

Testování hypotéz v regresních rovnicích s aplikací testů založených na metodě maximální věrohodnosti

K testování hypotéz v lineárních regresních modelech lze kromě „tradičních“ nástrojů statistické analýzy – individuálních t-statistik významnosti regresních koeficientů a souhrnného F-testu celkové shody modelu s daty - uplatnit též méně využívanou trojici testů vyvozených pro prostředí nasazení metody maximální věrohodnosti. Jde o:

test **věrohodnostního poměru**, test **Lagrangeových multiplikátorů** a **Waldův** test.

1A. Formulace regresního modelu

Mějme standardní zápis jednorovnicového ekonometrického (regresního) modelu:

$$(1a) \quad y_t = \beta_0 \cdot x_{t0} + \beta_1 \cdot x_{t1} + \beta_2 \cdot x_{t2} + \beta_3 \cdot x_{t3} + \dots + \beta_k \cdot x_{tk} + \varepsilon_t$$
$$t = 1, 2, \dots, T \quad (\text{počet pozorování vzorku})$$

neboli v maticové notaci

$$(1b) \quad y = X\beta + \varepsilon, \text{ kde}$$

y T – členný vektor vysvětlované proměnné $y = (y_1, y_2, \dots, y_T)$

X $T \times (k + 1)$ -rozměrná matice vysvětlujících proměnných $X = \{x_{tj}\}$
 $t = 1, 2, \dots, T; j = 1, 2, \dots, k, k + 1$

β $(k + 1)$ -členný vektor parametrů příslušných vysvětlujícím proměnným

ε T – členný vektor náhodných složek (disturbancí) regresní rovnice

1B. Hypotézy o parametrech regresního modelu

V případě oprávněné možnosti předpokládat sdružené normální rozdělení náhodných složek regresní rovnice (včetně snad i situací, kdy statistická analýza rozdělení pozorovaných dat nesvědčí o opaku) přichází výhodně v úvahu nasazení některého z následující trojice testů :

- A) **test věrohodnostního poměru** vyvinutý již v r.1938 S.S. Wilksem (dále **LR-test**)
- B) **Waldův test** formulovaný v r.1943 Abrahamem Waldem (dále **AW-test**)
- C) **test Lagrangeových multiplikátorů** použitý mj. Aitchesonem a Silveym v r.1958, předtím (r.1948) pak pod jiným názvem (**skórový test**) uplatněn R.Raem (dále **LM-test**) .

Společným znakem uvedené trojice testů je skutečnost, že se v nich operuje s reziduy získanými z **neomezeného regresního tvaru**, **omezeného regresního tvaru** (parametry jsou podrobeny testovaným omezením) eventuálně s oběma současně.

V následujícím textu užijeme pro rozlišení náhodných složek toto značení:

e - vektor reziduí vyvozených z odhadů parametrů restrikcemi **neomezeného** regresního vztahu (1a,b), tzn. odhady parametrů jsou spočteny metodou maximální věrohodnosti (v jednorovnicovém modelu shodných s odhady metodou nejmenších čtverců) při ignorování omezení (např. o nulové hodnotě) na regresní parametry.

e^* - vektor reziduí vyvozených z restrikcemi **omezeného** regresního vztahu (1a,b), kde odhady parametrů jsou spočteny metodou maximální věrohodnosti tak, že jsou respektována omezení (daná nulovou hypotézou) na parametry (např., že jsou některé z parametrů nulové).

Formulujeme standardní lineární regresní model ve tvaru (1)

$$y_t = \beta_0 \cdot x_{t0} + \beta_1 \cdot x_{t1} + \dots + \beta_{j-1} \cdot x_{t,j-1} + \beta_j \cdot x_{t,j} + \dots + \beta_k \cdot x_{tk} + \varepsilon_t$$

Nejčastějším případem bude test, zda určitá konkrétní (j -tá) proměnná má být zařazena do regresní rovnice nebo naopak zda má být vynechána. Příslušné omezení zde bude představováno podmínkou

$$\beta_j = 0,$$

které představuje testovanou (nulovou) hypotézu H_0 . Alternativní (oboustranná) hypotéza je pak dána jako

$$H_1 : \beta_j \neq 0.$$

Konkrétní testování bude provedeno tak, že spočteme dvě regrese. Jednu ve tvaru

$$(2a) \quad y_t = \beta_0 \cdot x_{t0} + \beta_1 \cdot x_{t1} + \beta_{j-1} \cdot x_{t,j-1} + \beta_j \cdot x_{t,j} + \dots + \beta_k \cdot x_{tk} + \varepsilon_t$$

tedy se zahrnutím j -té vysvětlující proměnné

$$(2b) \quad y_t = \beta_0 \cdot x_{t0} + \beta_1 \cdot x_{t1} + \beta_{j-1} \cdot x_{t,j-1} + \beta_{j+1} \cdot x_{t,j+1} + \dots + \beta_k \cdot x_{tk} + \varepsilon_t^*$$

tedy bez zahrnutí j -té vysvětlující proměnné.

Poznámka:

S ohledem na to, že ubrání kterékoliv vysvětlující proměnné (představující restrikcí $\beta_j = 0$) znamená vždy zvětšení (nebo aspoň nezmenšení) reziduálního rozptylu, bude platit

$$(3) \quad \sum_{t=1}^T e_t^2 \leq \sum_{t=1}^T e_t^{*2} \quad \text{neboli} \quad e'e \leq e^*e^*$$

2. Obecná definice testů

v dalším výkladu užijeme značení:

θ ... obecná (nespecifikovaná, ale pevná) hodnota parametru (bez omezení)

$\hat{\theta}$ odhad θ pořízený konzistentní odhadovou funkcí (nejčastěji ML-estimátor)

θ^0 hypotetická (hypotézou H_0 předpokládaná) hodnota parametru θ .

