

Demonstrature

Pripreme za II kolokvij iz STATISTIKE, dio II
Akademski godina 2007/2008

(1)

Analizira se odnos industrijske proizvodnje (varijabla y) i broja zaposlenih (varijabla x) u RH od 1990 – 1999. godine. Varijable su izražene kao indeksi na stalnoj bazi. Podaci o varijablama dani su u SLJRH 2000, na stranici 256. U analizi se primjenjuje model jednostavne linearne regresije.

a) Kako glasi regresijska jednadžba? Protumačite značenje regresijskog koeficijenta. b) Izračunajte vrijednost regresijske varijance, standardne devijacije i koeficijenta varijacije. Proporcija neprotumačene u ukupnoj sumi kvadrata je 0.3746. c) Koliki je koeficijent determinacije i linearne korelacije? d) Sastavite tabelu ANOVA. e) Odredite granice 95% intervala pouzdanosti parametra β . Standardna pogreška procjene jednaka je 0.1773. f) Provedite jednosmjerni test o značajnosti regresorske varijable. Razina signifikantnosti 5%.

Pri računanju koristite ove međurezultate:

$$\Sigma x_i = 1095.8 \quad \Sigma y_i = 1156.6 \quad \Sigma x_i y_i = 131\,546.4 \quad \Sigma x_i^2 = 127\,492.9 \quad \Sigma y_i^2 = 138\,753.3$$

(2)

U SLJRH za 2001. na stranici 591. nalaze se podaci o prometu i zaposlenima za obrtnike u distributivnoj trgovini po županijama u 2000. godini. Analizira se ovisnost prometa (zavisna varijabla u milijardama kuna) o broju zaposlenih (nezavisna varijabla u tisućama). Primjenom programa EXCEL dobiveni su slijedeći rezultati:

| | coefficients | standard error | tStat |
|-----------|--------------|----------------|------------|
| intercept | 0.084678403 | | 1.13209116 |
| x | 0.299866365 | 0.050286293 | |

| ANOVA | | | | |
|------------|----|-------------|----|---|
| | df | SS | MS | F |
| regression | 1 | 0.773805154 | | |
| residual | | | | |
| total | 19 | 1.1655 | | |

(df = stupnjevi slobode, SS = zbroj kvadrata, MS = sredine kvadrata, F = empirijski F omjer, intercept = konstantni član, standard error = standardna pogreška procjene parametra, tStat = empirijski t omjer)

a) Napišite jednadžbu s procijenjenim parametrima i protumačite značenje procjena u konkretnom slučaju. b) Odredite nedostajuće elemente u navedenom ispisu obrade programskom potporom. c) Izračunajte vrijednost procjene standardne devijacije regresije. Protumačite njeno značenje. d) Koliki je koeficijent determinacije i linearne korelacije? Interpretirajte te veličine. e) Odredite granice 95% intervala pouzdanosti procjene parametra β . f) Testirajte hipotezu o značajnosti regresije. Razina signifikantnosti 5%. g) Broj zaposlenih u Zagrebačkoj županiji iznosio je 3 tisuće, a promet 0.8 milijardi kuna. Kolika je regresijska vrijednost prometa, rezidualno odstupanje i standardizirano rezidualno odstupanje?

(3)

Analizira se industrijska proizvodnja u Republici Hrvatskoj u razdoblju 1988 – 1999. Varijable u modelu višestruke linearne regresije su:

y – industrijska proizvodnja

x_1 – broj zaposlenih radnika u industriji

x_2 – produktivnost rada

Sve varijable izražene su kao indeksi na stalnoj bazi, 1995=100. Podaci o varijablama dani su u SLJRH 2000, na stranici 256.

