alpha/2}\right)$ for the two-tailed test.

The calculated *p*–value is *different* for *every data sample*.

Most often used rule of thumb: *H_o* is rejected, if it holds:

$$p \leq (\alpha = 0.05)$$

In general: *p*–value is the probability of obtaining test statistic value at least as extreme as the one actually observed, under the assumption that the *null hypothesis is true*.

Meaning of Hypotheses Testing

Rejection of the null hypothesis means:

- The ***difference*** between the regression coefficient estimate and the value β_0, tested in the null hypothesis, is ***"statistically significant"*** (statistically significantly different from 0).
- It is unlikely that the difference was established (purely) by chance. This probability is equal to the level of statistical significance of the test α, or $\alpha/2$ for the two-tailed test.
- If the value of regression coefficient, tested in the null hypothesis, is equal to zero ($\beta_0 = 0$), then rejection of the null hypothesis means that the ***regression coefficient estimate*** is statistically significantly different from 0 (the estimated regression coefficient is ***"statistically significant"***).

Meaning of Hypotheses Testing

Statistical significance ⟷ **Economic / theoretical acceptability**

- **A statistically significant effect of an explanatory variable on the dependent variable, measured by the value of the *t*–statistic, does not yet mean the theoretical acceptability of the given explanatory variable.**
- **High value of the *t*–statistic does not mean that this explanatory variable is more important than another explanatory variable with a lower value of *t*–statistic. We can not make conclusions about this based on the value b_j either. For this purpose we use the standardised regression coefficients.**
- **We have to be aware that the value of *t*–statistic depends on several already mentioned determinants. This has to be taken into account when doing statistical inference.**

Statistical Inference

Conclusion based on sample data	Actual state of nature (on the population)	
	H_0 is true	H_0 is false
H_0 is rejected	False conclusion $Pr = \alpha$ (Type I Error)	Correct conclusion $Pr = 1 - \beta$ (Power of the test)
H_0 is not rejected	Correct conclusion $Pr = 1 - \alpha$	False conclusion $Pr = \beta$ (Type II Error)

Due to the presence of Type II Error, the null hypothesis is either rejected or not rejected, but not accepted. If the null hypothesis is rejected, a conclusion is made from the alternative hypothesis. The alternative hypothesis is neither accepted nor rejected, except on those occasions, when its distribution is explicitly known.

Statistical Inference in Econometrics

Goal: Application of the *linear regression model* for determining the validity of theory based on a sample of data.

The model is used for testing hypotheses about a given data generating process (DGP), where we consider only linear restrictions (for now).

The hypothesis has to be testable:

– we are able to write it down in terms of mathematical notation;

– we are able to test it within our model.

Statistical Inference in Econometrics

Nested and non-nested models: models are nested, if one can be obtained by applying a restriction on the other.

$$
\begin{aligned}
M_1 &= (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) \\
M_2 &= (\beta_1, 0\ , \beta_3, \beta_4, \beta_5) \\
M_3 &= (\beta_1, \beta_2, 0\ , \beta_4, \beta_5)
\end{aligned}
$$

Restricted models and non-restricted models:

– the restricted model (“null” model) is the model that has the null hypothesis already incorporated;

– the non-restricted model (“alternative” model) is the base model, i.e. the model from the alternative hypothesis.

Testing Linear Combinations of b

$$H_0 : \mathbf{c}^{\mathrm{T}}\boldsymbol{\beta} = r \quad ; \quad H_1 : \mathbf{c}^{\mathrm{T}}\boldsymbol{\beta} \neq r$$

$$t = \frac{\mathbf{c}^{\mathrm{T}}\mathbf{b} - r}{se(\mathbf{c}^{\mathrm{T}}\mathbf{b})} = \frac{\mathbf{c}^{\mathrm{T}}\mathbf{b} - r}{\sqrt{\operatorname{var}(\mathbf{c}^{\mathrm{T}}\mathbf{b})}}$$

$$\operatorname{var}(\mathbf{c}^{\mathrm{T}}\mathbf{b}) = \mathbf{c}^{\mathrm{T}}((\operatorname{var}-\operatorname{cov}(\mathbf{b}))\mathbf{c}$$

Testing Linear Combinations of **b**

$$\text{var}-\text{cov}(\mathbf{b}) = s_e^2\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}$$

$$var\left(ab_j + cb_l\right) = a^2\,var\left(b_j\right) + c^2\,var\left(b_l\right) + 2ac\,cov\left(b_j, b_l\right)$$

$$var\left(ab_j + cb_l + db_m\right) = a^2\,var\left(b_j\right) + c^2\,var\left(b_l\right) + d^2\,var\left(b_m\right) +$$
$$+2ac\,cov\left(b_j, b_l\right) + 2ad\,cov\left(b_j, b_m\right) + 2cd\,cov\left(b_l, b_m\right)$$