(θ v dalším může být jak skalární hodnota, tak k-členný vektor)

Nejjednodušší testovací problém je postaven tak, že data jsou generována sdruženou hustotou $f(y, \theta^0)$ odpovídající parametrům nulové hypotézy, resp. $f(y, \theta)$ při platnosti alternativní hypotézy ($\theta \neq \theta^0$). Toto je test *prosté* nulové hypotézy oproti *složené* alternativě. Logaritmovaná věrohodnostní funkce je zcela obecně definována jako¹

¹ Odlišnost v chápání sdružené hustoty a věrohodnostní funkce (obrácené pořadí zápisu proměnných a parametrů) není pouze ve formálním zápisu: na sdruženou hustotu pohlížíme jako na veličinu s prvně zadanými parametry, která vyjadřuje rozdělení proměnných; na věrohodnostní funkci naopak jako na veličinu, která zahrnuje (pevně dané) naměřené hodnoty proměnných a její maximalizaci provádíme vzhledem k (měnlivým) parametrům při pevně daných hodnotách pozorovaných proměnných.

$$L(\theta; y) = \ln f(y; \theta),$$

a tato je maximalizována v hodnotě $\hat{\theta}$, splňující podmínku

$$\frac{\partial L(\hat{\theta}; y)}{\partial \theta} = 0$$

Definujeme-li veličiny $s(\theta, y) = \frac{\partial L(\theta; y)}{\partial \theta}$ jako vektor skóru (jde o gradient log-
věrohodnostní funkce), pak k odhadům metodou maximální věrohodnosti směřujeme
tak, že položíme tento vektor skóru roven nule.

Rozptyl $V(\hat{\theta})$ lze snadno spočítat jako inverzi Fisherovy informační matice, neboli

$$(4a) \quad V(\hat{\theta}) = \Phi^{-1}(\theta) / T, \quad \text{přičemž}$$

$$(4a) \quad \Phi(\theta) = -E \frac{\partial^2 (\theta; y)}{\partial \theta \partial \theta' } / T$$

Waldův test je asymptotickou aproximací známých a oblíbených t a F testů
v ekonometrii: Lze ukázat, že jestliže $\hat{\theta}$ má asymptoticky normální rozdělení, a
jestliže lze $\Phi(\theta)$ konzistentně odhadnout pomocí $\Phi(\hat{\theta})$, pak statistika

$$(5) \quad \xi_W = T \cdot (\hat{\theta} - \theta)' \Phi(\hat{\theta}) (\hat{\theta} - \theta)$$

bude mít asymptoticky χ^2 – rozdělení o k stupních volnosti, za platnosti nulové
hypotéza.

Test věrohodnostního poměru je založen na rozdílu mezi maximy věrohodnostních
funkcí při nulové a při alternativní hypotéze. Za obecných podmínek má statistika

$$(6) \quad \xi_{LR} = -2 \cdot (L(\hat{\theta}; y) - L(\theta; y))$$

také asymptoticky χ^2 – rozdělení o k stupních volnosti, pokud platí nulová hypotéza.
Patrně první, kdo toto obecné limitní rozdělení odvodil, byl Wilks [1938].

Test Lagrangeových multiplikátorů je odvozen z principu maximalizace při
omezeních. Předpokládejme, že maximalizace (logaritmované) věrohodnostní funkce
při omezeních daných podmínkou $\theta = \theta^0$ vyžaduje zadat množinu Lagrangeových
multiplikátorů, které měří stínovou cenu omezení. Jestliže je tato cena vysoká,
omezení by mělo být odmítnuto jako nekonzistentní s hypotézou. Definujeme-li jako
veličinu H tzv. Lagrangian

$$(7) \quad H = L(\theta; y) - \lambda \cdot (\theta - \theta^0),$$

pak jsou podmínky 1.řádu dány vztahy

$$\frac{\partial L(\theta, y)}{\partial \theta} = \lambda, \quad \theta = \theta^0$$

Znamená to, že test založený na principu Lagrangeových multiplikátorů Aitchisona a Silvey [1958-59], je identický s testem založeným na skórech, tak jak byl tento původně navržen R. Raem [1948]. V obou případech lze rozdělení skóre snadno nalézt za nulové hypotézy, protože vektor skóre bude mít nulovou střední hodnotu a rozptyl rovný $\Phi(\theta).T$

Jestliže na skóre aplikujeme centrální limitní větu, pak výraz představovaný LM-metodou

$$(8) \quad \xi_{LM} = s(\theta^0, y)' \Phi^{-1}(\theta^0) \cdot s(\theta^0, y) / T$$

bude opět mít limitní χ^2 – rozdělení o k stupních volnosti, platí-li nulová hypotéza.

Všimněme si, že všechny tři principy jsou založeny na různých statistikách, které (každá ale jiným způsobem) „měří“ rozdíl mezi H^0 a H^1 :

- a) Waldův test je formulován v podmínkách rozdílu $\theta^0 - \bar{\theta}$
- b) LR test je formulován v podmínkách rozdílu $L(\theta^0; y) - L(\hat{\theta}; y)$
- c) LM test je formulován v podmínkách $s(\theta^0)$

obrázek pro $k = 1$

Idea LR- testu: - LR-test je **založen na vertikálním rozdílu** $L(\hat{\theta}; y) - L(\theta^0; y)$; pokud je hypotéza platná, pak vnesení omezení by nemělo vést k znetelné redukci hodnoty log-věrohodnostní funkce. Zřejmě vždy platí $L(\hat{\theta}; y) \geq L(\theta^0; y)$;

Idea Waldova testu: AW test je **založen na horizontálním rozdílu** $\theta^0 - \bar{\theta}$, neboť – pokud vezmeme v úvahu tvar omezení $\zeta(\hat{\theta}) = 0$ - jde o vyčíslení výrazu $\zeta(\theta^0)$. Pokud body $\theta^0, \hat{\theta}$ leží blízko u sebe, pak také $\zeta(\theta^0) \cong 0$, protože v bodě $\hat{\theta}$ dává omezení přesně nulovou hodnotu. Pokud $\zeta(\theta^0)$ dává velkou hodnotu, vede to k zamítnutí hypotézy H_0 .

Idea LM- testu: - LM-test je **založen na sklonu věrohodnostní funkce v bodě** θ^0 . Pokud hypotéza H_0 platí, pak by tečna k log-věrohodnostní funkci měla mít minimální sklon; k zamítnutí testu by naopak měla vést vysoká zjištěná hodnota tohoto sklonu.

