



Ekonometrija – I deo

Doktorske studije

Predavač: Aleksandra Nojković

Beograd, školska 2025/2026

Literatura

- **Mladenović, Z. i Petrović, P. (2020), *Uvod u ekonometriju*, Ekonomski fakultet, Beograd**
- **Mladenović, Z., Nojković, A. i Anić, A. (2021), *Zbirka rešenih zadataka iz ekonometrije*, Ekonomski fakultet, Beograd**
- Verbeek, M. (2017), *A Guide to Modern Econometrics*, 5th edition, John Wiley&Sons
- Asteriou, D. and Hall, S.H. (2021), *Applied Econometrics*, 4rd ed., Bloomsbury Academic
- Wooldridge, J.M. (2021) *Introductionary Econometrics: A Modern Approach*, 7th ed., Cengage Learning

Izvodi sa predavanja:

<http://ekonometrija.ekof.bg.ac.rs/doktorske.html>



Literatura (napredniji kurs)

- Greene, W.H. (2018), *Econometric Analysis*, 8th edition, Pearson
- Wooldridge, J. M. (2011), *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, The MIT Press
- Nojković, A. (2017). Odabrane ekonometrijske teme: metodologija i primena, Ekonomski fakultet, Beograd



Struktura predavanja

- Uvodna razmatranja
- Jednostavna regresiona analiza i metod običnih najmanjih kvadrata (metod ONK)
- Klasični jednostavni linearni regresioni model(KLRM)
- Statističko zaključivanje u KLRM
(podsećanje na neke teorijske raspodele)
- Svojstva ocena dobijenih metodom ONK
(podsećanje na svojstva ocena u malim i velikim uzorcima)

Neke definicije termina ekonometrija

- Naučna disciplina koja se bavi **merenjima u ekonomiji**
- Nastanak se vezuje osnivanje Ekonometrijskog društva 1930 godine (1933. god. pokreće se časopis "*Econometrica*")
- Schumpeter (1933): Ekonometrija je primena specifičnog metoda u opštem polju ekonomske nauke u naporu da se postignu numerički rezultati i verifikuju ekonomske teoreme
- Goldberger (1964): Ekonometrija je društvena nauka koja primenjuje metode matematičke statistike na ekonomske podatke u cilju analize valjanosti postavki ekonomske teorije



Još neke definicije termina ekonometrija

- Samuelson, Koopmans i Stone (1954): Ekonometrija je društvena nauka u kojoj se metode ekonomske teorije, matematike i teorije statističkog zaključivanja primenjuju u cilju analize ekonomskih pojava
- Kennedy (1998): Osnovni zadatak ekonometrije je oživljavanje teorijskih struktura



Osnovni ciljevi ekonometrije

- **Utvrđivanje kvantitativne zavisnosti** veličina u ekonomskoj relaciji
 - Modeliranje ekonomskih veličina: koliko se promeni jedna veličina sa promenom druge.
- **Ispitivanje** valjanosti postavki **ekonomske teorije**
 - Testiranje konkurentnih hipoteza.
- **Predviđanje** budućeg kretanja ekonomskih veličina na osnovu utvrđene kvantitativne veze.

Ekonometrijska istraživanja se zasnivaju na rezultatima sledećih naučnih disciplina:

1. Ekonomska teorija (matematička ekonomija): teorije i ideje su formulisane u formi matematičkih jednačina (bez podataka)
 2. Ekonomska statistika: prikupljanje i obrada podataka
 3. Matematička (teorijska) statistika: izvođenje zaključaka o ekonomskim odnosima **primenom statističkih metoda** (ocenjivanje i testiranje hipoteza) na konkretne podatke
- Ekonometrija je interakcija teorije, ekonomskih podataka i statističkih metoda



Veza ekonometrije sa drugim naučnim disciplinama

- Ekonomija vs. Ekonometrija: Tačna matematička forma modela, definisanje promenljivih i razmere reakcije
- Statistika vs. Ekonometrija: Ekonomske zavisnosti uključuju stohastički član
- Statističke metode su definisane pod pretpostavkom da slučajna greška zadovoljava određene pretpostvake

Razlika između ekonomije i ekonometrije

- Posmatramo npr. model tražnje za novcem:

$$m=f(Y,r),$$

gde je: m -realni novac, Y -realni dohodak i r – kamatana stopa. Teorija sugeriše: rast Y dovodi do rasta m , dok rast r dovodi do pada tražnje za novcem

Brojna pitanja:

- 1) Kako su definisane promenljive modela?
 - 2) Matematička forma modela: linearna ili ne?
 - 3) Ne znamo preciznu kvalitativnu zavisnost: poznat je tip reakcije, ali ne i razmere.
- Ekonometrija pruža odgovore na navedena pitanja

Razlika između statistike i ekonometrije

- U ekonomske relacije (deterministička veza) se uvodi slučajna komponenta (stohastička komponenta ili slučajna greška), koja obuhvata brojne nepredvidive uticaje
- U zavisnosti od slučajne greške, u kretanju realnog novca uvek će postojati **izvestan stepen slučajnih varijacija**. Parametri jednačine se **ne mogu tačno izmeriti**, već se **mogu oceniti** na bazi odgovarajućih statističkih metoda
- Statističke metode su definisane pod pretpostavkom da slučajna greška zadovoljava određene uslove (razvijene za kontrolisane eksperimente)
- Ukoliko slučajna greška ne zadovoljava zadate pretpostavke koriste se ekonometrijske metode



Ekonometrijske metode

- Prilagođavanje problemima ekonomskog života se sastoji u specifikovanju stohastičkih elemenata u ekonomskom ponašanju
- Korišćeni podaci se mogu interpretirati kao slučajan uzorak, na koji se primenjuju statistički metodi prilagođeni radu sa ekonomskim podacima (**korigovani metodi statističke nalaize**)
- **Suština ekonometrijskih metoda:** Analiziranje izvora i posledica narušavanja standardnih pretpostavki, testiranje i otklanjanje neželjenih uticaja na ocene modela. **Zasnivaju se na analizi reziduala**

Pravci razvoja ekonometrije

- Uslovljeni konkretnim empirijskim problemima
- Tradicionalna ekonometrija je u fokusu imala **makroekonomske modela** koji su se sastojali od **manjeg ili većeg broja jednačina**
- Poslednjih decenija 20-og veka razvoj se bazirao na metodologiji **analize vremenskih serija** (posebno važan je koncept kointegracije)
- Od 1970-ih razvoj **mikorekonometrije** (modeli sa specifičnom zavisnom promenljivom)
- Skorije, značajan razvoj **finansijske ekonometrije**
- Danas ekonometrija ima dominantnu ulogu u empirijskim istraživanjima gotovo svih oblasti ekonomije
- Dve osnovne grane: **teorijska i primenjena**



Metodologija (faze) ekonometrijskog istraživanja

1. Izbor teorijskog modela
2. Postavka ekonometrijskog modela
3. Prikupljanje podataka
4. Ocena parametara modela
5. Ispitivanje valjanosti ocenjenog modela
Vrednovanje dobijenih ocena: ekonomski, statistički i ekonometrijski kriterijumi
6. Predviđanje (vrednovanje moći predviđanja modela)



Vrste podataka

- Podaci vremenskih serija
 - Godišnji, kvartalni mesecni, dnevni, kako se obavi transakcija
- Podaci preseka (strukture, uporedni)
 - Vrednosti različitih promenljivih koje definišu strukturu u datom trenutku vremena
- Podaci panela
 - Kombinacija podataka vremenskih serija i podataka preseka
- Veštačke promenljive

Primer 1: Uporedni podaci (izvor: Wooldridge, 2021)

TABLE 1.1 A Cross-Sectional Data Set on Wages and Other Individual Characteristics					
obsno	wage	educ	exper	female	married
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
525	11.56	16	5	0	1
526	3.50	14	5	1	0

Primer 2: Uporedni podaci (izvor: Wooldridge, 2021)

TABLE 1.2 A Data Set on Economic Growth Rates and Country Characteristics

obsno	country	gpcrgdp	govcons60	second60
1	Argentina	0.89	9	32
2	Austria	3.32	16	50
3	Belgium	2.56	13	69
4	Bolivia	1.24	18	12
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
61	Zimbabwe	2.30	17	6

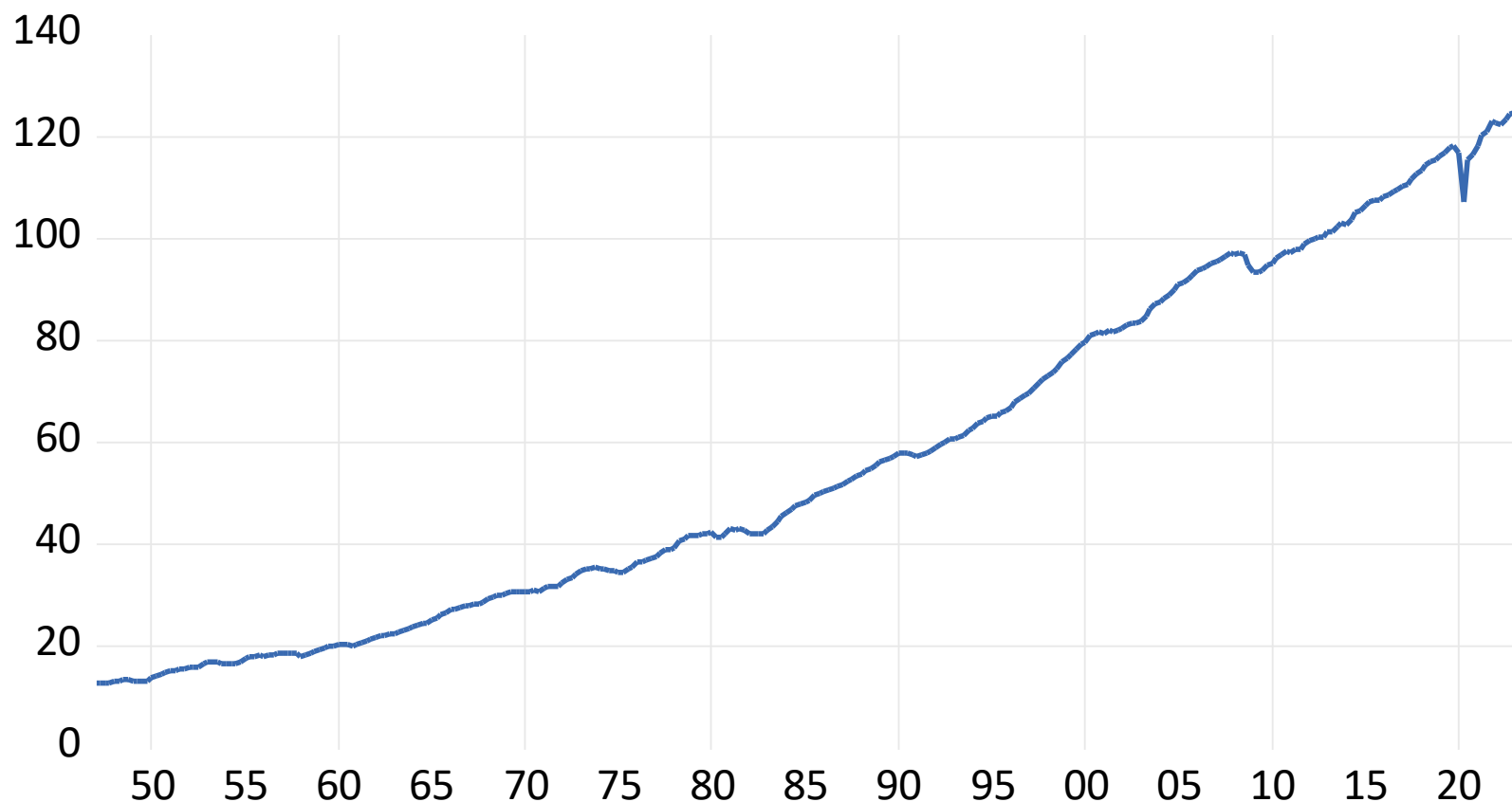
Primer 3: Podaci vremenskih serija (izvor: Wooldridge, 2021)

TABLE 1.3 Minimum Wage, Unemployment, and Related Data for Puerto Rico

obsno	year	avgmin	avgcov	prunemp	prgnp
1	1950	0.20	20.1	15.4	878.7
2	1951	0.21	20.7	16.0	925.0
3	1952	0.23	22.6	14.8	1015.9
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
37	1986	3.35	58.1	18.9	4281.6
38	1987	3.35	58.2	16.8	4496.7

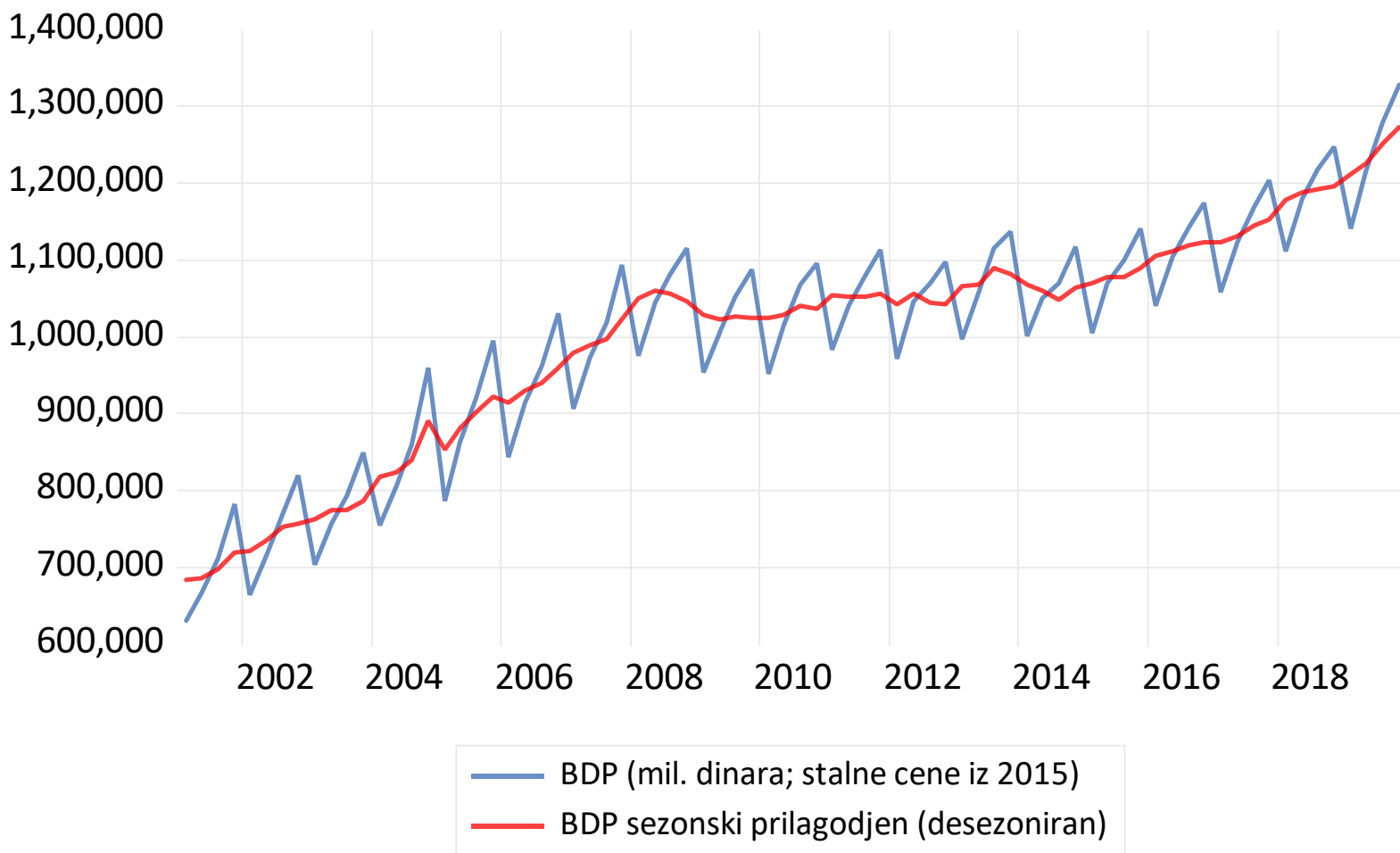
Primer 4: Podaci vremenskih serija BDP Američke provrede, desezoniran (1947Q1-2023Q2)

