

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

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

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

\mathbf{X} $T \times (k + 1)$ -rozměrná matice vysvětlujících proměnných $\mathbf{X} = \{x_{tj}\}$

$$t = 1, 2, \dots, T; j = 1, 2, \dots, k, k + 1$$

$\boldsymbol{\beta}$ $(k + 1)$ -členný vektor parametrů příslušných vysvětlujícím proměnným

$\boldsymbol{\varepsilon}$ T – členný vektor náhodných složek (disturbancí) regresní rovnice

1B. Hypotézy o parametrech regresního modelu

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

$$\beta_j = 0,$$

kteřé 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

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

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)

2. Obecná definice testů

Nejjednodušší testovací problém je postaven tak, že data jsou generována sdruženou hustotou $\mathbf{f}(\mathbf{y}, \theta^0)$ odpovídající parametrům nulové hypotézy, resp. $\mathbf{f}(\mathbf{y}, \theta)$ 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

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

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

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

Definujeme-li veličiny $\mathbf{s}(\theta, \mathbf{y}) = \frac{\partial \mathbf{L}(\theta; \mathbf{y})}{\partial \theta}$ 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

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

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

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.

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 hypotézy.**

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

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

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

Jestliže na skóry 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 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 $\theta^0 - \hat{\theta}$

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 znatelné 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** $\theta^0 - \hat{\theta}$, 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 zaznameneáme:

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

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

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

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

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

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

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

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 hodnot 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

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

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:

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

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 porizení odhadu σ^2 . Vektor skóru a informační matice odpovídající parametrům μ budou tyto:

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

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

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 $\bar{\theta}$, ale obecně ne v odhadu $\tilde{\theta}$ odpovídajícím nulové hypotéze.

Tři testové statistiky nabudou v tomto případě tvar:

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

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

(17) **LM-test**
$$\eta_{LM} = \tilde{\epsilon}'X_1 \left(X_1'X_1 - X_1'X_2 \cdot (X_2'X_2)^{-1} X_2'X_1 \right)^{-1} X_1'\tilde{\epsilon} / \tilde{\sigma}^2 \quad (6)$$

kde

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

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

-první [Txk₁] matice pozorování odpovídá parametrům podrobeným testu

-druhá [Tx(k-k₁)] 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)
$$\eta_{AW} = T \cdot \left(\frac{\tilde{\epsilon}'\tilde{\epsilon} - \hat{\epsilon}'\hat{\epsilon}}{\hat{\epsilon}'\hat{\epsilon}} \right) \quad \eta_{LM} = T \cdot \left(\frac{\tilde{\epsilon}'\tilde{\epsilon} - \hat{\epsilon}'\hat{\epsilon}}{\tilde{\epsilon}'\tilde{\epsilon}} \right)$$

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

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.

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

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

³ Důkaz provedli (pro zobecněný lineární regresní model)

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:

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

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

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

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

$$\eta_{LM}^* = \eta_{LR}^* - \frac{\eta_{LR}^{*2}}{2!} + \frac{\eta_{LR}^{*3}}{3!} - \frac{\eta_{LR}^{*4}}{4!} \dots$$

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 endogenních 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:

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

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

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

⁵ viz Engle [1979a]

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 \mathbf{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

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

Zde (jen zdánlivě) zmizel z testové statistiky odhadnutý rozptyl $\hat{\sigma}^2$; 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*:

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.