Všechny tři testy jsou za platnosti nulové hypotézy asymptoticky ekvivalentní, ale budou vykazovat rozdílné chování při použití vzorků o malém rozsahu. Bereme-li v úvahu obtížnost/snadnost vyčíslení, pak zaznamenáme:

Waldův test vyžaduje výpočet pouze neomezeného estimátoru

LM test vyžaduje výpočet pouze omezeného estimátoru

LR-test vyžaduje výpočet jak omezeného, tak neomezeného estimátoru

V určitých situacích může být jeden z estimátorů snadněji vyčíslitelný než ostatní: např. lineární model se odhadne snadno, ale stává se nelineárním, pokud do něj vneseme nelineární omezení. Pak je preferovatelná Waldova statistika. Naopak, restriktce někdy vedou k odstranění nelinearity, což by naznačovalo výhodnost užití LM-testu.

Věta: V případě, že uvažovaná věrohodnostní funkce má tvar kvadratické formy,

$$(9) \quad \text{zapsané ve tvaru} \quad L(\theta) = b - 1/2 \cdot (\theta - \hat{\theta})' A (\theta - \hat{\theta}),$$

kde A je symetrická pozitivně definitní matice, která může záviset na datech (a na známých parametrech), b je skalár a $\hat{\theta}$ je (jako odhad θ) funkcí dat, dávají všechny tři testy shodný výsledek.

Důkaz: Pro log-věrohodnostní funkci tvaru (9) zřejmě platí

$$(10A,B) \quad \frac{\partial L(\theta)}{\partial \theta} = -(\theta - \hat{\theta})' A = s(\theta), \quad \frac{\partial L^2(\theta)}{\partial \theta^2} = -A = -T\Phi(\theta)$$

$$\text{Pak} \quad \xi_W = (\theta^0 - \hat{\theta})' A (\theta^0 - \hat{\theta})$$

$$\xi_{LM} = s(\theta^0, y)' A^{-1} s(\theta^0, y) = (\theta^0 - \hat{\theta})' A (\theta^0 - \hat{\theta})$$

$$\xi_{LR} = -2 \cdot (L(\hat{\theta}, y) - L(\theta, y)) = -2(b - b + 1/2)b(\theta^0 - \hat{\theta})' A (\theta^0 - \hat{\theta}) = (\theta^0 - \hat{\theta})' A (\theta^0 - \hat{\theta})$$

protože $L(\hat{\theta}; y) = b$

Kdykoliv je skutečná hodnota θ rovná nebo blízká předpokládané θ^0 , pak bude věrohodnostní funkce v okolí θ^0 přibližně kvadratická pro velké výběry, tím, že matice A závisí pouze na θ^0 . V tom spočívá příčina asymptotické ekvivalence testů pro lokální alternativy za nulové hypotézy.

3. Nasazení testů v lineárním regresním modelu

Předpokládejme (jednorovnicový) lineární regresní model s jednou vysvětlovanou a k vysvětlujícími proměnnými, který zapíšeme (v pozorováních) stručnou formou

$$(11A) \quad y^* | X^* \approx N(X^* \beta, \sigma^2 I_T), \quad \text{kde}$$

y^* je $T \times 1$ vektor závisle proměnné

X^* je $T \times k$ matice nezávisle proměnných

a uvažujme testování hypotézy tvaru $R\beta = r$, kde

R je $k_1 \times k$ matice známých konstant

r je $k_1 \times 1$ vektor konstant (podmínek v mezeních)

Jestliže R má hodnotu k_1 (aby se nevyskytovala redundantní omezení), pak parametry a data (obojí současně) mohou být přetransformovány tak, že se původní test převede na test nepřítomnosti (některých) proměnných v regresní rovnici. V důsledku toho lze úlohu (11 A) reparametrizovat do zápisu

$$(11B) \quad y | X \approx N(X\mu, \sigma^2 I_T), \quad \text{kde}$$

nulová hypotéza má daleko jednodušší tvar. $\theta_0 = 0$

a kde transformované proměnné y, X jsou lineární kombinace původních y^*, X^* .

Testování tvaru (11B) je technicky jednodušší, než testování tvaru (11A). Testování druhé specifikace je zpravidla výhodnější i z důvodů zapracovaných algoritmů běžných ekonometrických software.

Logaritmovaná věrohodnostní funkce pro standardní model (1) má tvar:

$$(12) \quad L(\mu, y) = k - \frac{T}{2} \cdot \ln \sigma^2 - \frac{1}{2\sigma^2} \cdot (y - X\mu)' \cdot (y - X\mu) ,$$

kde k je konstanta. Pokud by σ^2 bylo známé, pak by díky Větě 1 bylo zajištěno, že by všechny tři testy byly identické.² Odtud plyne, že důležitým rozdílem mezi testovými statistikami bude způsob porovnání odhadu σ^2 . Vektor skóru a informační matice odpovídající parametrům μ budou tyto:

$$(14) \quad s(\mu, y) = \frac{X' \varepsilon}{\sigma^2} , \quad \varepsilon = y - X' \mu , \quad \Phi_{\mu\mu} = \frac{X' X}{\sigma^2 \cdot T}$$

a informační matice je blokově diagonální mezi μ a σ^2 . Všimněme si, že skóry jsou proporční korelačnímu koeficientu mezi reziduy a proměnnými X . Ten je přirozeně vždy nulový v $\hat{\theta}$, ale obecně ne v odhadu $\tilde{\theta}$ odpovídajícím nulové hypotéze.

Tři testové statistiky nabudou v tomto případě tvar:

$$(15) \text{ LR-test} \quad \eta_{LR} = T \cdot \ln \left(\frac{\tilde{\varepsilon}' \tilde{\varepsilon}}{\hat{\varepsilon}' \hat{\varepsilon}} \right)$$

$$(16) \text{ AW-test} \quad \eta_{AW} = (\mu_1^0 - \hat{\mu}_1)' (X_1' X_1 - X_1' X_2 \cdot (X_2' X_2)^{-1} X_2' X_1) (\mu_1^0 - \hat{\mu}_1) / \hat{\sigma}^2$$

$$(17) \text{ LM-test} \quad \eta_{LM} = \tilde{\varepsilon}' X_1 (X_1' X_1 - X_1' X_2 \cdot (X_2' X_2)^{-1} X_2' X_1)^{-1} X_1' \tilde{\varepsilon} / \tilde{\sigma}^2$$

kde

$$(18) \quad \hat{\varepsilon} = y - X\hat{\mu} \quad \tilde{\varepsilon} = y - X\tilde{\mu} \quad \tilde{\sigma}^2 = \frac{\tilde{\varepsilon}' \tilde{\varepsilon}}{T} \quad \hat{\sigma}^2 = \frac{\hat{\varepsilon}' \hat{\varepsilon}}{T}$$

matice X je rozdělena na $X = (X_1, X_2)$:

-první $[T \times k_1]$ matice pozorování odpovídá parametrům podrobeným testu

-druhá $[T \times (k - k_1)]$ matice pozorování přísluší parametrům nepodrobeným testu

Z lineární algebry projekcí vyplývá možnost přepisu Waldova a LM-testu do tvarů obsahujících pouze reziduální hodnoty:

$$(19A,B) \quad \eta_{AW} = T \cdot \left(\frac{\tilde{\varepsilon}' \tilde{\varepsilon} - \hat{\varepsilon}' \hat{\varepsilon}}{\hat{\varepsilon}' \hat{\varepsilon}} \right) \quad \eta_{LM} = T \cdot \left(\frac{\tilde{\varepsilon}' \tilde{\varepsilon} - \hat{\varepsilon}' \hat{\varepsilon}}{\tilde{\varepsilon}' \tilde{\varepsilon}} \right)$$

Připomeňme, že vždy platí: $\tilde{\varepsilon}' \tilde{\varepsilon} \geq \hat{\varepsilon}' \hat{\varepsilon}$, protože $\tilde{\varepsilon}' \tilde{\varepsilon}$ odpovídá **omezenému** (jde o SSE, pokud platí nulová hypotéza H_0) a $\hat{\varepsilon}' \hat{\varepsilon}$ **neomezenému součtu čtverců reziduí** (maximálně věrohodný odhad bez omezení na regresní parametry)

² Pak by totiž první dva členy v (11B) byly konstantní a $L(\mu; y)$ by byla funkce tvaru (9): kvadratická v μ

Odtud je mj. zřejmé, že

a) oba čitatele v (19A,B) jsou nezáporná čísla

b) argument v logaritmu výrazu (15) není menší než 1, výraz η_{LR} je tudíž nezáporný

c) čitatele v (19A) i (19B) jsou nezáporné, proto též η_{AW} i η_{LM} jsou nezáporná čísla.

Všechny testové statistiky mají χ^2 -rozdělení o $T - r$ stupních volnosti, kde

T je počet pozorování regresní rovnice a

r je počet restrikcí (omezení) kladených na vztahy mezi regresními parametry.

Z výše uvedeného vyplývá, že **při testování platnosti hypotézy** (na libovolné, ale pro každý test stejné hladině významnosti α) **můžeme obdržet na základě uvedených tří testů protichůdné výsledky**. Výsledkem může být v jednom případě zamítnutí, ve druhém či třetím nezamítnutí nulové hypotézy. To nelze považovat za nerigoróznost testovacího postupu, ale za projev prosté skutečnosti, že síla jednotlivých testů (chyba 2.druhého druhu vyjadřující pravděpodobnost nezamítnutí hypotézy, pokud nulová hypotéza neplatí) není u všech testů shodná (a může navíc záviset na vysvětlujících veličinách).

Tvrzení 1 Mezi uvedenými třemi statistikami platí vztahy:

$$\text{a) (20A)} \quad \eta_{LR} = T \cdot \ln \left(1 + \frac{\eta_{AW}}{T} \right)$$

$$\text{b) (20B)} \quad \eta_{LM} = \frac{\eta_{AW}}{1 + \frac{\eta_{AW}}{T}}$$

$$\text{c) (20C)} \quad \eta_{LR} = T \cdot \ln \left(\frac{1}{1 - \frac{\eta_{LM}}{T}} \right)$$

Ověření: provedeme jednoduchým dosazením výrazů v definicích (15), (19A), (19B):

Z (19A) máme $\eta_{AW} = T \cdot \left(\frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}} \right)$, takže

$$\ln \left(1 + \frac{\eta_{AW}}{T} \right) = \ln \left(1 + \frac{T \cdot \left(\frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}} \right)}{T} \right) = \ln \left(1 + \frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}} \right) = \ln \left(\frac{\tilde{e}'\tilde{e}}{\hat{e}'\hat{e}} \right) = \frac{\eta_{LR}}{T}, \text{ čímž}$$

jsme dokázali a). Dále máme – opět z definice η_{AW} v (19A):

$$\left(\frac{\eta_{AW}}{1 + \frac{\eta_{AW}}{T}} \right) = \left(\frac{T \cdot \frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}}}{1 + \frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}}} \right) = \left(T \cdot \frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\tilde{e}'\tilde{e}} \right) = \eta_{LM}, \text{ čímž jsme dokázali b)}$$

a konečně určením η_{LM} z (19B) dostaneme

$$\ln\left(\frac{1}{1 - \frac{\eta_{LM}}{T}}\right) = \ln\left(\frac{1}{1 - \frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\tilde{e}'\tilde{e}}}\right) = \ln\left(\frac{1}{\frac{\hat{e}'\hat{e}}{\tilde{e}'\tilde{e}}}\right) = \ln\left(\frac{\tilde{e}'\tilde{e}}{\hat{e}'\hat{e}}\right) = \frac{\eta_{LR}}{T}$$

Tvrzení 2 Mezi testem věrohodnostního poměru, testem Lagrangeových multiplikátorů a Waldovým testem (všemi zasazenými do prostředí lineárního regresního modelu ale i do některých obecnějších modelových schémat) platí obecná nerovnost (neovlivněná hodnotami datového vzorku):³

$$(21) \quad \eta_{AW} \geq \eta_{LR} \geq \eta_{LM}$$

Ověření⁴: provedeme porovnáním výrazů v (19A), (19B), (19C)::

a) Podle (19B) platí $\eta_{LM} = \frac{\eta_{AW}}{1 + \frac{\eta_{AW}}{T}}$, takže $\eta_{LM} \left(1 + \frac{\eta_{AW}}{T}\right) = \eta_{AW}$

Odtud vzhledem k nezápornosti η_{AW} plyne $\eta_{AW} \geq \eta_{LM}$.

b) Podobně z (19A) máme $\frac{\eta_{LR}}{T} = \ln\left(1 + \frac{\eta_{AW}}{T}\right)$ neboli po odlogaritmování

$$\exp\left(\frac{\eta_{LR}}{T}\right) = 1 + \frac{\eta_{AW}}{T}, \quad \text{resp.} \quad \exp(\eta_{LR}^*) = 1 + \eta_{AW}^*$$

označíme-li $\eta_{LR}^* = \frac{\eta_{LR}}{T}$ a podobně $\eta_{AW}^* = \frac{\eta_{AW}}{T}$.

