4EK211 Základy ekonometrie

Autokorelace náhodné složky

Cvičení 7

Gauss-Markovy předpoklady

Náhodná složka: Gauss-Markovy předpoklady

1. $E(\boldsymbol{u}) = 0$
   – průměrná hodnota náhodné chyby je nula

2. $E(\boldsymbol{u}\,\boldsymbol{u}^T) = \boldsymbol{\sigma}^2\,\boldsymbol{I}_n$
   – konečný a konstantní rozptyl = homoskedasticita
   → porušení: heteroskedasticita
   – náhodné složky jsou sériově nezávislé
   → porušení: autokorelace náhodné složky

3. $\boldsymbol{X}$ je nestochastická matice – $E(\boldsymbol{X}^T\boldsymbol{u}) = 0$
   – veškerá náhodnost $\boldsymbol{y}$ je obsažena v náhodné složce

4. $\boldsymbol{X}$ má plnou hodnost k
   – matice $\boldsymbol{X}$ neobsahuje žádné perfektně lineárně závislé sloupce pozorování vysvětlujících proměnných
   → porušení: multikolinearita

Autokorelace – obecně

- Podle 2. G.M. předpokladu mají být nediagonální prvky kovarianční matice náhodné složky, E($\boldsymbol{u}\,\boldsymbol{u}^{\mathrm{T}}$), nulové:

$$\mathbf{E}(\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}) = \begin{bmatrix} \boldsymbol{\sigma}^2 & 0 & 0 & \dots & 0 \\ 0 & \boldsymbol{\sigma}^2 & 0 & \dots & 0 \\ 0 & 0 & \boldsymbol{\sigma}^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \boldsymbol{\sigma}^2 \end{bmatrix} = \boldsymbol{\sigma}^2\,\mathbf{I}_{\boldsymbol{n}}$$

- porušení předpokladu – **nediagonální prky ≠ 0** → autokorelace
- náhodné složky u_t **nejsou** sériově nezávislé – závislost mezi hodnotami jedné proměnné (náhodné složky
- $cov(u_t;u_s) = \mathrm{E}(u_t;u_s) \neq 0$; alespoň pro jednu kombinaci: $t \neq s$.
- náhodnou složku lze modelovat pomocí její předchozí hodnoty
  - $u_t = \rho^* u_{t-1} + \varepsilon_t$ AR(1) proces
  - $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$ AR(2) proces

Autokorelace – příčiny a důsledky

Příčiny

- setrvačnost ekonomických veličin (zejm. případ ČR)
- chybná specifikace modelu (specifikační chyba se stává součástí náhodné složky)
- chyby měření vysvětlované proměnné jsou zahrnuty do náhodné složky modelu
- odhad modelu z dat obsahující zpožděné, zprůměrované, vyrovnané, intra nebo extrapolované proměnné

Důsledky

- odhady β zůstávají **nevychýlené a konzistentní**
- odhady β **nejsou** vydatné ani asymptoticky vydatné
- odhady β nemají minimální rozptyl
- **vychýlené odhady** rozptylu modelu (sigma) a směrodatných chyb bodových odhadů (s_{bi})
  - intervaly spolehlivosti nejsou směrodatné
  - statistické testy ztrácejí na síle

Test autokorelace 1. řádu – koeficient autokorelace

- testování vztahu:
  $\boldsymbol{u_t = \rho^* u_{t-1} + \varepsilon_t}$, $t = 1, 2, \ldots, T$
  $\boldsymbol{\rho}$ je z intervalu <-1,1>
- $\boldsymbol{\rho}$ je **koeficient autokorelace**
- $\boldsymbol{\varepsilon_t}$ je normálně rozdělená náhodná složka
- **vztah**: náhodné složky jsou generovány stacionárním autoregresním stochastickým procesem prvního řádu (AR1)

Vyhodnocení koeficientu $\boldsymbol{\rho}$

- $\boldsymbol{\rho} > 0$ … pozitivní autokorelace
- $\boldsymbol{\rho} < 0$ … negativní autokorelace
- $\boldsymbol{\rho} = 0$ … sériová nezávislost náhodných složek / autokorelace neexistuje