BDP privrede SAD-a (2012=100, SA)

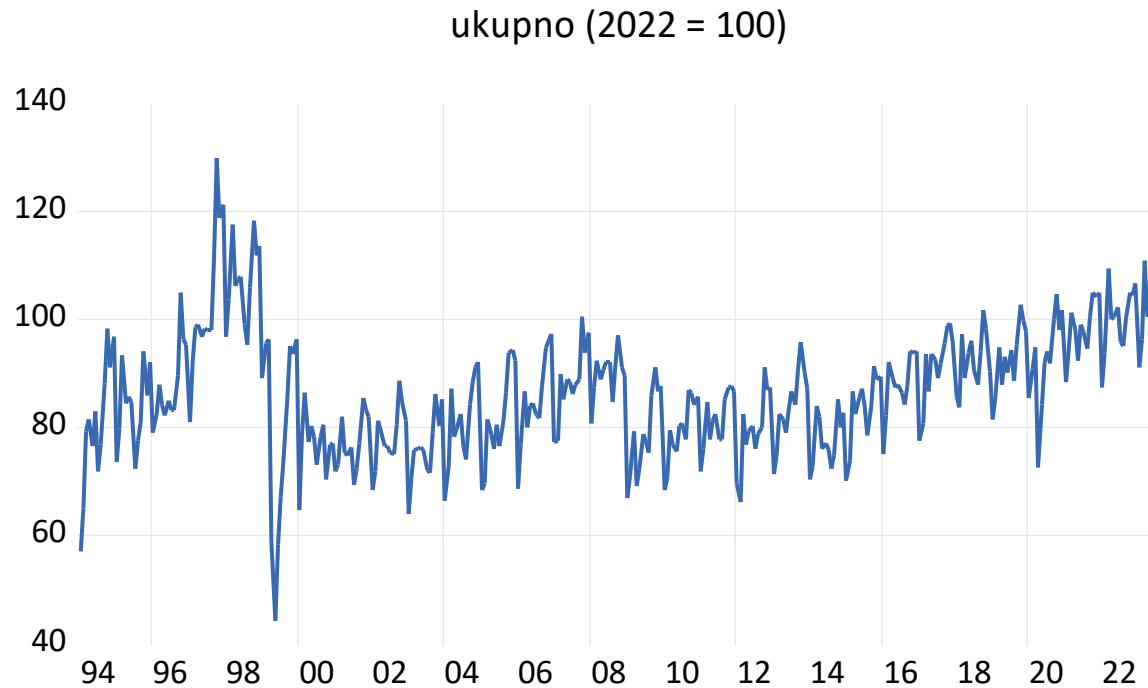


Primer 5: Podaci vremenskih serija

BDP privrede Srbije (2001Q1-2019Q4)

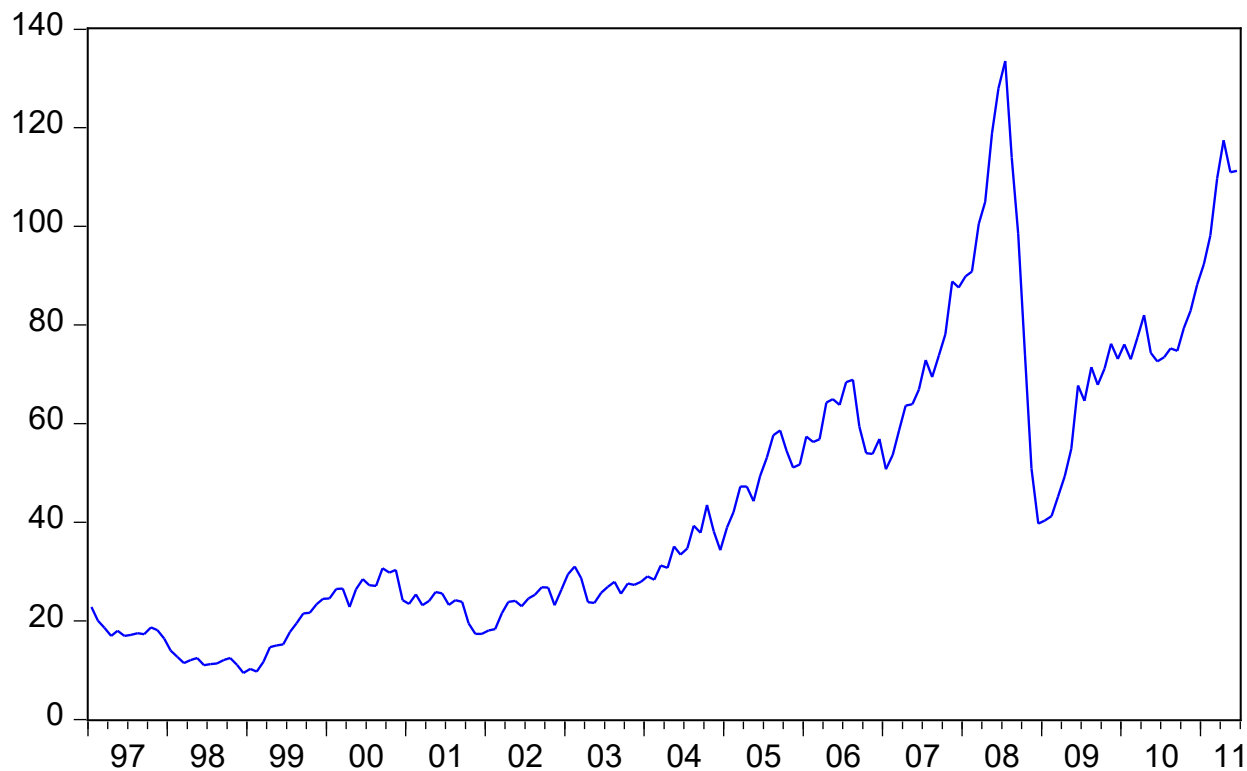


Primer 6. Veštačke promenljive (IIP privrede Srbije (1994M1-2023M7))



Primer 7. Veštačke promenljive (npr. svetska cena nafte; 1997:1-2011:6)

Svetska cena nafte (dolari po barelu)



Primer 8. Panel podaci (izvor: Wooldridge, 2021)

TABLE 1.5 A Two-Year Panel Data Set on City Crime Statistics

obsno	city	year	murders	population	unem	police
1	1	1986	5	350,000	8.7	440
2	1	1990	8	359,200	7.2	471
3	2	1986	2	64,300	5.4	75
4	2	1990	1	65,100	5.5	75
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
297	149	1986	10	260,700	9.6	286
298	149	1990	6	245,000	9.8	334
299	150	1986	25	543,000	4.3	520
300	150	1990	32	546,200	5.2	493

Primer 9. Panel 12 zemalja CEE (N=12) period 1995-2018 (T=24)

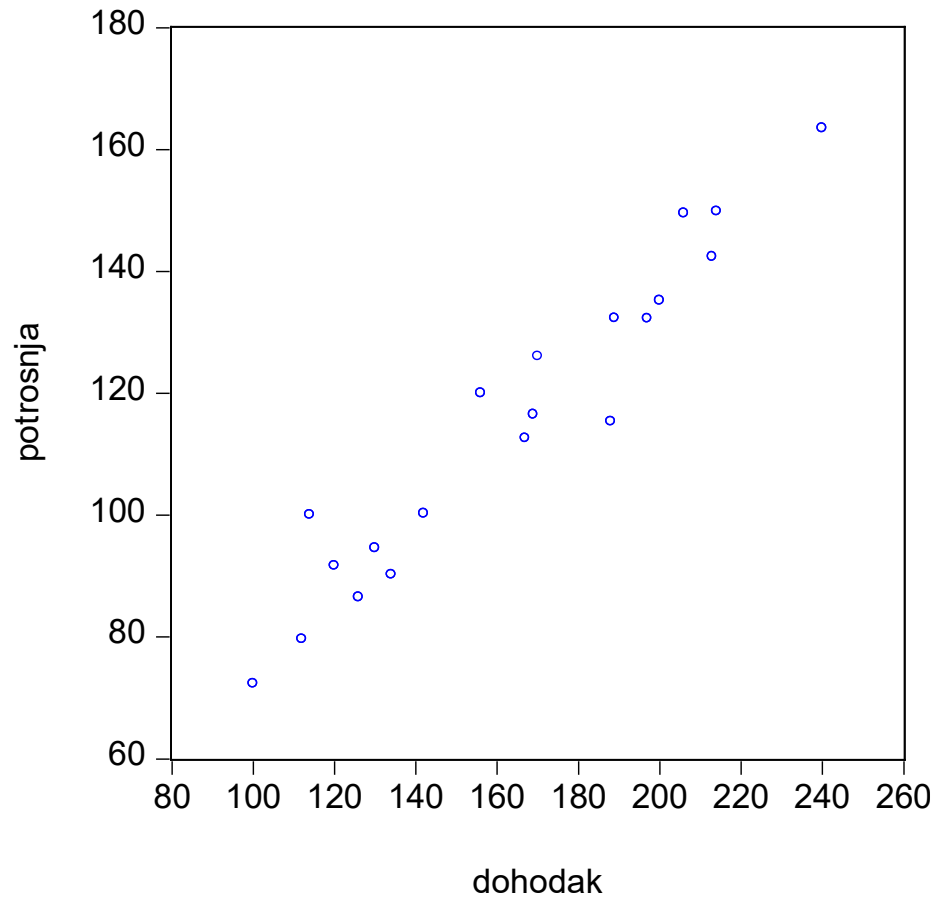




Primena jednostavne regresije

- Pretpostavimo da raspolažemo podacima o potrošnji i dohotku za određeni broj slučajnih ispitanika period i da želimo da otkrijemo prirodu njihove međusobne povezanosti (primer: Asteriou and Hall (2015))
- Cilj regresione analize jeste utvrđivanje prirode i forme povezanosti između promenljivih

Dijagram rasturanja tačka



Populaciona i uzoračka reg. prava

- Pretpostavljamo da je veza između potrošnje i dohotka pozitivna. Hoćemo da opišemo potrošnju kao funkciju dohotka.
 - Potrošnja: zavisna promenljiva/varijabla (Y)
 - Dohodak: nezavisna promenljiva/varijabla (X)
- U regresionoj analizi zavisna (Y) i nezavisna (X) *promenljiva* imaju **potpuno razlicitu poziciju** (razlika sa korelacionom analizom).
- Postoji **jednosmeran pravac uzrocnosti**: samo X *utice na* Y , dok Y *ne utice na* X .

Populaciona regresiona prava (jednačina)

- Populaciona regresiona prava označava **stvarnu vezu** između datih promenljivih, koja **karakteriše osnovni skup** (sadrži parametre β_0 i β):

$$Y = \underbrace{\beta_0 + \beta X}_{E(Y)} + \underset{\substack{\downarrow \\ \text{sl. greška}}}{\varepsilon},$$

gde se X i Y odnose na dohodak i potrošnju populacije.

- **Opšta veza** je predstavljena očekivanim nivom potrošnje $E(Y)$ za dati nivo dohotka X, odnosno prosečnom potrošnjom za veći broj pojedinaca sa istim dohotkom X.
- Stvarna potrošnja Y najčešće nije jednaka njenoj očekivanoj vrednosti, otuda prisustvo slučajne greške (ε) u modelu.



