Osnovi ekonometrije – Glava 8

Osnovne studije

Predavač: Aleksandra Nojković

Struktura predavanja

- **Narušavanje pretpostavki KLRM**

  - Heteroskedasticnost
  - Autokorelacija

Pretpostavke KLRM

1. $E(\varepsilon_i) = 0$

2**. Var($\boldsymbol{\varepsilon_i}$) = σ^2 = const.**

3. Cov $(\varepsilon_i, \varepsilon_j) = 0$ za *i* različito od *j*

4. Objašnjavajuće promenljive nisu određene stohastičkim članom

5. $\varepsilon_i \sim N(0,\sigma^2)$

6. Ne postoji tačna linearna zavisnost između objašnjavajućih promenljivih.

- **Narušavanje pretpostavki KLRM**

Pretpostavka 2: Var(ε_i) = σ^2 = const. Homoskedastičnost

- Homoskedastičnost: varijansa slučajne greške modela je konstantna za sve opservacije.

$$\mathrm{var}(\varepsilon_1) = \mathrm{var}(\varepsilon_2) = \ldots = \mathrm{var}(\varepsilon_n) = \sigma^2 = \mathrm{const}$$

- Heteroskedastičnost:

pretpostavka o homoskedastičnosti je narušena, što znači da se varijanse slučajnih greški razlikuju po pojedinim opservacijama:

$$\left.\begin{array}{l} \mathrm{var}(\varepsilon_1) = \sigma_1^2 \\ \mathrm{var}(\varepsilon_2) = \sigma_2^2 \\ \vdots \\ \mathrm{var}(\varepsilon_n) = \sigma_n^2 \end{array}\right\} \sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_n^2$$

Heteroskedastičnost

- Uvod: pojam heteroskedastičnosti
- Posledice: kakve su ocene dobijene metodom ONK?
- Otkrivanje prisustva heteroskedastičnosti
- Rešavanje problema heteroskedastičnosti

Pretpostavka 2: Var(ε_i) = σ^2 = const. Homoskedastičnost

- Homoskedastičnost: varijansa slučajne greške modela je konstantna za sve opservacije.

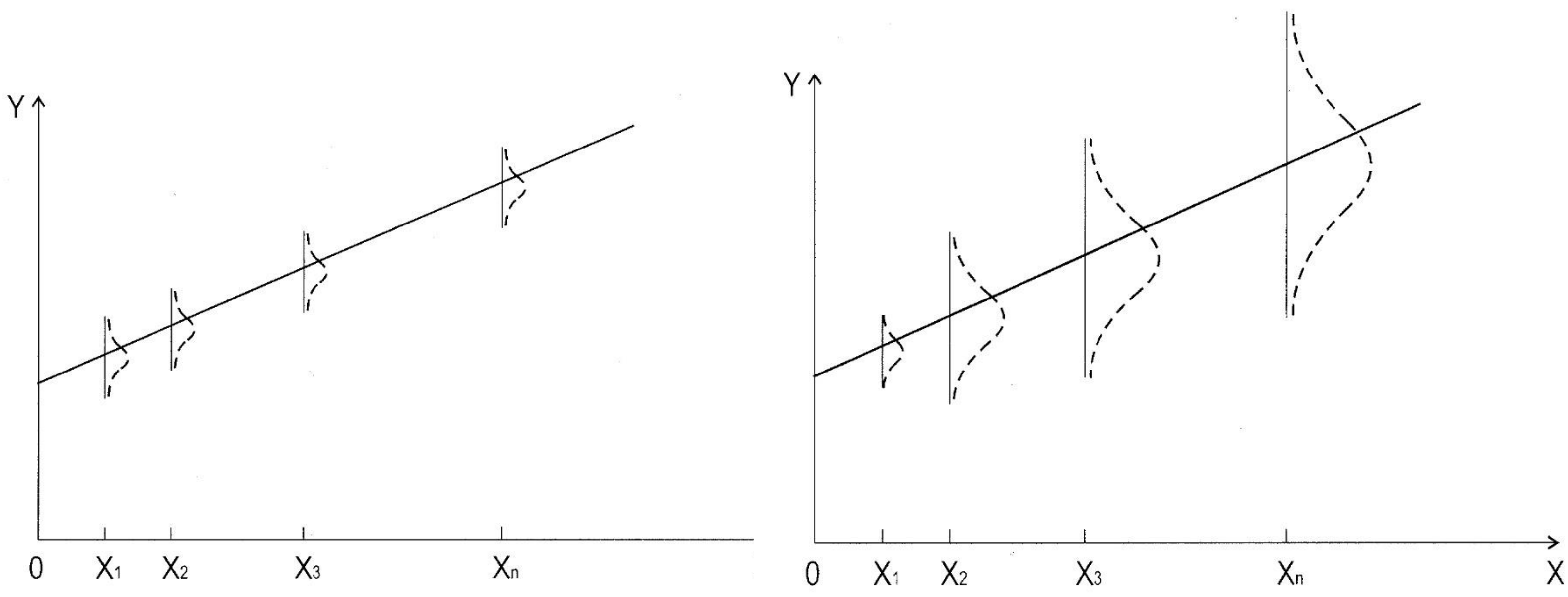
$$\mathrm{var}(\varepsilon_1) = \mathrm{var}(\varepsilon_2) = \ldots = \mathrm{var}(\varepsilon_n) = \sigma^2 = \mathrm{const}$$

- Heteroskedastičnost:

pretpostavka o homoskedastičnosti je narušena, što znači da se varijanse slučajnih greški razlikuju po pojedinim opservacijama:

$$\left.\begin{array}{l} \mathrm{var}(\varepsilon_1) = \sigma_1^2 \\ \mathrm{var}(\varepsilon_2) = \sigma_2^2 \\ \vdots \\ \mathrm{var}(\varepsilon_n) = \sigma_n^2 \end{array}\right\} \sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_n^2$$

Homosedastične (levo) i heteroskedastične (desno) greške



Posledice primene metoda ONK u prisustvu heteroskedastičnosti

- Primenom metoda ONK na model sa heteroskedastičnim greškama dobijaju se ocene koje nisu najbolje linearne nepristrasne ocene.
  - Ocene su nepristrasne
  - Ocene nisu efikasne–njihova varijansa nije najmanja moguća **(pokazati...).**

- Posledice:
  - Standardne greške ocena nisu precizna mera varijabiliteta ocena.
  - Standardne greške ocena najčešće potcenjuju stvarnu varijansu ocena parametara modela.
  - *t*-odnosi su nepouzdani.

