Univerza v Ljubljani

EKONOMSKA FAKULTETA

3. Preizkušanje domnev

doc. dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

Ljubljana, februar 2014

Univerza v Ljubljani

EKONOMSKA FAKULTETA

3.1 Izračunavanje intervala zaupanja za vrednosti regresijskih koeficientov

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

Motivacija

Temeljna naloga ekonometrije je soočanje obstoječe ekonomske teorije z dejanskimi dogajanji ob uporabi metode statističnega sklepanja.

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

Statistično sklepanje

Ničelna domneva → **Alternativna domneva**

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

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

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

Enostavna domneva ↔ **Sestavljena domneva**

Enostranski preizkus ↔ **Dvostranski preizkus**

Statistično sklepanje

Cilj: ZAVRNITEV H_o

? **Kaj obravnavati kot ničelno domnevo ?**

H_o ne zavrnemo ! → **H_o zavrnemo!**

Računalniški paketi vključujejo (neposredno) preverjanje domneve:

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

Enostavna domneva ! **Dvostranski preizkus !**

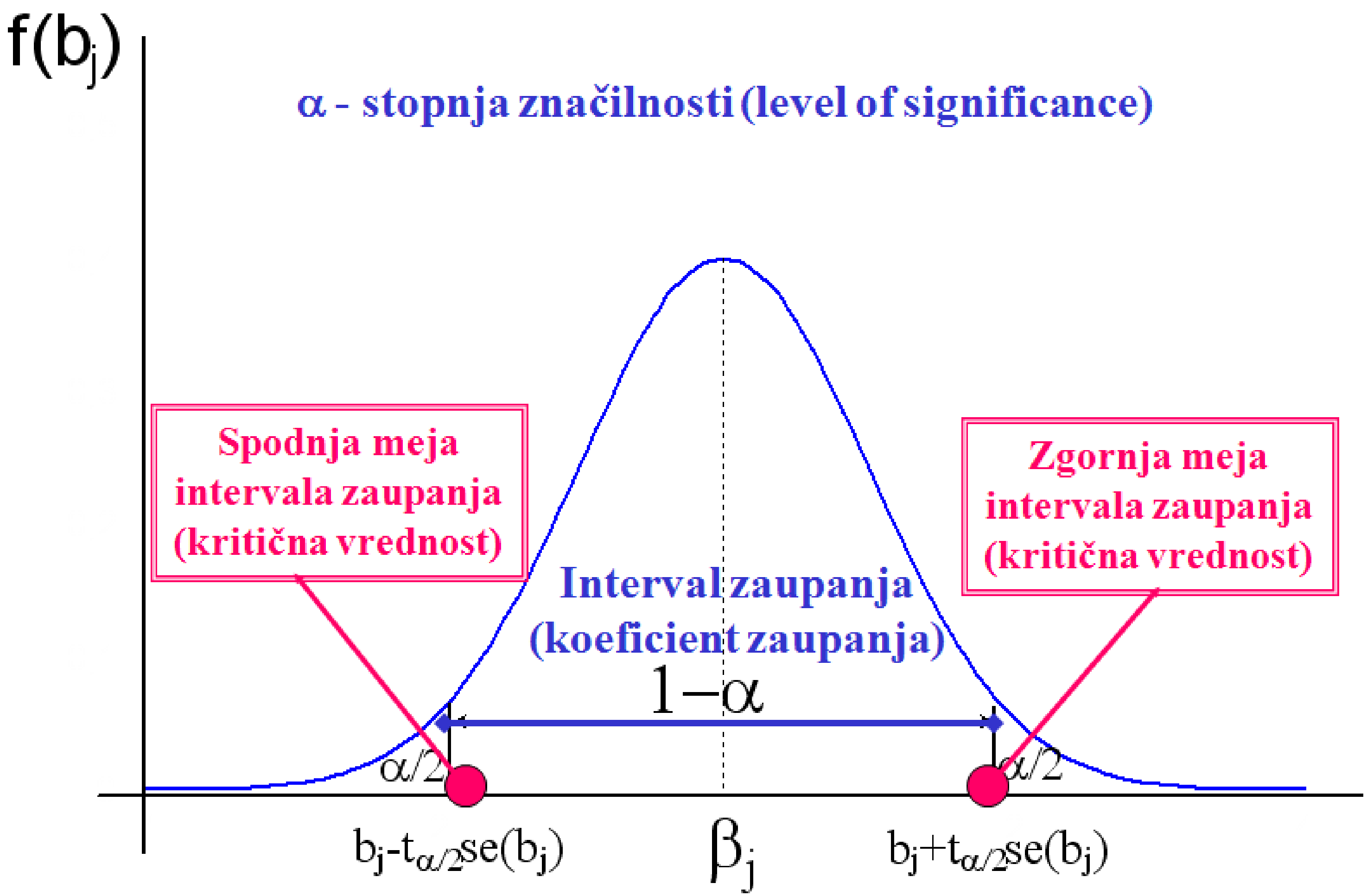
Izračunavanje intervala zaupanja

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

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

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

Določitev intervala zaupanja



Razlage intervala zaupanja

ODLOČITVENO PRAVILO:

A **Če je vrednost regresijskega koeficienta pri ničelni domnevi, to je β_0 , zajeta v izračunani interval, H_0 ne zavrnemo.**

B **Če vrednost β_0 ni zajeta v izračunani interval, H_0 zavrnemo.**

Razlage intervala zaupanja

- Verjetnost, da na podlagi vzorčnih podatkov (s prikazanim postopkom) oblikujemo interval, ki bo vključeval β_j, je enaka $(1 - \alpha)$.
- Interval je naključen, torej se od vzorčnih do vzorčnih podatkov spreminja.
- Zapisano velja le v povprečju, torej za veliko število izračunanih intervalov.
- Izračunanega intervala ne moremo razlagati, da vsebuje (vključuje) pravo vrednost z verjetnostjo $(1 - \alpha)$. Verjetnost, da interval vključuje β_j, je ali 1 ali 0.

Univerza v Ljubljani

EKONOMSKA FAKULTETA

3.2 Preverjanje statistične značilnosti regresijskih koeficientov

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

Domneve o posameznih β_j

Temeljna elementa:

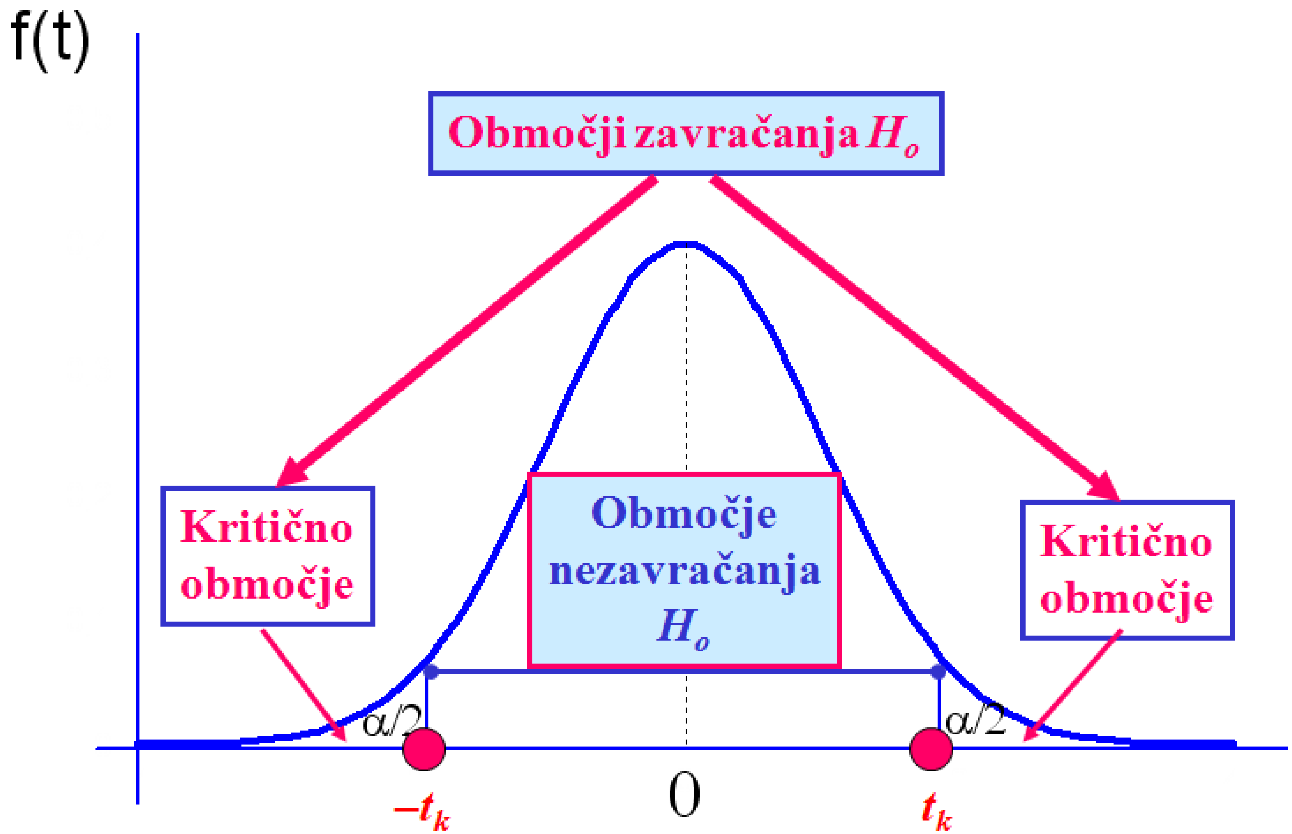
① **testna statistika**

② **vzorčna porazdelitev testne statistike ob upoštevanju H_o**

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

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

Domneve o posameznih β_j



Če se vrednost *t*–statistike nahaja v kritičnem območju, H_o zavrnemo pri stopnji značilnosti α !

Domneve o posameznih β_j

Shema preverjanja domnev o vrednostih regresijskih koeficientov na podlagi *t* - statistike

Vrsta domneve	H_0	H_1	H_0 zavrnemo, če
Dvostranski preizkus	$\beta_j = \beta_{j0}$	$\beta_j \neq \beta_{j0}$	$\lvert t \rvert \geq t_{\alpha/2}$
Enostranski preizkus	$\beta_j \leq \beta_{j0}$	$\beta_j > \beta_{j0}$	$t \geq t_{\alpha}$
Enostranski preizkus	$\beta_j \geq \beta_{j0}$	$\beta_j < \beta_{j0}$	$t \leq -t_{\alpha}$

Napaka I. vrste ! $\longleftrightarrow$ **Napaka II. vrste !**

Pravilo "palca":

H_0 zavrnemo pri $\alpha = 0,05$

- **$t = 2$; dvostranski preizkus**
- **$t = 1,65$; enostranski preizkus**

Natančna stopnja značilnosti

p–vrednost

Prava ali točna stopnja značilnosti

$$\Pr\left(\left|t\right| \geq t_{\alpha/2}\right) = p\text{–vrednost}$$

Običajno pravilo: ***H_o* zavrnemo, če velja:**

$$p\text{–vrednost} \leq (\alpha = 0{,}05)$$

STATA	P, Prob
SORITEC	Signf
SPSS	Sig.
EViews	Prob.
TSP	2 – tail Sig

Pomen preverjanja domnev

Zavrnitev ničelne domneve pomeni:

- Razlika med oceno regresijskega koeficienta in vrednostjo β_0, ki jo preizkušamo v ničelni domnevi, je statistično značilna (statistično značilno različna od vrednosti 0).
- Malo verjetno je, da smo to razliko ugotovili (povsem) naključno. Ta verjetnost je enaka stopnji značilnosti preizkusa α oziroma $\alpha/2$ pri dvostranskem preizkusu.
- Če je vrednost regresijskega koeficienta, ki jo preizkušamo v ničelni domnevi, enaka nič ($\beta_0 = 0$), potem zavrnitev ničelne domneve pomeni, da je ocena regresijskega koeficienta statistično značilno različna od 0 (ocenjeni regresijski koeficient je “statistično značilen”).