Razlozi prisustva slučajne greške

- 1) Postojanje faktora čiji su pojedinačni uticaji na kretanje potrošnje sporadični i neregularni. Greška sadrži njihovo zbirno dejstvo
- 2) Nepredvidivost ljudskog ponašanja
- 3) Greške u merenju datih promenljivih

Uzoracka regresiona prava (jednačina)

- Elementi uzorka obeležavaju se sa i ($i=1,2,\dots, n$), pa uzoracka regresiona prava postaje:

$$Y_i = \underbrace{\beta_0 + \beta X_i}_{\text{sist. deo}} + \underbrace{\varepsilon_i}_{\text{stoh. deo}}$$

- Uzoracka regresiona prava opisuje vezu prema datom uzorku:

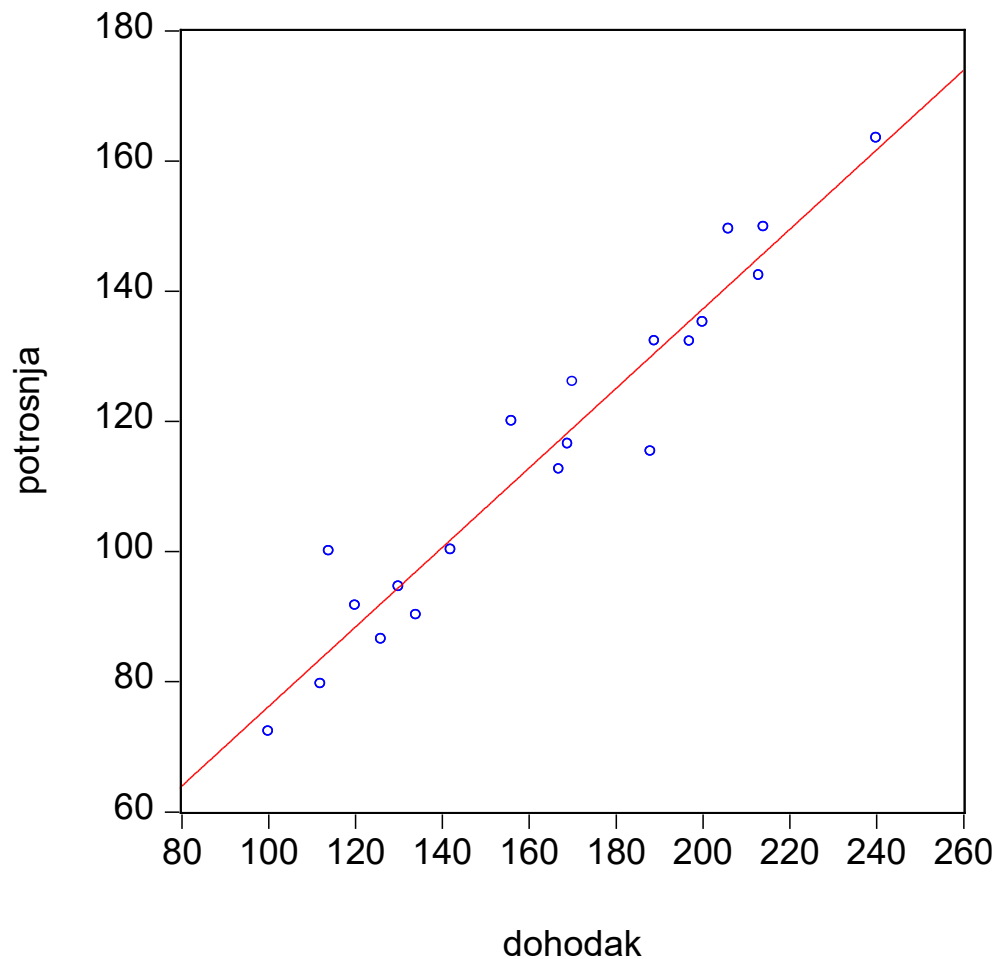
$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta} X_i = b_0 + bX_i$$

- Stvarni nivo zavisne promenljive je zbir ocenjenog nivoa i onoga što model nije ocenio (reziduala, obeležavaju kao $\hat{u}_i$)

$$Y_i = \hat{Y}_i + e_i$$

- Uzoracka regresiona prava (jednacina) se koristi za donošenje zaključaka o parametrima populacione regresione jednačine

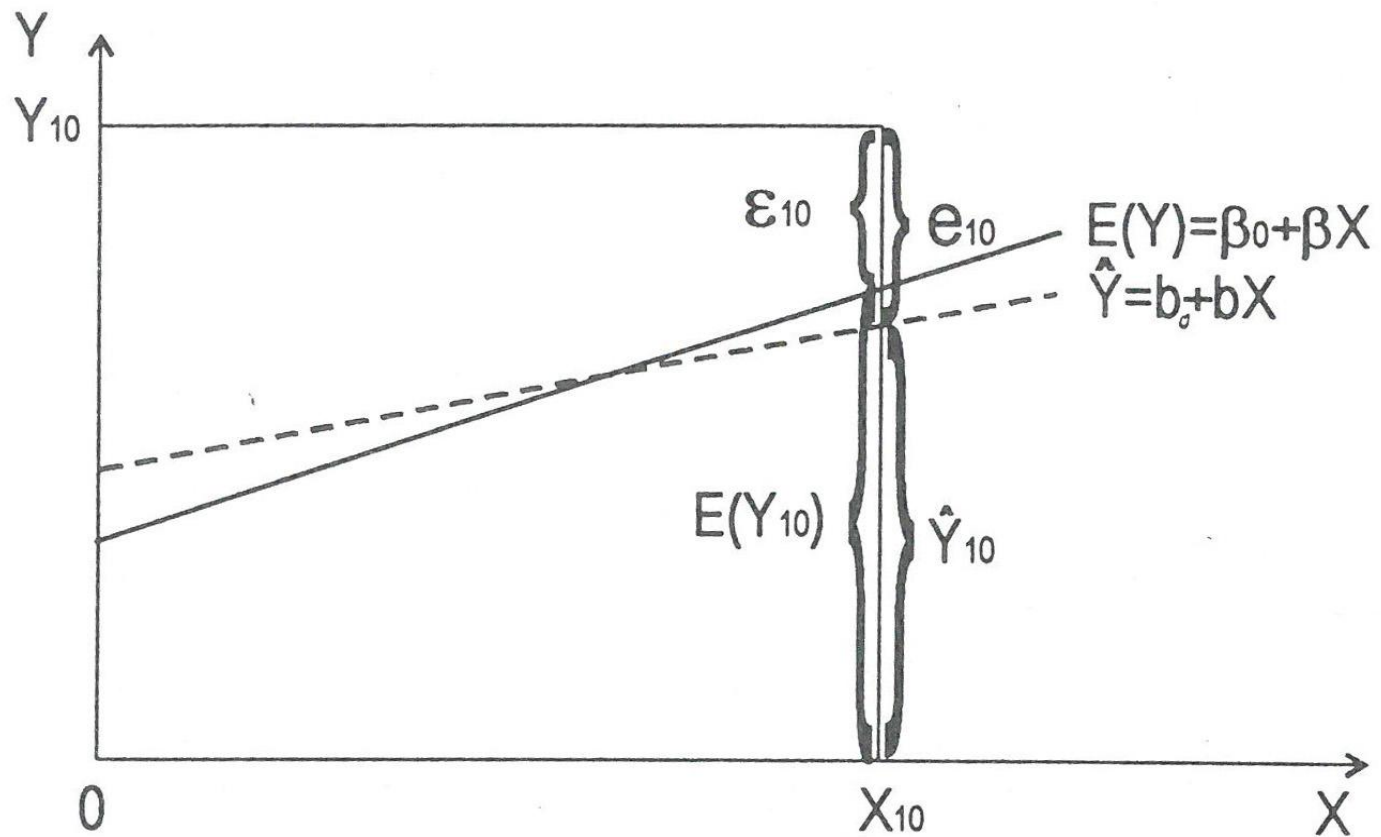
Dijagram rasturanja tačaka sa regresionom pravom



Populaciona i uzoračka reg. prava

- Razlika između reziduala (e_i ili $\hat{u}_i$) i slučajne greške (ε_i ili u_i)? – objasniti na grafiku!
- Za svaki podatak iz uzorka može se odrediti vrednost reziduala, dok je vrednost sl. greške nepoznata
- U praksi: **reziduali** se koriste kao **ocena** slučajnih greški

Populaciona i uzoračka regresiona prava



Jednostavna regresiona analiza

- *Regresiona analiza* predstavlja osnovni metodološki okvir ekonometrijskog modeliranja
- Jednostavan model:

$$Y_i = \underbrace{\beta_0 + \beta X_i}_{\text{sistematski deo}} + \underbrace{\varepsilon_i}_{\text{stohastički deo}} \quad \text{za } i=1,2,\dots, n.$$

gde je zavisna promenljiva Y za sve opservacije i iskazana kao funkcija samo jedne nezavisne promenljive (X_i) i greške modela ε_i (oznaka i u_i)



Metod običnih najmanjih kvadrata (metod ONK)

- Najčešće korišćen metod postavljanja prave i izbora regresionih koeficijenata jeste metod obicnih najmanjih kvadrata (ONK)
- Ideja metoda: minimizirati zbir kvadrata odstupanja podataka od prave

Ocene metodom ONK

- Izvođenje ocena...

- Ocene ONK:

$$\hat{\beta} = b = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

$$\hat{\beta}_0 = b_0 = \bar{Y} - b\bar{X}$$

Metodom ONK postavlja se prava za koju važi:

1)
$$\sum_{i=1}^n e_i = 0$$

2)
$$\sum_{i=1}^n X_i e_i = 0$$

Pokazati....

Ocena potrošne funkcije iz Eviews-a (Primer 1)

Dependent Variable: Y
Method: Least Squares
Date: 06/08/11 Time: 14:32
Sample: 1 20
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15.11641	6.565638	2.302352	0.0335
X	0.610889	0.038837	15.72951	0.0000
R-squared	0.932182	Mean dependent var	115.5160	
Adjusted R-squared	0.928415	S.D. dependent var	25.71292	
S.E. of regression	6.879603	Akaike info criterion	6.789639	
Sum squared resid	851.9210	Schwarz criterion	6.889212	
Log likelihood	-65.89639	Hannan-Quinn criter.	6.809076	
F-statistic	247.4176	Durbin-Watson stat	2.283770	
Prob(F-statistic)	0.000000			

Interpretacija ocena bo i b ?

- b označava prirast zavisne (potrošnje) po jedinici prirasta nezavisne (dohotka)
- bo označava nivo zavisne promenljive kada je nivo objašnjavajuće promenljive nula (dohodak je nula)
- Sa rastom dohotka za jednu jedinicu potrošnja raste za 0.61 jedinica
- Ukoliko je nivo dohotka nula, potrošnja iznosi 15.12 jedinica



Korelacija

- Nakon postavljanja regresione prave potrebno je odrediti u kojoj meri parovi podataka (X_i, Y_i) odstupaju od te prave.
- Potreba za sumarnom statistikom.
- Veći stepen objašnjenosti postoji u slučaju manje devijacije originalnih podataka od uzoračke regresione prave.

Korelacija (nastavak)

- Ukupno odstupanje zavisne promenljive od njene aritmetičke sredine $(Y_i - \bar{Y})$ se može razdvojiti kao:

$$\left(Y_i - \bar{Y} \right) = \underbrace{\left(\hat{Y} - \bar{Y} \right)}_{\text{objaš. reg. prava}} + \underbrace{e_i}_{\text{neobjaš. deo}}$$

- Ukupne varijacije (ukupni varijabilitet) zavisne prom:

$$\underbrace{\sum_{i=1}^n \left(Y_i - \bar{Y} \right)^2}_{\text{ukupne var ijacije}} = \underbrace{\sum_{i=1}^n \left(\hat{Y} - \bar{Y} \right)^2}_{\text{objaš. var ijacije}} + \underbrace{\sum_{i=1}^n e_i^2}_{\text{neobjaš. var ijacije}}$$

Pokazati...

Koeficijent determinacije R^2

- Dakle, ukupne varijacije se mogu predstaviti kao:

$$\sum_i \left(Y_i - \bar{Y} \right)^2 = \sum_i \left(\hat{Y}_i - \bar{Y} \right)^2 + \sum_i e_i^2$$

- Koeficijent determinacije predstavlja udeo objašnjenog u ukupnom varijabilitetu:

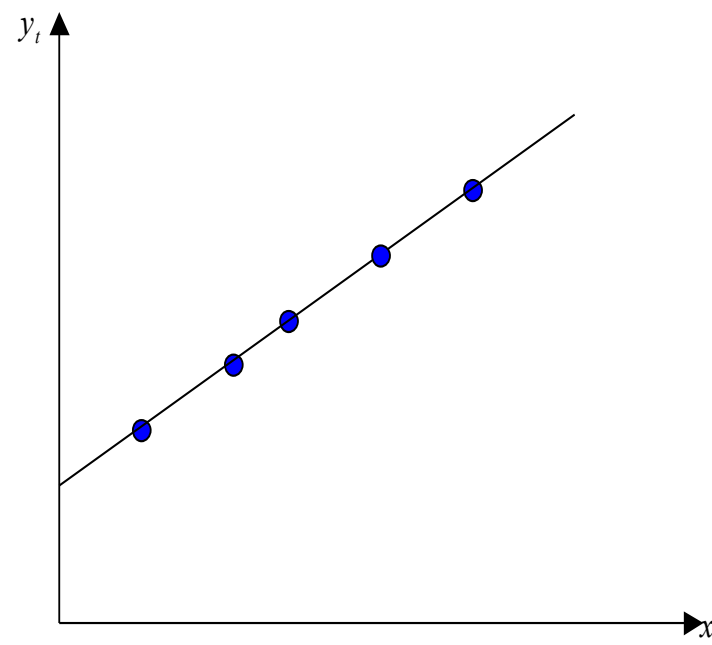
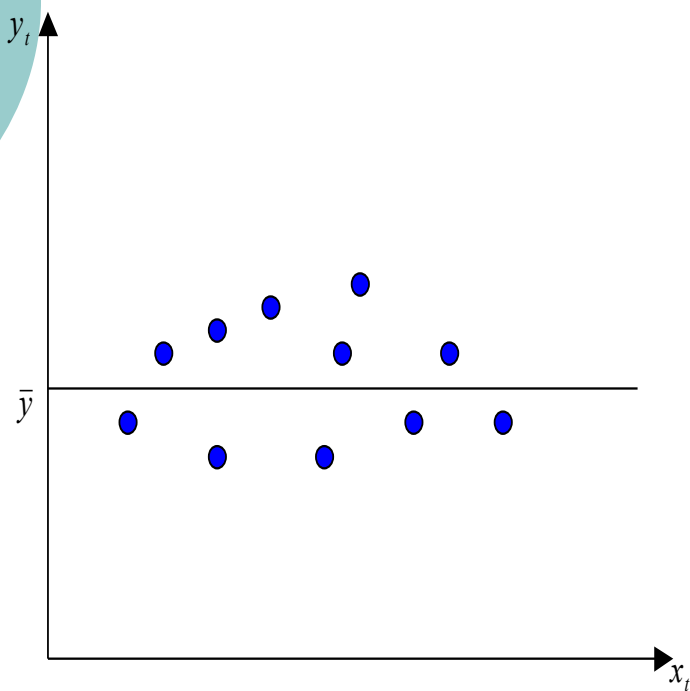
$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2},$$

odnosno:

$$R^2 = 1 - \frac{\sum_i e_i^2}{\sum_i \left(Y_i - \bar{Y} \right)^2}$$

- R^2 se uvek nalazi u intervalu od 0 do 1
- „Lažna” korelacija ili „besmislena” regresija

