

# Ekonometrija 1

## Šeste vaje:

### *Preverjanje domnev o linearni kombinaciji regresijskih koeficientov.*

Na šestih vajah bomo najprej preverjali ničelno domnevo, ki je sestavljena iz večih omejitev vrednosti regresijskih koeficientov. V nadaljevanju pa si bomo pogledali postopek za preverjanje smiselnosti širjenja modela ter s Chowovim testom preverili, ali se proučevana regresijska funkcija razlikuje med dvema podobdobjema.



**Primer 1:** Na podlagi podatkov za 125 gospodinjstev so ocenili naslednjo funkcijo povpraševanja po maslu:

$$\widehat{\ln Q} = 1,18 - 0,75 \ln CM + 0,35 \ln CMAR + 0,20 \ln CZP + 0,65 \ln RD \quad R^2 = 0,750$$
$$t: \quad (1,2) \quad (-2,4) \quad (1,3) \quad (0,9) \quad (4,1) \quad \sum (\ln Q_i - \overline{\ln Q})^2 = 118,5$$

Pri tem  $Q$  predstavlja obseg povpraševanja,  $CM$  ceno masla,  $CMAR$  ceno margarine,  $CZP$  cene osnovnih življenjskih potrebščin in  $RD$  razpoložljivi dohodek. Na podlagi ekonomske teorije so oblikovali naslednje tri omejitve:

- ♦ sprememba vseh pojasnjevalnih spremenljivk za en odstotek ne povzroči nobene spremembe v obsegu povpraševanja (funkcija je homogena stopnje 0);
- ♦ elastičnost povpraševanja glede na ceno margarine in elastičnost povpraševanja glede na ceno masla sta si različni le po predznaku;
- ♦ dohodkovna elastičnost povpraševanja po maslu je enaka 0,5.

Ker so želeli preveriti njihovo hkratno sprejemljivost, so ocenili naslednji regresijski model:

$$\ln \left( \frac{Q \cdot CZP^{0,5}}{RD^{0,5}} \right) = 1,12 + 0,25 \ln \left( \frac{CMAR}{CM} \right) \quad R^2 = 0,625 \quad s_e = 0,500$$
$$t: \quad (2,8) \quad (3,6)$$

- Zapišite omenjene omejitve in prikažite, kako so prišli do transformirane funkcije povpraševanja po maslu.
- Z ustreznim testom preverite hkratno sprejemljivost zapisanih omejitev.
- Zapišite prvotno funkcijo povpraševanja po maslu z regresijskimi koeficienti, dobljenimi ob upoštevanju omenjenih omejitev in jih razložite. ■

**Primer 2:** V okviru raziskave o značilnostih novih avtomobilov smo želeli ugotoviti, kateri dejavniki določajo njihovo najvišjo možno doseženo hitrost (*HITROST*; v km/h). Iz baze podatkov za različne tipe novih avtomobilov (Vir: Revija Avto+Šport, Ljubljana 1997) smo

jih na slučajen način izbrali sto (glej datoteko `hitrost.dta`) in se najprej odločili za ocenjevanje naslednjega regresijskega modela:

$$HITROST_i = \beta_1 + \beta_2 KW_i + \beta_3 CCM_i + u_i$$

pri čemer je  $KW$  (v kilovatih) moč motorja,  $CCM$  (v kubičnih centimetrih) pa njegova prostornina. Nato smo zgornji model razširili še z dvema pojasnjevalnima spremenljivkama: število vrtljajev na minuto ( $VRTLJAJI$ ) in višina avtomobila ( $VISINA$ ; v centimetrih). Ocenite oba modela, razložite dobljene rezultate in z ustreznimi testi presodite, ali bi razširjenemu modelu dali prednost pred osnovnim.

