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) $$\mathbf{y_t} = \beta_0.\mathbf{x_{t0}} + \beta_1.\mathbf{x_{t1}} + \beta_2.\mathbf{x_{t2}} + \beta_3.\mathbf{x_{t3}} + \dots\dots \beta_{\mathbf{k}}.\mathbf{x_{tk}} + \varepsilon_{\mathbf{t}}$$

$$\mathbf{t} = 1,2,\dots,\mathbf{T} \quad \text{(počet pozorování vzorku)}$$

neboli v maticové notaci

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

$\mathbf{y}$ .... $T-$členný vektor vysvětlované proměnné $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T)$
$\mathbf{X}$ .... $Tx(k+1)$-rozměrná matice vysvětlujících proměnných $\mathbf{X} = \{\mathbf{x_{tj}}\}$

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

β .... $(k+1)$-členný vektor parametrů příslušných vysvětlujícím proměnným
ε .... $T-$ členný vektor náhodných složek (disturbancí) regresní rovnice

1B. Hypotézy o parametrech regresního modelu

Formulujeme standardní lineární regresní model ve tvaru (1)

$$\mathbf{y_t} = \beta_0.\mathbf{x_{t0}} + \beta_1.\mathbf{x_{t1}} + \dots + \beta_{\mathbf{j}-1}.\mathbf{x_{t,j-1}} + \underline{\beta_{\mathbf{j}}.\mathbf{x_{t,j}}} + \dots\dots \beta_{\mathbf{k}}.\mathbf{x_{tk}} + \varepsilon_{\mathbf{t}}$$

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

$$\beta_{\mathbf{j}} = 0\,,$$

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

$$\mathbf{H}_1 : \beta_{\mathbf{j}} \neq 0.$$

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

(2a) $$\mathbf{y_t} = \beta_0.\mathbf{x_{t0}} + \beta_1.\mathbf{x_{t1}} + \beta_{\mathbf{j}-1}.\mathbf{x_{t,j-1}} + \underline{\beta_{\mathbf{j}}.\mathbf{x_{t,j}}} + \dots\dots \beta_{\mathbf{k}}.\mathbf{x_{tk}} + \varepsilon_{\mathbf{t}}$$

tedy se zahrnutím $\mathbf{j-té}$ vysvětlující proměnné

(2b) $$\mathbf{y_t} = \beta_0.\mathbf{x_{t0}} + \beta_1.\mathbf{x_{t1}} + \beta_{\mathbf{j}-1}.\mathbf{x_{t,j-1}} + \beta_{\mathbf{j}+1}.\mathbf{x_{t,j+1}} + \dots\dots \beta_{\mathbf{k}}.\mathbf{x_{k0}} + \varepsilon^*{}_{\mathbf{t}}$$

tedy bez zahrnutí $\mathbf{j-té}$ vysvětlující proměnné.

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

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

<u>*v dalším výkladu užijeme značení:*</u>

$\boldsymbol{\theta}$ ... obecná (*nespecifikovaná, ale pevná)* hodnota parametru (bez omezení)

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

$\boldsymbol{\theta}^0$ .....hypotetická (hypotézou H_0 předpokládaná) hodnota parametru $\boldsymbol{\theta}$.

($\boldsymbol{\theta}$ v dalším může být jak skalární hodnota, tak k-členný vektor)

<u>2. Obecná definice testů</u>

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 jako[1]

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

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

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

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

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

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

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

Waldův test je asymptotickou aproximací známých a oblíbených t a F testů

[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.

v ekonometrii: Lze ukázat, že jestliže $\hat{\theta}$ má asymptoticky normální rozdělení, a jestliže lze $\Phi(\theta)$ konzistentně odhadnout pomocí $\boldsymbol{\Phi}(\hat{\boldsymbol{\theta}})$, pak statistika

(5) $$\xi_{\mathbf{W}} = \mathbf{T}.(\hat{\theta}-\theta)'\Phi(\theta)(\hat{\theta}-\theta)$$

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) $$\xi_{\mathbf{LR}} = -2.(\mathbf{L}(\hat{\theta}\,;\mathbf{y}) - \mathbf{L}(\theta\,;\mathbf{y}))$$

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 $\theta = \theta$ 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. Lagrangián

(7) $$\mathbf{H} = \mathbf{L}(\theta\,;\mathbf{y}) - \lambda.(\theta-\theta^0),$$

pak jsou podmínky 1.řádu dány vztahy

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

To znamená, že test založený na principu Lagrangeových multiplikátorů Aitchesona 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 skórů snadno nalézt za nulové hypotézy, protože vektor skórů bude mít nulovou střední hodnotu a rozptyl rovný $\Phi(\theta).T$

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

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

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 $\theta^0 - \theta$

b) LR test je formulován v podmínkách rozdílu $L(\theta^0;y) - L(\hat{\theta}\,;y)$

c) LM test je formulován v podmínkách $s(\theta^0)$

<u>obrázek</u> 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 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, restrikce někdy vedou k odstranění nelinearit, což by naznačovalo výhodnost užití LM-testu.

