Klasični linearni regresioni model: predviđanje

Zorica Mladenović

1

Elementi statističkog zaključivanja u KLRM: Predviđanje

2

Uvod

- Na osnovu ocenjenih parametara KLRM moguće je predvideti kretanje izabrane zavisne promenljive.
- Podaci vremenskih serija: prognoziranje se odnosi **na buduće vrednosti zavisne promenljive.**
- Podaci preseka: predviđa se vrednost zavisne promenljive za onu **vrednost objašnjavajuće promenljive koja nije uključena u uzorak.**

3

Polazne osnove

- Prema n podataka vremenskih serija za Y_i i X_i ocenjen je model

$$\hat{Y}_i = b_0 + bX_i,\ i = 1,2,\ldots,n.$$

- Neka je data vrednost X_i u periodu $n+1$: X_{n+1}. Predviđena vrednost zavisne promenljive u $n+1$ je:

$$\hat{Y}_{n+1} = b_0 + bX_{n+1}.$$

- Stvarna vrednost zavisne promenljive u $n+1$ je:

$$Y_{n+1} = \beta_0 + \beta X_{n+1} + \varepsilon_{n+1}$$

4

Greška predviđanja (prognoze)

- Razlika između prognozirane i stvarne vrednosti u trenutku $n+1$ je **greška predviđanja – oznaka *gp***:

$$gp = \hat{Y}_{n+1} - Y_{n+1}$$
$$= (b_0 - \beta_0) + (b - \beta) X_{n+1} - \varepsilon_{n+1}$$

- Ova greška je rezultat
  - razlike između ocenjenih i stvarnih vrednosti parametara nagiba i slobodnog člana i
  - slučajne greške koja nastaje za X_{n+1}.

5

Svojstva greške predviđanja

$$gp = \hat{Y}_{n+1} - Y_{n+1}$$
$$= (b_0 - \beta_0) + (b - \beta) X_{n+1} - \varepsilon_{n+1}$$

1. To je slučajna promenljiva sa normalnom raspodelom.
2. Njena očekivana vrednost je nula.
3. Njena varijansa (oznaka σ^2_{gp}):

$$\sigma^2_{gp} = \sigma^2 \left(1 + \frac{1}{n} + \frac{(X_{n+1} - \overline{X})^2}{\sum x_i^2} \right)$$

- Prema tome:

$$gp : N(0, \sigma^2_{gp})$$

6

Varijansa greške predviđanja

$$\sigma^2_{gp} = E\left(\hat{Y}_{n+1} - Y_{n+1}\right)^2$$
$$= E\left((b_0 - \beta_0) + (b - \beta)X_{n+1} - \varepsilon_{n+1}\right)^2$$

- Javlja se kao posledica varijabiliteta:

1. Ocene b_0
2. Ocene b
3. Slučajne greške ε_{n+1}

   i
4. Kovarijanse b_0 i b.

7

Formiranje intervalne ocene za predviđenu vrednost Y_{n+1}

- Standardizacija greške predviđanja daje:

$$gp : N(0, \sigma^2_{gp})$$

$$\frac{gp}{\sigma_{gp}} = \frac{\hat{Y}_{n+1} - Y_{n+1}}{\sigma_{gp}} : N(0, 1)$$

- Varijansa greške predviđanja nije poznata. Zamenjuje se ocenom varijanse greške predviđanja:

$$s^2_{gp} = s^2\left(1 + \frac{1}{n} + \frac{\left(X_{n+1} - \overline{X}\right)^2}{\sum x_i^2}\right)$$

- Kvadratni koren: standardna greška predviđanja – s_{gp}
- Uključivanjem s_{gp} u gornji količnik dobijamo:

8

Formiranje intervalne ocene za predviđenu vrednost Y_{n+1} (II)

$$\frac{gp}{s_{gp}} = \frac{\hat{Y}_{n+1} - Y_{n+1}}{s_{gp}} : t_{n-2} \Rightarrow$$

$$P\left(-t_{n-2}(0.025) \le \frac{\hat{Y}_{n+1} - Y_{n+1}}{s_{gp}} \le t_{n-2}(0.025) \right) = 0.95$$

$$P\left(\hat{Y}_{n+1} - s_{gp} t_{n-2}(0.025) \le Y_{n+1} \le \hat{Y}_{n+1} + s_{gp} t_{n-2}(0.025) \right) = 0.95$$