Nyní uplatníme Taylorův rozvoj levé strany tohoto vztahu a porovnáme ho s pravou stranou:

$$\exp(\eta_{LR}^*) = 1 + \eta_{LR}^* + \left(\frac{\eta_{LR}^*}{2}\right)^2 + \dots + \dots = 1 + \eta_{AW}^*,$$

Zaznamenáme, že (po zrušení jedniček) porovnání členu η_{AW}^* s členem η_{LR}^* vede k relaci $\eta_{AW}^* \geq \eta_{LR}^*$ (a odtud k $\eta_{AW} \geq \eta_{LR}$), protože nekonečný součet na levé straně (od výrazu η_{LR}^{*2} včetně dále) obsahuje jen nezáporné členy.

c) Zbývá ukázat, že platí $\eta_{LR} \geq \eta_{LM}$. K tomu využijeme vztah (19C), podle něhož máme:

³ Důkaz provedli (pro zobecněný lineární regresní model)

⁴ Toto ověření *bez znalosti původního podkladu* provedl D.Moravanský, který tímto žádá čtenáře o případné poznámky k textu, pokud v něm shledá nejasnosti.

$$\eta_{LR} = T \cdot \ln \left(\frac{1}{1 - \frac{\eta_{LM}}{T}} \right) \text{ nebo také jinak } \eta_{LR}^* = \ln \left(\frac{1}{1 - \eta_{LM}^*} \right),$$

kde jsme opět označili $\eta_{LR}^* = \frac{\eta_{LR}}{T}$ a podobně $\eta_{LM}^* = \frac{\eta_{LM}}{T}$. Odtud máme po

odlogaritmování $\exp(\eta_{LR}^*) = \frac{1}{1 - \eta_{LM}^*}$ a po jednoduchých úpravách

$\exp(\eta_{LR}^*) = 1 + \eta_{LM}^* \exp(\eta_{LR}^*)$ neboli $\eta_{LM}^* = 1 - \exp(-\eta_{LR}^*)$. Opět rozvedeme pravou stranu pomocí Taylorova rozvoje

$$\exp(-\eta_{LR}^*) = 1 - \eta_{LR}^* + \frac{\eta_{LR}^{*2}}{2!} - \frac{\eta_{LR}^{*3}}{3!} + \frac{\eta_{LR}^{*4}}{4!} \dots, \quad \text{Odtud máme}$$

$$\eta_{LM}^* = \eta_{LR}^* - \frac{\eta_{LR}^{*2}}{2!} + \frac{\eta_{LR}^{*3}}{3!} - \frac{\eta_{LR}^{*4}}{4!} \dots \text{ Zanedbáme-li členy od řádu 3 včetně dále,}$$

je odtud zřejmé, že $\eta_{LR}^* \geq \eta_{LM}^*$ a tedy i $\eta_{LR} \geq \eta_{LM}$.

Důsledek tvrzení 2:

Z relace (21) bezprostředně vyplývá tento důsledek: Kdykoliv vede závěr z testování testem Lagrangeových multiplikátorů k zamítnutí nulové hypotézy, poskytne tentýž výsledek též testování pomocí Waldova a LR- testu. Naopak, kdykoliv vede k přijetí (k nezamítnutí) nulové hypotézy Waldův test, dospějeme ke stejnému závěru testováním LR i LM testem.

Nerovnost (21) nicméně neříká nic o relativních přednostech testů (při platnosti různých alternativ), protože se vztahuje toliko k testování za (platnosti) nulové hypotézy.

Znamená to, že jestliže hladina α zvolená pro zamítnutí/nezamítnutí Waldovým testem má velikost/sílu 5%, pak pro LR a ML testy budou mít velikost/sílu menší než 5%. Jejich zřejmě slabší síla výpovědí je prostě důsledkem volbou konzervativnější (blíže k 0) velikosti α . Pokud se však hladiny sil poopraví na tutéž velikost, pak už nerovnost v prosté podobě (21) neplatí. Jak ukázali mj. Rothenberg [1979] a Evans a Savin [1983]: když jsou hladiny testů zhruba vyrovnány, pak jsou jejich síly přibližně tytéž.

Jak již bylo řečeno, výsledky testování mohou být vzájemně konfliktní: mohou přitom silně záviset na zvolené hladině významnosti α : Tak např. Waldův test může zamítnat nulovou hypotézu a LM statistika ji přijímat s pravděpodobností 95% (tj. hladině významnosti $\alpha = 5\%$), ale na jiných hladinách (odpovídajících 90% nebo 99% pravděpodobnostem) může být výsledek testování oběma testy ve vzájemné shodě.

4. Nasazení testů v kontextu uplatnění instrumentálních proměnných

Dostatečně obecným prostředím za účelem formulace výše uvedeného testového „tria“ v prostředí soustavy simultánních strukturních rovnic je kontext metody instrumentálních proměnných IV:

Připomeňme, že pokud jde o značení, znamenají :

y_i vektor vysvětlované běžné endogenní proměnné i-té rovnice

Y_i matice (vysvětlujících) běžných endogenních proměnných i-té rovnice

X_i matice (vysvětlujících) predeterminovaných proměnných i-té rovnice

Z_i sdružení všech vysvětlujících proměnných $Z_i = (Y_i, X_i)$

X matice (vysvětlujících) predeterminovaných proměnných celé soustavy

β_i, γ_i vektor parametrů příslušných běžným endogenním a predeterminovaným proměnným i-té rovnice (jejich sloučením dostaneme vektor δ_i)

ε_i vektor náhodných složek i-té rovnice

e_i vektor reziduí i-té rovnice

Zápis i-té strukturní (regresní) rovnice tedy v této symbolice vypadá následovně:

$$(22a) \quad y_i = Y_i \beta_i + X_i \gamma_i + \varepsilon_i = Z_i \delta_i + \varepsilon_i ,$$

kde vektor náhodných složek má rozdělení $\varepsilon_i \approx N(0, \sigma^2 I_T)$

