

Dodatne vežbe: Specifikacija i izbor modela

Nastavnik: Aleksandra Nojković, nojkovic@gmail.com

1. Na osnovu 48 kvartalnih podataka o tražnji (Y_t) i ceni u dinarima (X_{1t}) datog proizvoda ocenjen je model i dobijeni sledeći rezultati:

$$\hat{Y}_t = 12,617 - 0,772X_{1t}; \quad \sum_{t=1}^{48} e_t^2 = 0,90.$$

(0,186)

Uvođenjem nove objašnjavajuće promenljive koja se odnosi na dohodak (X_{2t}) ocenjen je model:

$$\hat{Y}_t = 8,94 - 0,558X_{1t} + 0,313X_{2t}; \quad \sum_{t=1}^{48} e_t^2 = 0,757.$$

(0,179) (0,130)

$$\sum_{t=1}^{48} x_{1t}^2 = 0,565; \quad \sum_{t=1}^{48} x_{2t}^2 = 3,266; \quad \sum_{t=1}^{48} x_{1t} x_{2t} = -0,387;$$

- a) Koji od dva modela bolje aproksimira date podatke?
- b) Objasniti zbog čega dolazi do razlike u oceni uz X_1 u dve navedene specifikacije. Izračunati pristrasnost dobijene ocene nagiba u jednostavnom modelu. Šta se dešava sa njenom standardnom greškom?

Rešenje:

- a) Najjednostavnije je izračunati i uporediti standardne grešku dva modela.

$$s_I^2 = \frac{\sum e_t^2}{n - k_I} = \frac{0,90}{48 - 2} = 0,0196.$$

$$s_{II}^2 = \frac{\sum e_t^2}{n - k_{II}} = \frac{0,757}{48 - 3} = 0,0168.$$

Na osnovu kriterijuma manje standardne greške modela, zaključujemo da je ispravna (tačna) II specifikacija koja uključuje obe objašnjavajuće promenljive. Takođe, rezultat t-testa potvrđuje opravdanost uključivanja promenljive X_{2t} u model ($t_{b_2} = 2.41 > t_{45(0,025)}$).

- b) Reč je o problemu izostavljanja relevantne objašnjavajuće promenljive u prvom modelu. Ova greška specifikacije dovodi do pristrasnosti ocenu uz promenljivu koja je zadržana u modelu (uz X_{1t}). Pristrasnost ocene u pogrešno specifikovanom modelu se može izračunati

kao: $\text{Pr}(b_1) = E(b_1) - \beta_1 = \beta_2 p^*$, gde je p^* označena ocena nagiba iz regresije izostavljene promenljive (X_{2t}) na promenljivu koja je ostala u modelu (X_{1t}):

$$\widehat{X}_{2t} = p_0 + p^* X_{1t}$$

$$p^* = \frac{\sum x_{1t} x_{2t}}{\sum x_1^2} = \frac{-0,387}{0,565} = -0,685.$$

U prvom modelu ocena uz promenljivu X_{1t} je pristrasna naniže, odnosno

$$\beta_2 p^* = 0,313 (-0,685) = -0,214.$$

Reč je negativnoj pristrasnosti ocene uz promenljivu X_{1t} , odnosno ocena u pogrešno specificovanom modelu je -0,772, što je potcenjena ocena u poređenju sa ocenom -0,558 tačno specificovanog modela (-0,558-0,214=-0,772). Pored pristrasnosti, ova greška specificacije povećava standardnu grešku ocene uz promenljivu X_{1t} u pogrešno specificovanom modelu ($s_{b_1(T)} = 0,179 < s_{b_1(\perp)} = 0,186$).

2. Na bazi 10 opservacija ocenjene su dve alternativne specificacije modela:

$$(I) \quad \hat{Y} = 1,609 + 0,810X_1 + 0,915X_2 + 0,139X_3; R_I^2 = 0,8726, s_I = 1,8770.$$

(0,146) (0,261) (0,395)

$$(II) \quad \hat{Y} = 3,067 + 0,797X_1 + 0,893X_2; R_{II}^2 = 0,8700, s_{II} = 1,7556.$$

(0,133) (0,237)

Izabрати ispravnu specificaciju modela. Koja greška specificacije je napravljena u pogrešno specificovanom modelu? Diskutovati njene posledice na ocene dobijene metodom ONK.

