

CZESŁAW DOMAŃSKI

Uniwersytet Łódzki

UWAGI O PROBLEMACH AUTOKORELACJI W PROCESIE EKONOMETRYCZNEGO MODELOWANIA I PROGNOZOWANIA

1. Uwagi wstępne

Rozważmy model ekonometryczny postaci:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad (1)$$

gdzie:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x_{10} & x_{11} & \cdots & x_{1k} \\ x_{20} & x_{21} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_{n0} & x_{n1} & \cdots & x_{nk} \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \boldsymbol{\alpha}^T = [\alpha_0, \alpha_1 \cdots \alpha_k].$$

Zwykle, w badaniach ekonometrycznych przyjmuje się następujące założenia modelu (1):

- stabilność,
- liniowość,
- elementy macierzy \mathbf{X} nie są losowe,
- $\text{rz}\mathbf{X} = k + 1 < n$,
- $\mathbf{E}\boldsymbol{\varepsilon} = \mathbf{0}$,
- $\mathbf{D}^2\boldsymbol{\varepsilon}$ jest macierzą diagonalną (brak autokorelacji),

- diagonalne elementy macierzy $D^2\varepsilon$ są sobie równe (homoskedastyczność),
- wektor losowy ε ma n -wymiarowy rozkład normalny.

Założmy teraz, że proces stochastyczny $\{\varepsilon_t\}$, $t = 1, 2, \dots, n$ jest stacjonarny. Oznacza to między innymi, że:

$$\text{cov}(\varepsilon_t, \varepsilon_{t-s}) = \text{cov}(\varepsilon_u, \varepsilon_{u-s}) = c_s$$

dla $t, u, s = 1, 2, \dots, n$ przy $s \leq t, s \leq u$.

Przyjmując, że $\rho_s = c_s / c_0$, oczywiście $c_0 = \sigma^2$, mamy:

$$D^2\varepsilon = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \cdots & \rho_{n-2} \\ \dots & \dots & \dots & \dots \\ \rho_{n-1} & \rho_{n-3} & \cdots & \rho_1 \\ \rho_{n-1} & \rho_{n-2} & \cdots & 1 \end{bmatrix}. \quad (2)$$

Z równości (2) można wyprowadzić równanie:

$$\varepsilon_n = \gamma_1 \varepsilon_1 + \gamma_2 \varepsilon_2 + \dots + \gamma_{n-1} \varepsilon_{n-1} + \eta_n \quad (3)$$

gdzie:

$$\text{cov}(\eta_n, \varepsilon_t) = 0 \text{ dla } t = 1, 2, \dots, n-1,$$

a współczynniki korelacji cząstkowej γ_i są funkcjami współczynników korelacji całkowitej ρ_i .

Równanie (2) przedstawia liniowy schemat autoregresji rzędu $n-1$ procesu $\{\varepsilon_t\}$. W badaniach empirycznych na ogół stosuje się schematy niższych rzędów. Często zachodzi konieczność redukcji liczby nieznanymi parametrów, a ponadto z doświadczeń wynika, że pominięte w modelu zmienne reprezentowane przez odchylenia losowe na ogół wykazują znaczną inercję w czasie. Kolejno po sobie występujące składniki losowe ε_t są zatem dość silnie skorelowane, natomiast zależności pomiędzy odległymi składnikami są stosunkowo słabe. Zatem, zwykle rozpatruje się schematy autoregresji:

$$\varepsilon_t = \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2} + \dots + \gamma_m \varepsilon_{t-m} + \eta_t, \quad (4)$$

$t = m+1, m+2, \dots, n$ gdzie $m < n-1$, przy dodatkowych założeniach o procesie η_t :

$$\text{cov}(\eta_t, \eta_u) = \begin{cases} 0 & \text{dla } t \neq u \\ \sigma_\eta^2 & \text{dla } t = u \end{cases}. \quad (5)$$

Najczęściej rozpatrywany jest model (4) przy $m = 1$. Zamieniając symbol γ_1 na ρ w modelu (4) dla $m = 1$, mamy:

$$\varepsilon_t = \rho\varepsilon_{t-1} + \eta_t, \quad (6)$$

przy czym proces η_t spełnia warunek (5). Zauważmy, że ρ jest współczynnikiem korelacji pierwszego rzędu w procesie ε_t , a więc $|\rho| < 1$ oraz $\text{cov}(\varepsilon_t, \varepsilon_{t-s}) = \sigma_\varepsilon \rho^s$ dla $t = 1, 2, \dots, n, s = 1, 2, \dots, t-1$. Zatem ρ^s jest współczynnikiem korelacji rzędu s procesu ε_t . Stąd zamiast równania (2) możemy zapisać:

$$D^2\varepsilon = \sigma^2\Omega(\rho), \quad (7)$$

$$\Omega(\rho) = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{n-2} & \rho^{n-3} & \rho^{n-4} & \dots & \rho \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix} \quad (8)$$

Dla $|\rho| < 1$ macierz $\Omega(\rho)$ jest dodatnio określona, przy czym oczywiście $\Omega(0) = I$.

