

Klasični linearni regresioni model

Brooks, *Introductory econometrics for finance*, 2002, CUP
modifikacije: Zorica Mladenović
Mladenović i Petrović, *Uvod u ekonometriju*, 2007/10, EF

1

Struktura predavanja

- Jednostavna regresiona analiza i metod običnih najmanjih kvadrata (metod ONK)
- Klasični jednostavni linearni regresioni model (KLRM)
- Svojstva ocena dobijenih primenom metoda ONK u KLRM
- Statističko zaključivanje u KLRM
- Klasični višestruki linearni regresioni model

2

- **Jednostavna regresiona analiza i metod
običnih najmanjih kvadrata (metod ONK)**

3

1. Jednostavna regresiona analiza

- Regresiona analiza predstavlja osnovni metodološki okvir ekonometrijskog modeliranja.
- Pretpostavimo da raspolažemo podacima o inflaciji i deprecijaciji deviznog kursa za određeni period vremena i da želimo da otkrijemo prirodu njihove međusobne povezanosti.
- Cilj regresione analize jeste utvrđivanje prirode i forme povezanosti između promenljivih.

4

Oznake

- Razlikujemo dva tipa promenljivih:
 - Zavisna promenljiva: Y
 - Nezavisne promenljive: $X_1, X_2, \dots, X_k$ (ukupno k).
- Alternativni nazivi za Y i X :

Y	X
zavisna promenljiva	nezavisna promenljiva
regresant	regresor
	objašnjavajuća promenljiva
	ekspanatorna promenljiva
- U ovom trenutku fokusiramo se na situaciju kada postoji samo jedna objašnjavajuća promenljiva – **jednostavna regresiona analiza.**

Razlika između regresione i korelacione analize

- Ako kažemo da su Y i X korelisane promenljive, to znači da ih tretiramo na simetričan način. Ne insistiramo na pravcu uzročnosti.
- U regresionoj analizi zavisna (Y) i nezavisna (X) promenljiva imaju potpuno različitu poziciju.
 - Promenljiva Y je stohastičkog tipa, što znači da je slučajna promenljiva koju karakteriše određena raspodela.
 - Promenljiva X uzima fiksirane vrednosti iz ponovljenih uzoraka. Ona nije stohastičke prirode.
 - Postoji jednosmeran pravac uzročnosti: samo X utiče na Y , dok Y ne utiče na X .

6

Primeri jednostavnih regresija

- Da li je inflacija isključivo određena deprecijacijom deviznog kursa?
- Da li nivo izvoza zavisi od nivoa industrijske proizvodnje?
- U kojoj meri je nivo izvoza determinisan nivoom realnog deviznog kursa?
- Da li se tražnja za novcem može objasniti na osnovu deprecijacije deviznog kursa?

7

Primena jednostavne regresije

- Posmatramo mesečne podatke (u %) o inflaciji i deprecijaciji deviznog kursa privrede Srbije za prvih sedam meseci 2004. godine:

MESEC	1	2	3	4	5	6	7
Inflacija	0.40	0.90	0.40	0.80	1.00	1.39	1.39
Deprecijacija	0.45	1.20	0.86	1.09	1.31	1.29	1.29

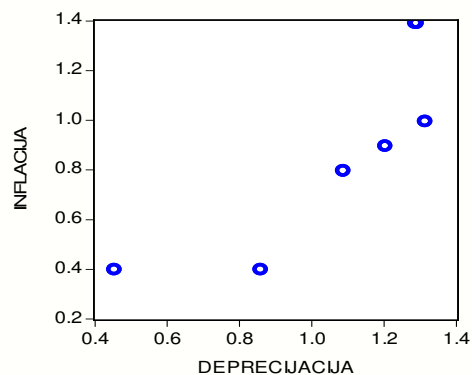
8

Primena jednostavne regresije (II)

- Pretpostavljamo da je veza između inflacije i deprecijacije deviznog kursa pozitivna. Hoćemo da opišemo inflaciju kao funkciju od deprecijacije deviznog kursa:
 - Inflacija: zavisna promenljiva (Y)
 - Deprecijacija: nezavisna promenljiva (X)
- Prvi korak: grafički prikaz parova podataka (X_t, Y_t) , $t=1,2,3,4,5,6,7$.
- Parovi: $(0.45, 0.40), \dots, (1.29, 1.39)$.

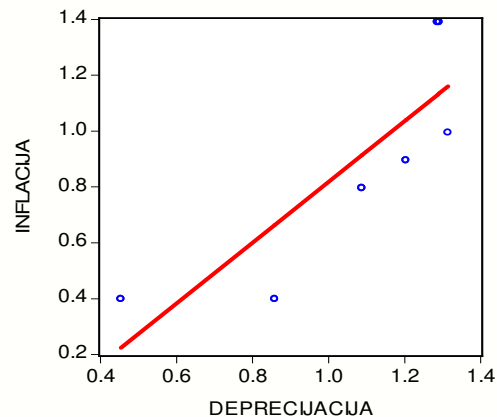
9

Grafički prikaz: dijagram rasturanja (raspršenosti) tačaka



10

Dijagram rasturanja tačaka sa pravom linijom



11

Postavljanje prave

- Namera nam je da postavimo pravu tako da najbolje aproksimira skup podataka.
- Postaviti pravu znači odrediti njene parametre:
$$Y=a+bX$$
- Jednačina $Y=a+bX$ je deterministička: za dati nivo deprecijacije uvek znamo nivo inflacije.
- Da li je to realno? Ne. Zato dodajemo slučajan član u u jednačinu:

$$Y_t = \alpha + \beta X_t + u_t$$

i $t = 1, 2, 3, 4, 5, 6, 7$.

12

Zašto uključujemo slučajan član?

- Postoje faktori čiji su pojedinačni uticaji na kretanje izabrane zavisne promenljive sporadični i neregularni. Njihovo uključivanje u model kao objašnjavajućih promenljivih bi bilo nepotrebno. Umesto toga, možemo smatrati da slučajna greška sadrži njihovo zbirno dejstvo.
- Čak i kada bi sve relevantne promenljive bile obuhvaćene analizom, postojanje slučajne greške može se objasniti slučajnošću - nepredvidivošću ljudskog ponašanja.
- Slučajna greška može izraziti i grešku u merenju promenljivih (zaokruživanje i sl.).