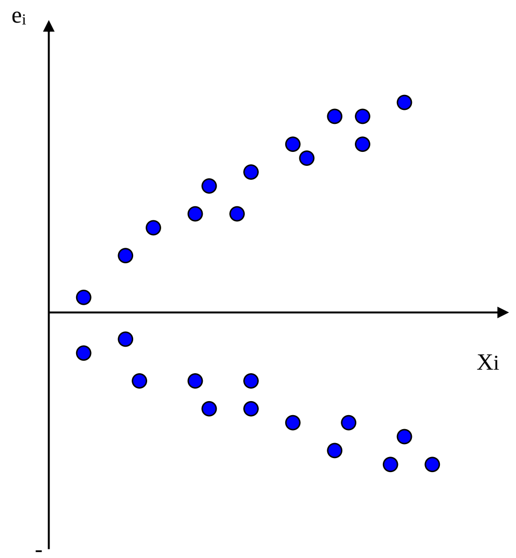
Kako se otkriva prisustvo heteroskedastičnosti u modelu?

1. Neformalni (grafički) metodi

2. Formalni metodi (testiranje)

Neformalni (preliminarni) metodi

Grafički prikazi: dijagram rasturanja reziduala (apsolutne vrednosti reziduala ili njihovih kvadrata) u odnosu na neku od objašnjavajućih promenljivih ili prema ocenjenoj vrednosti Yi (lin. kombinaciija svih objašnjavajućih promenljivih).



Testiranje postojanja heteroskedastičnosti (formalni testovi)

- Goldfeld-Kvantov (engl. Goldfeld-Quandt) test

- Glejzerov (engl. Glejser) test

- Vajtov (engl. White) test

Nulta hipoteza: $var(\varepsilon_i)=\delta^2=const.$, odnosno slucajne greške imaju stabilnu varijansu (**greške su homoskedastične**) ili **Alternativna hipoteza**: $var(\varepsilon_i)=\delta_i^2\neq const.$, odnosno varijansa slucajne greške nije konstantna (**greške su heteroskedastične**).

Goldfeld-Quandt-ov test

- Algoritam:

1. Pretpostavimo da je polazni model oblika:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i.$$

2. Opservacije poređati prema rastućem redosledu nezavisne promenljive.
3. Izostaviti jedna broj (c) centralnih opservacija (oko četvrtina).
4. Obaviti odvojeno regresije za prvih i poslednjih (n-c)/2 opsrevacija.
5. Statistika testa je: $F = \dfrac{\Sigma e_2^2}{\Sigma e_1^2} \sim F_{(n-c-2k)/2}^{(n-c-2k)/2},$

pri čemu se indeks 1 odnosi na reziduale dobijene za niže vrednosti regresora, a indeks 2 za više.

- Pogodan za modele sa malim brojem param. i velike uzorke.

Glejser-ov test

- <u>Algoritam:</u>

1. Iz polazne regresije računaju se reziduali e_i:
2. Ocenjuju se sledeće regresije:

$$\left|e_i\right| = c_0 + c_1 X_i^h + greška.$$

(parametar *h* najčešće:1,-1,1/2 i 2).

3. Testa se statistička značajnost ocene parametra c_1 primenom t-testa.
4. Upoređuju se koef. determinacije dobijeni za različite vrednosti *h*, a sam karakter heteroskedastičnosti određuje se prema regresiji sa najvećim R^2.

White-ov test

- Osnove testa:

Nulta hipoteza: slucajne greške imaju stabilnu varijansu
Alternativna hipoteza: varijansa slucajne greške je zavisna od objašnjavajućih promenljivih, njihovih kvadrata i međuproizvoda.

- Algoritam:

1. Pretpostavimo da je polazni model oblika:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i.$$

2. Ocenjujemo model iz 1., dobijamo reziduale i potom ocenjujemo pomocnu regresiju:

$$e_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{1i}^2 + \alpha_4 X_{2i}^2 + \alpha_5 X_{1i} X_{2i} + v_t$$

White-ov test (nastavak)

