

5. Napovedovanje na podlagi regresijskega modela

doc. dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

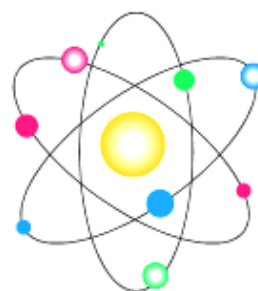


Ljubljana, februar 2014

Motivacija

New York State Criminal Code, Section 899:
**“Za osebo, ki se pretvarja, da napoveduje prihodnost,
je predvidena kazen 250 \$ in/ali 6 mesecev zapora...”**

A. H. Studenmund, H. J. Cassidy:
Using Econometrics, str. 369



**“Napovedovanje je kot poskus vožnje avtomobila
z zavezanimi očmi ob sledenju navodil osebe,
ki gleda skozi zadnje okno.”**

Anonimen avtor

Temeljni pojmi

NAPOVEDOVANJE OZ. PREDVIDEVANJE
vrednosti odvisne spremenljivke
na podlagi ocenjenega regresijskega modela

POGOJNO NAPOVEDOVANJE
(Conditional forecast)

Pri tem se moramo zavedati:



Slučajnostne narave modela.



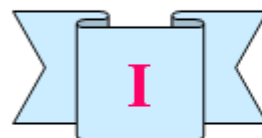
Dejstva, da so regresijski koeficienti le ocene pravih vrednosti.



Možnosti napak pri specifikaciji modela.

Temeljni pojmi

PREDPOSTAVKI:



$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$E(\mathbf{y}|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}$$

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$$



$$\mathbf{x}_0^T = (1, x_{20}, x_{30}, \dots, x_{k0})$$

Napoved povprečne vrednosti

A NAPOVED POVPREČNE VREDNOSTI (Mean prediction)

I Napoved (točkovna napoved) (Point estimate)

$$E(y_0 | \mathbf{x}_0) = \mathbf{x}_0^T \boldsymbol{\beta}$$

$$\hat{y}_0 = \mathbf{x}_0^T \mathbf{b}$$

$$\begin{aligned} E(\hat{y}_0 - E(y_0 | \mathbf{x}_0)) &= E(\mathbf{x}_0^T (\mathbf{b} - \boldsymbol{\beta})) = \\ &= \mathbf{x}_0^T E(\mathbf{b} - \boldsymbol{\beta}) = 0 \end{aligned}$$

Napoved je nepristranska!

Napoved povprečne vrednosti

II Zanesljivost (kvaliteta) napovedi - Intervalna napoved (Confidence interval)

$$\begin{aligned}\text{var}(\hat{y}_0) &= E \left[(\mathbf{x}_0^T \mathbf{b} - \mathbf{x}_0^T \boldsymbol{\beta})^2 \right] = \\ &= \mathbf{x}_0^T (\text{var} - \text{cov}(\mathbf{b})) \mathbf{x}_0 = \\ &= \mathbf{x}_0^T (s_e^2 (\mathbf{X}^T \mathbf{X})^{-1}) \mathbf{x}_0 \\ \text{var}(\hat{y}_0) &= s_e^2 (\mathbf{x}_0^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_0)\end{aligned}$$

$$se(\hat{y}_0) = s_e \sqrt{\mathbf{x}_0^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_0}$$

$$t = \frac{\hat{y}_0 - E(y_0 | \mathbf{x}_0)}{se(\hat{y}_0)}$$

$$\Pr(\hat{y}_0 - t_{\alpha/2} se(\hat{y}_0) \leq E(y_0 | \mathbf{x}_0) \leq \hat{y}_0 + t_{\alpha/2} se(\hat{y}_0)) = 1 - \alpha$$

Napoved posamezne vrednosti

B NAPOVED POSAMEZNE VREDNOSTI (Individual prediction)

I Napoved (točkovna napoved) (Point estimate)

$$y_0 = \mathbf{x}_0^T \boldsymbol{\beta} + u_0$$

$$\hat{y}_0 = \mathbf{x}_0^T \mathbf{b}$$

$$\begin{aligned} E(\hat{y}_0 - y_0) &= E(\mathbf{x}_0^T (\mathbf{b} - \boldsymbol{\beta})) - E(u_0) = \\ &= \mathbf{x}_0^T E(\mathbf{b} - \boldsymbol{\beta}) - E(u_0) = 0 \end{aligned}$$

Napoved je nepristranska!