Izračunana je slijedeća regresijska jednadžba:

$$\hat{y} = -116.1489 + 1.1204x_1 + 1.0100x_2$$

$$(8.7816) \quad (0.0291) \quad (0.0623)$$

a) Objasnite konkretno značenje regresijskih koeficijenata. b) Zbroj kvadrata odstupanja regresijskih vrijednosti zavisne varijable od prosječne vrijednosti te varijable iznosi 17 215.5339, a zbroj kvadrata odstupanja vrijednosti zavisne varijable od regresijskih vrijednosti zavisne varijable iznosi 104.2261. Sastavite tabelu ANOVA. c)

Koliki je koeficijent determinacije i linearne korelacije? d) Izračunajte vrijednost regresijske varijance, standardne devijacije i koeficijenta varijacije ($\bar{y}=130$). e) Odredite granice 95% intervala pouzdanosti parametra β_2 . f) Provedite jednosmjerni test o značajnosti prve regresorske varijable u modelu. Razina signifikantnosti 5%.

(4)

Analizira se promet za obrtnike u distributivnoj trgovini u 20 županija RH u 2000. godini. Varijable u modelu višestruke linearne regresije su:

y – promet u milijardama kuna, zavisna varijabla

x_1 – broj prodavaonica u tisućama

x_2 – broj zaposlenih u tisućama

Podaci s varijablama dani su u SLJRH 2001, na stranici 591. Primjenom programa EXCEL dobiveni su slijedeći rezultati:

| | coefficients | standard error | tStat |
|-----------|--------------|----------------|-------------|
| intercept | | 0.076644541 | 1.084130724 |
| x_1 | 0.156577844 | | 0.42818631 |
| x_2 | | 0.21042329 | 1.009881755 |

(intercept=konstantni član, standard error=standardna pogreška procjene parametra, tStat=empirijski t omjer)

- a) Napišite kako glasi procijenjena regresijska jednadžba. Objasnite konkretno značenje regresijskih parametara. b) Odredite nedostajuće elemente u navedenom ispisu obrade programskom potporom. c) Provedite test o značajnosti druge nezavisne varijable u modelu. Razina signifikantnosti 5%. Obavezno formulirajte hipoteze. d) Izračunajte granice 90% intervala procjene uz prvu regresorsku varijablu i objasnite njegovo značenje. e) Modelom je protumačeno 66.75% svih odstupanja, dok suma kvadrata odstupanja regresijskih vrijednosti od prosjeka iznosi 0.777984475. Koliko iznosi procjena standardne devijacije regresije? f) Odredite korigirani koeficijent determinacije.

(5)

- a) Prevezeni putnici u zračnom prijevozu u RH, u tisućama

| godina | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 |
|---------------------------|------|------|------|------|------|------|------|------|
| prevezeni putnici, u tis. | 238 | 507 | 661 | 679 | 824 | 866 | 920 | 926 |

Izvor: SLJRH 2000, str. 303

$$\sum y_t = 5621 \quad \sum x_t y_t = 29\,115 \quad \sum (y_t - \hat{y}_t)^2 = 47\,478.82$$

Kako glasi jednadžba linearnog trenda (uz jednadžbu navedite oznake)? Koliki su standardna devijacija i koeficijent varijacije trenda? Koliki je prema trendu očekivani broj prevezenih putnika u 2005. godini?

- b) Dinamika proizvodnje proizvoda A opisuje se jednadžbom jednostavnog eksponencijalnog trenda:

$$\hat{y} = 52.78 \cdot 1.53^x$$

$x=1, 30.06.1990.$

jedinica za x je 1 godina

jedinica za y je 1 tona

Odredite prognostičku vrijednost za 2003. godinu.

(6)

Vrijednost dionica (u USD) kompanije na burzi u razdoblju 1994 – 2000 iznosila je:

| godina | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 |
|------------|-------|-------|--------|--------|--------|--------|--------|
| vrijednost | 50.20 | 70.50 | 101.00 | 150.60 | 220.20 | 320.30 | 450.70 |

- a) Prikažite navedeni niz grafički. b) Izračunajte verižne indekse zadane serije, te na temelju njih zaključite o prikladnosti izbora eksponencijalnog trenda kao izraza tendencije kretanja vrijednosti dionica. Protumačite značenje indeksa vezanog za 2000. godinu. c) Napišite kako glasi jednadžba eksponencijalnog trenda vrijednosti dionica u standardnom obliku. Uz jednadžbu navedite sve potrebne oznake. Interpretirajte izračunane parametre. Pri računanju koristite slijedeće međurezultate:

$$\sum x_t \log y_t = 65.45 \quad \sum \log y_t = 15.23$$