13

Osnovni skup (populacija) i uzorak

- **Osnovni skup** je skup svih jedinica posmatranja.
- **Uzorak je podskup osnovnog skupa.**
 - Uzorak je slučajan ako svaka jedinica osnovnog skupa ima jednaku verovatnoću da bude izvučena kao element uzorka.
 - To što je neka od jedinica osnovnog skupa postala deo uzorka ne menja verovatnoću da druga jedinica bude izvučena kao element uzorka.
- Definicija uzorka: skup od T nezavisnih i jednako raspodeljenih slučajnih promenljivih.

14

Populaciona i uzoračka regresiona prava (jednačina)

- Populaciona regresiona prava označava stvarnu stohastičku vezu između datih promenljivih (sadrži stvarno α i β):

$$Y_t = \alpha + \beta X_t + u_t$$

- Uzoračka regresiona prava opisuje vezu prema datom uzorku:

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t$$

- Stvarni nivo zavisne promenljive je zbir ocenjenog nivoa i onoga što model nije ocenio (reziduala)

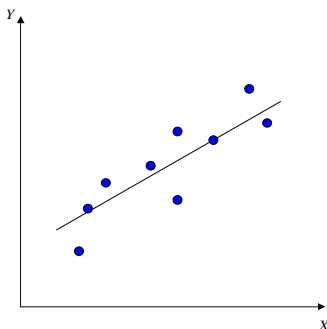
$$Y_t = \hat{Y}_t + \hat{u}_t$$

- Uzoračka regresiona prava (jednačina) se koristi za donošenje zaključaka o parametrima populacione regresione jednačine.

15

Određivanje pozicije prave (regresionih koeficijenata)

- Kako određujemo vrednosti $\hat{\alpha}$ i $\hat{\beta}$?
- Kriterijum: biramo $\hat{\alpha}$ i $\hat{\beta}$ tako da je prava najmanje moguće udaljena od tačaka dijagrama rasturanja
- *Drugim rečima: da je odstupanje prave od tačaka minimalno*



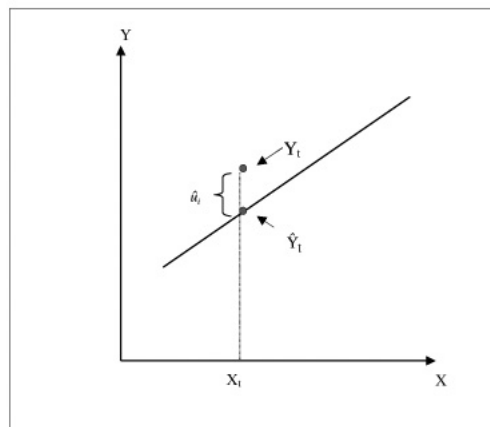
16

1.1 Metod običnih najmanjih kvadrata (metod ONK)

- Najčešće korišćen metod postavljanja prave i izbora regresionih koeficijenata jeste metod običnih najmanjih kvadrata (ONK).
- Ideja metoda: minimizirati zbir kvadrata odstupanja podataka od prave.
- Oznake:
 - Y_t - stvarna vrednost u trenutku t
 - $\hat{Y}_t$ - vrednost Y_t koja je ocenjena regresionom pravom
 - $\hat{u}_t$ - razlika stvarne i ocenjene vrednosti, rezidual, $Y_t - \hat{Y}_t$

17

Stvarna i ocenjena vrednost zavisne promenljive



18

Izvođenje ONK ocena

- Potrebno je minimizirati zbir (tzv. rezidualnu sumu kvadrata):

$$\hat{u}_1^2 + \hat{u}_2^2 + \dots + \hat{u}_7^2 = \sum_{t=1}^7 \hat{u}_t^2$$

- Šta je $\hat{u}_t$? To je razlika $\hat{u}_t = Y_t - \hat{Y}_t$

- Naći minimum $\sum_{t=1}^7 \hat{u}_t^2$
je ekvivalentno određivanju minimuma $\sum_{t=1}^7 (Y_t - \hat{Y}_t)^2$

19

Izvođenje ocena metoda ONK (II)

- Kako je $\hat{Y}_t = \hat{\alpha} + \hat{\beta}X_t$, možemo definisati funkciju

$$L(\hat{\alpha}, \hat{\beta}) = \sum_{t=1}^7 (Y_t - \hat{Y}_t)^2 = \sum_{t=1}^7 (Y_t - \hat{\alpha} - \hat{\beta}X_t)^2$$

- Potrebno je minimizirati funkciju L u odnosu na $\hat{\alpha}$ i $\hat{\beta}$:

$$\frac{\partial L}{\partial \hat{\alpha}} = -2 \sum_t (Y_t - \hat{\alpha} - \hat{\beta}X_t) = 0 \quad (1)$$

$$\frac{\partial L}{\partial \hat{\beta}} = -2 \sum_t X_t (Y_t - \hat{\alpha} - \hat{\beta}X_t) = 0 \quad (2)$$

- Iz (1): $\sum_t (Y_t - \hat{\alpha} - \hat{\beta}X_t) = 0 \Rightarrow \sum_t Y_t - T\hat{\alpha} - \hat{\beta}\sum_t X_t = 0$.

$$\frac{\sum_{t=1}^T Y_t}{T} = \bar{Y} \Rightarrow \sum_{t=1}^T Y_t = T\bar{Y}, \quad \sum_{t=1}^T X_t = T\bar{X}$$

20

Izvođenje ONK ocena (III)

- Možemo pisati: $T\bar{Y} - T\hat{\alpha} - \hat{\beta}T\bar{X} = 0 \Rightarrow \bar{Y} - \hat{\alpha} - \hat{\beta}\bar{X} = 0$ (3)

- Iz (2): $\sum_t X_t (Y_t - \hat{\alpha} - \hat{\beta}X_t) = 0$ (4)

- Iz (3): $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$ (5)

- Zamenom u (4) za $\hat{\alpha}$ iz (5):

$$\sum_t X_t (Y_t - \bar{Y} + \hat{\beta}\bar{X} - \hat{\beta}X_t) = 0$$

$$\sum_t X_t Y_t - \bar{Y} \sum_t X_t + \hat{\beta}\bar{X} \sum_t X_t - \hat{\beta} \sum_t X_t^2 = 0$$

$$\sum_t X_t Y_t - T\bar{Y}\bar{X} + \hat{\beta}T\bar{X}^2 - \hat{\beta} \sum_t X_t^2 = 0$$

21

Izvođenje ONK ocena (IV)

- Rešavanjem po $\hat{\beta}$ dobijamo:

$$\hat{\beta}(T\bar{X}^2 - \sum_t X_t^2) = T\bar{Y}\bar{X} - \sum_t X_t Y_t$$

- Dakle, ocene su:

$$\hat{\beta} = \frac{\sum_t X_t Y_t - T\bar{X}\bar{Y}}{\sum_t X_t^2 - T\bar{X}^2} \text{ i } \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