2. Testy weryfikujące hipotezę o braku autokorelacji

Najstarszym testem nieparametrycznym weryfikującym hipotezę $H_0 : \rho = 0$ jest test Neumana (1941) oparty na statystyce postaci:

$$Q = \frac{n \sum_{t=2}^n (v_{t-1} - v_t)^2}{(n-1) \sum_{t=1}^n (v_t - \bar{v})^2}, \quad (9)$$

gdzie v_1, v_2, \dots, v_n są niezależnymi zmiennymi losowymi o tym samym rozkładzie normalnym, a

$$\bar{v} = \frac{1}{n} \sum_{t=1}^n v_t.$$

Autor testu wykazał, że statystyka (9) jest zbieżna asymptotycznie do rozkładu normalnego

$$N\left(\frac{2n}{n-1}; \frac{4n^2(n-1)}{(n+1)(n-3)^3}\right).$$

Hart (1942), opierając się na wynikach von Neumanna, zbudował tablice kwantyli statystyki (9). W przypadku weryfikacji hipotezy o nieskorelowaniu ε_t w modelu ekonometrycznym w miejsce v_t przyjmujemy oceny e_t - elementów wektora reszt.

Najczęściej stosowany jest test d Durбина-Watsona (1950, 1951, 1971) oparty na statystyce:

$$d = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}, \quad (10)$$

weryfikujący hipotezę, że zakłócenia generowane przez stacjonarny proces autoregresyjny pierwszego stopnia postaci (6), gdzie zakłócenia η_t mają niezależne, identyczne rozkłady normalne $N(0, \sigma_\eta^2)$.

Z założenia $E(\eta) = 0$ wynika konieczność wprowadzenia wyrazu wolnego do wyjściowego równania.

Weryfikujemy, jak poprzednio, hipotezę $H_0 : \rho = 0$ przeciwko hipotezie alternatywnej $H_1 : \rho > 0$ albo $H_1 : \rho < 0$.

Test ten jest, jak wiadomo, niekonkluzywny, tzn. nie dla każdej wartości statystyki można podjąć decyzję odrzucenia hipotezy H_0 lub H_1 .

Kolejnym testem weryfikującym hipotezę H_0 jest test Geary'ego (1970), którego funkcja testowa przyjmuje postać:

$$\tau = \sum_{t=2}^n k_t, \quad k_t = \begin{cases} 1, & \text{gdy } \varepsilon_t \varepsilon_{t-1} \leq 0 \\ 0, & \text{gdy } \varepsilon_t \varepsilon_{t-1} > 0 \end{cases} \quad (11)$$

którą słownie można wyrazić jako liczbę zmian znaku w ciągu reszt modelu (6).

W literaturze można znaleźć również inne testy dla rozważanego problemu (por. np. Domański 1990, Domański, Pruska 2000).

Przeprowadziliśmy badanie mocy testu Geary'ego po uprzedniej randomizacji. W modelu (6) przyjęliśmy $\rho = 0$ oraz rozkład normalny zmiennej η_t . Zastosowaliśmy następującą zmodyfikowaną procedurę. Niech

$$\tau_\alpha = \max\{r : P(\tau \leq r) \leq \alpha\}. \quad (12)$$

Odrzucamy H_0 , gdy $\tau > \tau_\alpha + 1$, przyjmujemy H_0 , gdy $\tau \leq \tau_\alpha$, odrzucamy H_0 z prawdopodobieństwem randomizacyjnym $p^R(\alpha)$, gdy $\tau = \tau_\alpha + 1$, przy czym

$$p^R(\alpha) = \frac{\alpha - p(\tau \leq \tau_\alpha)}{\alpha - p(\tau = \tau_\alpha + 1)}, \quad (13)$$

gdzie α jest wybranym poziomem istotności, τ_α oznacza całkowitą wartość krytyczną, tzn. wartość krytyczną testu niezrandomizowanego.

W celu oszacowania $p^R(\alpha)$ wygenerowaliśmy 10 000 prób modelu (6) dla kolejnych $n = 10, 15, \dots, 30$. Uwzględniając tę procedurę, wyniki badania mocy testu Geary'ego przedstawione są w tablicy 1.