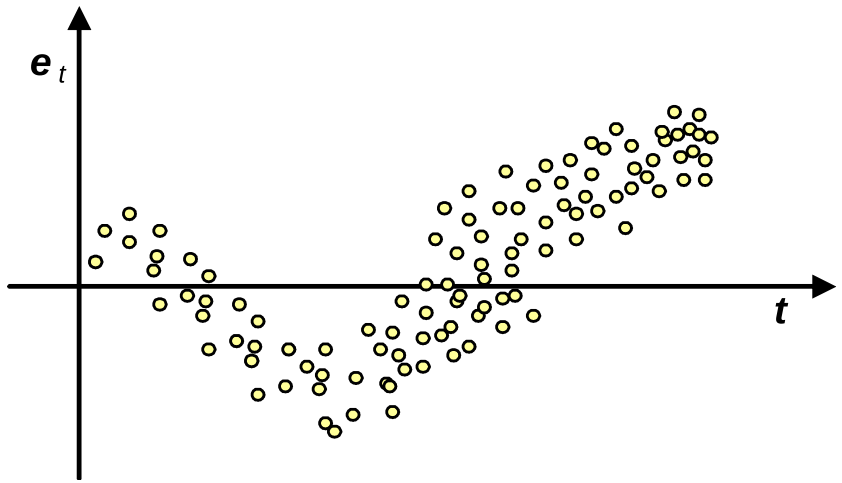
Testovaná hypotéza

$\boldsymbol{H_0}$: rezidua mají zcela náhodný charakter, tj. $\boldsymbol{\rho} = 0$ (sériová nezávislost nezávislost náhodných složek / autokorelace neexistuje)

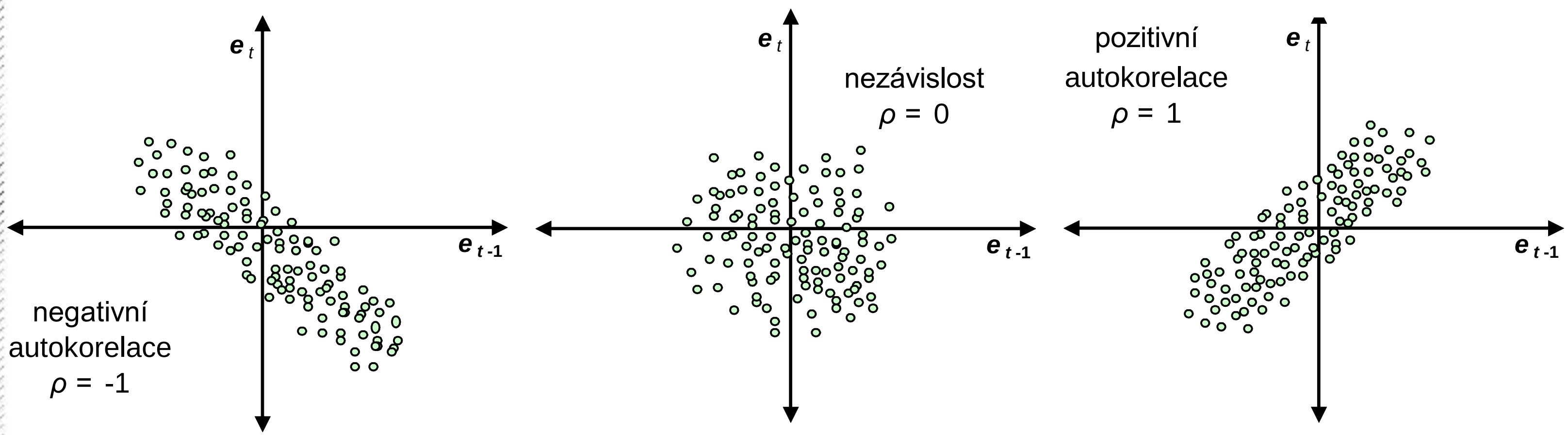
$\boldsymbol{H_1}$: rezidua nemají zcela náhodný charakter, tj. $\boldsymbol{\rho} \neq 0$

Test autokorelace 1. řádu – grafický test

Graficky – graf vývoje reziduí v čase



graf vztahu reziduí a zpožděných reziduí



Test autokorelace 1. řádu – *d*-statistika

- nejznámější test: Durbin-Watsonova *d*-statistika – tj. hodnota **DW** ve výstupu PcGive i GRETL
- hodnoty $\boldsymbol{u_t}$ nejsou známy, proto se vychází z jejich odhadu – tj. z reziduí $\boldsymbol{e_t}$
- testuje se vztah: $\boldsymbol{e_t = r^* e_{t-1} + v_t}$,kde $\boldsymbol{r}$ je odhad $\boldsymbol{\rho}$, značíme $\boldsymbol{\hat{\rho}}$
- platí

$$\hat{\rho} = r \approx 1 - (d / 2)$$

$$d = \frac{\sum_{t=2}^{T} (e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

- *d*-statistika: symetrické rozdělení <0,4> se střední hodnotou 2
- závisí na: n = počet pozorování
  k = počet vysvětlujících proměnných
  α = hl. významnosti - hodnoty *d* tabelovány