**a) Postopek preverjanja domneve z  $F$ -testom:**

1. Zapišemo in ocenimo razširjeni (novi) model:

$$HITROST_i = \beta_1 + \beta_2 KW_i + \beta_3 CCM_i + \beta_4 VRTLJAJI_i + \beta_5 VISINA_i + u_i$$

$$\widehat{HITROST}_i = b_1 + b_2 KW_i + b_3 CCM_i + b_4 VRTLJAJI_i + b_5 VISINA_i$$

2. Ničelno domnevo zapišemo v obliki, primerni za njeno vključitev v razširjeni (novi) model:

$$H_0 : \begin{cases} \beta_4 = 0 \\ \beta_5 = 0 \end{cases} \quad H_1 : \text{vsaj ena omejitev ne velja}$$

3. Ničelno domnevo vključimo v razširjeni model, s čimer dobimo osnovni model:

$$HITROST_i = \beta_1 + \beta_2 KW_i + \beta_3 CCM_i + u_i$$

4. Ocenimo osnovni model:

$$\widehat{HITROST}_i = b_1 + b_2 KW_i + b_3 CCM_i$$

5. Izračunamo  $F$ -statistiko s pomočjo ustreznega obrazca:

$$F = \frac{(NVK_O - NVK_N)/(k_N - k_O)}{NVK_N/(n - k_N)} =$$

- $NVK_N$  : nepojasnjena vsota kvadratov ocenjenega razširjenega (novega) modela;
- $NVK_O$  : nepojasnjena vsota kvadratov ocenjenega osnovnega modela;
- $n$  : velikost vzorca (število opazovanih enot);
- $k_N$  : število parametrov v razširjenem (novem) modelu;
- $k_O$  : število parametrov v osnovnem modelu;
- $k_O - k_R = m$  : število omejitev o regresijskih koeficientih.

$$F = \frac{(R_N^2 - R_O^2)/(k_N - k_O)}{(1 - R_N^2)/(n - k_N)} =$$

$R_N^2$  : determinacijski koeficient ocenjenega razširjenega (novega) modela;

$R_O^2$  : determinacijski koeficient ocenjenega osnovnega modela.

6. Izvedemo ustrezen sklep (na podlagi primerjave izračunane  $F$ -statistike in njene ustrezne kritične vrednosti ali pa na podlagi  $p$ -vrednosti):

**b) Postopek preverjanja domneve z  $LM$ -testom:**

1. Ocenimo osnovni regresijski model in izračunamo njegove ostanke:

$$\widehat{HITROST}_i = b_1 + b_2 KW_i + b_3 CCM_i$$

2. Ocenimo pomožni regresijski model, v katerem so ostanki iz prejšnje točke odvisna spremenljivka, pojasnjevalne pa so vse tiste, ki imajo to vlogo tudi v razširjenem modelu:

$$\hat{e}_i = g_1 + g_2 KW_i + g_3 CCM_i + g_4 VRTLJAJI_i + g_5 VISINA_i$$

3. Izračunamo vrednost  $LM$ -statistike s pomočjo ustreznega obrazca:

$$LM = nR^2 =$$

$R^2$  : determinacijski koeficient pomožnega regresijskega modela;

$n$  : število opazovanih enot.

4. Izvedemo ustrezen sklep (na podlagi primerjave izračunane  $LM$ -statistike (porazdeljuje se v  $\chi^2$ -porazdelitvi) in ustrezne kritične vrednosti ali pa na podlagi  $p$ -vrednosti):

c) **Postopek preverjanja domneve z LR–testom:**

1. Zapišemo in ocenimo osnovni regresijski model ter izračunamo vrednost logaritma funkcije verjetja,  $\ln L_O$ :

$$HITROST_i = \beta_1 + \beta_2 KW_i + \beta_3 CCM_i + u_i$$

$$\widehat{HITROST}_i = b_1 + b_2 KW_i + b_3 CCM_i$$

2. Zapišemo in ocenimo razširjeni (novi) model ter izračunamo vrednost logaritma funkcije verjetja,  $\ln L_N$ :

$$HITROST_i = \beta_1 + \beta_2 KW_i + \beta_3 CCM_i + \beta_4 VRTLJAJI_i + \beta_5 VISINA_i + u_i$$

$$\widehat{HITROST}_i = b_1 + b_2 KW_i + b_3 CCM_i + b_4 VRTLJAJI_i + b_5 VISINA_i$$