Abychom se však v následujícím obešli bez dvojího indexování (budeme ho potřebovat pro odlišení skupiny parametrů s omezeními od skupiny parametrů nepodléhajících omezením), zapíšeme (21a) jednoduše (bez indexu i) jako

$$(22b) \quad y = Y\beta + X\gamma + \varepsilon = Z\delta + \varepsilon ,$$

Definujeme-li matici P jako $P = X(X'X)^{-1}X'$ a formulujeme-li hypotézu o nulovosti (některých) parametrů ve vektoru δ jako $\delta_1 = 0$ (přičemž zbývající část vektoru parametrů δ_2 není omezeními dotčena) , bude mít testová statistika odpovídající LM-testu tvar:

$$(23) \quad \xi_w = \frac{\hat{\delta}_1' (Z_1' P Z_1 - Z_1' P Z_2 (Z_2' P Z_2)^{-1} Z_2' P Z_1) \hat{\delta}_1}{\hat{\sigma}^2} ,$$

kde $\hat{\delta} = (Z' P Z)^{-1} Z' P y$, $\hat{e} = y - Z \hat{\delta}$, $\hat{\sigma}^2 = \frac{\hat{e}' \hat{e}}{T}$.

Vektor vysvětlujících proměnných $Z = (Z_1, Z_2)$ jsme rozdělili souhlasně s dělením vektoru parametrů δ na $\delta = (\delta_1, \delta_2)$, tzn. proměnné v Z_1 budou ty, u nichž stojící parametry jsou dotčeny omezeními, zatímco v Z_2 budou ty, jímž příslušející parametry nepodléhají omezením.

Čitatel (23) může být přepsán v podmínkách reziduí z omezené regrese využívající téže matice G . Pokud vezmeme $\hat{\delta}_2 = (Z_2' P Z_2)^{-1} Z_2' P y$ a dále definujeme $\tilde{e} = y - Z_2 \tilde{\delta}_2$, může být Waldova statistika vyjádřena jako

$$(23) \quad \xi_w = \frac{T \cdot (\tilde{e}' P \tilde{e} - \hat{e}' P \hat{e})}{\hat{e}' P \hat{e}},$$

Protože P je idempotentní matice, mohou být oba součty čtverců v čitateli (23) vyčísleny provedením regrese příslušných reziduí na matici instrumentálních proměnných X a vzetím příslušných vysvětlených součtů čtverců. Jejich diference je také získatelná jako rozdíl sum čtverců reziduí z 2. stupně regrese při nasazení 2SLS odhadové metody.

Pokud se instrumenty při přechodu od nulové k alternativní hypotéze nezmění, pak nevnikne žádná obtíž ve formulaci Waldova testu. Pokud se však tento soubor instrumentů změní, pak je Waldův test použitelný (jen) s výběrem instrumentů odpovídajících alternativě. Dalo by se usuzovat, že za této situace by se mohl hodit LM test využívající omezenější okruh instrumentů, ale není tomu tak: přinejmenším ne v jeho původní jednoduché podobě: Pokud jsou totiž oba soubory instrumentů odlišné, pak lze LM test odvodit [učinil tak Engle 1979], ale takto odvozený test nemá žádoucí jednoduchý tvar - např. tvar obdobný (19b).⁵

V obecnější situaci, kdy zápisem obdobným (22) vyjádříme sevřený tvar soustavy simultánních rovnic, bude kovarianční matice mít obecný tvar $\Phi = \Sigma \otimes I_T$, kde Σ je kovarianční matice náhodných složek rovnic soustavy (vyjádřených v pevném čase t)

Soubor predeterminovaných proměnných P lze pro takovou soustavu zapsat jako $I_T \otimes X$. Pokud nyní vezmeme odhad $\hat{\Sigma}$ jako odhadnutou kovarianční matici náhodných složek rovnic soustavy za alternativy, potom můžeme zapsat odhadovou funkci třístupňové metody nejmenších čtverců 3SSL jako

$$(24) \quad \hat{\delta} = (Z' P Z)^{-1} Z' P y,$$

Pak lze ukázat, že Waldův test lze zapsat (díky asymptotické ekvivalenci metod 3SLS a FIML) jako

$$(25) \quad \xi_w = \hat{\delta}_1' \left(Z_1' P Z_1 - Z_1' P Z_2 \cdot (Z_2' P Z_2)^{-1} Z_2' P Z_1 \right) \hat{\delta}_1,$$

kterýžto výraz může být reformulován do podoby

$$(26) \quad \xi_w = T \cdot (\tilde{e}' P \tilde{e} - \hat{e}' P \hat{e}),$$

Zde (jen zdánlivě) zmizel odhadnutý rozptyl $\hat{\sigma}^2$ z testové statistiky; ve skutečnosti se však $\hat{\sigma}^2$ nachází v obsahu vektoru $\hat{\delta}_1$ (v němž je tentokrát zahrnuta i informace z matice $\hat{\Sigma}$). I v tomto případě je tedy rozdíl v (26) tvořen rozdílem mezi součty čtverců reziduí *spočtených* (při resp. bez respektování zadaných omezení na parametry) *nyní třetím stupněm metody 3SLS*:

⁵ Viz Engle (1979a)

Literatura

- (1) Maddala G.S. : Introduction to Econometrics: London, Macmillan P.C. 1988.
- (2) Engle R.,F.: "Wald, Likelihood Ratio and Lagrange Multiplier test in Econometrics". In : Intriligator, M.D., Griliches, Z : Handbook of Econometrics Vol.II, Ch.13. North-Holland Amsterdam, 1986.
- (3) Berndt, E.R. and Savin N.E.: "Conflict Among Criteria for Testing Hypotheses in a the Multivariate Linear Regression Model". Econometrica 45/1977 s.1263-1278.
- (4) Greene,W.,H. : Econometric Analysis. 4th edition. Prentice Hall, New Jersey . 2000.

5. Dodatek (jako základ pro další úpravy textu)

Z historického hlediska je nejstarším testem test věrohodnostního poměru založeného na úvahách Neymana a Pearsona [1928]. Věrohodnostní poměr je, obecně řečeno, podíl tvaru

$$\lambda = \frac{\max L(\theta) \text{ pres hodnoty } \theta \text{ specifikované hypotézou } H_0}{\max L(\theta) \text{ pres všechny hodnoty } \theta}$$

$L(\theta)$ je věrohodnostní funkce = sdružená hustota rozdělení (obvykle náhodných složek) .

Z definice je patrné, že hodnota λ musí ležet v intervalu $\langle 0,1 \rangle$, protože jmenovatel (neomezené maximum) musí být větší nebo roven čitateli (omezené maximum).