VĚTA V případě, že uvažovaná věrohodnostní funkce má tvar kvadratické formy,

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

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

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

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

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

protože $L(\hat{\theta}\ ;y) = b$ □.

Kdykoliv je skutečná hodnota θ rovná nebo blízká předpokládané $\boldsymbol{\theta^0}$, pak bude věrohodnostní funkce v okolí $\boldsymbol{\theta^0}$ přibližně kvadratická pro velké výběry, tím, že matice A závisí pouze na $\boldsymbol{\theta^0}$. 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) $$y^*\middle|\boldsymbol{X}^* \approx N\left(\boldsymbol{X}^*\beta,\sigma^2.I_T\right)$$, kde

y^* je T x 1 vektor závisle proměnné

$\boldsymbol{X}^*$ je T x k matice nezávisle proměnných

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

R je k_1 x k matice známých konstant

r je k_1 x 1 vektor konstant (podmínek v mezeních)

Jestliže R má hodnost 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) $$y\middle|\boldsymbol{X} \approx N\left(\boldsymbol{X}\mu,\sigma^2.I_T\right)$$, kde

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

a kde transformované proměnné $y,\boldsymbol{X}$ jsou lineární kombinace původních $y^*,\boldsymbol{X}^*$.

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) $$L(\mu,y) = k - \frac{T}{2}.\ln\sigma^2 - \frac{1}{2\sigma^2}.(y-\boldsymbol{X}\mu).(y-\boldsymbol{X}\mu)$$,

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 pořízení odhadu σ^2. Vektor skórů a informační matice odpovídající parametrům μ budou tyto:

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

[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 $\boldsymbol{X}$. 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) **LR-test** $$\eta_{LR} = T.\ln\left(\frac{\tilde{e}'\tilde{e}}{\hat{e}'\hat{e}}\right)$$

(16) **AW-test** $$\eta_{AW} = \left(\mu_1^0 - \hat{\mu}_1\right)\left(\boldsymbol{X}_1'\boldsymbol{X}_1 - \boldsymbol{X}_1'\boldsymbol{X}_2.(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{X}_1\right)\left(\mu_1^0 - \hat{\mu}_1\right)/\hat{\sigma}^2$$

(17) **LM-test** $$\eta_{LM} = \tilde{\varepsilon}'\boldsymbol{X}_1\left(\boldsymbol{X}_1'\boldsymbol{X}_1 - \boldsymbol{X}_1'\boldsymbol{X}_2.(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{X}_1\right)^{-1}\boldsymbol{X}_1'\tilde{\varepsilon}/\tilde{\sigma}^2 \qquad (6)$$

kde

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

matice $\boldsymbol{X}$ je rozdělena na $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$:

-první [Txk_1] matice pozorování odpovídá parametrům podrobeným testu

-druhá [$Tx(k-k_1)$] 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) $$\eta_{AW} = T.\left(\frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\hat{e}'\hat{e}}\right) \qquad \eta_{LM} = T.\left(\frac{\tilde{e}'\tilde{e} - \hat{e}'\hat{e}}{\tilde{e}'\tilde{e}}\right)$$

Připomeňme, že v této notaci vždy platí: $\tilde{e}'\tilde{e} \geq \hat{e}'\hat{e}$, protože $\tilde{e}'\tilde{e}$ odpovídá ***omezenému*** (jde o SSE, pokud platí nulová hypotéza H_0) a $\hat{e}'\hat{e}$ ***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:

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

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

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

<u>Ověření</u> provedeme jednoduchým dosazením výrazů v definicích (15) ,(19A), (19B):