3. Nulta hipoteza se svodi na:

$$\mathrm{H}_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_5 = 0$$

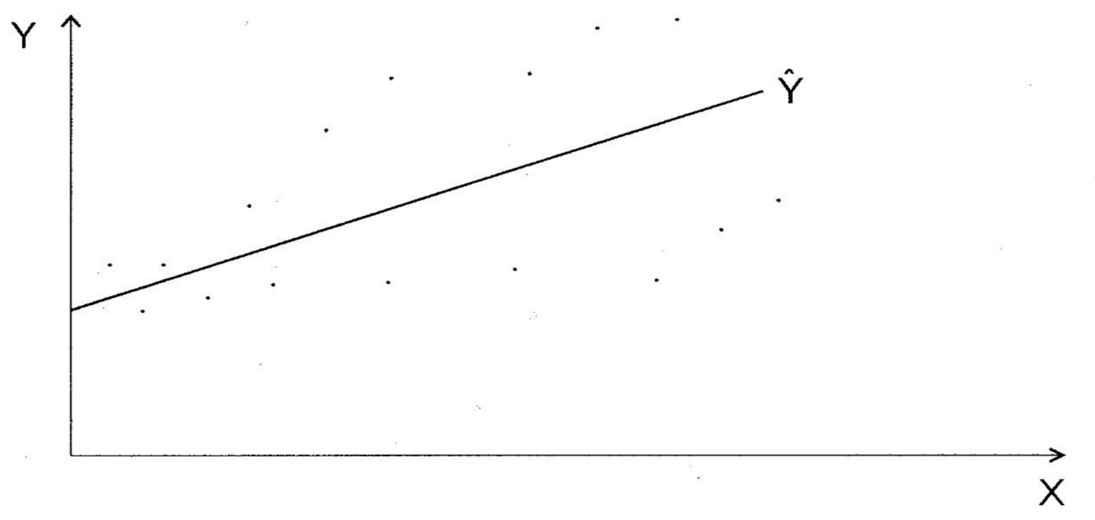
4. Odredujemo koeficijent determinacije R^2 *iz pomocne regresije i potom* ga množimo obimom uzorka *n*. *Whiteova test-statistika:* $WH=nR^2 \sim \chi^2$ sa *m stepeni slobode,* gde je *m* broj objašnjavajućih promenljivih pomoćne regresije bez slobodnog člana (*m=5*).

5. Ako je izracunata vrednost test-statistike veca od odgovarjuće kriticne vrednosti χ^2 testa na datom nivou znacajnosti tada se odbacuje nulta hipoteza o odsustvu heteroskedasticnosti.

- Istovetan postupak testiranja se sprovodi i u varijanti testa bez uključivanja nivoa promenljivih i/ili međuproizvoda u pomoćnu regresiju (predloženo za modele sa velikim brojem objašnjavajućih promenljivih i/ili za male uzorke, na razmatranom primeru m=2 ili m=4).

Kako se eliminiše uticaj heteroskedastičnosti (I) ?

- Primenjuje se metod ponderisanih najmanjih kvadrata (metod uopštenih najmanjih kvadrata).
- Ideja: u postupku minimiziranja sume kvadrata reziduala, onim rezidualima koji su po apsolutnoj vrednosti veći daje se manji ponder i obratno.



Kako se eliminiše uticaj heteroskedastičnosti (II)?

- **Prvi slučaj** se odnosi na situaciju kada su varijanse slučajne greške poznate i međusobno različite:

$$\left.\begin{aligned} \operatorname{var}(\varepsilon_1) &= \sigma_1^2 \\ \operatorname{var}(\varepsilon_2) &= \sigma_2^2 \\ &\vdots \\ \operatorname{var}(\varepsilon_n) &= \sigma_n^2 \end{aligned}\right\} \sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_n^2$$

- Potrebna transformacija polaznog modela sastoji se u tome da se opservacije svih promenljivih dele sa standardnom devijacijom δ_i (bez većeg praktičnog značaja).

- U ovom modelu nova slučajna greška poseduje stabilnu varijansu **(pokazati).**

Kako se eliminiše uticaj heteroskedastičnosti (III)?

Pretpostavimo da postoji zavisnost varijanse slučajne greške od objašnjavajuće promenljive x_i:

$$\operatorname{var}(\varepsilon_i) = k^2 X_i^2, \mathrm{k} = \mathrm{const}$$

- **Drugi način** je da sve promenljive modela delimo sa merom varijabiliteta, X_i:

$$\frac{Y_i}{X_i} = \beta_0 \frac{1}{X_i} + \beta_1 \frac{X_i}{X_i} + \frac{\varepsilon_i}{X_i}$$

- U ovom modelu nova slučajna greška je $\frac{\varepsilon_i}{X_i}$.

Njena varijansa je stabilna:

$$\operatorname{var}\left(\frac{\varepsilon_i}{X_i}\right) = \frac{\operatorname{var}(\varepsilon_i)}{X_i^2} = \frac{k^2 X_i^2}{X_i^2} = k^2 = const.$$

Kako se eliminiše uticaj heteroskedastičnosti (IV)?

- **Treći način** se koristi kada raspolažemo sa velikom brojem podataka na osnovu kojih je moguće doći do neke pretpostavke o δ_i^2.

- Ako pretpostavimo da varijansa slučajne greške sledi šemu:

$$\sigma_i == C\ E(Y_i) = C(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}), \quad C = const.$$

model sa homoskedastičnim greškama dobija se deljenjem sa $E(Y_i)$ – **pokazati.**

- Pri tome, očekivana vrednost Y_i aproksimira se ocenjenom vrednošću $\hat{Y}_i$.

Alternativni pristupi eliminisanja efekata heteroskedastičnosti

1. Koristimo logaritmovane vrednosti podataka.

2. Prilikom računanja standardnih grešaka ocena pravimo korekciju koju je predložio Vajt (engl. White). Na ovaj nacin se obezbeđuje efikasnost i konzistentnost standardnih grešaka ocena parametara u modelu sa heteroskedastičnim greškama.

. Ovo je najzastupljeniji pristup u empirijskoj analizi poslednjih godina.

Struktura predavanja

- **Narušavanje pretpostavki KLRM**

  - Autokorelacija

    - Pojam autokorelacije
    - Posledice autokorelacije
    - Testiranje
    - Otklanjanje posledica autokorelacije

Pretpostavke KLRM

1. $E(\varepsilon_i) = 0$
2. $Var(\varepsilon_i) = \sigma^2 < \infty$
3. **Cov ($\boldsymbol{\varepsilon_i}$, $\boldsymbol{\varepsilon_j}$) = 0 za *i* različito od *j***
4. Objašnjavajuće promenljive nisu određene stohastičkim članom
5. $\varepsilon_i \sim N(0,\sigma^2)$
6. Ne postoji tačna linearna zavisnost između objašnjavajućih promenljivih.