Ekstremni slučajevi: $R^2 = 0$ i $R^2 = 1$



D. Hendry (*Economica* 47, 1980): primer besmislene regresije

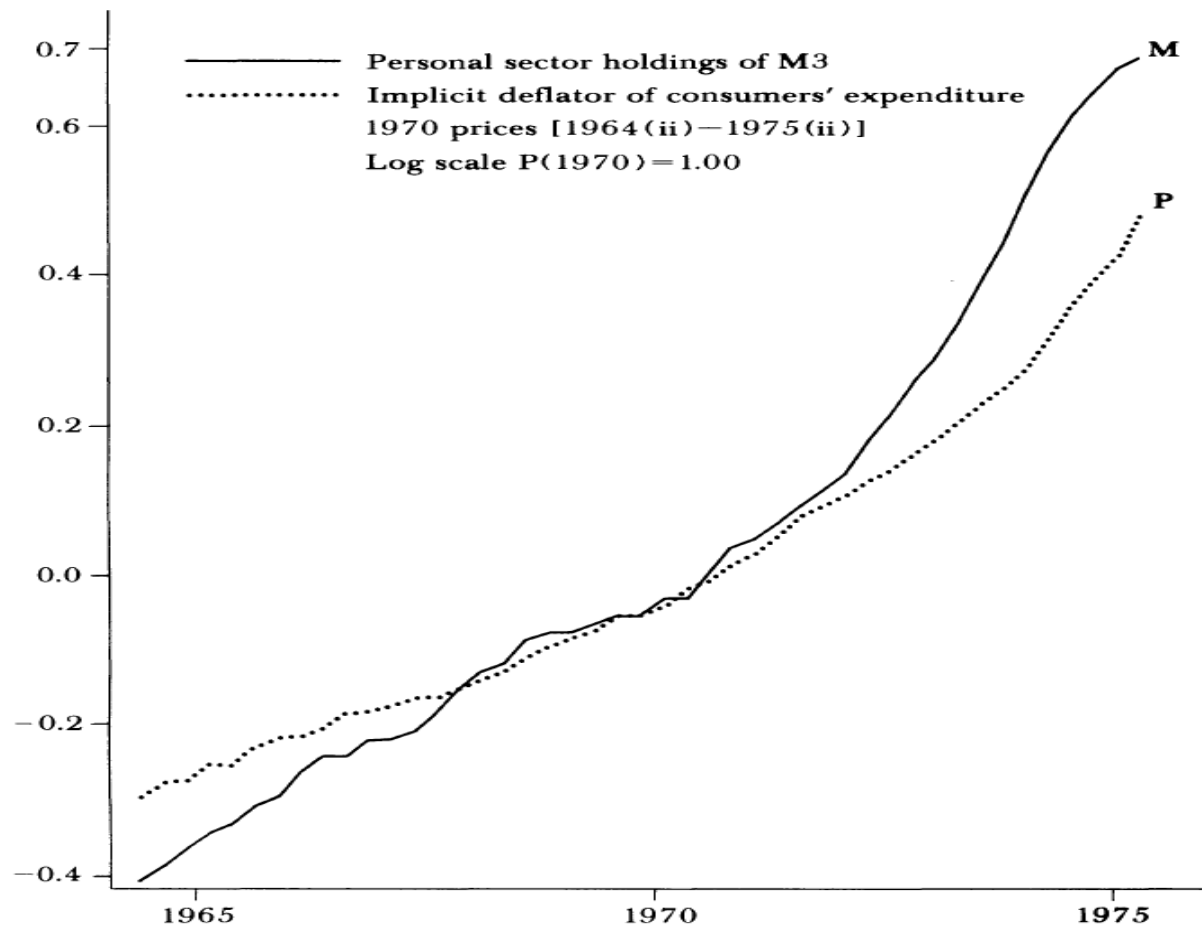


FIGURE 1

Besmislena regresija (II)

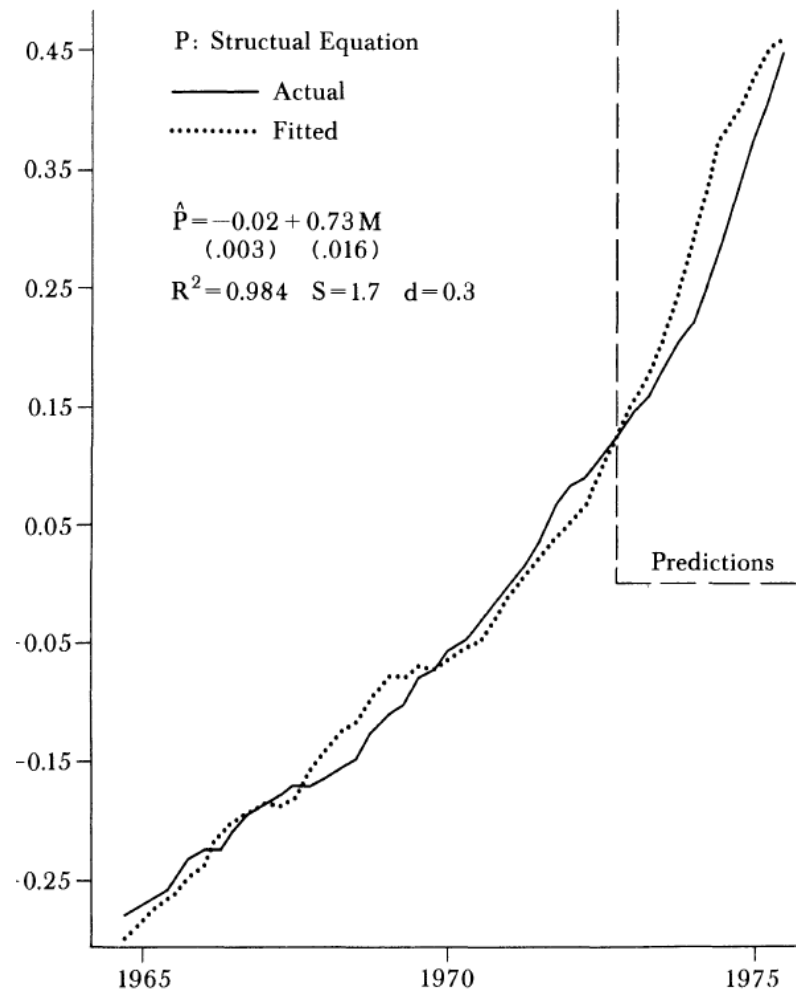


FIGURE 3

Besmislena regresija (IV)

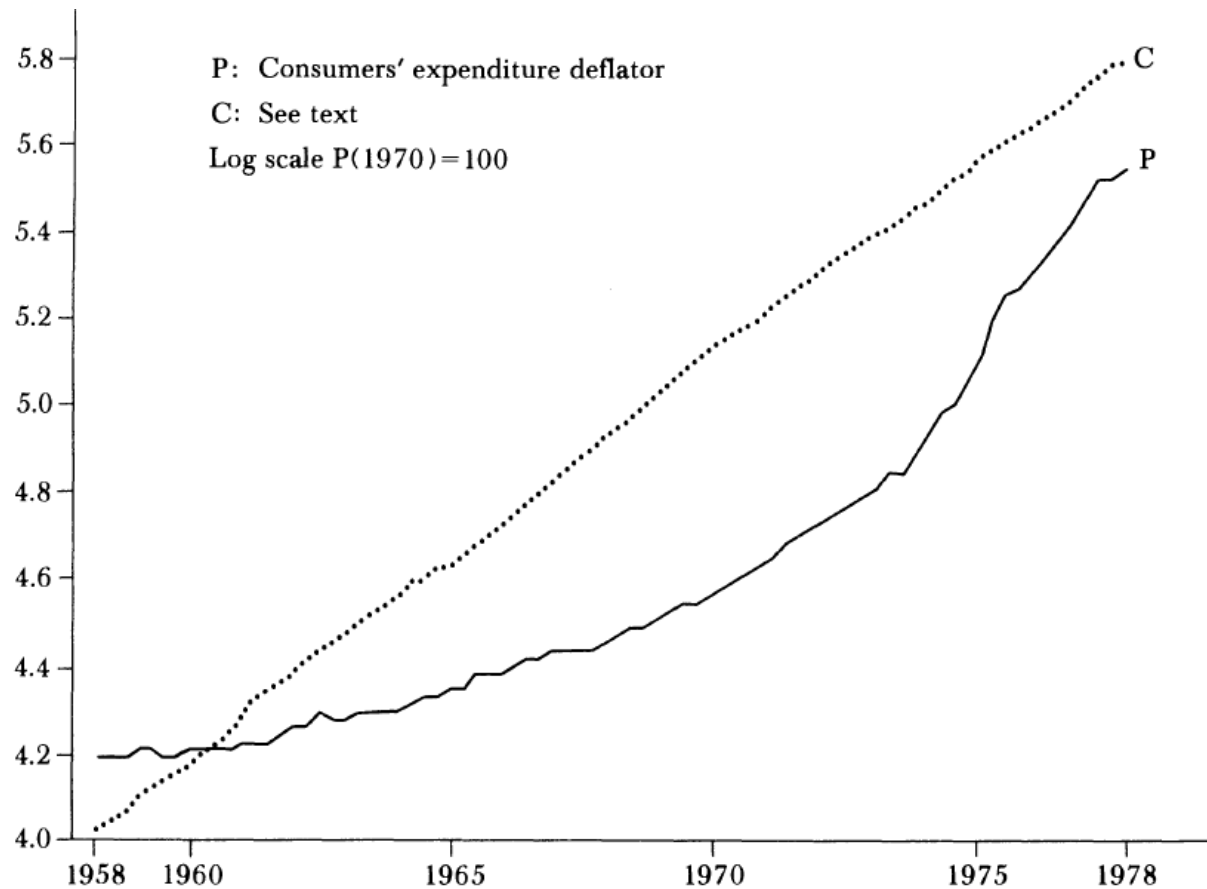


FIGURE 5

Besmislena regresija (III), pri čemu je C – kumulativna količina padavina

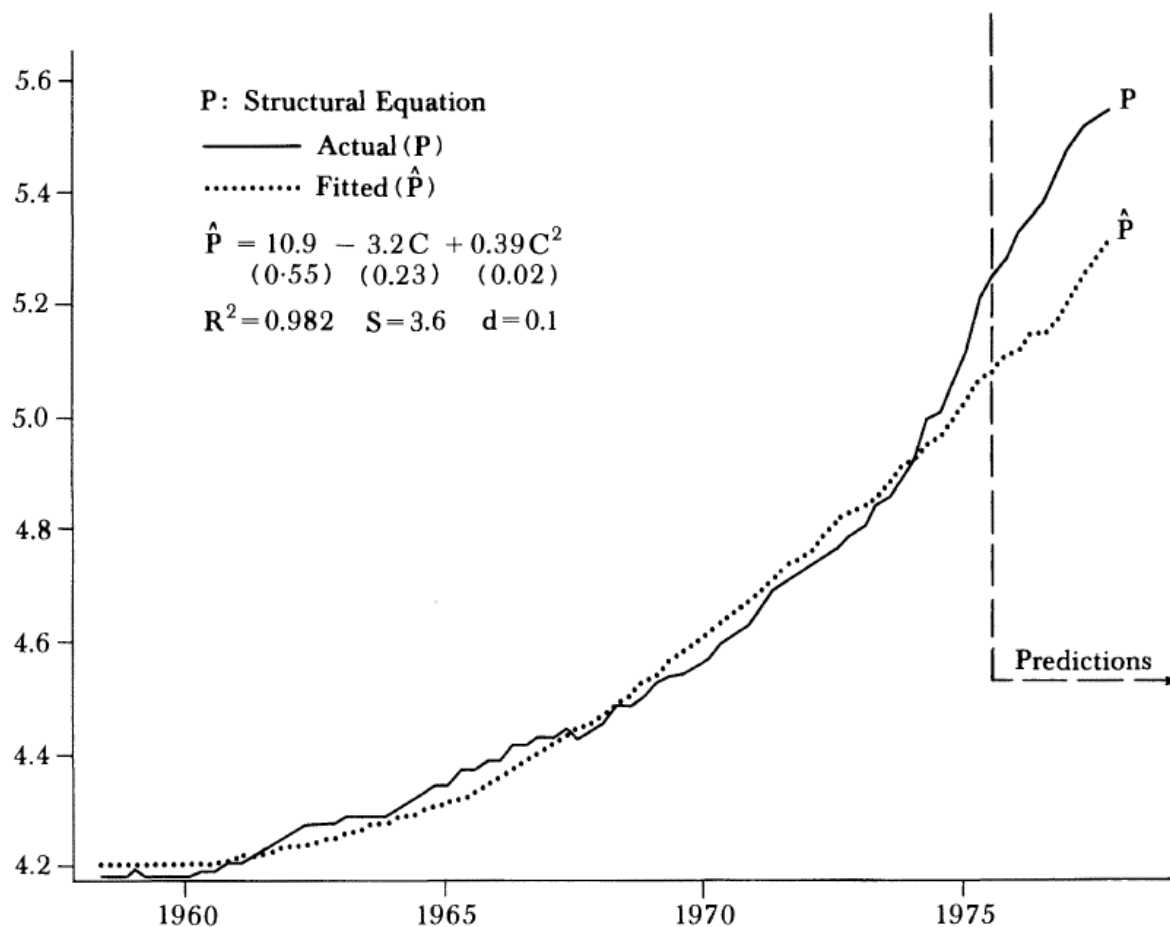


FIGURE 7

Rezidualna suma kvadrata

○ Pokazati...

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n y_i^2 - b^2 \sum_{i=1}^n x_i^2$$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n y_i^2 - b \sum_{i=1}^n x_i y_i.$$

Koeficijet korelacije

- Mera linearne korelisanosti slučajnih promenljivih:

$$r_{XY} = \frac{\hat{\text{cov}}(X, Y)}{\sqrt{\hat{\text{var}}(X) \hat{\text{var}}(Y)}}.$$

- Znak koeficijenta korelacije odgovara znaku ocene b (pokazati...)
- U jednostavnoj regresiji vazi:

$$r_{XY}^2 = R.^2$$

Pokazati...

Koeficijent korelacije vs koef. determinacije

- Koeficijent korelacije određuje stepen linearne zavisnosti između X i Y, ali ne i **pravac uzročnosti!**
- Koeficijent determinacije se koristi u regresionoj analizi – određen je smer uzročnosti.
- U AVS obratiti pažnju na pojavu **lažne korelacije!!!**

Pretpostavke klasičnog linearnog regresionog modela (KLRM)

○ Pretpostavke KNLRM/KLRM o ε_i :

1. $E(\varepsilon_i)=0$, za svako i .
2. $Var(\varepsilon_i)=E(\varepsilon_i^2)=\sigma^2$, za svako i .
3. $Cov(\varepsilon_i, \varepsilon_j)=E(\varepsilon_i \varepsilon_j)=0$, za svako i, j , tako da $i \neq j$.
4. $E(\varepsilon_i X_i)=0$, za svako i .
5. $\varepsilon_i \sim N(0, \sigma^2)$.