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

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

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

$$\left(\frac{\eta_{AW}}{1+\frac{\eta_{AW}}{T}}\right) = \left(\frac{T.\frac{\tilde{e}'\tilde{e}-\hat{e}'\hat{e}}{\hat{e}'\hat{e}}}{1+\frac{\tilde{e}'\tilde{e}-\hat{e}'\hat{e}}{\hat{e}'\hat{e}}}\right) = \left(T.\frac{\tilde{e}'\tilde{e}-\hat{e}'\hat{e}}{\tilde{e}'\tilde{e}}\right) = \eta_{LM}$$,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)
$$\eta_{\mathbf{AW}} \geq \eta_{\mathbf{LR}} \geq \eta_{\mathbf{LM}}$$

---

[3] Důkaz provedli (pro zobecněný lineární regresní model)

<u>Ověření</u>[4]*: provedeme porovnáním výrazů v (19A), (19B), (19C)::*

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

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

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

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

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

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

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

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

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

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

kde jsme opět označili $\eta_{LR}{}^{*} = \dfrac{\eta_{LR}}{T}$ a podobně $\eta_{LM}{}^{*} = \dfrac{\eta_{LM}}{T}$. Odtud máme po odlogaritmování $\exp\left(\eta_{LR}{}^{*}\right) = \dfrac{1}{1-\eta_{LM}{}^{*}}$ a po jednoduchých úpravách $\exp\left(\eta_{LR}{}^{*}\right) = 1+\eta_{LM}{}^{*}\exp\left(\eta_{LR}{}^{*}\right)$ neboli $\eta_{LM}{}^{*} = 1-\exp\left(-\eta_{LR}{}^{*}\right)$. Opět rozvedeme pravou stranu pomocí Taylorova rozvoje

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

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

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

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

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 zamítat 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) $$y_i = Y_i \beta_{.i} + X_i \gamma_{.i} + \varepsilon_i = Z_i \delta_{.i} + \varepsilon_i \quad ,$$

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

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) $$\xi_{\mathrm{w}} = \frac{\hat{\delta}_1'.\left(\mathbf{Z}_1'\mathbf{P}.\mathbf{Z}_1 - \mathbf{Z}_1'\mathbf{P}.\mathbf{Z}_2.(\mathbf{Z}_2'\mathbf{P}.\mathbf{Z}_2)^{-1}\mathbf{Z}_2'\mathbf{P}.\mathbf{Z}_1\right).\hat{\delta}_1}{\hat{\sigma}^2} \quad ,$$

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

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 $\hat{\delta}_2 = (\mathbf{Z}_2'\mathbf{P}\mathbf{Z}_2)^{-1}\mathbf{Z}_2'\mathbf{P}\mathbf{y}$ a dále definujeme $\tilde{\mathbf{e}} = \mathbf{y} - \mathbf{Z}_2.\tilde{\delta}_2$, může být Waldova statistika vyjádřena jako

(23) $$\xi_{\mathrm{w}} = \frac{\mathbf{T}.(\tilde{\mathbf{e}}'\mathbf{P}\tilde{\mathbf{e}} - \hat{\mathbf{e}}'\mathbf{P}\hat{\mathbf{e}})}{\hat{\mathbf{e}}'\mathbf{P}\hat{\mathbf{e}}} \quad ,$$

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 2SLSL 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 $\boldsymbol{P}$ 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) $$\hat{\boldsymbol{\delta}} = \left(\boldsymbol{Z' PZ}\right)^{-1} \boldsymbol{Z' Py},$$

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

(25) $$\xi_w = \hat{\boldsymbol{\delta}}_1'.\left(\mathbf{Z}_1'\mathbf{P}.\mathbf{Z}_1 - \mathbf{Z}_1'\mathbf{P}.\mathbf{Z}_2.\left(\mathbf{Z}_2'\mathbf{P}.\mathbf{Z}_2\right)^{-1}\mathbf{Z}_2'\mathbf{P}.\mathbf{Z}_1\right)\hat{\boldsymbol{\delta}}_1 ,$$

kterýžto výraz může být reformulován do podoby

(26) $$\xi_w = \mathbf{T}.\left(\tilde{\mathbf{e}}'\mathbf{P}\tilde{\mathbf{e}} - \hat{\mathbf{e}}'\mathbf{P}\hat{\mathbf{e}}\right),$$

Zde (jen zdánlivě) zmizel z testové statistiky odhadnutý rozptyl $\hat{\sigma}^2$; ve skutečnosti se však $\hat{\sigma}^2$ nachází v obsahu vektoru $\hat{\boldsymbol{\delta}}_1$ (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.

[5] Viz Engle (1979a)