- Ocena $\hat{\beta}$ se može dobiti i na nešto drugačiji način:

22

Izvođenje ONK ocena (V)

$$\hat{\beta} = \frac{\sum_{t=1}^T X_t Y_t - T \bar{X} \bar{Y}}{\sum_{t=1}^T X_t^2 - T \bar{X}^2} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{\sum_{t=1}^T (X_t - \bar{X}) Y_t}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

$$\begin{aligned} \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y}) &= \sum_{t=1}^T X_t Y_t - \sum_{t=1}^T X_t \bar{Y} - \sum_{t=1}^T \bar{X} Y_t + T \bar{X} \bar{Y} \\ &= \sum_{t=1}^T X_t Y_t - T \bar{X} \bar{Y} - T \bar{X} \bar{Y} + T \bar{X} \bar{Y} \\ &= \sum_{t=1}^T X_t Y_t - T \bar{X} \bar{Y}. \end{aligned}$$

23

Izvođenje ONK ocena (VI)

$$\hat{\beta} = \frac{\sum_{t=1}^T X_t Y_t - T \bar{X} \bar{Y}}{\sum_{t=1}^T X_t^2 - T \bar{X}^2} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{\sum_{t=1}^T (X_t - \bar{X}) Y_t}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

$$\begin{aligned} \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y}) &= \sum_{t=1}^T (X_t - \bar{X}) Y_t - \sum_{t=1}^T (X_t - \bar{X}) \bar{Y} \\ &= \sum_{t=1}^T (X_t - \bar{X}) Y_t - \bar{Y} \sum_{t=1}^T (X_t - \bar{X}) \\ &= \sum_{t=1}^T (X_t - \bar{X}) Y_t. \end{aligned}$$

$$\sum_{t=1}^T (X_t - \bar{X}) = \sum_{t=1}^T X_t - T \bar{X} = T \bar{X} - T \bar{X} = 0.$$

24

Izvođenje ONK ocena (VII)

- Drugi parcijalni izvodi su uvek pozitivni:

$$\frac{\partial^2 L}{\partial \hat{\alpha}^2} = 2T, \frac{\partial^2 L}{\partial \hat{\beta}^2} = 2 \sum_{t=1}^T X_t^2, \frac{\partial^2 L}{\partial \hat{\alpha} \partial \hat{\beta}} = 2 \sum_{t=1}^T X_t$$

25

Kakva je interpretacija $\hat{\alpha}$ i $\hat{\beta}$?

$$\hat{\beta} = \frac{\sum_{t=1}^T X_t Y_t - T \bar{X} \bar{Y}}{\sum_{t=1}^T X_t^2 - T \bar{X}^2} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{0.6526}{0.6002} = 1.087$$

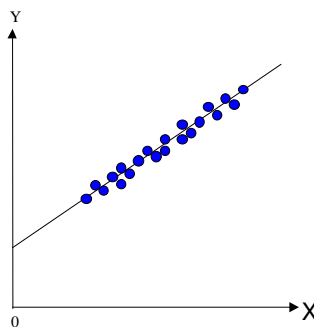
$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} = 0.8971 - 1.087 \cdot 1.0700 = -0.266$$

- $\hat{\beta}$ označava prirast zavisne promenljive po jedinici prirasta nezavisne.
- $\hat{\alpha}$ označava nivo zavisne promenljive kada je nivo objašnjavajuće nula.
- Ako dinar deprecira za 1 jedinicu (a pošto su jedinice procenti, za 1 procentni poen), tada će cene porasti za 1.09 procentnih poena.
- Ukoliko je stopa rasta kursa 0, tada je inflacija -0.27%.

26

Preciznost ocene slobodnog člana

- Ocena slobodnog člana predstavlja očekivanu vrednost Y kada je X jednako nuli.
- Treba biti pažljiv prilikom njene interpretacije posebno kada nema dovoljno podataka koji su blizu y-ose.
- Ukoliko je takvih podataka malo onda je ova ocena neprecizna.



27

1.2 Linearnost

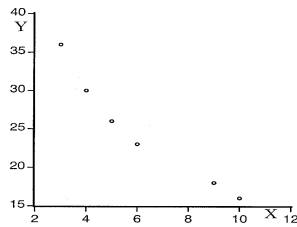
- Primena metoda ONK zahteva da model bude linearan, što znači da parametri modela figurišu na linearan način (α i β). Model ne mora da bude linearan po promenljivima (Y i X).
- Postoje jednostavne nelinearne forme koje su od interesa u ekonomskim analizama, a koje se jednostavnim transformacijama mogu prevesti na linearne.
- Primer 1: Dvojno-logaritamski model (log-log model)

$$Y_t = \beta_0 X_t^\beta e^{u_t} \Leftrightarrow \underbrace{\ln Y_t}_{Y_t^*} = \underbrace{\ln \beta_0}_{\alpha} + \underbrace{\beta \ln X_t}_{X_t^*} + u_t$$

$$Y_t^* = \alpha + \beta X_t^* + u_t.$$

28

Log-log (dvojno-logaritamski) model



- Y : 16, 18, 23, 18, 26, 30, 36
- X : 10, 9, 6, 9, 5, 4, 3
- Y – tražnja, X - cena
- β je proporcionalna promena Y (%) koja je rezultat proporcionalne promene X (%).
- Ako se X promeni za 1%, Y će se promeniti za $\beta\%$.
- β je elastičnost Y u odnosu na X .
- Ocena: -0.65, cenovna elastičnost tražnje

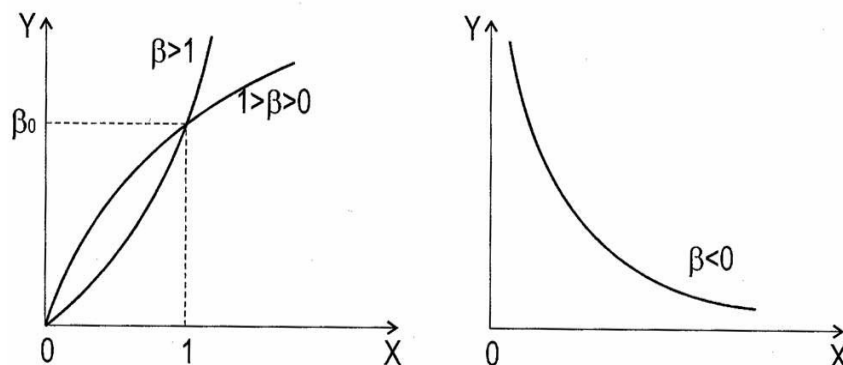
$$Y_t = \beta_0 X_t^\beta \Rightarrow$$

$$\frac{\partial Y_t}{\partial X_t} = \beta_0 \beta X_t^{\beta-1} = \frac{\beta_0 \beta X_t^\beta}{X_t} = \frac{\beta Y_t}{X_t}$$

$$\Rightarrow \beta = \frac{\partial Y_t}{\partial X_t} \frac{X_t}{Y_t} = \frac{\partial Y_t / Y_t}{\partial X_t / X_t}$$