Hypotézu H_0 zamítáme, jestliže $\lambda \leq c_\alpha$, kde α je konstanta definovaná tak, aby chyba 1.druhu byla rovna α .

LR test má několik důležitých vlastností: Je **nestranný** a **konzistentní**.⁶ Dále, veličina

$2\ln(\lambda)$ má χ^2 - rozdělení o r stupních volnosti, kde $\frac{r}{\sigma}$ je počet parametrů, pro které H_0 specifikuje předepsané hodnoty.

Příklad: jestliže uvažujeme normální rozdělení se střední hodnotou μ a rozptylem σ^2 a nechť má nulová hypotéza tvar $H_0 : \mu = 2$ (přičemž velikost rozptylu σ^2 není nijak specifikována)

Jestliže $L(\mu, \sigma^2)$ je věrohodnostní funkce, pak definujeme věrohodnostní poměr jako

$$\lambda = \frac{\max L(2, \sigma^2)}{\max L(\mu, \sigma^2)},$$

přičemž veličina $-2\ln(\lambda)$ má asymptoticky (pro $T \rightarrow \infty$) χ^2 - rozdělení o 1 stupni volnosti.

(Věrohodnostní funkce normálního rozdělení má tvar

$$L(\mu, \sigma^2) = \frac{1}{(2\pi)^{T/2} \sigma^T} \cdot \exp\left(-\frac{\sum_{t=1}^T (x_t - \mu)^2}{2\sigma^2}\right)$$

⁶ Bude nutné doplnit, co je rozuměno konzistencí a nestranností testu.

Definice slabé exogenity

Množina (vysvětlujících) proměnných x_t v parametrizovaném modelu se nazývá **slabě exogenní**, jestliže celý model může být zapsán pomocí marginální hustoty proměnných $f(x_t)$ a podmíněné hustoty pro $f(y_t|x_t)$ tak, že odhad parametrů podmíněné hustoty není méně vydatný než odhad celé množiny parametrů obsažených ve sdružené hustotě. Toto bude splněno, jestliže žádný z parametrů podmíněné hustoty $f(y_t|x_t)$ se neobjeví v marginální hustotě pro $f(x_t)$.

Definice predeterminovanosti

Množina (vysvětlujících) proměnných x_t v parametrizovaném modelu se nazývá **predeterminovaná**, jestliže x_t jsou nezávislé na všech následujících strukturálních náhodných složkách ε_{t+s} pro $s > 0$. Takto definované predeterminované proměnné lze považovat (*přínejmenším asymptoticky*) jako by byly exogenní v tom smyslu, že pokud se vyskytnou jako regresory (může jít o zpožděné hodnoty závisle proměnné), získáme vždy konzistentní odhady parametrů.

Definice Grangerovy (ne)kauzality

Grangerova (ne)kauzalita (chápaná jako jistý druh statistické zpětné vazby) není přítomna tehdy, když se rovnají podmíněné hustoty

$$f(x_t|x_{t-1}, y_{t-1}) = f(x_t|x_{t-1})$$

Definice vyjadřuje, že v podmíněné hustotě $f(x_t|x_{t-1}, y_{t-1})$ nepřinášejí zpožděné veličiny y_{t-1} žádný další informační přínos k vysvětlení chování závisle proměnných x_t jiný než ten, který přinášejí zpožděné hodnoty proměnných x_{t-1} .

Definice silné exogenity

Množina (vysvětlujících) proměnných x_t se nazývá **silně exogenní**, jestliže tato množina splňuje podmínky **slabé exogenity** a jestliže **proměnné** y_{t-1} **nepůsobí v Grangerově smyslu na proměnné** x_t .

pramen: Engle,R, Granger,C., Richard,J.: Exogeneity. *Econometrica* 51/1983 p.277-304

6. Konkrétní výsledky testování nestrannosti expektace daňových výnosů

V našem případě, je-li hypotéza o regresních parametrech formulována jako $\beta_0=0$, $\beta_1=1$, dostáváme pro jednotlivé situace tyto výsledky :

1) test u daně z příjmů fyzických osob : $b_0 = - 0,6363$, $b_1 = 1,0337$

Výsledek F-testu nulové hypotézy ($\Phi_{2,6}^*=0,1783$) vede k závěru, že tuto hypotézu nelze zamítnout na žádné konvenční hladině významnosti. Stejný závěr poskytuje i testování pomocí zmíněných tří testů založených na metodě maximální věrohodnosti, neboť žádná empiricky spočtených testových hodnot (0,4488 pro LM, 0,4618 pro LR resp. 0,4754 pro AW) se nijak nepřibližuje ani na hladině významnosti 0,10 kritickým hodnotám χ^2 rozdělení o 2 stupních volnosti.- $\chi^2_{2}(0,10) = 10,6446$. Lze vyvodit, že posuzování odhadu skutečnosti u této daňové kategorie decizním subjektem je objektivní a není systematicky zkreslené.

2) test u daně z příjmů právnických osob : $b_0 = 0,1116$, $b_1 = 0,4701$

Vzhledem k velmi nízké hodnotě koeficientu determinace ($R^2 = 0,09$) pro formulovanou regresní rovnici lze učinit jen opatrné závěry : F-statistika ($\approx 0,39$) leží opět zřetelně pod všemi kritickými hodnotami příslušného F-rozdělení, analogicky ani empirická hodnota Waldova testu ($\eta_{AW}^* = 1,04$) se ani zdaleka neblíží hodnotě “nejmírnější” průkaznosti χ^2 -statistického testu : opět $\chi^2_{2}(0,10) = 10,6446$.

Opět lze podmíněně – s ohledem na velmi špatné výsledky regresního vystižení – vyvodit, že odhad vývoje daně z příjmů právnických osob není poznamenán systematickým vychylováním ani úrovně ani sklonu, resp. přesněji, že výsledky statistické analýzy (provedené v mezích možností dostupných dat) neprokazují opak.