Trinity of Tests in Econometrics

1. **Wald Test (*F*–test)**
2. **Lagrange Multiplier Test (*LM*–test)**
3. **Likelihood Ratio Test (*LR*–test)**

Fundaments of *F*–test and *LM*–test

BASE MODEL

RSS_B

$$RSS_B = \mathbf{e}_B^T \mathbf{e}_B$$

RESTRICTED MODEL

RSS_R

$$RSS_R = \mathbf{e}_R^T \mathbf{e}_R$$

$$RSS_R \geq RSS_B$$

$$R_R^2 \leq R_B^2$$

?

Are the differences in *RSS* statistically significant?

Wald Test (*F*–test)

$$F = \frac{(RSS_R - RSS_B)/m}{RSS_B/(n-k_B)} \sim F_{(m,\, n-k_B)}$$

Degrees of freedom:

$m = (n - k_R) - (n - k_B) = k_B - k_R$

$$F = \frac{((1-R_R^2)-(1-R_B^2))/m}{(1-R_B^2)/(n-k_B)} = \frac{(R_B^2-R_R^2)/m}{(1-R_B^2)/(n-k_B)} \sim F_{(m,\, n-k_B)}$$

I. $F < F_c$ **at given statistical significance level;**
H_0 **is not rejected**

II. $F \geq F_c$ **at given statistical significance level;**
H_0 **is rejected**

Wald Test (*F*–test)

Comparable multiple determination coefficient of the restricted model, when the dependent variable in the restricted model is defined differently than in the base (unrestricted) model:

$$R_R^2 \quad \Rightarrow \quad R_R^{*2} = 1 - \frac{RSS_R}{TSS_B}$$

Wald Test (*F*–test)

APPLICATION OF THE *F*–STATISTIC (WALD TEST)

❶ **Testing hypotheses on particular values of regression coefficients or linear combinations of regression coefficients (equivalent to the *t*–statistic)**

$$F = \frac{(RSS_R - RSS_B)/m}{RSS_B/(n-k_B)} \sim F_{(m,\, n-k_B)}$$

Wald Test (*F*–test)

❷ Simultaneous testing of multiple hypotheses on regression coefficients

A Testing statistical significance of model reduction "General to Simple"

$$H_o: \beta_l = 0 \\ \beta_{l+1} = 0 \\ \vdots \\ \beta_{l+(m-1)} = 0$$

H_1 : At least one value is not equal to 0

$$F = \frac{(RSS_R - RSS_B)/m}{RSS_B/(n-k_B)} \sim F_{(m,n-k_B)}$$

Wald Test (*F*–test)

B Testing statistical significance of model expansion "Simple to General"

$$H_o: \beta_{k+1} = 0$$
$$\beta_{k+2} = 0$$
$$\vdots$$
$$\beta_{k+m} = 0$$

H_1 : At least one value is not equal to 0

$$F = \frac{(RSS_B - RSS_N)/m}{RSS_N/(n-k_N)} \sim F_{(m,\, n-k_N)}$$

Lagrange Multiplier (LM) Test

Example of model reduction:

$$y_i = \beta_1 + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \beta_{k+1} x_{k+1,i} + \cdots + \beta_{k+m} x_{k+m,i} + u_i$$

1

$$H_o : \beta_{k+1} = 0; \quad \beta_{k+2} = 0; \quad \ldots; \quad \beta_{k+m} = 0$$

H_1 : At least one value is not equal to 0

2

Estimate the restricted regression model and calculate e_i

$$y_i = g_1 + g_2 x_{2i} + \cdots + g_k x_{ki} + e_i \quad \Rightarrow \quad e_i$$

Lagrange Multiplier (LM) Test

3 **Estimate the auxiliary regression:**

$$\hat{e}_i = g_1 + g_2 x_{2i} + \cdots + g_k x_{ki} + g_{k+1} x_{k+1,i} + \cdots + g_{k+m} x_{k+m,i} \quad \Rightarrow \quad R^2$$

4 $$LM = nR^2 \sim \chi_m^2$$

5 $$LM \geq \chi_m^2$$

H_o **is rejected**

Likelihood Ratio (LR) Test

BASE MODEL

$\ln L_B$

RESTRICTED MODEL

$\ln L_R$

$$\ln L = -\frac{n}{2}\left[\ln(2\pi) + \ln\left(\frac{RSS}{n}\right) + 1\right]$$

?

Are the differences in ln L statistically significant?

$$LR = -2\left(\ln L_R - \ln L_B\right) \sim \chi^2_m$$

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

3. Hypotheses Testing

Prof. Dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Ljubljana, October 2025