29

Log-log model (II)



30

Razlika između linearnog i nelinearnog modela

- Primer II: inverzni model:

$$Y_t = \beta_0 + \frac{\beta}{X_t} + u_t$$

se jednostavno ocenjuje primenom metoda ONK kada se redefiniše objašnjavajuća promenljiva na sledeći način:

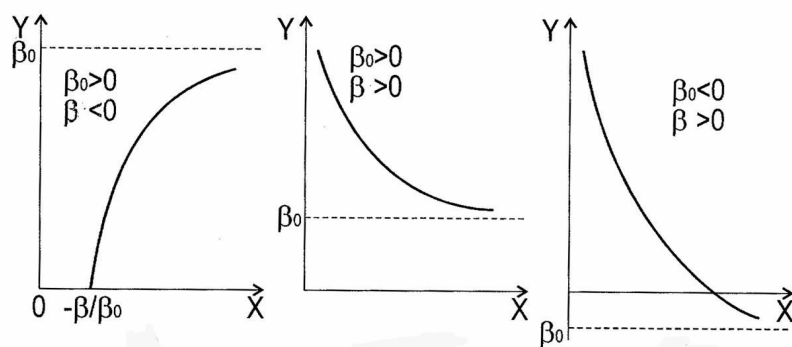
$$Z_t = \frac{1}{X_t}$$

- Primer III: stvarni nelinearni model:

$$Y_t = \alpha + X_t^\beta + u_t$$

31

Još malo o inverznom modelu



$$Y_t = \beta_0 + \frac{\beta}{X_t} + u_t, \frac{\partial Y_t}{\partial X_t} = -\frac{\beta}{X_t^2} \Rightarrow \partial Y_t = -\frac{\beta}{X_t} \frac{\partial X_t}{X_t}.$$

32

Inverzni model - primeri

- Levi grafik: Engelova kriva potrošnje
 - (Y – potrošnja određenog proizvoda, X – dohodak)
 - Ispod izvesnog nivoa dohotka potrošnja nije moguća.
 - Postoji saturacioni nivo potrošnje: nezavisno od nivoa dohotka potrošnja se ne ostvaruje iznad gornjeg praga.
- Desni grafik: Filipsova kriva
 - (Y – stopa rasta plata, X – stopa nezaposlenosti)
 - Postoji asimetrična reakcija plata na promenu nezaposlenosti na različitim nivoima nezaposlenosti.
 - Ako je nivo nezaposlenosti ispod prirodne stope (presek krive sa x-osom), tada jedinična promena nezaposlenosti dovodi do snažnije reakcije plata nego kada je nezaposlenost iznad prirodnog nivoa (X veće od tačke preseka krive sa x-osom)

Primer nelinearnog modela

$$Y_t = \alpha + X_t \beta + u_t$$

- U uslovima hiperinflacije ponuda novca je ogromna. Međutim, tražnja za novcem drastično opada sa rastom cena, odnosno padom vrednosti domaće valute. Realni novac je funkcija inflacije ili deprecijacije kursa.
- Linearni model ne može dovoljno dobro da obuhvati pad realnog novca koji je funkcija rasta nominalnog kursa. Može se koristiti model oblika:

$$md = f(dep), \quad f \text{ je monotono opadajuća funkcija}$$

$$md = \alpha_0 + \alpha_1 dep^\pi, \quad \alpha_1 < 0$$

Reakcija md (tražnje za novcem) se opisuje parametrom koji opada (po apsolutnoj vrednosti) sa rastom kursa (dep) za $\pi < 1$:

$$\alpha_1 \pi (dep)^{\pi-1}$$

34

Primer nelinearnog modela (II)

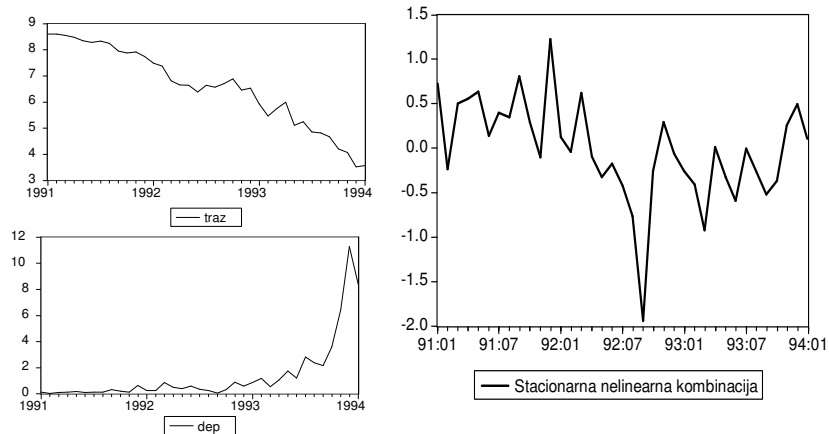
- Za celi period YU hiperinflacije (1991 – 1994) utvrdili smo da su tražnja za novcem i stopa deprecijacije povezane na dugi rok.
- Ocenjena je sledeća zavisnost (ali ne primenom metoda ONK):

$$md = 14.2 - 8.15dep^{0.13}$$

- Izvor: Petrović and Mladenović (2000) , *Journal of Money, Credit and Banking*.