Pretpostavka 3: Cov (ε_i , ε_j) = 0 za $i \neq j$ Odsustvo autokorelacije

- Odsustvo autokorelacije: slučajne greške su nekorelisane
  - Cov (ε_i , ε_j) = 0 za $i \neq j$

  Nema pravilnosti u korelacionoj strukturi slučajnih greški.
- Postoji autokorelacija: slučajne greške koje su uređene tokom vremena su korelisane
  - Cov (ε_i , ε_j) $\neq$ 0 za $i \neq j$

  Slučajne greške slede prepoznatljiv obrazac u kretanju.
- Najčešća se javlja u analizi vremenskih serija:
  - Cov (ε_t , ε_{t-s}) $\neq$ 0 za $s=1,2,\ldots$

Zašto se javlja autokorelacija?

1. Trajni efekat egzogenih šokova na kretanje ekonomskih vremenskih serija
   - Primer: obustava rada i ocenjivanje zavisnosti ostvarene proizvodnje od količine uloženog rada.
2. Inercija u kretanju ekonomskih veličina.
3. Modifikacija polaznih podataka
   - Neki kvartalni podaci se dobijaju kao prosek tromesečnih vrednosti.

- Autokorelacija može biti “prava” i “lažna”
  - “Prava”: posledica prirode podataka
  - “Lažna”: model je pogrešno postavljen (izostavljanje prom., pogrešna funkcionalna forma)
- Autokorelacija može biti pozitivna ili negativna (koef. korelacije između sukcesivnih vrednosti = autokor. koef. prvog reda, AR(1) šema - **pokazati...**).

Kovarijntna matrica (pokazati!)

$$E(\varepsilon\varepsilon') = \sigma^2\Omega = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho^{t-1} \\ \rho & 1 & \cdots & \rho^{t-2} \\ \vdots & \vdots & & \vdots \\ \rho^{t-1} & \rho^{t-2} & \cdots & 1 \end{bmatrix}.$$

Posledice autokorelacije

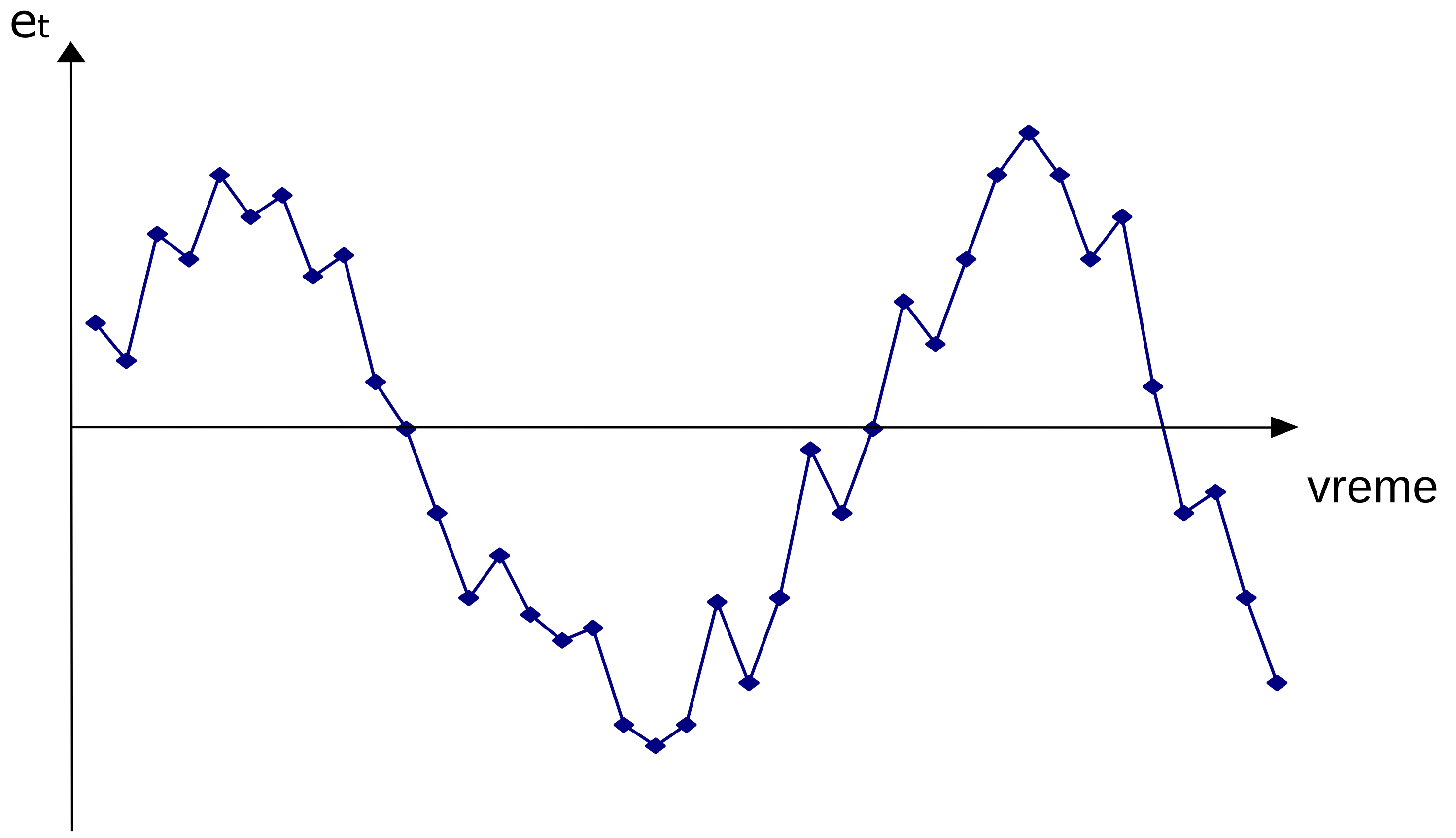
- Ocene ONK su nepristrasne, ali neefikasne.
- Ocena varijanse slučajne greške je pristrasna.
- R^2 nije valjan pokazatelj kvaliteta regresije.
- Rezultati t i F testa su pristrasni i nepouzdani.
- Intervali poverenja su neprecizni.
- Predviđanje je nepouzdano.
- **Pokazati...**

Kako se proverava postojanje autokorelacije?