3) test u daně přidané hodnoty : $b_0 = - 2,8389$, $b_1 = 0,9957$

Ani v tomto případě neprokazuje výsledek statistického testování hypotézy $\beta_0=0$, $\beta_1=1$ zřetelný náznak ve prospěch alternativy (lze to ostatně očekávat již z pohledu na regresní parametry, z nichž β_1 je velmi blízký 1 (0,996) a β_0 je sice od nuly poněkud vzdálen (-2,84), avšak je statisticky zcela nevýznamný). Proto F-test formulované hypotézy poskytující hodnotu jen 0,575 je značně vzdálen od kritické hodnoty $F(0,1)$, která je 3,46. Ke stejnému závěru dospějeme u kterémkoliv z ML-testů, kde nejnižší hodnota $\eta_{LR}^* = 1,041$ (daná testem Lagrangeových multiplikátorů) leží hluboko pod hodnotou $\chi^2_{2}(0,10)$. Ani zde tedy nelze prokázat záměrné vychylování v korekcích cílení daňových výnosů.

4) test u spotřební daně : $b_0 = 1,6599$, $b_1 = 0,4664$

Přestože první pohled na regresní koeficienty b_0 , b_1 by – zejména u druhého – nenasvědčoval blízkosti hodnot, které jsou přijímány pro test nulové hypotézy, způsobuje opět špatná kvalita vystižení regresního vztahu zvoleným regresorem možnost, že hypotéza $\beta_0 = 0$, $\beta_1 = 1$ nebude zamítnuta. Koeficient determinace R^2 je totiž jen 0,275 a ani jedna z obou z t-statistik parametrů nepřekračuje příslušnou kritickou hodnotu (při $\alpha = 0,05$ je $t_6(0,05) = \dots$). Opět jak F-test s hodnotou 1,4928, tak všechny testy LM, LR a AW nevedou k zamítnutí vyslovené hypotézy ani na nejmírnější významnostní hladině. Systematické zkreslení tedy není nijak indikováno.

5) test u výdajů na sociální a zdravotní pojištění : $b_0 = 20,7024$, $b_1 = -0,2247$

V tomto případě zaznamenáváme v regresním odhadu hodnoty výrazně vzdálené od testovaných. Pohled na příslušné t-statistiky 2,766 resp. -0,560 naznačuje v prvním případě zřetelnou statistickou významnost prvního regresora (jedničkového členu), nehledě na téměř zanedbatelnou nízkou hodnotu koeficientu determinace $R^2 = 0,043$. Lze konstatovat, že výsledek testování hypotézy $\beta_0 = 0$, $\beta_1 = 1$ je tentokrát na rozhraní přijetí či zamítnutí : F-test

dává pozitivní (ve smyslu zamítavý) výsledek na nejslabší hladině významnosti 0,10: $4,157 > 3,463$ (na této hladině se tedy hypotéza zamítá), přičemž z trojice navazujících testů je pozitivní hodnota Waldova testu ($\eta^*_{AW} = 11,086$), která je větší než příslušná kritická hodnota $\chi^2_2(0,10) = 10,6446$. Při volbě vyšší úrovně signifikance nebo z jiného trojice ML-testů by však ani zde k zamítnutí nulové hypotéza nedošlo.

Jde tedy o první případ, kdy lze zaznamenat určitý náznak možného systematického vychýlení (dle významnosti parametrů především u úroňové konstanty) a že daňový decizor či analytik možná přihlíží k předchozím chybám v odhadech tohoto typu daňových výnosů.

6) test u daně z nemovitostí

$$b_0 = 0,1861 \quad , \quad b_1 = -0,0906$$

Situaci poněkud podobnou předchozí nacházíme i zde : Výsledky regresní analýzy (při naprosto špatné míře vystižení rovnice daty : R^2 je jen 0,0012) udávají hodnoty parametrů regresní přímky velmi nízké. Pro parametr sklonu přímky, který je sice téměř nulový ale statisticky zcela nesignifikantní, z toho lze stěží něco vyvodit, přesto náznak slabounké signifikance úroňové konstanty (t-statistika = 1,52) spolu s blízkostí 0 parametru sklonu vedou k připuštění možnosti, že se jisté systematické vychylující prvky mohou v uvažování daňového decizora či korektora mohou projevit : Opět jak hodnota F-testu (4,344) – překročením kritické hodnoty na nejnižší hladině 0,01, která je 3,463 - tak i Waldova testu (jako jediného z trojice) – překročením kritické hodnoty $\chi^2_2(0,10)$ na stejné hladině významnosti - naznačují možnou přítomnosti slabé systematické chyby.

7) test u souhrnu daňových/pojistných příjmů státního rozpočtu : $b_0 = 21,4497$, $b_1 = -0,2218$

Zde stojí za pozornost, že i když se – s výjimkou pojistného – u všech ostatních daňových kategorií – neprojevil žádný náznak systematické chyby, zde je její možná přítomnost indikována nejsilněji. Opatrné hodnocení výsledků (s hodnotami spočtených regresních parametrů hodně vzdálenými od hypotetických) při znatelné významnosti nenulovosti konstantního členu a přes zanedbatelnou kvalitu regresního odhadu ($R^2 = 0,05$) vedou k opět vzájemně konzistentním závěrům, že :

- Výsledek poskytnutý F-testem (6,50) vede k zamítnutí nulové hypotézy tentokrát nejen na hladině 0,1, ale i na “tradiční” hladině 0,05 ($F(0,05) = 5,143$)
- Výsledek indikovaný ML-testy udává, že “nejméně průkazný” Waldův test (s hodnotou $\eta^*_{AW} = 17,33$) je signifikantní dokonce i na nejpřísnější hladině 0,01 ($\chi^2_2(0,10) = 16,81$), zatímco oba ostatní signalizují nevýznamnost (resp. v případě LR testu náznak možné slabé významnosti s přiblížením se k 90% pravděpodobnostní průkaznosti). Lze konstatovat, že při této “velmi špatné kvalitě regresního vystižení a velké vzdálenosti odhadnutých hodnot od hypotetických/testovaných” může dojít ke značným rozdílům v hodnocení reality jednotlivými z trojice ML-testů.

Propočty, které byly podkladem pro vyslovení předchozích závěrů byly provedeny algoritmy naprogramovanými RNDr. Daliborem Moravským ,CSc. tvořícími součást interního programového vybavení ELIMO zpracovaného v jazyce MATLAB. Je nicméně třeba mít na paměti, že s ohledem na malý počet pozorování časových řad, které byly podkladem regresních propočtů, a na obtížnou verifikovatelnost normality rozdělení náhodných složek při tak krátkých časových řadách, je třeba zachovat jistou obezřetnost při případné snaze o zobecnění těchto (předběžných) výsledků.