35

Primer nelinearnog modela (III)



36

- **Klasični jednostavni linearni regresioni model (KLRM)**

37

2. Pretpostavke klasičnog linearnog regresionog modela (KLRM)

- Model koji smo do sada razmatrali je deo specifikacije koja se naziva klasični linearni regresioni model.
- Budući da Y_t zavisi od slučajne greške u_t , potrebno je da definišemo pretpostavke kojima se opisuju svojstva u_t .
- Pretpostavke KLRM o u_t :
- | <u>Notacija</u> | <u>Interpretacija</u> |
|--|---|
| 1. $E(u_t) = 0$ | Očekivana vrednost sl. greške je nula. |
| 2. $\text{Var}(u_t) = \sigma^2$ za svako t | Varijansa sl. greške je konstantna za sve vrednosti X_t (greške su homoskedastične) |
| 3. $\text{Cov}(u_t, u_j) = 0, i \neq j$. | Sl. greške su međusobno nekorelisane (ne postoji autokorelacija) |
| 4. $\text{Cov}(u_t, X_t) = 0$ za svako t | Ne postoji korelacija između sl. greške i objašnjavajuće promenljive X_t |

38

Pretpostavke KLRM (II)

- Alternativna interpretacija uslova 4.: promenljive X_t nisu slučajne, već uzimaju fiksirane vrednosti iz ponovljenih uzoraka.
- Postoji i peta pretpostavka koja je neophodna u postupku statističkog zaključivanja, odnosno testiranja vrednosti o parametrima populacione regresione jednačine α i β .
- Pretpostavka 5.
5. u_t je normalno raspodeljena slučajna promenljiva.
- Imajući u vidu prethodne pretpostavke o nultoj srednjoj vrednosti i stabilnoj varijansi:

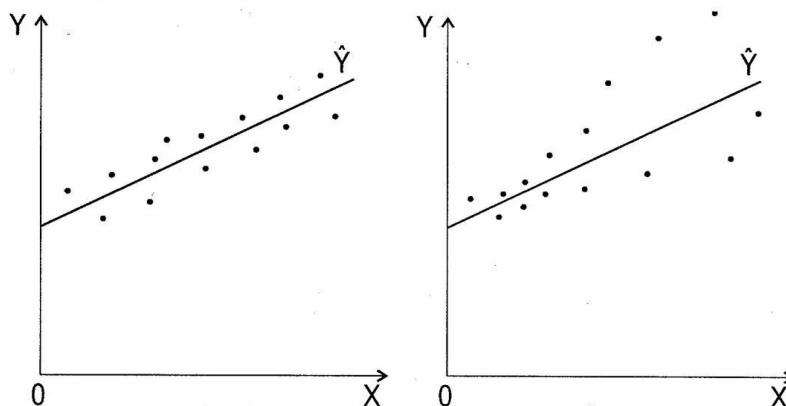
$$u_t \sim N(0, \sigma^2)$$

39

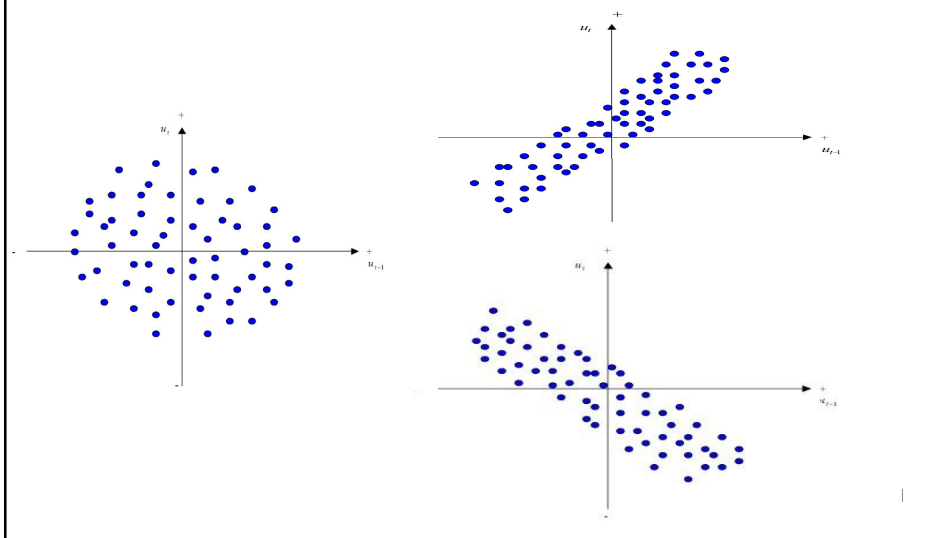
Pretpostavka 2.

Levi grafik: homoskedastičnost

Desni grafik: heteroskedastičnost



Pretpostavka 3.
(nema autokorelacije, pozitivna
i negativna autokorelacija)



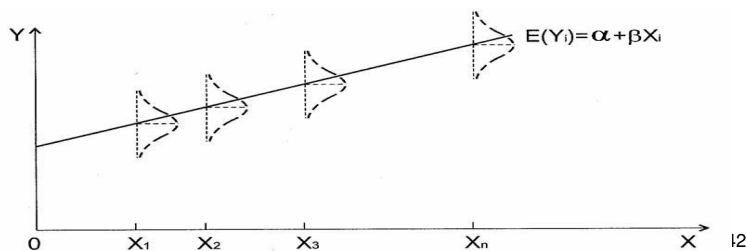
Pretpostavke 1., 2. i 5.:
i Y ima normalnu raspodelu

$$E(u_t) = 0 \Rightarrow E(Y_t) = E(\alpha + \beta X_t + u_t) = \alpha + \beta X_t.$$

$$\text{var}(u_t) = \sigma^2 \Rightarrow \text{var}(Y_t) = E(Y_t - E(Y_t))^2$$

$$\text{var}(Y_t) = E(\alpha + \beta X_t + u_t - \alpha - \beta X_t)^2 = E(u_t)^2 = \sigma^2.$$

$$u_t \sim N(0, \sigma^2) \Rightarrow Y_t \sim N(\alpha + \beta X_t, \sigma^2).$$



Implikacije navedenih pretpostavki

$$u_t \sim N(0, \sigma^2), Y_t \sim N(\alpha + \beta X_t, \sigma^2),$$

$$\hat{\beta} = \frac{\sum_{t=1}^T (X_t - \bar{X}) Y_t}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{\sum_{t=1}^T w_t Y_t}{\sum_{t=1}^T w_t^2}, \quad w_t = \frac{(X_t - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

- Ocena $\hat{\beta}$ je linearna funkcija slučajne promenljive Y_t
- Posledice:
 - Ocena $\hat{\beta}$ je slučajna promenljiva
 - Ocena $\hat{\beta}$ ima normalnu raspodelu.

43

- **Svojstva ocena dobijenih primenom metoda ONK u KLRM**

44

3. Svojstva ocena koje su dobijene primenom metoda ONK

- Ako su zadovoljene pretpostavke KLRM od 1. do 4. tada se primenom metoda ONK dobijaju najbolje linearne nepristrasne ocene (NLNO)

Šta to znači?