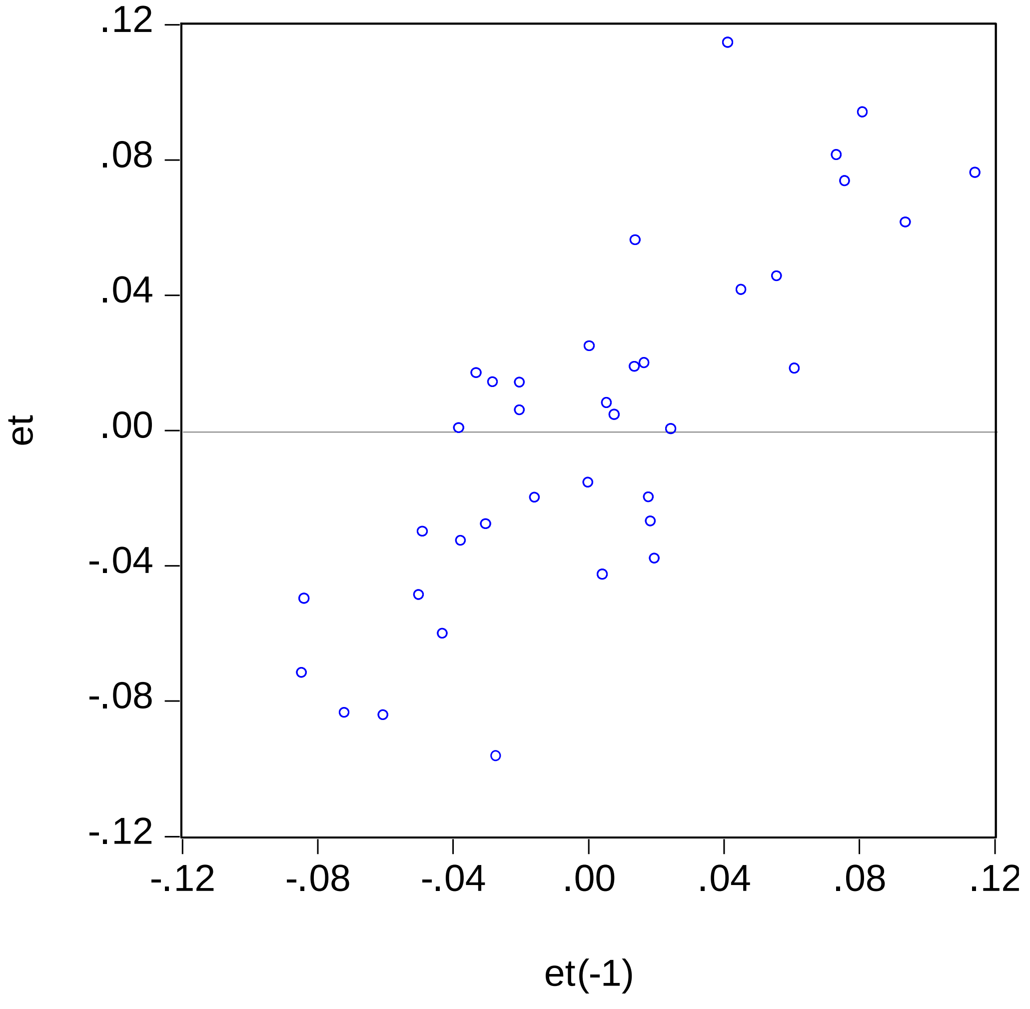
1. Neformalni (grafički) metodi

2. Formalni metodi (testiranje)

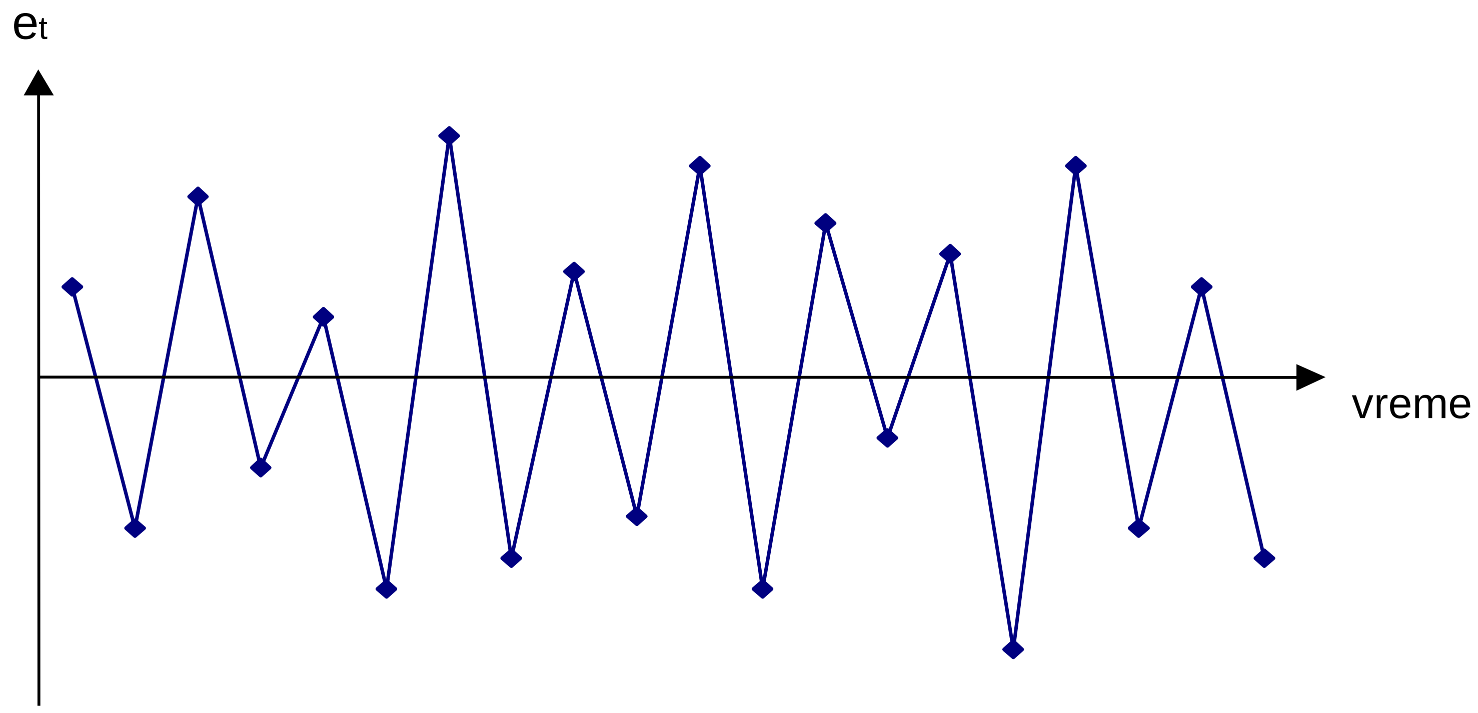
Pozitivna autokorelacija (reziduali