3. Izračunamo vrednost LR–statistike s pomočjo ustreznega obrazca:

$$LR = -2(\ln L_O - \ln L_N) =$$

$\ln L_O$  : vrednost logaritma funkcije verjetja v osnovnem modelu;

$\ln L_N$  : vrednost logaritma funkcije verjetja v razširjenem (novem) modelu.

4. Izvedemo ustrezen sklep (na podlagi primerjave izračunane LR–statistike (porazdeljuje se v  $\chi^2$ –porazdelitvi) in ustrezne kritične vrednosti ali pa na podlagi  $p$ –vrednosti):

***Izpis rezultatov obdelav v programskem paketu Stata:***

**. regress hitrost kw ccm**

Source	SS	df	MS	Number of obs =	100
Model	64645.1698	2	32322.5849	F( 2, 97) =	183.36
Residual	17099.0202	97	176.278559	Prob > F =	0.0000
Total	81744.19	99	825.698889	R-squared =	0.7908
				Adj R-squared =	0.7865
				Root MSE =	13.277

hitrost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
kw	.6865858	.040164	17.09	0.000	.6068714 .7663003
ccm	-.017643	.002115	-8.34	0.000	-.0218407 -.0134453
_cons	172.4126	4.342353	39.70	0.000	163.7942 181.0309

**. \* F-test**

**. regress hitrost kw ccm vrtljaji visina**

Source	SS	df	MS	Number of obs =	100
Model	73156.6452	4	18289.1613	F( 4, 95) =	202.32
Residual	8587.54479	95	90.3952083	Prob > F =	0.0000
Total	81744.19	99	825.698889	R-squared =	0.8949
				Adj R-squared =	0.8905
				Root MSE =	9.5076

hitrost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
kw	.3932561	.048096	8.18	0.000	.2977735 .4887387
ccm	-.0056628	.0022004	-2.57	0.012	-.0100312 -.0012945
vrtljaji	.0061643	.0021917	2.81	0.006	.0018133 .0105153
visina	-.6696778	.0830941	-8.06	0.000	-.8346404 -.5047152
_cons	244.018	19.37144	12.60	0.000	205.5608 282.4752

**. display ((17099.0202-8587.54479)/2) / (8587.54479/(100-5))**

47.07924

**. display Ftail(2,95,47.07924)**

6.204e-15

**. \* LM-test**

**. regress hitrost kw ccm**

Source	SS	df	MS	Number of obs =	100
Model	64645.1698	2	32322.5849	F( 2, 97) =	183.36
Residual	17099.0202	97	176.278559	Prob > F =	0.0000
Total	81744.19	99	825.698889	R-squared =	0.7908
				Adj R-squared =	0.7865
				Root MSE =	13.277

hitrost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
kw	.6865858	.040164	17.09	0.000	.6068714 .7663003
ccm	-.017643	.002115	-8.34	0.000	-.0218407 -.0134453
_cons	172.4126	4.342353	39.70	0.000	163.7942 181.0309

```
. predict test, resid
```

```
. regress test kw ccm vrtljaji visina
```