- Ocena: $\hat{\beta}$ je ocena stvarne vrednosti parametra β
- Linearna: $\hat{\beta}$ je linearna funkcija raspoloživih podataka
- Nepistrasna: u proseku ocena $\hat{\beta}$ je jednaka parametru β
- Najbolja: ocena je efikasna (nepistrasna ocena sa najmanjom varijansom)

45

Nepistrasnost/Efikasnost/Konzistentnost

- Nepistrasnost
Ocene metoda ONK su nepristrasne. To znači da su ocene u proseku jednake parametrima koji se ocenjuju: $E(\hat{\alpha})=\alpha$ i $E(\hat{\beta})=\beta$.
- Efikasnost
Ocene metoda ONK su efikasne ocene. Ocena je efikasna ako je nepristrasna i ako ne postoji druga nepristrasna ocena koja poseduje manju varijansu. To je nepristrasna ocena sa najmanjom mogućom varijansom.
- Konzistentnost
Ocene metoda ONK su konzistentne. To znači da sa porastom obima uzorka ocena konvergira u verovatnoći ka stvarnoj vrednosti parametra. Nepistrasna ocena je konzistentna ako njena varijansa teži nuli sa porastom obima uzorka.

$$\lim_{T \rightarrow \infty} Pr \left[\left| \hat{\beta} - \beta \right| > \varepsilon \right] = 0 \quad \text{za svako } \varepsilon > 0$$

46

Kako merimo preciznost ocena?

- Svaki drugi uzorak daje nove ocene parametara $\hat{\alpha}$ i $\hat{\beta}$. Ako se sa promenom uzorka ocene malo razlikuju, onda one imaju malu varijansu i obratno.
- Preciznost ocene se meri na osnovu ocene varijanse ocena.
- Kvadratni koren iz ocene varijanse je standardna greška ocene.
- Da bi se izračunale standardne greške ocena potrebno je prethodno oceniti varijabilitet slučajne greške modela.
- U pitanju je ocena parametra σ^2 .

47

Ocena varijanse slučajne greške modela σ^2

- Varijansa slučajne greške u_t je:
$$\text{Var}(u_t) = E[(u_t) - E(u_t)]^2 = \sigma^2$$

odnosno:
$$\text{Var}(u_t) = E(u_t^2)$$
- Ako bi slučajne greške bile poznate tada bi ocenu varijanse dobili na sledeći način: :

$$s^2 = \frac{1}{T} \sum u_t^2$$

- Međutim, ne znamo vrednosti u_t . Ali, poznate su nam vrednosti reziduala $\hat{u}_t$
$$s^2 = \frac{1}{T} \sum \hat{u}_t^2$$

Ova ocena je pristrasna ocena parametra σ^2 .

48

Ocena varijanse slučajne greške modela (II)

- Nepristrasna ocena σ^2 je:

$$s^2 = \frac{1}{T-2} \sum \hat{u}_t^2$$

gde je $\sum \hat{u}_t^2$ rezidualna suma kvadrata, a T je obim uzorka.
Kvadratni koren, s , je **standardna greška regresije, odnosno standardna devijacija reziduala.**

- Sada možemo da analiziramo ocene varijansi ocena parametara $\hat{\alpha}$ i $\hat{\beta}$
- Oznake: $SE(\hat{\alpha})$ i $SE(\hat{\beta})$ su odgovarajuće standardne greške ocena koje se dobijaju kao kvadratni koren iz ocena varijansi.

49

Ocene varijansi ocena parametara $\hat{\alpha}$ i $\hat{\beta}$

$$var(\hat{\alpha}) = E(\hat{\alpha} - E(\hat{\alpha}))^2 = \dots = \sigma^2 \left(\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^T (X_t - \bar{X})^2} \right)$$

$$var(\hat{\beta}) = E(\hat{\beta} - E(\hat{\beta}))^2 = \dots = \frac{\sigma^2}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

$$var(\hat{\alpha}) = s^2(\hat{\alpha}) = s^2 \left(\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^T (X_t - \bar{X})^2} \right) \Rightarrow SE(\hat{\alpha}) = s \sqrt{\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^T (X_t - \bar{X})^2}}$$

$$var(\hat{\beta}) = s^2(\hat{\beta}) = \frac{s^2}{\sum_{t=1}^T (X_t - \bar{X})^2} \Rightarrow SE(\hat{\beta}) = s \sqrt{\frac{1}{\sum_{t=1}^T (X_t - \bar{X})^2}}$$

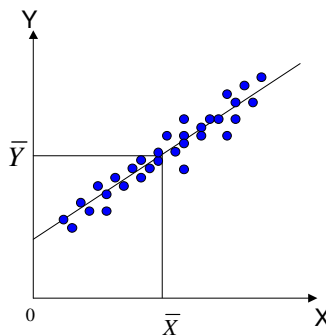
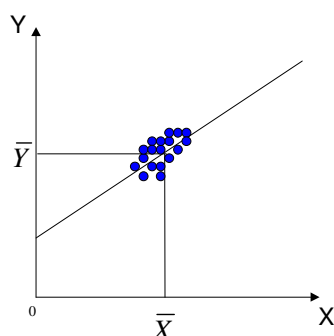
50

Standardne greške ocena parametara zavise od sledećih faktora:

1. Varijabilitet modela (s^2 ili s). Što je veći varijabilitet modela, to je veći stepen raspršenosti slučajne greške modela, a time i veći varijabilitet zavisne promenljive Y . Rezultat: neprecizne ocene parametara.
2. Suma kvadrata odstupanja X od aritmetičke sredine. U pitanju je mera varijabiliteta objašnjavajuće promenljive. Veća vrednost ove sume utiče na povećanje preciznosti ocena, odnosno na pad njihovog varijabiliteta.
3. Obim uzorka T . Javlja se eksplicitno u imeniocu formule za standardnu grešku slobodnog člana i implicitno u imeniocu formule za obe ocene kroz zbir kvadrata odstupanja X od aritmetičke sredine. Veći obim uzorka pruža više informacija. Time se smanjuje varijabilitet ocena parametara.
4. Standardna greška ocene slobodnog člana zavisi i od aritmetičke sredine podataka za X . Podaci su udaljeniji od y -ose što je vrednost ove aritmetičke sredine veća. Rezultat: nepreciznija ocena slobodnog člana.

51

Šta se dešava ako je suma $\sum (x_i - \bar{x})^2$ relativno mala ili relativno velika?