zadržavaju isti znak u nizovima)



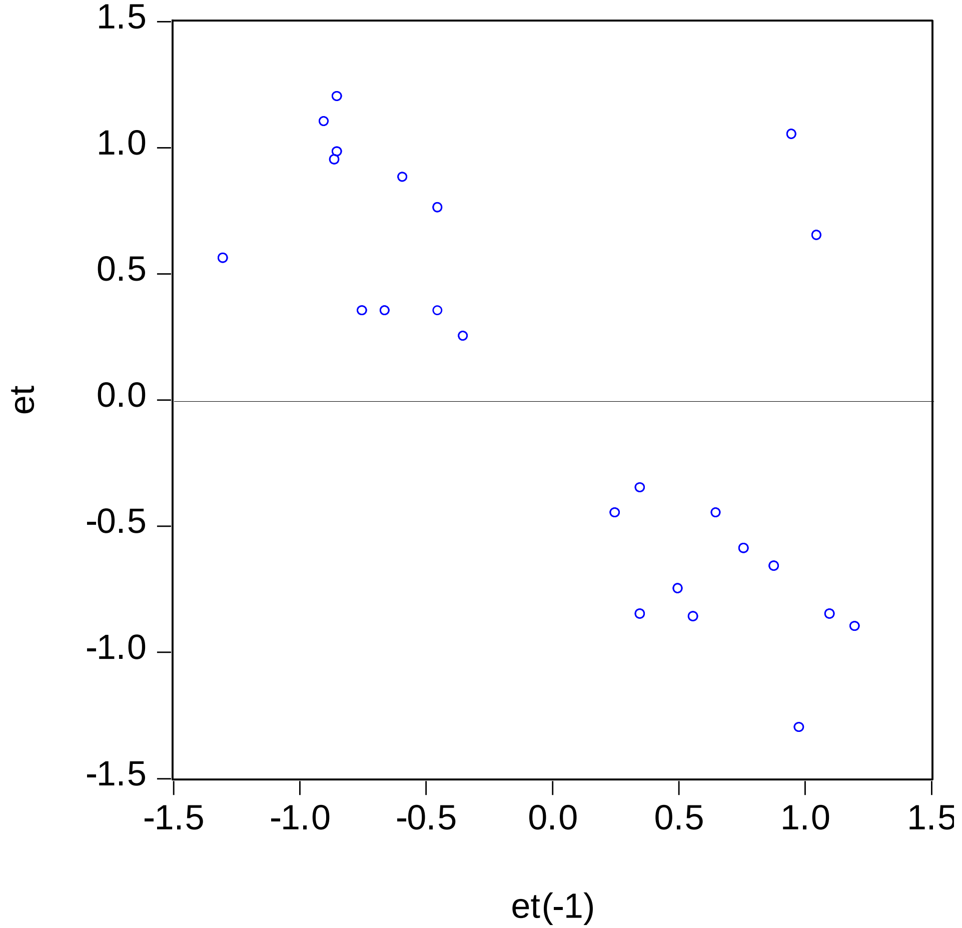
Pozitivna autokorelacija
(reziduali u funkciji sopstvenih prethodnih vrednosti grupisani u I i III kvadrantu)