Pomen preverjanja domnev

Statistična značilnost ⟷ **Ekonomsko-teoretična sprejemljivost**

- **Statistično značilen vpliv pojasnjevalne spremenljivke na odvisno spremenljivko, potrjen z vrednostjo *t* - statistike, ne pomeni hkrati tudi teoretične sprejemljivosti pojasnjevalne spremenljivke.**

- **Visoka vrednost *t* - statistike ne pomeni, da je ta pojasnjevalna spremenljivka pomembnejša od druge z nižjo vrednostjo *t* - statistike. O tem ne moremo sklepati niti na podlagi vrednosti b_j . O tem lahko sklepamo le na podlagi vrednosti "beta koeficientov".**

- **Zavedati se moramo, da je vrednost *t* - statistike odvisna od več že omenjenih faktorjev. To moramo upoštevati pri presoji primernosti regresijskega modela kot celote in samih *t* - statistik.**

Statistično sklepanje

Ugotovitev na podlagi vzorčnih podatkov	Dejansko stanje (na populaciji)	
	H_0 je pravilna	H_0 je napačna
H_0 zavrnemo	sklep je napačen $Pr = \alpha$ (napaka I. vrste)	sklep je pravilen $Pr = 1 - \beta$ (moč preizkusa)
H_0 sprejmemo	sklep je pravilen $Pr = 1 - \alpha$	sklep je napačen $Pr = \beta$ (napaka II. vrste)

Zaradi prisotnosti napake II. vrste ničelne domneve ne sprejemamo; zgolj zavračamo ali pa ne zavračamo. V kolikor zavrnemo ničelno domnevo, sprejmemo sklep iz alternativne domneve. Alternativne domneve same niti ne sprejemamo niti ne zavračamo, razen v redkih primerih, ko eksplicitno poznamo njeno porazdelitev.

Statistično sklepanje v ekonometriji

Cilj: Uporaba linearnega regresijskega modela za ugotavljanje veljavnosti teorije na podlagi vzorca podatkov.

Model uporabljamo za preizkušanje domnev o danem procesu generiranja podatkov (DGP), pri čemer se omejimo na uporabo linearnih omejitev.

Domneva mora biti preverljiva:
– mora se jo dati oblikovati oziroma matematično zapisati;
– mora se jo dati preizkusiti znotraj našega modela.

Statistično sklepanje v ekonometriji

Gnezdeni in negnezdeni modeli: modela sta gnezdena, če lahko enega dobimo z vključitvijo omejitve v drugega.

$$M_1 = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$$
$$M_2 = (\beta_1, 0\ , \beta_3, \beta_4, \beta_5)$$
$$M_3 = (\beta_1, \beta_2, 0\ , \beta_4, \beta_5)$$

Modeli z omejitvami in modeli brez omejitev:

– model z omejitvami ("ničelni" model) je model, v katerega ustrezno vključimo ničelno domnevo;

– model brez omejitev ("alternativni" model) je osnovni model oziroma model iz alternativne domneve.

Domneve o linearni kombinaciji **b**

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

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

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

Domneve o linearni kombinaciji **b**

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

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

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

Trojica klasičnih testov v ekonometriji

1. **Waldov test (*F*–test)**
2. **Test Lagrangevega multiplikatorja (*LM*–test)**
3. **Test verjetnostnega razmerja (*LR*–test)**

Waldov test (*F* – test)

OSNOVNI MODEL

NVK_o

$$NVK_o = \mathbf{e}_o^{\mathrm{T}} \mathbf{e}_o$$

MODEL Z RESTRIKCIJAMI

NVK_R

$$NVK_R = \mathbf{e}_R^{\mathrm{T}} \mathbf{e}_R$$

$$NVK_R \geq NVK_O$$

$$R_R^2 \leq R_O^2$$

Ali so razlike v NVK značilne ali neznačilne?