52

Primer: izračunavanje ocena parametara modela i odgovarajućih standardnih grešaka

- Pretpostavimo da smo izračunali sledeće vrednosti iz uzorka od 22 podatka:
- Odgovarajući međurezultati su:

$$\sum X_t Y_t = 830102, T = 22, \bar{X} = 416.5, \bar{Y} = 86.65,$$

$$\sum X_t^2 = 3919654, \sum \hat{u}_t^2 = 130.6.$$

- Ocene:

$$\hat{\beta} = \frac{830102 - (22 * 416.5 * 86.65)}{3919654 - 22 * (416.5)^2} = 0.35$$

$$\hat{\alpha} = 86.65 - 0.35 * 416.5 = -59.12$$

- Ocenjen model:

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t$$

$$\hat{Y}_t = -59.12 + 0.35 X_t$$

53

Primer (II)

- Standardna greška regresije, s :

$$s = \sqrt{\frac{\sum \hat{u}_t^2}{T-2}} = \sqrt{\frac{130.6}{20}} = 2.55$$

$$SE(\hat{\alpha}) = 2.55 * \sqrt{\frac{3919654}{(22 * 3919654) - (22 * 416.5)^2}} = 3.35$$

$$SE(\hat{\beta}) = 2.55 * \sqrt{\frac{1}{3919654 - (22 * 416.5^2)}} = 0.0079$$

- Uobičajeno se svi dobijeni rezultati zapisuju na sledeći način:

$$\hat{Y}_t = -59.12 + 0.35 X_t$$

$$(3.35) \quad (0.0079)$$

54

- **Statističko zaključivanje u KLRM**

55

4. Statističko zaključivanje u KLRM

- Izvođenje zaključaka o svojstvima parametara osnovnog skupa na osnovu ocenjenih regresionih parametara.

Primer: Ocenjen je model oblika:

$$\hat{Y}_t = -59.12 + 0.35X_t$$

(3.35) (0.0079)

- Ocena 0.35 je (tačkasta) nepoznatog parametra nagiba. Koliko je ta ocena pouzdana?
- Odgovor na to pitanje daje standardna greška ocene.

56

Testiranje hipoteze: osnovni elementi

- Interesuje nas da li parametar nagiba uzima tačno određenu vrednost.
- Postavljamo dve hipoteze: nultu (oznaka H_0) i alternativnu hipotezu (oznaka H_1).
- Nulta hipoteza je iskaz čiju valjanost ispitujemo, odnosno testiramo. Alternativna hipoteza obuhvata sva alternativna tvrđenja.
- Na primer, interesuje nas da li se zavisna promenljiva menja u istom obimu kao i objašnjavajuća, odnosno da li je β jednako 1. Koristimo sledeću notaciju:

$$H_0 : \beta = 1$$

$$H_1 : \beta \neq 1$$

57

U pitanju je tzv. dvostrani test.

Kako ostvariti diskriminaciju između hipoteza? Raspodela verovatnoće ocena dobijenih metodom ONK

- Ocene koje su dobijene primenom metoda ONK su i same normalno raspodeljene:

$$u_t \sim N(0, \sigma^2), Y_t \sim N(\alpha + \beta X_t, \sigma^2),$$

$$\hat{\beta} \sim N(\beta, \text{var}(\hat{\beta}))$$

$$\hat{\alpha} \sim N(\alpha, \text{var}(\hat{\alpha}))$$

- Šta ako slučajna greška modela nema normalnu raspodelu? Da li bi i onda ocene imale normalnu raspodelu?
- Da, ako važe ostale pretpostavke KLRM i uzorak je dovoljno velikog obima.

58

Raspodela verovatnoće ocena dobijenih metodom ONK (II)

- Standardizovanjem slučajnih promenljivih $\hat{\alpha}$ i $\hat{\beta}$ dobijamo:

$$\frac{\hat{\alpha} - \alpha}{\sqrt{\text{var}(\hat{\alpha})}} \sim N(0,1) \quad \text{i} \quad \frac{\hat{\beta} - \beta}{\sqrt{\text{var}(\hat{\beta})}} \sim N(0,1)$$

- Međutim, varijanse ocena su nepoznate veličine. Ako ih zamenimo odgovarajućim ocenama, tada dobijamo slučajne promenljive sa t -raspodelom (proveriti!)

$$\frac{\hat{\alpha} - \alpha}{SE(\hat{\alpha})} \sim t_{T-2} \quad \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \sim t_{T-2}$$

59

Testiranje hipoteza: algoritam

- Posmatramo model oblika

$$Y_t = \alpha + \beta X_t + u_t \quad \text{za } t=1,2,\dots,T$$

- Testiramo validnost hipoteze:

$$H_0: \beta = \beta^*, \quad H_1: \beta \neq \beta^*$$

- Koraci u postupku testiranja:

1. Ocenjujemo: $\hat{\alpha}$, $\hat{\beta}$, $SE(\hat{\alpha})$, $SE(\hat{\beta})$ na poznati način.

2. Računamo test-statistiku koristeći sledeću formulu:

$$\text{test - statistika} = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \sim t_{T-2}$$

gde je β^* vrednost β u uslovima važenja nulte hipoteze.

60

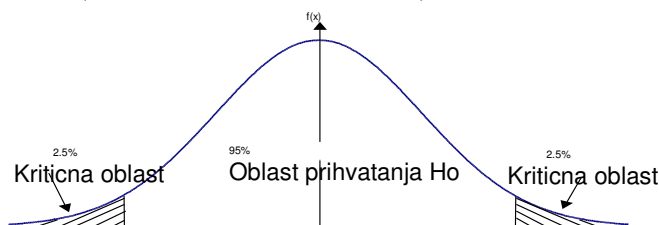
Testiranje hipoteza: algoritam (II)

3. Sastavni deo testiranja hipoteze je izbor **nivoa značajnosti**, koji se često označava sa α . To je verovatnoća odbacivanja nulte hipoteze u situaciji kada je ona tačna. Uobičajeno se koristi nivo značajnosti 5%.