Tablica 1

Porównanie mocy testów: Durбина-Watsona, Geary'ego oraz t-Studenta

n	ρ	Durbina-Watsona	Geary'ego	t-Studenta
10	0,0	0,06	0,05	0,05
	0,1	0,08	0,07	0,06
	0,3	0,17	0,12	0,11
	0,5	0,28	0,18	0,17
	0,7	0,38	0,24	0,24
	0,9	0,46	0,29	0,30
15	0,0	0,06	0,05	0,05
	0,1	0,09	0,08	0,07
	0,3	0,23	0,16	0,13
	0,5	0,43	0,27	0,21
	0,7	0,62	0,40	0,32
	0,9	0,74	0,51	0,41
20	0,0	0,05	0,05	0,05
	0,1	0,10	0,08	0,07
	0,3	0,31	0,18	0,13
	0,5	0,58	0,35	0,22
	0,7	0,79	0,54	0,35
	0,9	0,89	0,68	0,49
25	0,0	0,05	0,05	0,05
	0,1	0,11	0,08	0,07
	0,3	0,36	0,21	0,13
	0,5	0,68	0,42	0,24
	0,7	0,89	0,66	0,38
	0,9	0,96	0,80	0,53
30	0,0	0,05	0,05	0,05
	0,1	0,12	0,09	0,08
	0,3	0,44	0,25	0,14
	0,5	0,77	0,49	0,23
	0,7	0,94	0,73	0,38
	0,9	0,99	0,89	0,55

Źródło: Obliczenia własne.

Statystykę Durбина-Watsona próbowano uczynić konkluzywną kilkoma sposobami, na przykład poprzez aproksymację wartości krytycznych za pomocą rozkładu beta lub wielomianów Jacobiego, lub przez wyznaczenie gęstości rozkładu form kwadratowych rozkładów normalnych. Jednak momenty rozkładu statystyki d są bardzo skomplikowane i trudne do zastosowań komputerowych. W związku z tym,

przeprowadziliśmy symulacyjną ocenę wartości krytycznych d_α testu Durбина-Watsona dla kolejnych $n = 10, 15, \dots, 30$, a następnie zastosowaliśmy wygodną do aplikacji komputerowych formułę aproksymacyjną dla wartości krytycznych:

$$\tilde{\tau}_\alpha = \gamma_{-2}(\alpha)/n^2 + \gamma_{-1}(\alpha)/n + \gamma_0(\alpha) + \gamma_1(\alpha) \cdot n + \gamma_2(\alpha)n^2. \quad (14)$$

Wyniki badania mocy testu Durбина-Watsona dla takiego samego modelu jak w przypadku testu Geary'ego zawiera tablica 1. W ostatniej kolumnie, dla porównania, przedstawiono również moc znanego testu t-Studenta dla współczynnika korelacji $\rho = 0,0, 0,3, 0,5, 0,7$ i $0,9$.

Na podstawie przeprowadzonego badania metodą Monte Carlo (liczba powtórzeń 10^4), można wyciągnąć następujące wnioski:

1. Test Durбина-Watsona jest znacznie mocniejszy od dwóch pozostałych testów.
2. Moc testu Durбина-Watsona jest niewielka dla małych prób, nawet dla silnej autokorelacji $\rho = 0,9$, np. dla $n = 10$ jego empiryczna moc kształtuje się na poziomie 45%. Potwierdza to ogólne przekonanie o tym, że autokorelacja jest trudna do wykrycia.
3. Rozmiar randomizowanego testu Geary'ego jest zgodny z przyjętym poziomem istotności.

Bibliografia

- Domański Cz., (1990), *Testy statystyczne*, PWE, Warszawa
- Domański Cz., Pruska K., (2000), *Nieklasyczne metody wnioskowania statystycznego*, PWE, Warszawa
- Durbin G. S., Watson J., (1950), *Testing for serial correlation in least squares regression I*, „Biometrika” 37, s. 409–428
- Durbin G. S., Watson J., (1951), *Testing for serial correlation in least squares regression II*, „Biometrika” 38, s. 159–178
- Durbin G. S., Watson J., (1951), *Testing for serial correlation in least squares regression III*, „Biometrika” 58, s. 1–19
- Geary R. C. (1970), *Relative efficiency of count of sign changes for assessing residual autoregression in least squares regression*, „Biometrika” 57, s. 123–127
- Hart B. I. (1942), *Significance levels for the ratio on the mean-square successive difference to the variance*, AMS 13, s. 445–447
- Neuman von J. (1941), *Distribution of the mean square successive difference to the variance*, AMS 12, s. 367–395