Source	SS	df	MS	Number of obs =	100
Model	8511.47534	4	2127.86884	F( 4, 95) =	23.54
Residual	8587.54479	95	90.3952083	Prob > F =	0.0000
				R-squared =	0.4978
				Adj R-squared =	0.4766
				Root MSE =	9.5076
Total	17099.0201	99	172.717375		

test	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
kw	-.2933297	.048096	-6.10	0.000	-.3888123 -.1978472
ccm	.0119802	.0022004	5.44	0.000	.0076118 .0163485
vrtljaji	.0061643	.0021917	2.81	0.006	.0018133 .0105153
visina	-.6696778	.0830941	-8.06	0.000	-.8346404 -.5047152
_cons	71.60545	19.37144	3.70	0.000	33.14827 110.0626

```
. scalar lm=100*0.4978
```

```
. display "chi2(2) = " lm  
chi2(2) = 49.78
```

```
. display "Prob > chi2 = " chi2tail(2,lm)  
Prob > chi2 = 1.550e-11
```

```
. * LR-test
```

```
. qui regress hitrost kw ccm  
. estimates store mo
```

```
. qui regress hitrost kw ccm vrtljaji visina  
. estimates store mn
```

```
. lrtest mo mn
```

```
Likelihood-ratio test  
(Assumption: mo nested in mn)
```

```
LR chi2(2) = 68.87  
Prob > chi2 = 0.0000
```



**Primer 3:** Na podlagi podatkov za ZDA (glej datoteko denar\_povprasevanje.dta) smo ocenjevali naslednjo funkcijo povpraševanja po denarju (za razdobje 1960–1993):

$$\ln MI_t = \beta_1 + \beta_2 \ln BDP_t + \beta_3 OM_t + u_t$$

Pri tem imajo uporabljene oznake spremenljivk naslednji pomen:

*MI*: Realna količina denarja v obtoku (v mlrd USD);

*BDP*: Realni bruto domači proizvod (v mlrd USD);

*OM*: Obrestna mera za primarni denar (v odstotkih).

S pomočjo Chowovega testa preverite ali je proučevana funkcija povpraševanja po denarju v ZDA odvisna od tega, katera stranka je bila na oblasti (neprava spremenljivka  $DPOL_t$ , ki zavzame vrednost 1, če so bili na oblasti demokrati in 0, če so bili republikanci).

***Izpis rezultatov obdelav v programskem paketu Stata:***

```
. gen lml=log(ml)
. gen lbdp=log(bdp)
```

```
. regress lml lbdp om
```

Source	SS	df	MS	Number of obs =	34
Model	.48046176	2	.24023088	F( 2, 31) =	128.10
Residual	.058134101	31	.001875294	Prob > F =	0.0000
				R-squared =	0.8921
				Adj R-squared =	0.8851
Total	.538595861	33	.016321087	Root MSE =	.0433

lml	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lbdp	.6090349	.0383392	15.89	0.000	.5308415 .6872283
om	-.0386984	.0039026	-9.92	0.000	-.0466578 -.0307389
_cons	1.841136	.2897466	6.35	0.000	1.250194 2.432079

```
. regress lml lbdp om if dpol==0
```

Source	SS	df	MS	Number of obs =	21
Model	.28532717	2	.142663585	F( 2, 18) =	133.15
Residual	.019286695	18	.001071483	Prob > F =	0.0000
				R-squared =	0.9367
				Adj R-squared =	0.9296
Total	.304613865	20	.015230693	Root MSE =	.03273

lml	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lbdp	.5779606	.0357665	16.16	0.000	.5028179 .6531033
om	-.037186	.0036035	-10.32	0.000	-.0447567 -.0296152
_cons	2.094345	.2782536	7.53	0.000	1.509756 2.678934

```
. regress lml lbdp om if dpol==1
```

Source	SS	df	MS	Number of obs =	13
Model	.170258925	2	.085129463	F( 2, 10) =	38.90
Residual	.021885775	10	.002188577	Prob > F =	0.0000
				R-squared =	0.8861
				Adj R-squared =	0.8633
Total	.1921447	12	.016012058	Root MSE =	.04678

lml	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lbdp	.6862782	.0786221	8.73	0.000	.5110972 .8614591
om	-.0550313	.0089648	-6.14	0.000	-.0750061 -.0350565
_cons	1.311925	.5822084	2.25	0.048	.0146836 2.609166

```
. display Ftail(3,28,3.8450099)
.02011808
```