Nivo značajnosti određuje veličinu oblasti prihvatanja, odnosno neprihvatanja validnosti nulte hipoteze. Oblast odbacivanja nulte hipoteze je **kritična oblast testa**.

$$H_0: \beta = \beta^* \Rightarrow \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \sim t_{T-2} \Rightarrow P\left(-t_{T-2}(\alpha/2) \leq \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \leq t_{T-2}(\alpha/2)\right) = 1 - \alpha$$

$$\alpha = 0.05, P\left(-t_{T-2}(0.025) \leq \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \leq t_{T-2}(0.025)\right) = 0.95.$$



61

Testiranje hipoteza: algoritam (III)

4. Definišemo pravilo odlučivanja: kriterijum po kojem odbacujemo nultu hipotezu.

$$\frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \in (\pm t_{T-2}(0.025)) \Rightarrow H_0 \text{ prihvatamo kao tačnu}$$

$$\frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \notin (\pm t_{T-2}(0.025)) \Rightarrow H_0 \text{ odbacujemo kao netačnu uz nivo značajnosti 5\%}$$

Alternativna notacija

$$\left| \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \right| > t_{T-2}(0.025) \Rightarrow H_0 \text{ odbacujemo kao netačnu uz nivo značajnosti 5\%}$$

62

Testiranje hipoteza: algoritam (IV)

5. Konačno sprovodimo testiranja. Ako izračunata test-statistika leži u oblasti prihvatanja nulte hipoteze, tada se nulta hipoteza ne odbacuje. Obratno, ako izračunata test-statistika pripada kritičnoj oblasti testa, tada nultu hipotezu odbacujemo za dati nivo značajnosti.

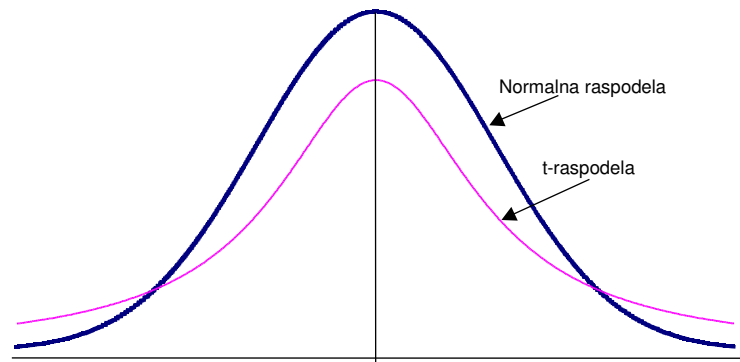
63

Odnos t i standardizovane normalne raspodele

- Obe raspodele su simetrične u nuli sa oblikom zvona.
- Varijansa standardizovane slučajne promenljive (pa i normalne) je 1.
- Varijansa slučajne promenljive sa t -raspodelom je $T/(T-2)$, što je uvek veće od 1.
- Slučajna promenljiva sa t -raspodelom poseduje veći stepen raspršenosti od normalne standardizovane slučajne promenljive.
- Sa povećanjem obima uzorka količnik $T/(T-2)$ postaje blizak vrednosti jedan, tako da se i t -raspodela može aproksimirati normalnom raspodelom.

64

Odnos t i standardizovane normalne raspodele (II)



65

Odnos t i standardizovane normalne raspodele (III)

- U asimptotskom slučaju, t -raspodela sa beskonačnim brojem stepeni slobode je standardizovana normalna raspodela: $t(\infty) = N(0,1)$
- Primeri iz statističkih tabela:

Nivo značajnosti	$N(0,1)$	$t(40)$	$t(4)$
10%	1.64	1.68	2.13
5%	1.96	2.02	2.78
1%	2.58	2.70	4.60

- Još jednom podsećanje: koristimo t raspodelu umesto normalne zato što ocenjujemo nepoznatu varijansu slučajne greške modela.

66

Primer testiranja hipoteza

- Podsećamo na ocenu modela:

$$\hat{Y}_t = -59.12 + 0.35 X_t$$

$(0.56) \quad (0.0079)$

- Testiramo valjanost nulte hipoteze:

$$H_0: \beta = 1 \text{ protiv alternativne } H_1: \beta \neq 1$$

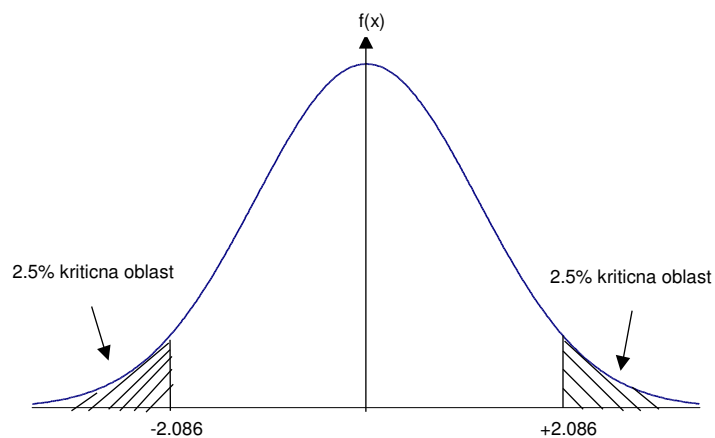
- Potrebna nam je kritična vrednost t raspodele za 20 stepeni slobode i nivo značajnosti 5%. Budući da je test dvostran i da je ukupna veličina kritične oblasti 5%, koristimo sledeću notaciju:

$$t_{20}(0.025) \text{ ili } t_{20}(2.5\%)$$

- Tablice: $t_{20}(0.025)=2.086$

67

Određivanje kritične oblasti testa



68

Testiranje hipoteze

- Hipoteze:

$$H_0 : \beta = 1$$

$$H_1 : \beta \neq 1$$

- Izračunata test-statistika:

$$t = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} = \frac{0.35 - 1}{0.0079} = -82.28$$

- Kako je

$$|-82.28| > 2.086$$

odbacujemo hipotezu H_0 na datom nivou značajnosti.

69

Testiranje drugih hipoteza

- Može nas interesovati sledeće: $H_0 : \beta = 0$ ili $H_0 : \beta = 2$.

- $H_0 : \beta = 0$
 $H_1 : \beta \neq 0$

$$H_0 : \beta = 2$$

$$H_1 : \beta \neq 2$$

70

Specijalni tip hipoteze: *t*-odnos

- Opšti oblik testa koji smo koristili je:

$$\text{test statistika} = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})}$$

- Pretpostavimo da nas interesuje $H_0 : \beta = 0$ protiv $H_1 : \beta \neq 0$. Ako je tačna nulta hipoteza, tada objašnjavajuća promenljiva ne utiče na kretanje zavisne promenljive. Na ovaj način proveravamo opravdanost postavke modela.
- Test se naziva ***t*-odnos**, zato što za $\beta = 0$, test statistika postaje odnos ocene i odgovarajuće standardne greške ocene:

$$\text{test statistika} = t\text{-odnos} = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

$$t = \frac{0.35}{0.0079} = 44.30, 44.30 > 2.086$$

$\Rightarrow H_1 : \beta \neq 0$ prihvata se kao tačno.

71

- **Klasični višestruki linearni regresioni model**

eko21-10.ppt

72