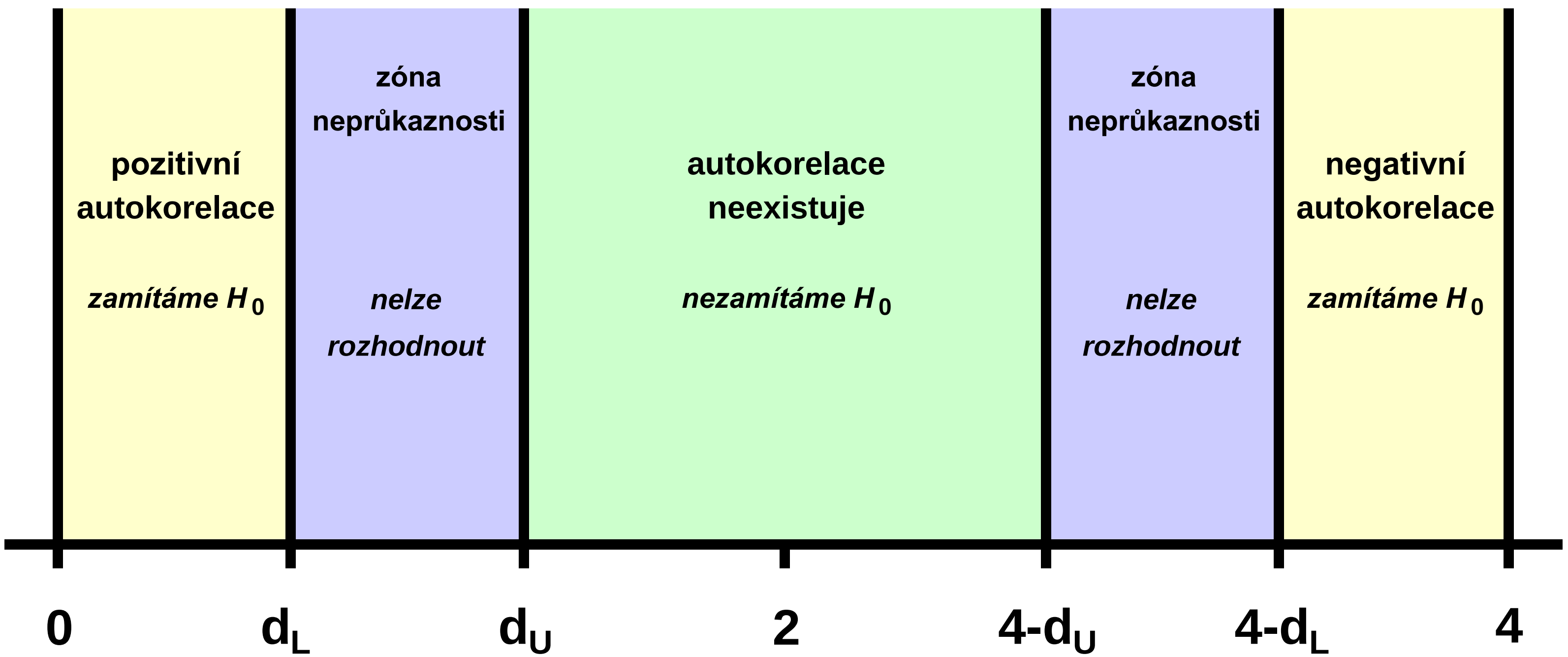
Testovaná hypotéza

$\boldsymbol{H_0}$: rezidua nejsou zatížena autokorelací prvního řádu, AR(1), $\boldsymbol{\rho} = 0$

$\boldsymbol{H_1}$: rezidua nemají zcela náhodný charakter, tj. $\boldsymbol{\rho} \neq 0$

Test autokorelace 1. řádu – *d*-statistika

- vypočtenou hodnotu *d*-statistiky porovnáváme s tabulkovou hodnotou d_L a d_U a vyhodnotíme dle obrázku ↓
- index L = lower, index U = upper
- Pro použití DW d-tesu potřebujeme dostatečný počet stupňů volnosti.



Test autokorelace 1. řádu – příklad

Soubor: CV7_PR1.xls

Data: ***y* =** reálná mzda, USA 1959-2006

***x* =** produktivita práce, USA 1959-2006

Zadání: Odhadněte závislost reálné mzdy (***y***) na produktivitě práce (***x***).

Otestujte autokorelaci:

- graficky
- přes *d*-statistiku pro α = 0,05.