Uz verovatnocu 95% ocekujemo da se

stvarna vrednost Y_{n+1} nalazi u intervalu :

$$\left(\hat{Y}_{n+1} \pm s_{gp} t_{n-2}(0.025) \right)$$

9

Intervalna ocena predviđene vrednosti i standardna greška predviđanja

- Intervalna ocena **primarno zavisi od standardne greške** (ocene varijanse) predviđanja.
  - Intervalna ocena je uža, odnosno preciznija za manju vrednost standardne greške predviđanja.
  - Intervalna ocena je šira, odnosno nepreciznija za veću vrednost standardne greške predviđanja

10

Intervalna ocena predviđene vrednosti i standardna greška predviđanja II

- Standardna greška predviđanja zavisi od:
  - varijabiliteta slučajne greške modela
  - obima uzorka
  - varijabiliteta objašnjavajuće promenljive
  - razlike između vrednosti objašnjavajuće promenljive za koju se predviđa i aritmetičke sredine polaznih vrednosti objašnjavajuće promenljive.

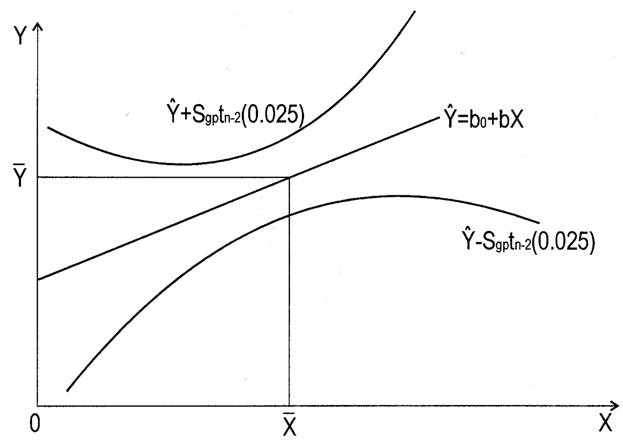
11

Intervalna ocena predviđene vrednosti i standardna greška predviđanja III

- Standardna greška predviđanja je **manja** ukoliko je:
  - **manji varijabilitet** slučajne greške modela
  - **veći** obim uzorka
  - **veći** varijabilitet objašnjavajuće promenljive
  - **manje odstupanje** vrednosti objašnjavajuće promenljive od aritmetičke sredine polaznih vrednosti objašnjavajuće promenljive.
    - Intervalna ocena predviđanja uzima najnižu vrednost za $X_{n+1} = \overline{X}$

12

Grafički prikaz intervalne ocene predviđene vrednosti Y_{n+1}



13

Ilustracija, zadatak 13, glava 1, Zbirka

Tražnja za novcem (mil. dol.)	36	50	46	30	20	35	37	61
Kamatna stopa (%)	6.3	4.6	5.1	7.3	8.9	5.3	6.7	3.5

Prema podacima ocenjen je model:

$$\hat{Y}_i = \underset{(5.94)}{81.60} - \underset{(0.96)}{7.08}\, X_i$$

Predvideti nivo tražnje za novcem u ekonomiji čija je kamatna stopa 8.1% i obrazovati interval poverenja predviđanja uz verovatnoću 95%.

14

Ilustracija (II)

1. Prognozira na vrednost

za $X_{n+1} = 8.1$:

$$\hat{Y}_{n+1} = 81.60 - 7.08 \cdot 8.1 = 24.25$$

2. Ocena varijanse

greške predvidjanja

$$s_{gp}^2 = s^2\left(1 + \frac{1}{n} + \frac{\left(X_{n+1} - \overline{X}\right)^2}{\sum x_i^2}\right) = 25.245 \Rightarrow s_{gp} \approx 5.$$

$$s^2 = \frac{\sum e_i^2}{n-2} = \frac{112.0761}{6} = 18.679$$

$$\overline{X} = 5.9625$$

$$\sum x_i^2 = 20.17875$$

15

Ilustracija (III)

3. Obrazovanje intervalne ocene

prognoze uz $p = 0.95$

$$\left(\hat{Y}_{n+1} \pm s_{gp} t_{n-2}(0.025)\right)$$

$$\left(24.25 \pm 5 \cdot 2.45\right)$$

$$\left(12.0, 36.5\right)$$

16