Waldov test (*F* – test)

$$F = \frac{(NVK_R - NVK_O)/m}{NVK_O/(n-k_O)} \sim F_{(m,(n-k_O))}$$

Stopinje prostosti:

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

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

I. $F < F_k$ **pri dani stopnji značilnosti;**
H_0 **ne zavrnemo**

II. $F \geq F_k$ **pri dani stopnji značilnosti;**
H_0 **zavrnemo**

Waldov test (*F* – test)

Primerljivi multipli determinacijski koeficient modela z omejitvami, ko je odvisna spremenljivka ocenjenega modela z omejitvami definirana drugače, kot v osnovnem modelu:

$$R_R^2 \quad \Rightarrow \quad R_R^{*2} = 1 - \frac{NVK_R}{SVK_O}$$

Waldov test (*F* – test)

UPORABE *F* – STATISTIKE (WALDOVEGA TESTA)

❶ **Prevera domnev o posamezni vrednosti ali linearni kombinaciji vrednosti regresijskih koeficientov (enakovredna uporaba *t* – statistike)**

$$F = \frac{(NVK_R - NVK_O)/m}{NVK_O/(n-k_O)} \sim F_{(m,(n-k_O))}$$

Waldov test (*F* – test)

❷ Hkratna prevera več domnev o vrednostih regresijskih koeficientov

A Smotrnost izključevanja pojasnjevalnih spremenljivk "General to Simple"

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

H_1 : Vsaj ena od vrednosti ni enaka 0

$$F = \frac{(NVK_R - NVK_O)/m}{NVK_O/(n-k_O)} \sim F_{(m,(n-k_O))}$$

Waldov test (*F* – test)

B Smotrnost vključevanja dodatnih pojasnjevalnih spremenljivk "Simple to General"

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

H_1 : Vsaj ena od vrednosti ni enaka 0

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

Test Lagrangevega multiplikatorja (LM)

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

1

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

H_1 : Vsaj ena od vrednosti ni enaka 0

2

Ocenimo regresijski model z omejitvami in izračunamo e_i

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

Test Lagrangevega multiplikatorja (LM)

3

Ocenimo pomožno regresijo (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

$$LM = nR^2 \quad \sim \quad \chi^2_m$$

5

$$LM \geq \chi^2_m$$

***H_o* zavrnemo**

Test verjetnostnega razmerja (LR)

OSNOVNI MODEL	MODEL Z RESTRIKCIJAMI
$\ln L_O$	$\ln L_R$

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

Ali so razlike v ln L značilne ali neznačilne?

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

Kateri klasični test izbrati?

Odgovor je odvisen od tega, cenilko katerega modela, tj. z omejitvami ali brez, je lažje izračunati:

- test verjetnostnega razmerja (*LR* – test) zahteva ocenjevanje *obeh modelov* (M_O in M_R);
- Waldov test (*F* – test) zahteva zgolj ocenjevanje *modela brez omejitev* (M_O);
- test Lagrangevega multiplikatorja (*LM* – test) zahteva zgolj ocenjevanje *modela z omejitvami* (M_R).

Chowov test strukturne stabilnosti

Prevera enakosti regresijskih koeficientov dveh ali več regresijskih modelov

1

PRM: $N = N_1 + N_2$

$$N_1 : y = \beta_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u_1$$
$$N_2 : y = \gamma_1 + \gamma_2 x_2 + \cdots + \gamma_k x_k + u_2$$

VRM: $n = n_1 + n_2$

$$n_1 : y = b_1 + b_2 x_2 + \cdots + b_k x_k + e_1$$
$$n_2 : y = g_1 + g_2 x_2 + \cdots + g_k x_k + e_2$$

2

$$H_o : \beta_1 = \gamma_1, \beta_2 = \gamma_2, \ldots, \beta_k = \gamma_k$$

H_1 : regresijska modela se razlikujeta

3

$$F = \frac{(NVK - NVK_1 - NVK_2)/k}{(NVK_1 + NVK_2)/(n_1 + n_2 - 2k)}$$

$$F = \frac{(\mathbf{e}^T\mathbf{e} - \mathbf{e}_1^T\mathbf{e}_1 - \mathbf{e}_2^T\mathbf{e}_2)/k}{(\mathbf{e}_1^T\mathbf{e}_1 + \mathbf{e}_2^T\mathbf{e}_2)/(n_1 + n_2 - 2k)}$$

4

$F > F_{k\,(k,\, n1 + n2 - 2k)}$ → H_o **zavrnemo**

Univerza v Ljubljani

EKONOMSKA FAKULTETA

3. Preizkušanje domnev

doc. dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

Ljubljana, februar 2014