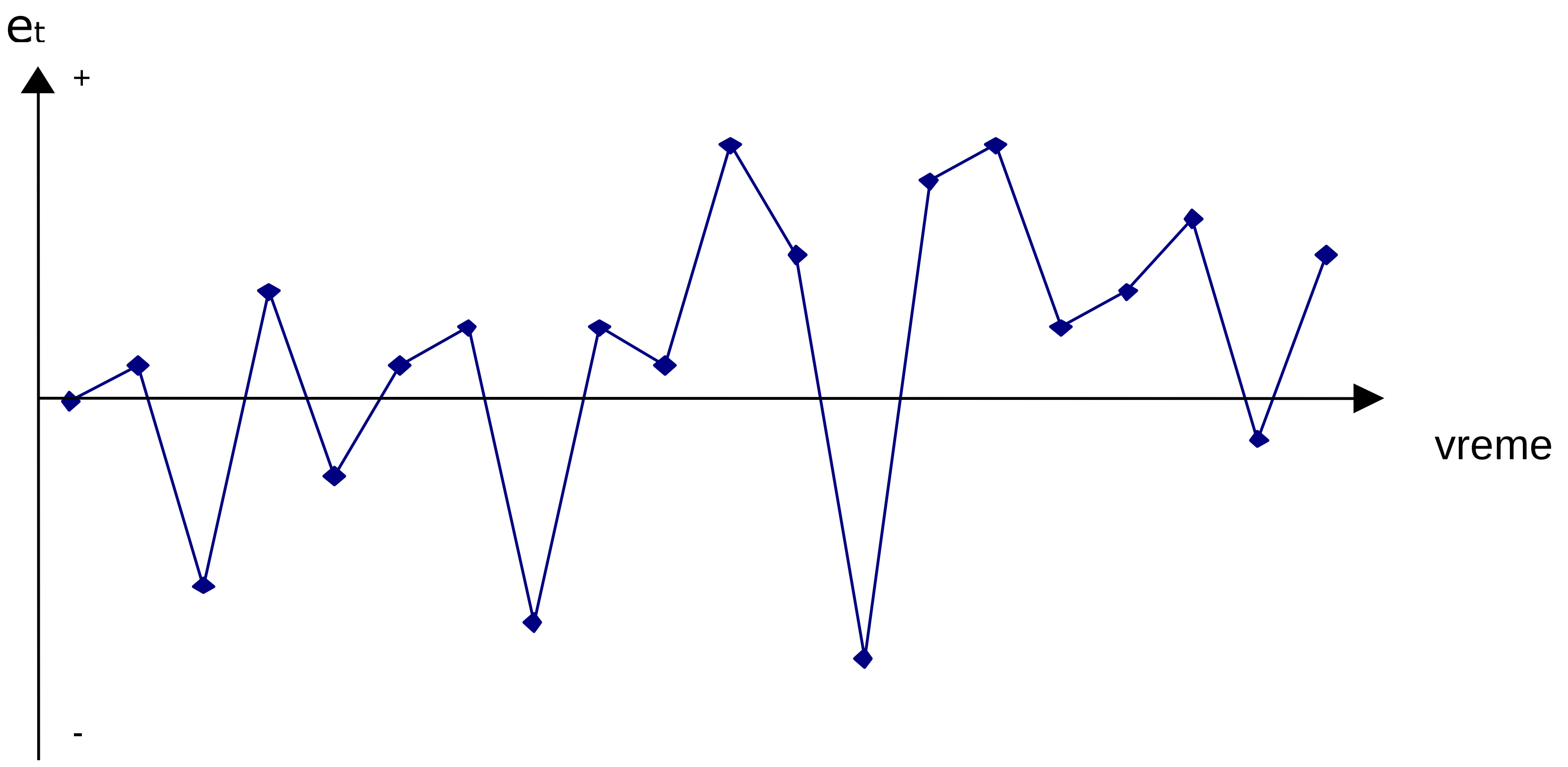
Negativna autokorelacija (reziduali naizmenično menjaju znak)



Negativna autokorelacija (reziduali u funkciji sopstvenih prethodnih vrednosti grupisani u II i IV kvadrantu)



Ne postoji autokorelacija (reziduali ne pokazuju pravilnost promene tokom vremena)



Ispitivanje postojanja autokorelacije: Darbin-Votsonov (engl. Durbin-Watson) test

- Darbin-Votsonov test (oznaka: DW ili d) se koristi za proveru postojanja autokorelacije prvog reda:

  $\varepsilon_t = \rho\, \varepsilon_{t\text{-}1} + v_t$

  gde je $v_t \sim N(0, \sigma_v^2)$ i ρ je autokorelacioni koeficijent prvog reda, koji se nalazi u intervalu (-1,+1).

  ρ = 0 ne postoji autokorelacija,

  ρ = 1, ekstremna pozitivna autokorelacija

  ρ = -1, ekstremna negativna autokorelacija

  0< ρ<*1, pozitivna autokorelacija*

  -1< ρ*<0, negativna autokorelacija*

  Relevantne hipoteze:

  H_0 *:* ρ=0 *(nema autokorelacije)*

  H_1 *:* $\rho\neq$0 *(postoji autokorelacija prvog reda)*

DW test (II):

$$DW = \frac{\sum_{t=2}^{n}\left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{n} e_t^2}$$

$$DW = \frac{\sum_{t=2}^{n}\left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{n} e_t^2} = \frac{\sum_{t=2}^{n} e_t^2 - 2\sum_{t=2}^{n} e_t\, e_{t-1} + \sum_{t=2}^{n} e_{t-1}^2}{\sum_{t=1}^{n} e_t^2},$$

$$\sum_{t=1}^{n} e_t^2 \approx \sum_{t=2}^{n} e_t^2 \approx \sum_{t=2}^{n} e_{t-1}^2$$

$$DW \approx 2\left(1 - \frac{\sum_{t=2}^{n} e_t\, e_{t-1}}{\sum_{t=1}^{n} e_t^2}\right)$$

$$DW \approx 2(1-\hat{\rho})$$

$\hat{u}_t$ - reziduali iz modela ciju autokorelaciju ispitujemo

$\hat{\rho}$ - ocena autokorelacionog koeficijenta prvog reda.