Odatle: Y_i je **slučajna promenljiva**;

$$Y \sim N(\beta_0 + \beta X_i, \sigma^2)$$

Promenljive u KLRM

- 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***



Značenje pretpostavki KLRM

- 1) Slučajna greška u proseku ne utiče na kretanje zavisne promenljive Y_i
- 2) Homoskedastičnost – konstantna varijansa slučajne greške
- 3) Greške nisu međusobno korelisane (neautokorelisane) za različite opservcije

Značenje pretpostavki KLRM (nastavak)

- 4) Nekorelisanost objašnjavajuće prom. X_i i slučajne greške ε_i upućuje na zaključak da je X_i egzogena promenljiva (nije stohastičke prirode) i uzima fiksirane vrednosti iz ponovljenih uzoraka
- 5) ε_i obuhvata dejstvo velikog broja međusobno nezavisnih nepredvidivih uticaja. CGT-zbir velikog broja međusobno nezavisnih slučajnih promenljivih je slučajna promenljiva, čija raspodela teži normalnoj sa povećanjem broja sabiraka (od vitalne važnosti za **donošenje statističkih zaključaka**)

Implikacije navedenih pretpostavki

- Ocena b je linearna funkcija slučajne promenljive Y_i .
- Posledice:
 - **Ocena b je slučajna promenljiva**
 - **Ocena b ima normalnu raspodelu**

Kako merimo preciznost ocena?

- Svaki drugi uzorak daje nove ocene parametara b_o i b . 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 $\hat{\sigma}^2$ (s^2)

Ocena varijanse slučajne greške modela (σ^2) i ocene varijansi ocena parametara b_0 i b

- Nepristrasna ocena σ^2 je:

$$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$$

- Sada možemo da analiziramo ocene varijansi ocena parametara b_0 i b :

$$\hat{\text{var}}(b_0) = s_{b_0}^2 = s^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum x_i^2} \right)$$

$$\hat{\text{var}}(b) = s_b^2 = \frac{s^2}{\sum x_i^2}$$

Standardne greške ocena parametara biće manje ukoliko je:

- Manji stepen stohastičnosti između X_i i Y_i , odnosno manja ocena varijanse s^2 (manji varijabilitet modela).
- Veći varijabilitet objašnjavajuće promenljive X_i (suma kvadrata odstupanja X od aritmetičke sredine).
- Veći uzorak (n).
- 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).

Statističko zaključivanje u KLRM

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

- Primer: Ocenjen je model oblika:

$$\hat{Y}_i = 15.12 + 0.61X_i$$

(6.57) (0.04)

- Ocena 0.61 je (tačkasta) nepoznatog parametra nagiba. Koliko je ta ocena pouzdana?
- Odgovor na to pitanje daje standardna greška ocene.

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:

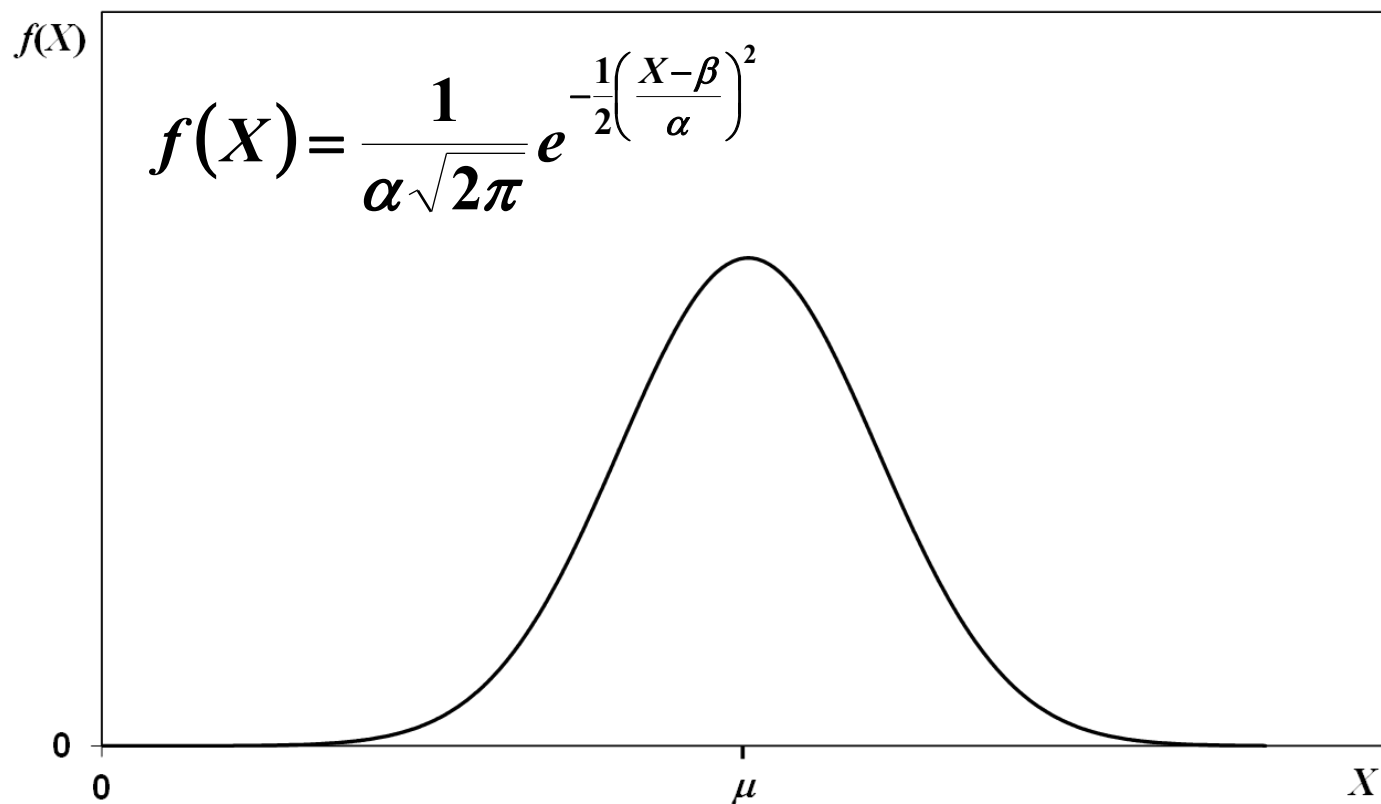
$$H_0 : \beta = 1$$

$$H_1 : \beta \neq 1$$

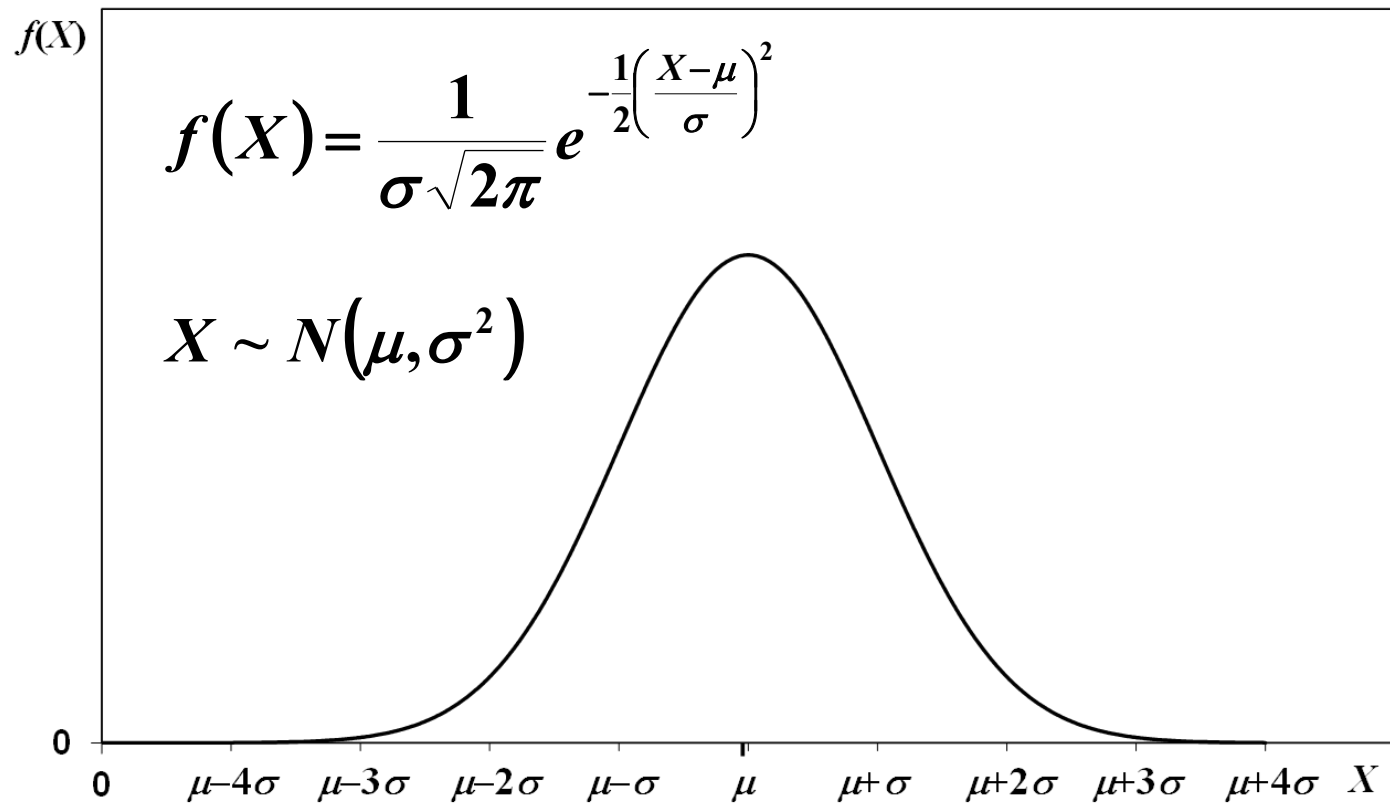
Neke teorijske raspodele (ponavljanje)

- U postupku statističkog zaključivanja korišćićemo sledeće četiri raspodele: normalnu, χ^2 -raspodelu, t-raspodelu i F-raspodelu

