

Ekonometrija 1

Četrte vaje:

Neposredni, posredni in skupni vplivi pojasnjevalnih spremenljivk na odvisno. Frisch-Waugh teorem. Razlaga regresijskih koeficientov v različnih funkcijskih oblikah regresijskih modelov.

Na četrth vajah si bomo pogledali, kako lahko na podlagi razumevanja oziroma logike skupnih, posrednih in neposrednih vplivov izračunamo neznane ocene regresijskih koeficientov. Ilustrirali bomo tudi veljavnost Frisch-Waugh teorema, v drugem delu vaj pa bomo na dveh konkretnih primerih ponovili razlago regresijskih koeficientov v različnih funkcijskih oblikah regresijskih modelov, kjer bomo med drugim dopustili tudi nelinearen vpliv pojasnjevalne spremenljivke na odvisno spremenljivko.



Primer 1: Z metodo najmanjših kvadratov sta bili ocenjeni naslednji regresijski funkciji:

$$\widehat{OP} = a_1 + 0,92RD \quad \text{in} \quad \widehat{CZPI} = c_1 + 0,78RD,$$

pri čemer OP predstavlja osebno porabo, $CZPI$ indeks cen življenjskih potrebščin in RD razpoložljivi dohodek. Ugotovite vrednost regresijskega koeficienta b_2 v regresijski funkciji:

$$\widehat{OP} = b_1 + b_2RD + b_3CZPI$$

$t: \quad (4,64) \quad (1,95)$

če je $\text{var}(b_3) = 0,09$.



Primer 2: Želeli smo analizirati, kako odstotek prebivalcev, starejših od 65 let ($STAR$) in poraba cigaret na prebivalca ($TOBAK$; izražena v številu zavitkov cigaret) vplivata na število umrlih na 100.000 prebivalcev ($SMRTNOST$). V vzorec smo izbrali 35 držav in ocenili različne funkcijske oblike regresijskih modelov (datoteka `tobak.dta`), ki so prikazane v spodnjem izpisu rezultatov.

Zapišite ocenjene regresijske modele in razložite dobljene rezultate. Izpeljite in pojasnite, kakšen pomen imajo ocenjeni parcialni regresijski koeficienti v teh modelih.

Izpis rezultatov obdelav v programskem paketu Stata:

```
. gen lsmrtnost=log(smrtnost)
. gen lstar=log(star)
. gen ltobak=log(tobak)
```

```
. regress smrtnost star tobak
```

Source	SS	df	MS	Number of obs =	35
Model	316665.524	2	158332.762	F(2, 32) =	33.26
Residual	152312.816	32	4759.7755	Prob > F =	0.0000
				R-squared =	0.6752
				Adj R-squared =	0.6549
Total	468978.34	34	13793.4806	Root MSE =	68.991

smrtnost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
star	52.79328	6.971745	7.57	0.000	38.5923 66.99426
tobak	1.441107	.501654	2.87	0.007	.4192715 2.462943
_cons	54.76546	104.3763	0.52	0.603	-157.8422 267.3731

```
. regress lsmrtnost lstar ltobak
```

Source	SS	df	MS	Number of obs =	35
Model	.526056623	2	.263028311	F(2, 32) =	47.67
Residual	.176566817	32	.005517713	Prob > F =	0.0000
				R-squared =	0.7487
				Adj R-squared =	0.7330
Total	.70262344	34	.020665395	Root MSE =	.07428

lsmrtnost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lstar	.7856827	.0935545	8.40	0.000	.5951183 .9762471
ltobak	.2834557	.0692842	4.09	0.000	.1423285 .424583
_cons	3.449101	.3855704	8.95	0.000	2.663719 4.234482

```
. regress lsmrtnost star tobak
```

Source	SS	df	MS	Number of obs =	35
Model	.477335891	2	.238667945	F(2, 32) =	33.90
Residual	.225287549	32	.007040236	Prob > F =	0.0000
				R-squared =	0.6794
				Adj R-squared =	0.6593
Total	.70262344	34	.020665395	Root MSE =	.08391

lsmrtnost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
star	.0634508	.0084789	7.48	0.000	.0461798 .0807219
tobak	.0020003	.0006101	3.28	0.003	.0007576 .0032431
_cons	5.746105	.1269412	45.27	0.000	5.487535 6.004676

```
. regress smrtnost lstar ltobak
```

Source	SS	df	MS	Number of obs =	35
Model	341892.883	2	170946.441	F(2, 32) =	43.04
Residual	127085.457	32	3971.42053	Prob > F =	0.0000
				R-squared =	0.7290
				Adj R-squared =	0.7121
Total	468978.34	34	13793.4806	Root MSE =	63.019

smrtnost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lstar	652.2562	79.37034	8.22	0.000	490.5841	813.9283
ltobak	202.1874	58.77972	3.44	0.002	82.45706	321.9178
_cons	-1718.884	327.1125	-5.25	0.000	-2385.19	-1052.577

. regress lsmrtnost star ltobak

Source	SS	df	MS	Number of obs = 35		
Model	.504679945	2	.252339972	F(2, 32) =	40.79	
Residual	.197943495	32	.006185734	Prob > F =	0.0000	
				R-squared =	0.7183	
				Adj R-squared =	0.7007	
Total	.70262344	34	.020665395	Root MSE =	.07865	

lsmrtnost	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
star	.0614618	.0079709	7.71	0.000	.0452256	.077698
ltobak	.2987248	.0731981	4.08	0.000	.1496251	.4478244
_cons	4.584218	.3565487	12.86	0.000	3.857952	5.310484



Primer 3: Na podlagi podatkov za 100 podjetij, ki se ukvarjajo s proizvodnjo visoko razvitih elektronskih naprav, so raziskovalci proučevali, kako je vrednost prodaje (*PROD*; v mio. denarnih enotah) odvisna od izdatkov za oglaševanje (*OGL*; v 1.000 denarnih enotah), izdatkov za raziskave in razvoj (*RR*) in odstotka zaposlenih s fakultetno izobrazbo (*IZOB*).

Med drugim so ocenili tudi naslednji regresijski funkciji, v katerih spremenljivka OGI^2 predstavlja kvadrat izdatkov za oglaševanje:

$$\widehat{PROD} = 12,33 + 0,46OGL - 0,002OGL^2 + 1,203 \ln RR + 0,004IZOB ;$$

$$\ln \widehat{PROD} = 1,78 + 0,125OGL - 0,005OGL^2 + 1,122 \ln RR + 0,008IZOB .$$

a) Pojasnite na podlagi prve ocenjene funkcije, kaj je značilno za spremembo obsega prodaje, če se obseg izdatkov za oglaševanje poveča za enoto, spremenljivki *RR* in *IZOB* pa ostaneta nespremenjeni? Za koliko se spremeni vrednost prodaje, če se izdatki za oglaševanje zvišajo iz 4.000 na 5.000 denarnih enot, spremenljivki *RR* in *IZOB* pa ostaneta nespremenjeni?

b) Razložite oceni parcialnih regresijskih koeficientov $b_4 = 1,122$ in $b_5 = 0,008$ v drugi ocenjeni funkciji. Kaj vsebinsko predstavljata razloženi oceni?

c) Komentirajte pravilnost naslednje trditve: “Zgornja regresijska modela sta nelinearna v parametrih, zato ju ni bilo možno neposredno oceniti z metodo najmanjših kvadratov.”