Rešenje:

U specificaciju (I) uključena je irelevantna promenljiva X_3 , odnosno tačna specificacija modela podrazumeva uključivanje samo dve objašnjavajuće promenljive X_1 i X_2 . Da je model (II) superiorniji možemo zaključiti na bazi manje standardne greške regresije ($s_{II} = 1,7556 < s_I = 1,8770$), veće vrednosti korigovanog koeficijenta determinacije ($\overline{R_{II}^2} = 0,833 > \overline{R_I^2} = 0,809$, izračunati!!!), kao i statističke nesignifikantnosti parametra uz promenljivu X_3 ($t_{b_3} = 0,352 < t^*$).

Greška specifikacije u I modelu (uključivanje irelevantne objašnjavajuće promenljive) ne dovodi do pristrasnosti ocena dobijenih metodom ONK (**pokazati!!!**), ali su standardne greške u pogrešno specifikovanom modelu veće ($s_{b_1}(\perp) = 0,146 > s_{b_1}(T) = 0,133$; $s_{b_2}(\perp) = 0,261 > s_{b_2}(T) = 0,237$).

3. Definirati hipotezu i sprovesti testiranje primenom RESET testa ako je data ocena pomoćne regresije ocenjena na bazi 100 opservacija:

$$\hat{Y}_i = 7.78 + 4.06X_{1i} - 2.28X_{2i} + 0.78\hat{Y}_i^2 + 1.26\hat{Y}_i^3.$$

(2.01) (0.98) (0.14) (2.32)

Navesti prednosti i neostatke ovog testa specifikacije modela.

Rešenje:

Smisao RESET testa se svodi na proveru opravdanosti uključivanja promenljivih \hat{Y}^2 i \hat{Y}^3 u osnovnu regresiju koja sadrži dve objašnjavajuće promenljive X_1 i X_2 . Ukoliko parametre uz ove dodate promenljive \hat{Y}^2 i \hat{Y}^3 obeležimo sa α_1 i α_2 , formalno proveravamo sledeće tvrđenje:

$$H_0: \alpha_1 = 0 \text{ i } \alpha_2 = 0$$

$$H_1: \alpha_1 \neq 0 \text{ i/ili } \alpha_2 \neq 0.$$

Dakle, značajnost makar jedne od ovih promenljivih ukazuje na grešku specifikacije modela. U ovom slučaju, na bazi sprovedenog t-testa, zaključujemo da je parametara α_1 statistički značajno različit od nule ($t_{\hat{\alpha}_1} = 5,57 > t^*$), te da u modelu postoji greška specifikacije.

O prednostima i nedostacima testa – videti Udžbenik, Glava 9!

4. Empirijsku raspodelu vremenske serija reziduala obima 100 karakteriše vrednost Jarque-Berra (JB) test-statistike 8,63. Definisati hipotezu i doneti zaključak o empirijskoj raspodeli reziduala razmatranog modela.

Rešenje:

Pretpostavka KLRM o normalnoj raspodeli slučajne greške proverava se primenom JB testa. Formalno proveravamo važenje sledeće pretpostavke o empirijskoj raspodeli reziduala modela:

$$H_0: \alpha_3 = 0 \text{ i } \alpha_4 = 3 \text{ (empirijska raspodela odgovara normalnoj)}$$

$$H_1: \alpha_3 \neq 0 \text{ i/ili } \alpha_4 \neq 0 \text{ (empirijska raspodela odstupa od normalne).}$$

U ovom slučaju izračunata vrednost test statistic je $JB = 8,63 > \chi_{2(0,05)}^2 = 5,99$, tako da odbacujemo nultu hipotezu i zaključujemo da slučajna greška nema normalnu raspodelu.