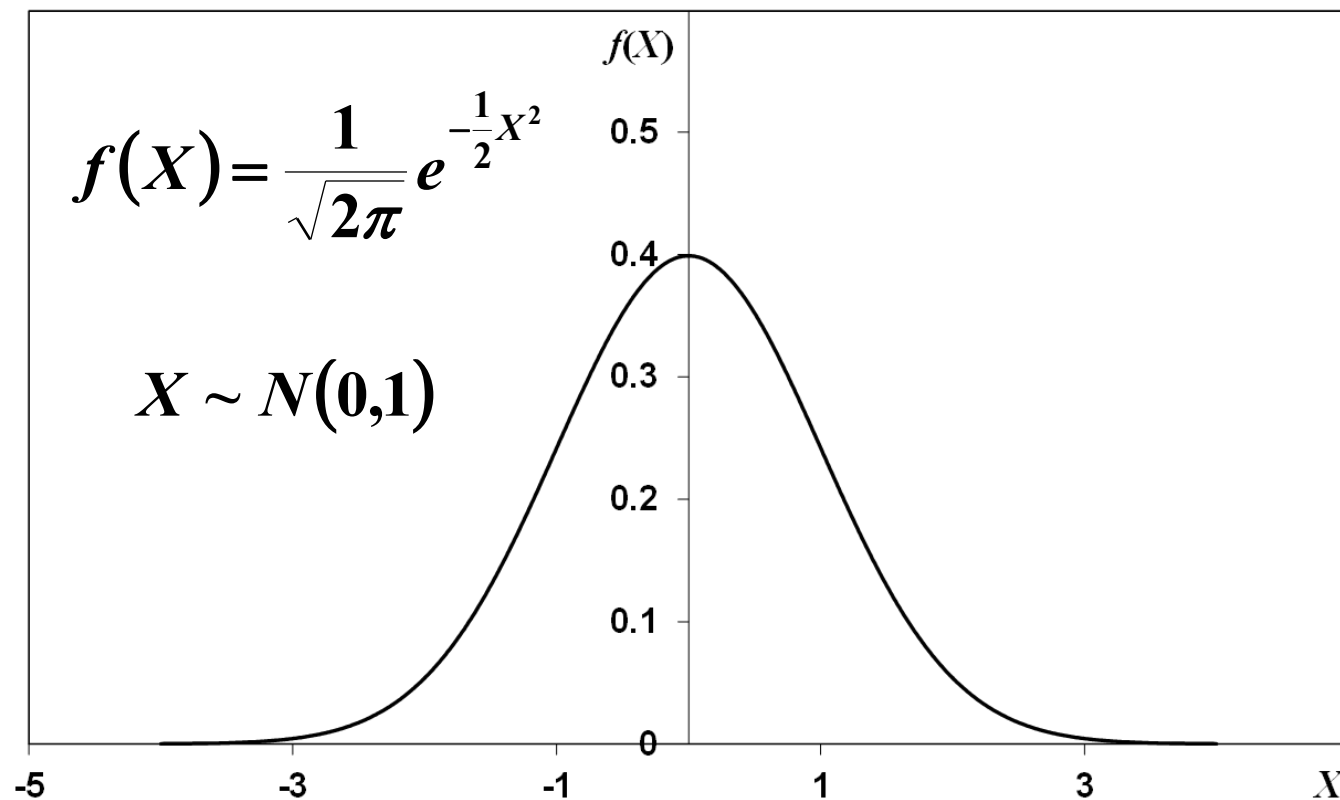
Funkcija gustine normalne raspodele



Normalna raspodela



Standardizovana normalna raspodela



χ^2 -raspodela

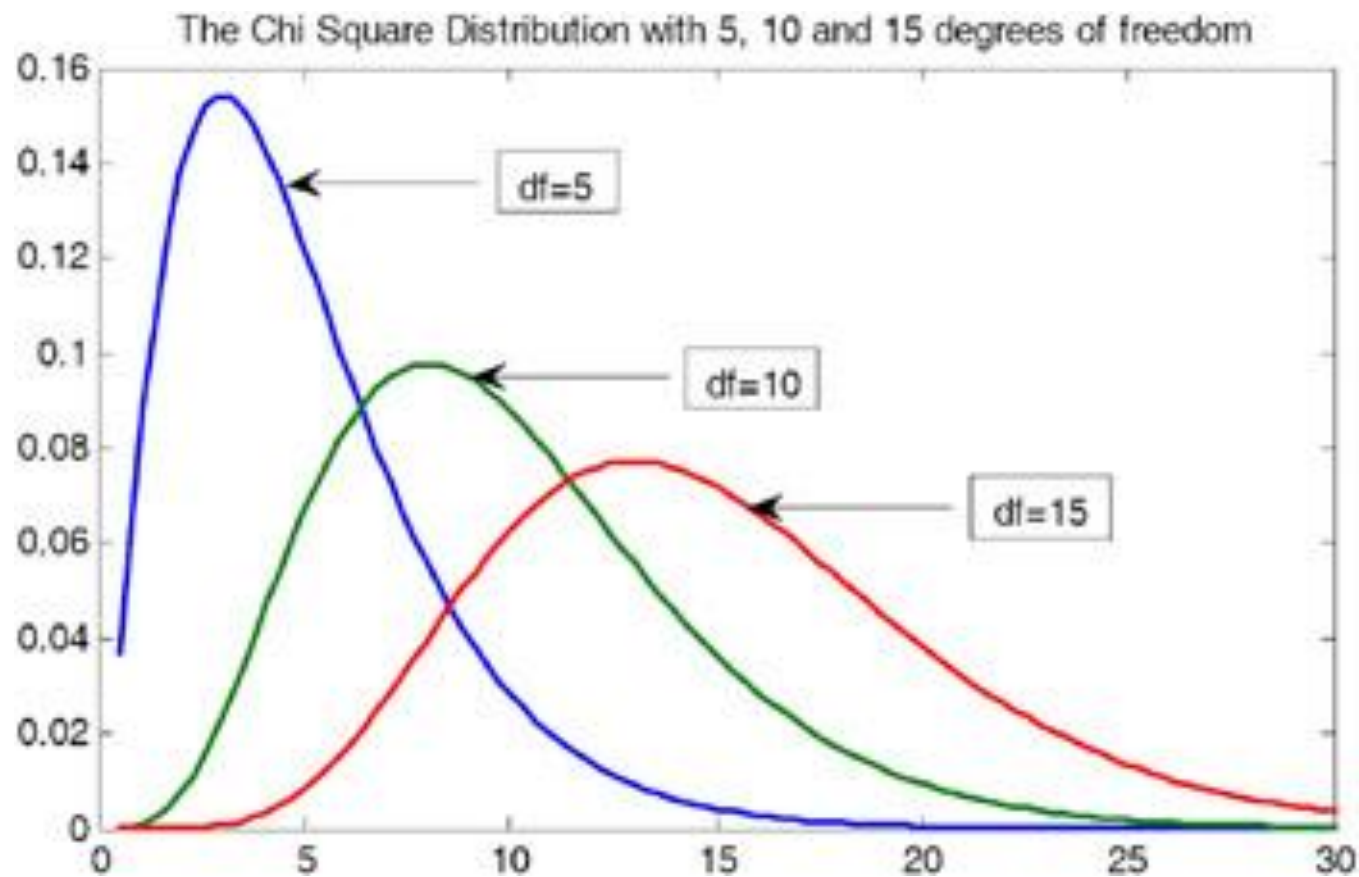
- Definiše se na sledeći način. Neka su $Z_1, Z_2, \dots, Z_n$ nezavisne i standardizovane normalno raspodeljene slučajne promenljive, onda slučajna promenljiva definisana kao njihov zbir kvadrata:

$$Z_1^2 + Z_2^2 + \dots + Z_n^2 = \sum_{i=1}^n Z_i^2,$$

ima χ^2 raspodelu sa n stepeni slobode

- Uobičajeno se koristi notacija: $\sum_{i=1}^n Z_i^2 \sim \chi_n^2$, pri čemu ovu raspodelu karakteriše očekivana vrednost n i varijansa jednaka $2n$.

Funkcija gustine χ^2 raspodele



Studentova t - raspodela

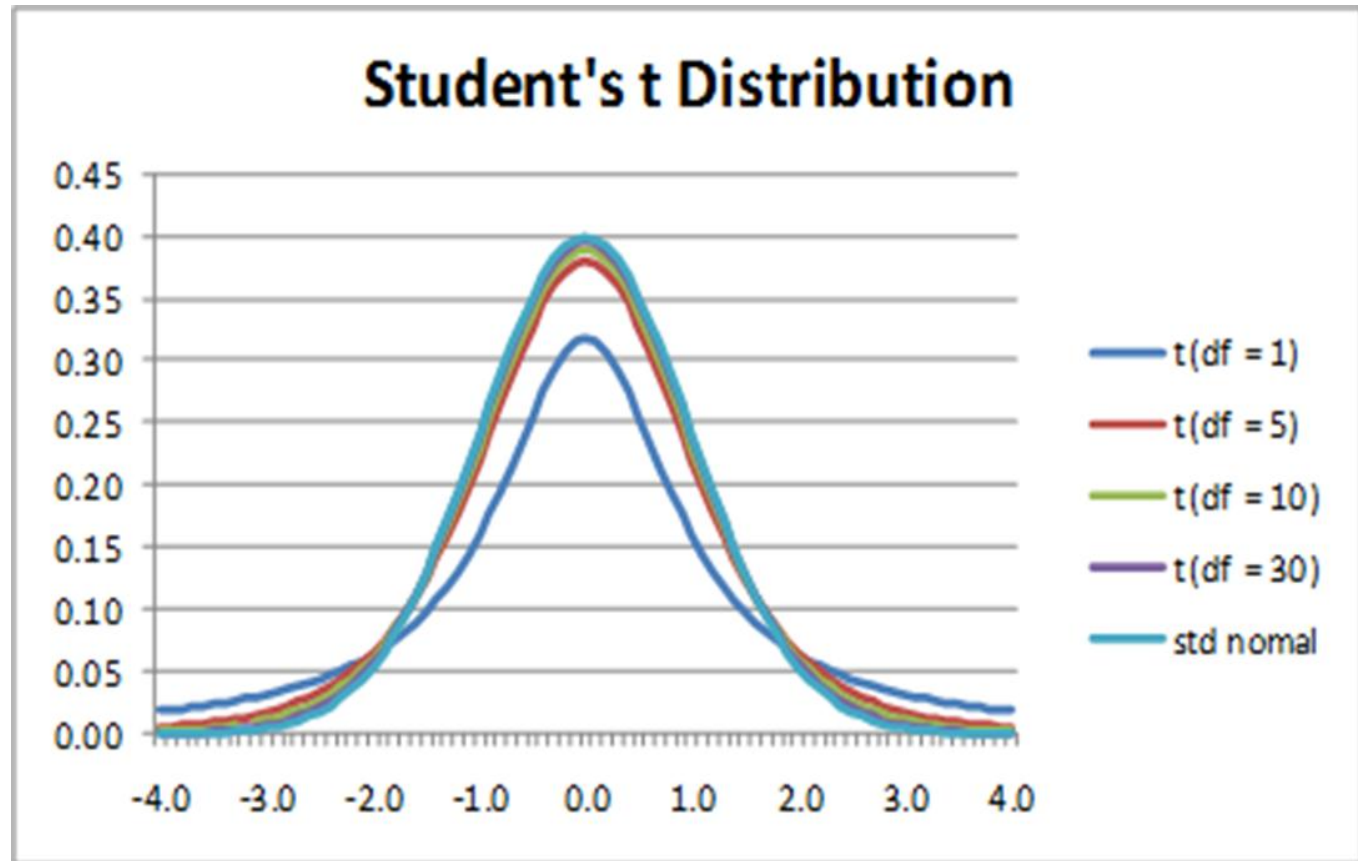
- Neka Z_0 poseduje standardizovanu normalnu raspodelu, $Z_0 \sim N(0, 1)$, i $\sum_{i=1}^n Z_i^2 \sim \chi_n^2$, pri čemu su slučajne promenljive nezavisne, tada količnik:

$$t = \frac{Z_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n Z_i^2}},$$

poseduje t -raspodelu sa n stepeni slobode. Poput standardizovane normalne raspodele, t - raspodela je simetrična oko nule, ali ima "teže" repove (uočljivo za malo n)

- Dakle, očekivana vrednost t -raspodele je 0, dok je varijansa $n/(n-2)$

Funkcija gustine t - raspodele



F – raspodela

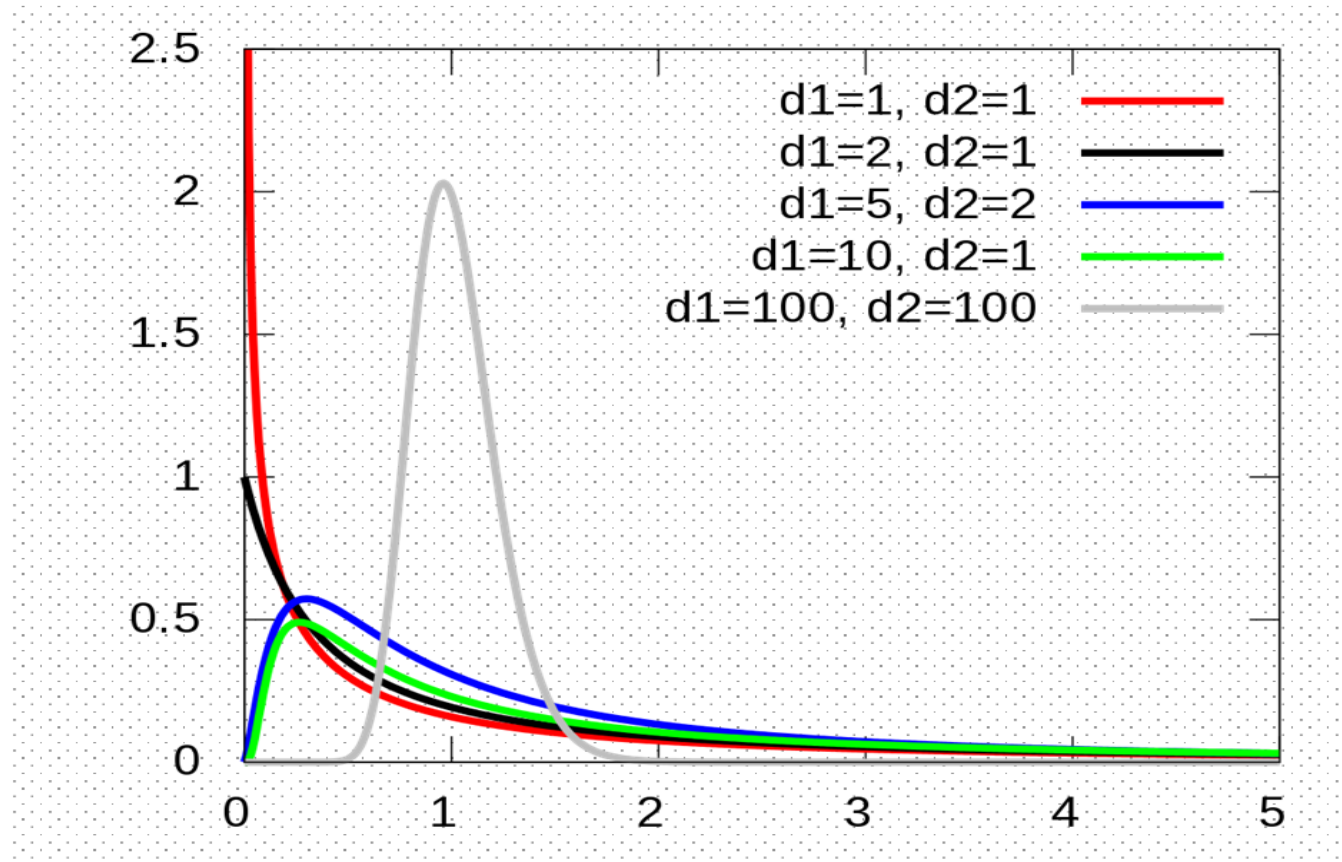
- Neka je $U \sim \chi_m^2$ i $V \sim \chi_n^2$, pri čemu su U i V nezavisne slučajne promenljive, tada sledeći količnik:

$$F = \frac{U/m}{V/n}$$

poseduje **F-raspodelu** sa m (d_1) i n (d_2) stepeni slobode (zapis: F_n^m)

- F – raspodela je količnik dve nezavisne χ^2 raspodele, podeljene odgovarajućim brojem stepeni slobode

Funkcija gustine F - raspodele



Raspodela verovatnoće ocena dobijenih metodom ONK

- Standardizovanjem slučajnih promenljivih b i b_0 dobijamo:

$$\frac{b - \beta}{\sqrt{\text{var}(b)}} : N(0,1), \frac{b_0 - \beta_0}{\sqrt{\text{var}(b_0)}} : N(0,1)$$

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

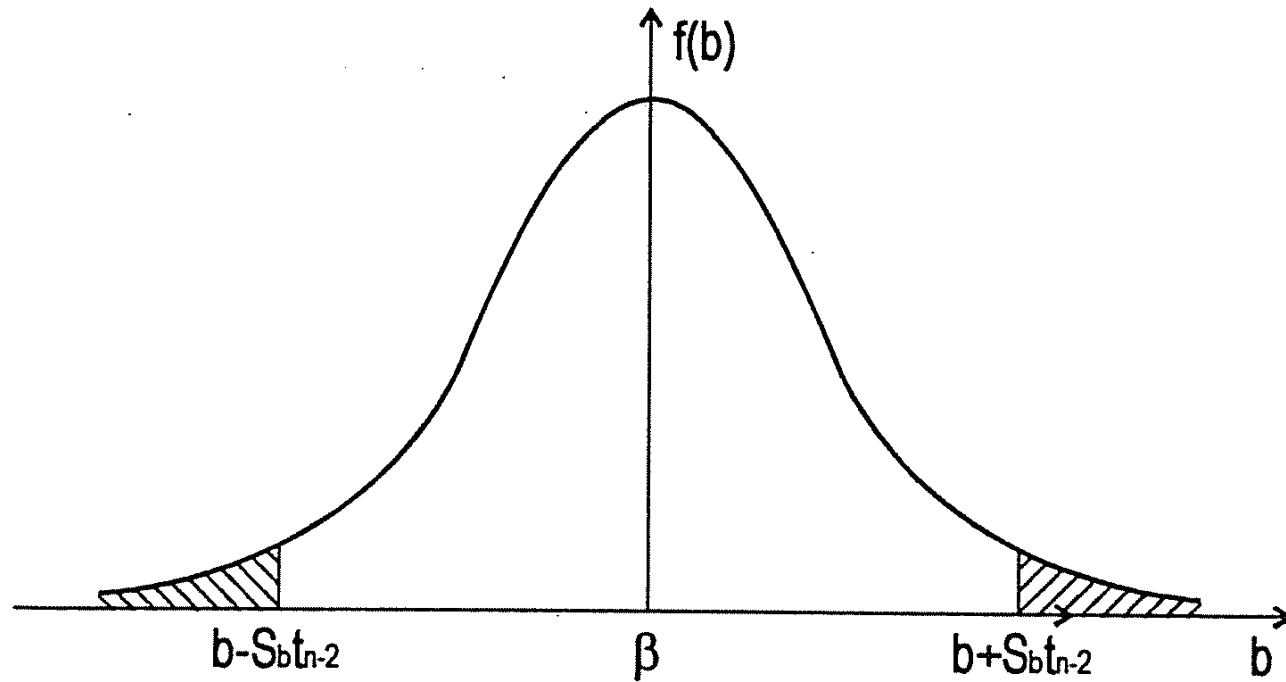
$$\frac{b - \beta}{s_b} : t(n-2), \frac{b_0 - \beta_0}{s_{b_0}} : t(n-2).$$

Interval poverenja za nepoznate parametre

- Na osnovu rezultata o t -raspodeli, moguće je odrediti granice intervala poverenja za parametre β_0 i β sa odgovarajućom verovatnoćom.
- Intervali poverenja nepoznatih parametara β_0 i β na nivou značajnosti α su :

$$b - s_b t_{n-2}(\alpha/2) \leq \beta \leq b + s_b t_{n-2}(\alpha/2)$$
$$b_0 - s_{b_0} t_{n-2}(\alpha/2) \leq \beta_0 \leq b_0 + s_{b_0} t_{n-2}(\alpha/2).$$

Interval poverenja za β



Testiranje hipoteza: algoritam

- Posmatramo model oblika:

$$Y_i = \beta_0 + \beta X_i + \varepsilon_i, \text{ za } i = 1, 2, \dots, n.$$

- Testiramo validnost hipoteze:

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

- Koraci u postupku testiranja:

1. Ocenjujemo: b, b_0, s_b, s_{b_0} na poznati način.
2. Racunamo test-statistiku koristeći sledeću formulu:

$$\text{test} - \text{statistika} = \frac{b - \beta^*}{s_b} : t(n - 2),$$

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

Testiranje hipoteza: algoritam (nastavak)

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

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

$$\left| \frac{b - \beta^*}{s_b} \right| > t_{(n-2)}(0.025),$$

Odbacujemo H_0 kao netačnu uz nivo značajnosti 5%.