DW test (III)

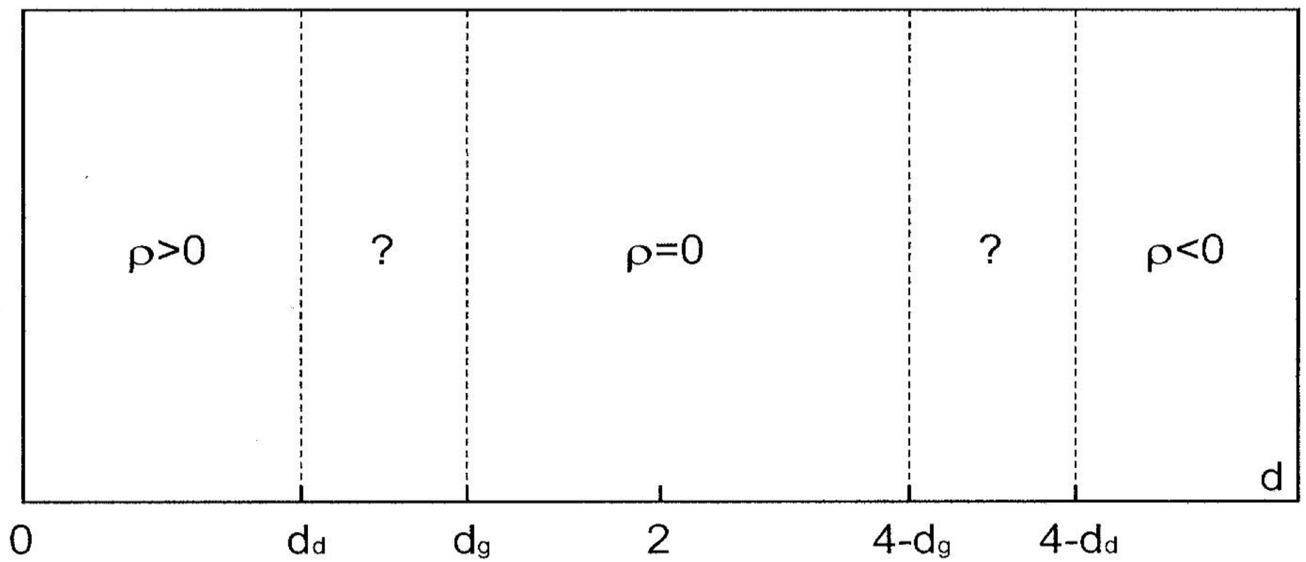
- U postupku testiranja koriste se kritične vrednosti koje su autori testa označili kao donja i gornja kritična vrednost (*E(d)*, kao i sama raspodela *sl. prom. d* zavise od podataka nezavisnih promenljivih u uzorku).

- Donja kritična vrednost: *dd*,
- Gornja kritična vrednost: *dg*.

- Kritične vrednosti zavise od obima uzorka i broja objašnjavajućih promenljivih.

- **Objasniti postupak testiranja**...

Primena DW testa



Ograničenja u primeni DW testa

- Ograničenja u primeni:

1. Postoje situacije kada se primenom testa ne može doneti precizan zaključak.
2. Test je definisan samo za model sa slobodnim članom.
3. Testom se <u>ne može proveriti postojanje autokorelacije većeg reda (napomena: Wallis-ova d_4 statistika).</u>
4. Test nije pouzdan u situaciji kada se kao objašnjavajuća promenljiva javlja zavisna sa docnjom:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 y_{t-1} + \varepsilon_t$$

Durbin-ova (1970) h statistika

- Za modele sa pomaknutom zavisnom promenljivom predložena je sledeća modifikacija:

$$h = \left(1 - \frac{d}{2}\right)\sqrt{\frac{T}{1 - Ts_{b2}^{2}}},$$

pri čemu je d vrednost DW statistike, S_{b2} ocenjena st.gr. ocene parametra uz Y_{t-1}, a T je veličina uzorka.

- Statistika poseduje normalnu standardizovanu raspodelu (pod pretpostavkom da važi H_0); zaključak se donosi poredjenjem sa vrednošću 1.96.

Opšti test autokorelacije: Brojš-Godfrijev (engl. Breusch-Godfrey) test

- U opštem slucaju autokorelacija može biti reda *m:*

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \ldots + \rho_m\varepsilon_{t-m} + v_t, \quad v_t \sim N(0, \sigma_v^2).$$

- Nulta i alternativna hipoteza

H_0 : $\rho_1 = \rho_2 = \ldots = \rho_k = 0$ *(ne postoji autokorelacija)*

H_1 : bar jedan od parametara je razlicit od nule (postoji autokorelacija

- Algoritam testiranja:

1. Pretpostavimo da je polazni model oblika:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

2. Ocenjujemo model iz 1., dobijamo reziduale i potom ocenjujemo pomocnu regresiju:

$$e_t = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \rho_1 e_{t-1} + \rho_2 e_{t-2} + \ldots + \rho_m e_{t-m} + v_t,$$

3. Odredujemo koeficijent determinacije R^2 *iz pomocne regresije i potom ga* množimo obimom uzorka *T. To je ($T R^2$) Brojš-Godfrijeva test-statistika. Može* se pokazati da važi: $T R^2 \sim \chi^2$ sa *m stepeni* slobode, pri uslovu istinitosti nulte hipoteze.

Kako se eliminiše uticaj autokorelacije?

- Korekcija polaznog modela u pravcu transformisanja promenljivih (**pokazati...).**

- Korekcija polaznog modela u pravcu eksplicitnog ukljucivanja dinamike – dinamicki modeli.

- Korekcija standardnih grešaka ocena kako bi odražavale stvarni varijabilitet ocena parametara: Njui-Vestova korekcija (engl. Newey-West).