Modifikovaný test autokorelace 1. řádu – Durbin *h*

- Durbin *h* (*h*-statistika)
- při výskytu zpožděné endogenní proměnné v modelu v roli exogenní proměnné
- např.: $y = f(y_{t\text{-}1}, x_2, x_3) + u$

$$\boldsymbol{h} = (1 - 0{,}5 * \boldsymbol{d})\sqrt{\frac{\boldsymbol{n}}{1 - \boldsymbol{n} * \boldsymbol{s}^2_{b(y_{t-1})}}}$$

- h ~ t-rozdělení, asympoticky: $h \sim N(0,1)$
- pro $(n\text{-}k\text{-}1) > 30$ lze užít tabulky $N(0,1)$ rozdělení a pracovat s příslušnými kvantily:
  - α = 0,1 → h^* = kvantil 1,64
  - α = 0,05 → h^* = kvantil 1,96
  - α = 0,01 → h^* = kvantil 2,57

Testovaná hypotéza
$\boldsymbol{H_0}$: nevyskytuje se autokorelace typu AR(1), tj. $\boldsymbol{\rho} = 0$
$\boldsymbol{H_1}$: rezidua nemají zcela náhodný charakter, tj. $\boldsymbol{\rho} \neq 0$

Vyhodnocení např. pro h^* = 1,96:

- $-1{,}96 \leq h \leq 1{,}96$ → nezamítám H_0 o neexistenci autokorelace
- $h < -1{,}96$ → odmítám H_0 ve prospěch negativní autokorelace
- $h > 1{,}96$ → odmítám H_0 pozitivní autokorelace

Test autokorelace 1. řádu – příklad

Soubor: CV7_PR1.xls

Data: ***y* =** reálná mzda, USA 1959-2006

***x* =** produktivita práce, USA 1959-2006

Zadání: Odhadněte závislost reálné mzdy (***y***) na reálné mzdě zpožděné o 1 období ($\boldsymbol{y}_{t-1}$) a produktivitě práce (***x***).

Otestujte autokorelaci přes Durbinovo *h* (*h*-statistiku) pro α = 0,05.

Breusch–Godfrey (BG) test autokorelace vyšších řádů

- **Asymptotický** test autokorelace náhodných složek.

- Pro první i vyšší řády autokorelace: AR(q).

  Máme-li podezření, že náhodná složka u_t se řídí AR(q) procesem

  $$u_t = r_1 u_{t-1} + r_2 u_{t-2} + \dots + r_q u_{t-q} + e_t$$

- **Popis BG testu – příklad:**

  Vyjdeme z obvyklého LRM, např. pro dvě exogenní proměnné:

  $$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t \,, \qquad t = 1,2,\dots,T$$

  odhad MNČ zapíšeme ve tvaru:

  $$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + e_t \,.$$

Breusch–Godfrey (BG) test autokorelace vyšších řádů

- **Popis BG testu – příklad (pokračování):**

- Sestavíme pomocnou regresi BG testu: **rezidua e_t se snažíme vysvětlit pomocí všech *k* proměnných původního modelu a na základě všech *q* uvažovaných autoregresních prvků e_{t-1} až e_{t-q} .**

  Například pro k = 2 a q = 3 má pomocná regrese tento tvar:

$$e_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \delta_1\, e_{t-1} + \delta_2\, e_{t-2} + \delta_3\, e_{t-3} + \omega_t$$

- Pomocnou rovnici odhadneme MNČ a vypočteme koeficient vícenásobné determinace R_e^2 u této pomocné rovnice. Čím menší, R_e^2 tím lépe – dostatečně malý koeficient ukazuje na statistickou nevýznamnost pomocné regrese a tedy na neexistenci autokorelace AR(q).

- Konstruujeme testovací statistiku: $LM = n^* R_e^2 \sim \chi^2(q)$
- kde $n = T - q$
  n - počet disponibilních pozorování pomocné regrese,
  T - počet pozorování původního modelu,
  q – testovaný max. řád autokorelace náhodné složky.

Breusch–Godfrey (BG) test autokorelace vyšších řádů

- **Popis BG testu – příklad (dokončení):**

- Testované hypotézy:

H_0: $\delta_1 = \delta_2 = ... = \delta_q = 0$;

(tj. náhodná složka u_t není autokorelovaná řádu 1 až q.)

H_1: $\neg H_0$.

- **Vyhodnocení BG testu:**

$LM \geq \chi^{2\,*}(q)$ ... zamítáme H_0 ve prospěch autokorelace AR(q),

$LM < \chi^{2\,*}(q)$ ... nepodařilo se zamítnout H_0.

BG test je v ekonometrických SW vestavěný (GRETL, Eviews, R, …)

- Řídíme se podle „p-value“:

  [p-value] $\leq \alpha$

- je-li hodnota [p-value] nižší nebo rovna zvolené hladině významnosti, odmítáme H_0 ve prospěch autokorelace AR(q).