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

Specijalni tip hipoteze: *t*-odnos

- Pretpostavimo da nas interesuje:

$$H_0: \beta = 0, H_1: \beta \neq 0.$$

Ako je tačna nulta hipoteza, tada objašnjavajuća promenljiva ne utiče na kretanje zavisne promenljive. Na ovaj način proveravamo opravdanost postavke modela.

- U tom slučaju opšti oblik test statistike postaje ***t*-odnos**, zapravo odnos ocene i odgovarajuće standardne greške ocene:

$$\text{test} - \text{statistika} = t - \text{odnos} = \frac{b}{s_b} : t(n-2).$$

Ispitivanje kvaliteta regresije na osnovu koeficijenta determinacije:

- Hipoteze od interesa:

$$H_0 : R^2 = 0$$

$$H_1 : \text{hipoteza } H_0 \text{ nije tacna} \Leftrightarrow H_1 : R^2 \neq 0$$

- Relevantna statistika:

$$F_{n-2}^{2-1} = \frac{R^2 / (2-1)}{(1-R^2) / (n-2)} = t_b^2 \quad .$$

- Pravilo odlučivanja: Ako je izračunata vrednost date statistike veća od kritične vrednosti F -raspodele sa 1 i $n-2$ stepeni slobode, tada se nulta hipoteza odbacuje uz izabrani nivo značajnosti



Svojstva ocena na malim uzorcima (uzoračka svojstva ocena)

- Nepristrasnost
- Efikasnost
- Linearnost

Nepristrasnost/Efikasnost/Linearnost

- Nepristrasnost

Ocene metoda ONK su nepristrasne. To znaci da su ocene **u proseku jednake** parametrima koji se ocenjuju:

$$E(b_0)=\beta_0 \text{ i } E(b)=\beta$$

- 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.

- Linearnost

Linearna je ocena koja se dobija kao linearna funkcija elemenata iz uzorka (uslov koji pojednostavljuje problem određivanja efikasne ocene).

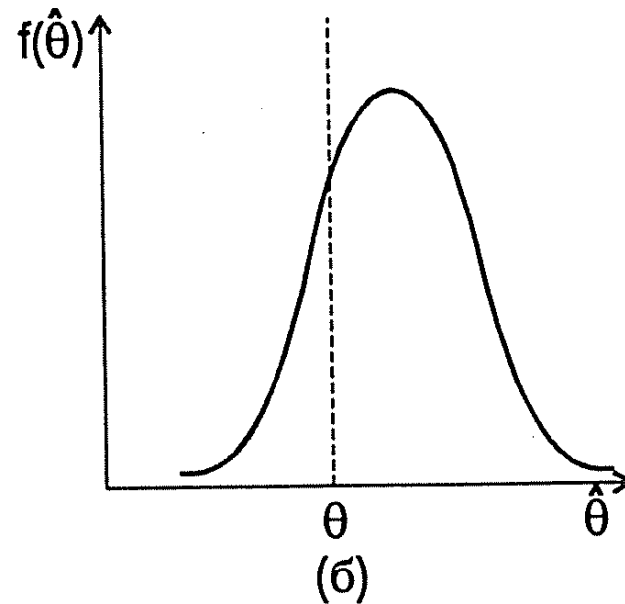
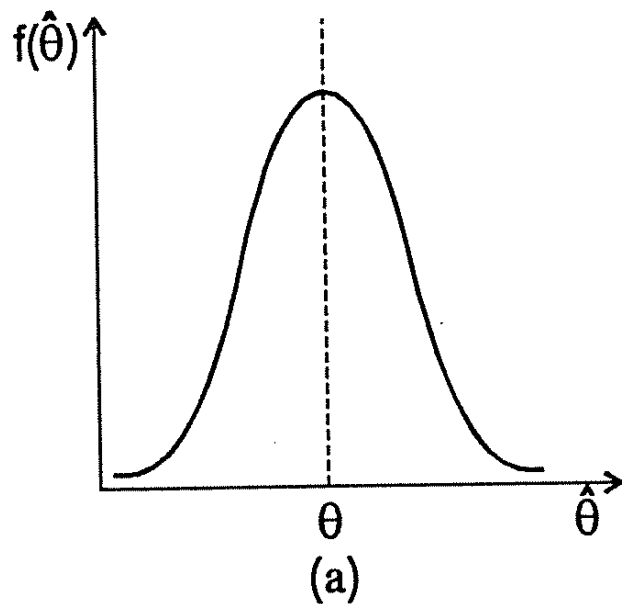
1) Nepristrasnost (centriranost)

- Ocena je nepristrasna ako je njena srednja vrednost jednaka parametru koji se ocenjuje.

$$E\left(\hat{\theta}\right) = \theta.$$

- Pojedinačno se svaka ocenjena vrednost izračunata na bazi većeg broja uzoraka istog obima može razlikovati od stvarne vrednosti parametra θ .

F-ja gustine nepristrasne i pristrasne ocene



Varijansa ocene

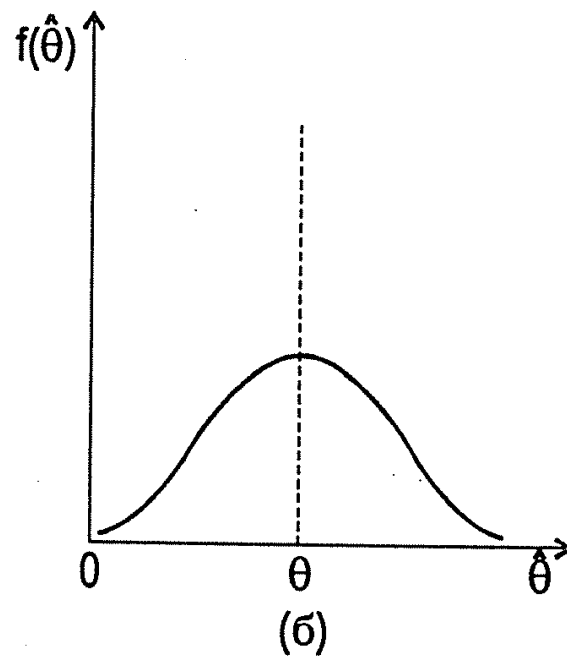
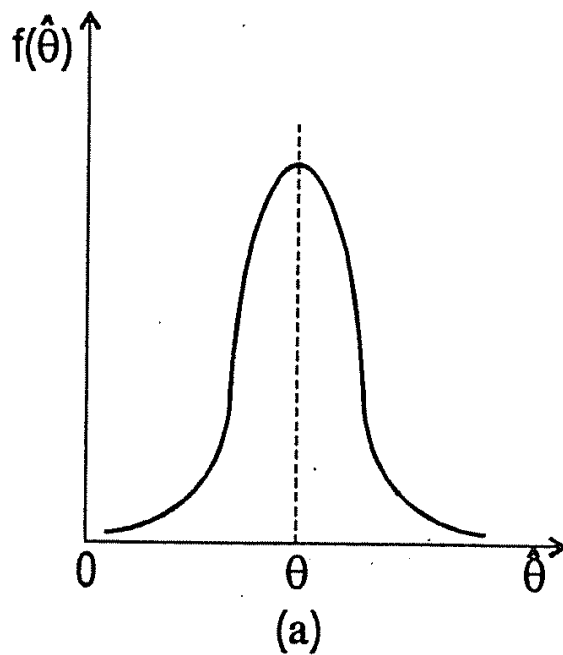
- U praksi se ocena nepoznatog parametra dobija na bazi jednog uzorka.
- Postoji verovatnoća da se i nepristrasna ocena $\hat{\theta}$ u izvesnoj meri razlikuje od θ .
- Pitanje **širine intervala** u čijim granicama varira ova ocena.
- Varijansa ocene se definiše:

$$\text{var}\left(\hat{\theta}\right) = E\left(\hat{\theta} - E\left(\hat{\theta}\right)\right)^2.$$

2) Efikasnost

- Ocena $\hat{\theta}$ je efikasna ocena parametra θ , ako je:
 - a) data ocena nepristrasna i
 - b) ne postoji druga nepristrasna ocena sa manjom varijansom.
- Efikasna je ona ocena koja ima najmanju disperziju oko prave vrednosti parametra.
- **Efikasna ocena** se često naziva i najbolja nepristrasna ocena (**najbolja** je ona ocena koja **ima najmanju varijansu**).

F-ja gustine relativno efikasne i neefikasne ocene



3) Linearnost

- Uslov koji pojednostavljuje problem određivanja efikasne ocene.
- Linearna je ocena koja se dobija kao linearna funkcija elemenata iz uzorka:

$$\hat{\theta} = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n,$$

gde su $\alpha_1, \alpha_2, \dots, \alpha_n$ parametri.

Najbolja linearna nepristrasna ocena (NLNO)

- Ocena $\hat{\theta}$ je NLNO ukoliko zadovoljava sledeće uslove:
 - a) linearna ocena,
 - b) nepristrasna ocena i
 - c) ocena sa najmanjom varijansom.
- Engl. akronim BLUE (engl. *Best Linear Unbiased Estimator*).

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: b je ocena stvarne vrednosti parametra β
 - Linearna: b je linearna funkcija raspoloživih podataka
 - Nepriistrasna: u proseku ocena b je jednaka parametru β
 - Najbolja: ocena je efikasna (nepriistrasna ocena sa najmanjom varijansom)

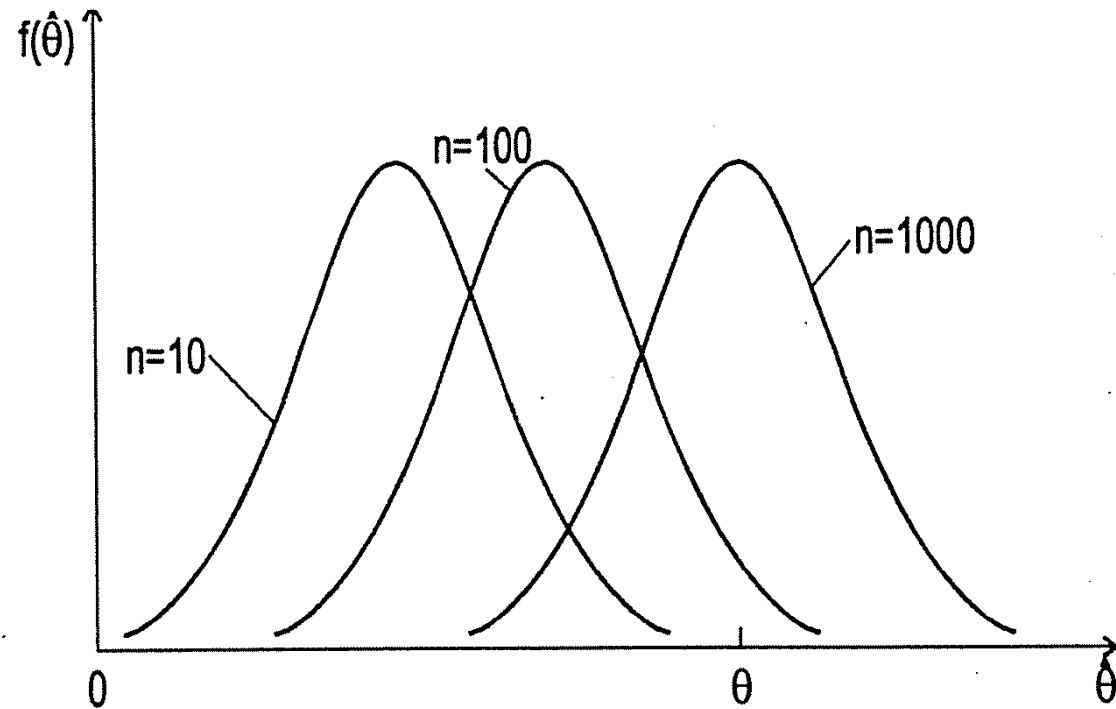
Asimptotska svojstva ocena

- **Asimptotska nepristrasnost:** $E(b) \rightarrow \beta$, za $n \rightarrow \infty$.
 - Ocena je asim. nepristrasna ako postaje nepristrasna sa rastom uzorka
- **Konzistentna** - ako konvergira u verovatnoći ka pravoj vrednosti parametra, kada n teži ka nuli:

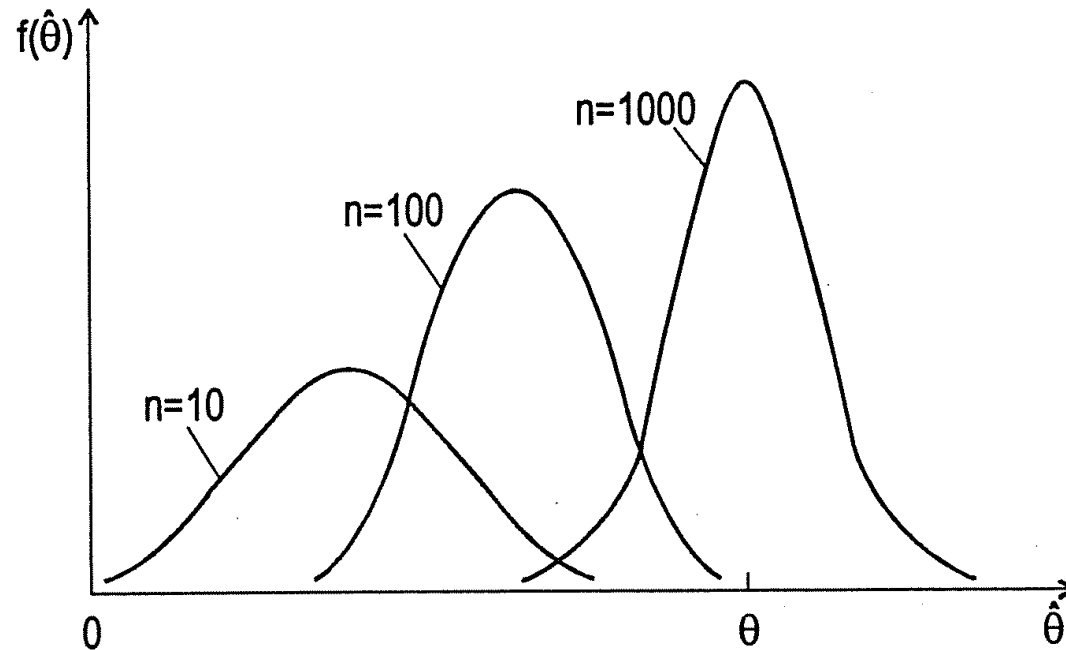
$$p \lim(b) = \beta.$$

- Konzistentna: varijansa i pristrasnost (SKG) teže ka nuli kada obim uzorka teži ka beskonačnosti
- **Asimptotska efikasnost:** konzistentna ocena sa najmanjom asimptotskom varijansom (najbrže konvergira u verovatnoći ka β)

Asimptotski nepristrasna ocena

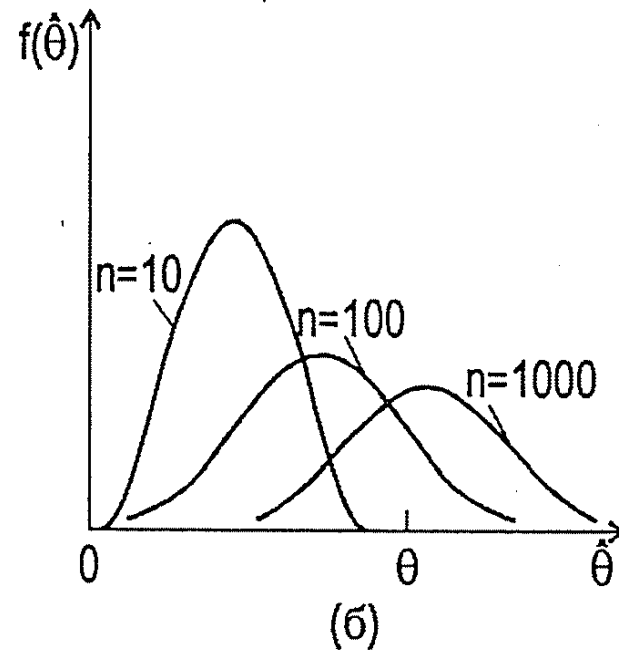
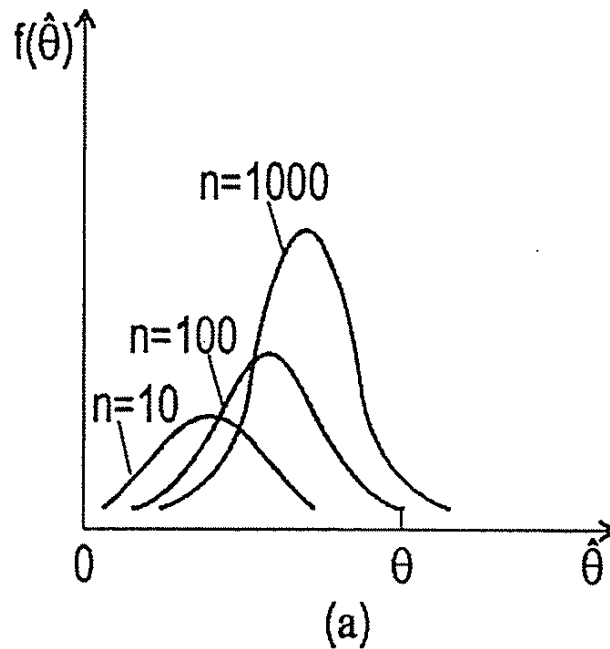


F-ja gustine konzistentne ocene

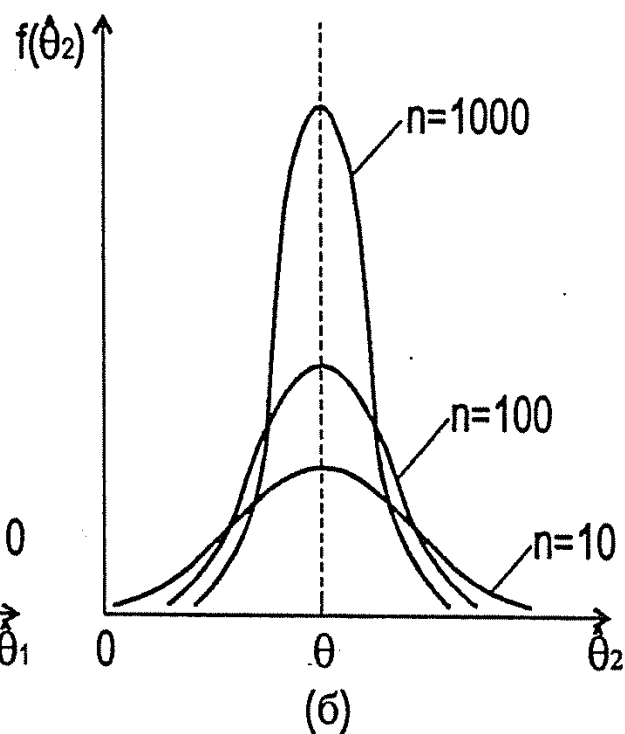
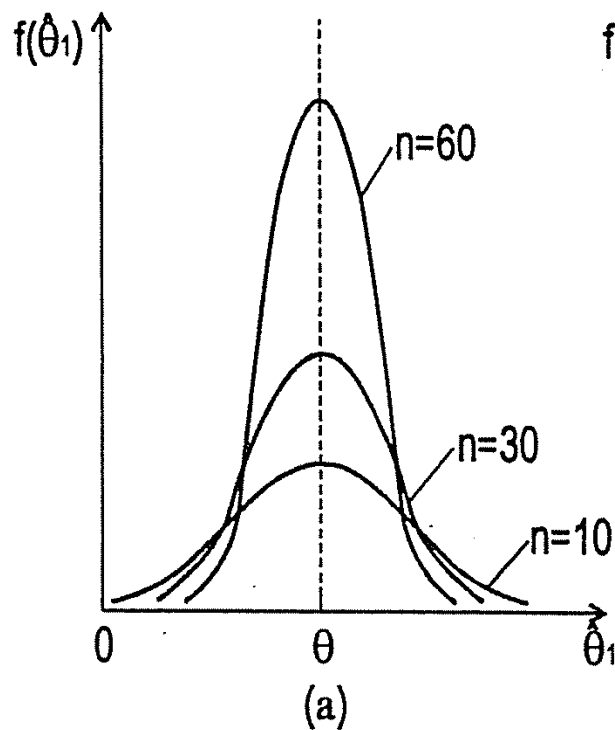


F-ja gustine nekonzistentne ocene

- Ocena je nekonzistentna ($\text{plim}(\hat{\theta}) \neq \theta$) kada sa rastom uzorka
 - a) smanjuje varijansa, ali ne i pristrasnost ili
 - b) smanjuje pristrasnost, ali ne i varijansa



Uzoračke raspodele nepristrasnih ocena $\hat{\theta}_1$ i $\hat{\theta}_2$

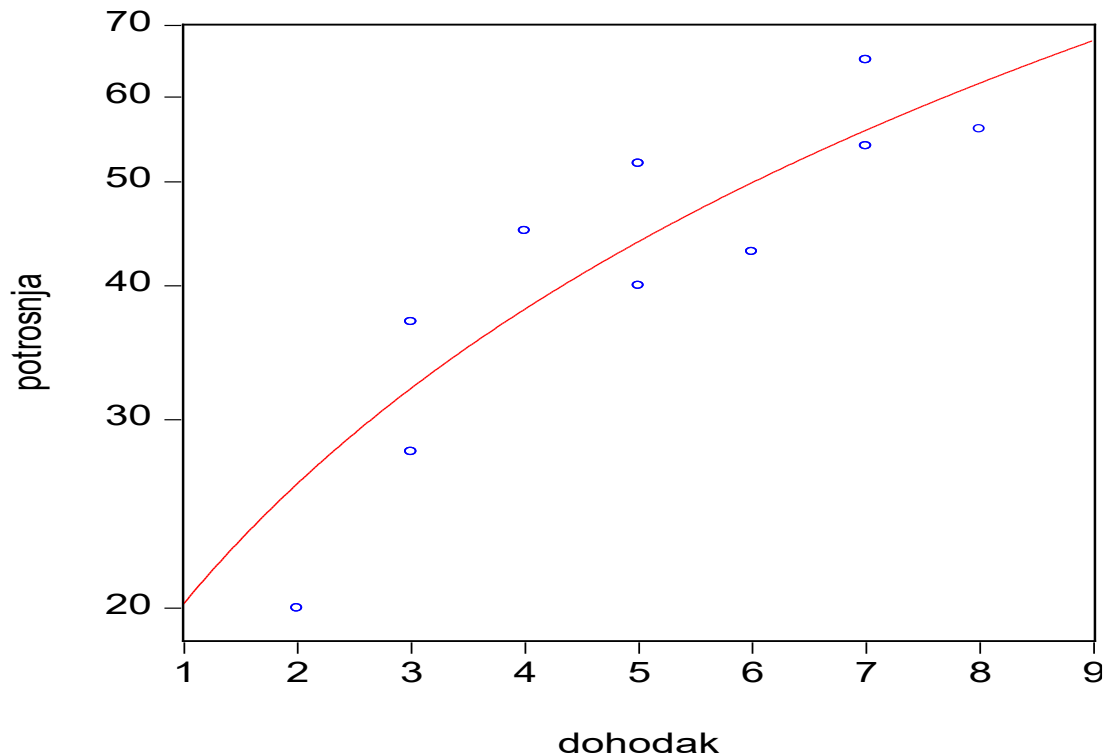


Osobine ocena dobijenih metodom NK i metodom MV

- Ocene dobijene metodom NK imaju ***sve poželjne osobine u malim uzorcima***.
- Metod maksimalne verodostojnosti (MV) daje ocene parametara koje imaju poželjne asimptotske osobine: konzistentost i asim. efikasnost.
- Metod MV se koristi kada su na raspolaganju veliki uzorci i kada se pretpostavka o normalnoj distribuciji grešaka može smatrati opravdanom (u opštem slučaju ***pristrasne u malim uzorcima***).

Primer nelinearne zavisnosti

- Ispitujemo zavisnost između potrošnje i dohotka.
- Dijagram rasturanja sugerije da je veza ove dve promenljive nelinearne prirode?



Dvojno-logaritamski (log-log) model

- Najveću primenu ima nelinearni model oblika:

$$Y = \beta_0 X^\beta ,$$

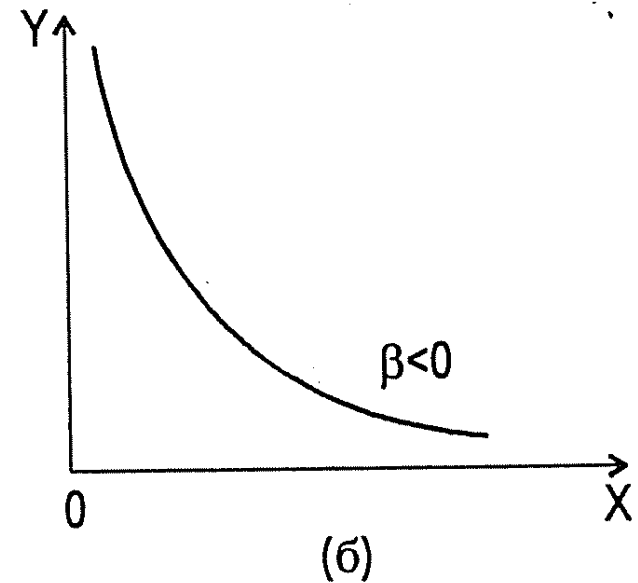
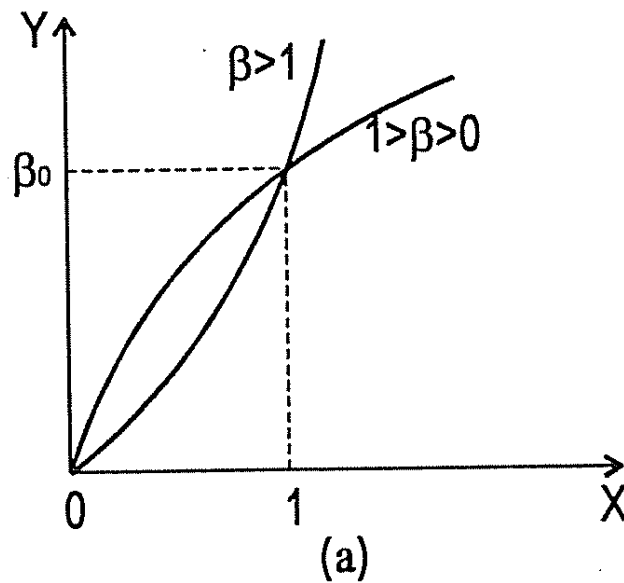
gde su β_0 i β parametri.

- Logaritmovanjem i uvođenjem odgovarajućih smena model se prevodi na linearni oblik:

$$\underbrace{\ln Y}_{Y^*} = \underbrace{\ln \beta_0}_{\beta_0^*} + \beta \underbrace{\ln X}_{X^*} .$$

- Postupak ocenjivanja...
- Interpretiranje dobijene vrednosti R^2 ...

Grafički prikaz log-log modela za različite vrednosti parametra β



Interpretacija parametra β log-log modela

- Parametar β predstavlja koeficijent elastičnosti Y u odnosu na X :

$$\beta = \frac{\partial Y / Y}{\partial X / X} = \frac{\text{relativna (procentua ln a) promena } Y}{\text{relativna (procentua ln a) promena } X}$$

- **Rast X za 1% dovodi do promene (rasta ili pada) Y za β %**
- Primena u ekonomskoj analizi: zavisnost između tražnje i cena, potrošnje i dohotka, Cobb-Douglas-ova proizvodna funkcija i dr.