Napoved posamezne vrednosti

II Zanesljivost (kvaliteta) napovedi - Intervalna napoved (Prediction interval)

$$\begin{aligned}\text{var}(\hat{y}_0 - y_0) &= E\left[\left((\hat{y}_0 - y_0) - E(\hat{y}_0 - y_0)\right)^2\right] = \\ &= E\left[\left(\mathbf{x}_0^T(\mathbf{b} - \boldsymbol{\beta}) - u_0\right)^2\right] = \mathbf{x}_0^T(\text{var} - \text{cov}(\mathbf{b}))\mathbf{x}_0 + \sigma^2 = \\ &= \mathbf{x}_0^T(s_e^2(\mathbf{X}^T\mathbf{X})^{-1})\mathbf{x}_0 + \sigma^2 \\ \text{var}(\hat{y}_0 - y_0) &= s_e^2\left[\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 + 1\right]\end{aligned}$$

$$se(\hat{y}_0 - y_0) = s_e \sqrt{\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 + 1}$$

$$\Pr(\hat{y}_0 - t_{\alpha/2} se(\hat{y}_0 - y_0) \leq y_0 \leq \hat{y}_0 + t_{\alpha/2} se(\hat{y}_0 - y_0)) = 1 - \alpha$$

Standardna napaka ocene napovedi

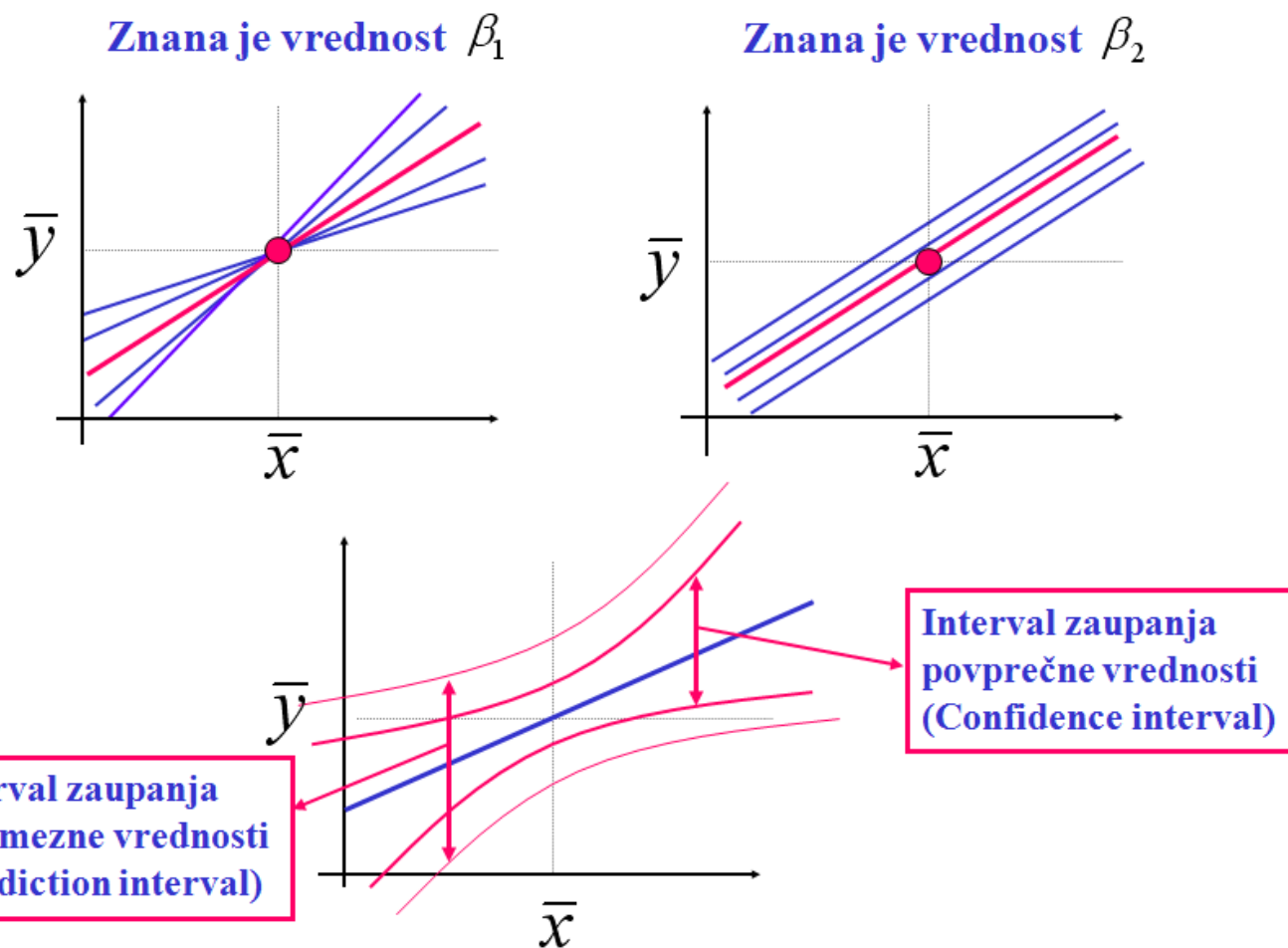
Standardna napaka ocene napovedi povprečne vrednosti odvisne spremenljivke za bivariatni regresijski model:

$$se(\hat{y}_0) = s_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x - \bar{x})^2}}$$

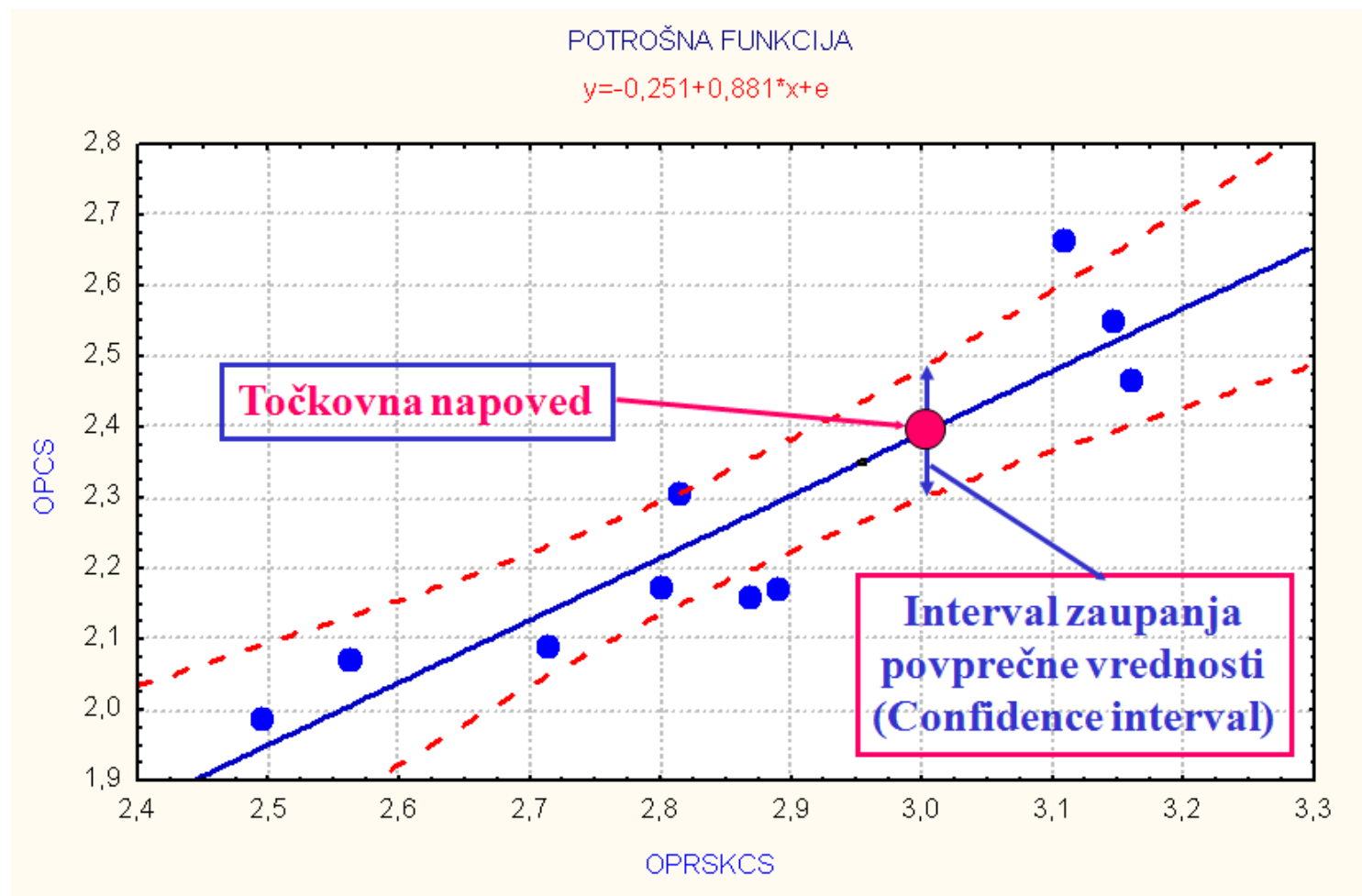
Odvisna je od:

- Standardne napake ocene regresije
- Velikosti vzorca
- Variance pojasnjevalne spremenljivke
- Razlike vrednosti pojasnjevalne spremenljivke x_0 od njene povprečne vrednosti \bar{x}

Napovedi in intervali zaupanja



Napovedi in intervali zaupanja



Uporaba nepravih spremenljivk

Izračun točkovne napovedi in standardne napake napovedi z uporabo nepravih spremenljivk

A Vpeljemo nepravo spremenljivko, ki ima pri vseh n opazovanih enotah vrednost 0.

B Opazovanim enotam dodamo eno enoto, pri kateri imajo spremenljivke naslednje vrednosti:

$n + 1$:	y	x_1	x_2, \dots, x_k	$x_{k+1} = D$
Vrednosti:	0	1	x_{20}, \dots, x_{k0}	1

C Ocenimo naslednji razširjeni model:

$$\hat{y} = b_1 + b_2 x_2 + \dots + b_k x_k + b_{k+1} D$$

Uporaba nepravih spremenljivk

Pri tem velja:

- ★ **Vrednosti regresijskih koeficientov b_1, \dots, b_k ostajajo nespremenjene.**
- ★ $b_{k+1} = -\hat{y}_0$; **točkovna napoved z nasprotnim predznakom.**
- ★ $se(b_{k+1}) = se(\hat{y}_0 - y_0)$; **standardna napaka napovedi posamezne vrednosti.**

V razmislek...



“Če bi pri izdelavi napovedi imel izbiro med Bobom Solowom in ekonometričnim modelom, bi izbral Boba Solowa; vendar bi imel raje Boba Solowa z ekonometričnim modelom, kot pa brez njega.”

Paul A. Samuelson

5. Napovedovanje na podlagi regresijskega modela

doc. dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si



Ljubljana, februar 2014