

Weryfikacja założeń KMRL

Część 2

Test Durbina-Watsona (1950)

- James Durbin, Geoffrey S. Watson (1950). "Testing for serial correlation in least squares regression." . Biometrika. 37 (3-4): 409–428.
- Weryfikowana jest hipoteza

$$H_0 : \text{cov}(\varepsilon_t, \varepsilon_{t-1}) = 0$$

$$H_1 : \text{cov}(\varepsilon_t, \varepsilon_{t-1}) \neq 0$$

- Statystyka testowa

$$DW = \frac{\sum_{i=2}^T (e_t - e_{t-1})^2}{\sum_{i=1}^T e_t^2} = 2(1 - \rho_{\varepsilon_t \varepsilon_{t-1}}) - \frac{e_1^2 + e_T^2 - 2e_1 e_0}{\sum_{i=1}^T e_t^2}$$

Test Durбина-Watsona

- Dla dużej liczebności próby

$$DW \xrightarrow{P} 2(1 - \rho_{\varepsilon_t \varepsilon_{t-1}})$$

- Rozkład statystyki testowej

- 1 jeżeli zakładana jest dodatnia autokorelacja, wtedy $DW < 2$,
oraz

- a) $DW < d_L$, odrzucamy hipotezę zerową,
- b) $d_L < DW < d_U$ brak konkluzji,
- c) $DW > d_U$ nie ma podstaw do odrzucenia H_0 .

- 2 jeżeli zakładana jest ujemna autokorelacja, wtedy $DW > 2$,
oraz

- a) $DW > 4 - d_L$, odrzucamy hipotezę zerową,
- b) $4 - d_U < DW < 4 - d_L$ brak konkluzji,
- c) $DW < 4 - d_U$ nie ma podstaw do odrzucenia H_0 .

- 3 jeżeli $DW = 2$ to brak jest autokorelacji.

Test Durbina-Watsona

- Wady testu
 - 1 wskazuje jedynie obecność autokorelacji pierwszego rzędu
 - 2 wartości krytycznych nie można uzyskać analitycznie
 - 3 obszar braku konkluzji
 - 4 niska moc testu

Test Durбина - Watsona - przykład

```
. reg gnp armed_forces employment
```

| Source | SS | df | MS | |
|----------|------------|----|------------|------------------------|
| Model | 1.4336e+11 | 2 | 7.1679e+10 | Number of obs = 16 |
| Residual | 4.8328e+09 | 13 | 371753044 | F(2, 13) = 192.81 |
| Total | 1.4819e+11 | 15 | 9.8794e+09 | Prob > F = 0.0000 |
| | | | | R-squared = 0.9674 |
| | | | | Adj R-squared = 0.9624 |
| | | | | Root MSE = 19281 |

| gnp | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|--------------|-----------|-----------|--------|-------|----------------------|
| armed_forces | -.6047395 | 8.043962 | -0.08 | 0.941 | -17.98266 16.77318 |
| employment | 27.89106 | 1.593964 | 17.50 | 0.000 | 24.44751 31.33461 |
| _cons | -1432485 | 96466.32 | -14.85 | 0.000 | -1640888 -1224083 |

```
. estat dwatson
```

```
Durbin-Watson d-statistic( 3, 16) = 1.537074
```

Test Breuscha-Godfrey'a (1978)

- Trevor S. Breusch (1978). "Testing for Autocorrelation in Dynamic Linear Models". Australian Economic Papers. 17: 334–355.
- Leslie G. Godfrey (1978). "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables". Econometrica. 46: 1293–1301.
- Test wykorzystuje mnożniki Lagrange'a.
- Jest w stanie wskazać na obecność autokorelacji wyższych rzędów.

Test Breuscha-Godfrey'a (1978)

- Weryfikowana jest hipoteza

H_0 : brak autokorelacji

H_1 : $\varepsilon_i = AR(p) \vee \varepsilon_i = MA(p)$

- W obu przypadkach statystyka testowa ma identyczną postać

$$LM = TR_0^2 \quad (1)$$

Test Breuscha-Godfrey'a (1978)

- Jej wartość można obliczyć dwoma metodami:
 - 1 Sposób 1. szacowana jest wartość parametrów równania regresji
 - przeprowadzamy regresję wektora reszt na macierz zmiennych pomocniczych

$$e_t = \gamma_0 + \gamma_1 e_{t-1} + \gamma_2 e_{t-2} + \dots + \gamma_p e_{t-p} + \xi_t$$

- obliczana jest wartość współczynnika $LM = TR_0^2$. Statystyka testowa ma rozkład $\chi^2(p)$ ($F(p, N - p)$)
- 2 Sposób 2. Procedura rozpoczyna się od wyjściowego modelu
 - do oryginalnego równania regresji dodawane jest p kolumn, zawierających opóźnione reszty

$$y_t = X_t \beta + \gamma_1 e_{t-1} + \gamma_2 e_{t-2} + \dots + \gamma_p e_{t-p} + \psi_t$$

- weryfikowana jest łączną istotność opóźnionych reszt za pomocą statystyki $LM = TR_0^2$. Ma ona rozkład $\chi^2(p)$ ($F(p, N - p)$)

Test Breuscha - Godfreya - przykład

```
. estat bgodfrey, lags(1 2 3)
```

Breusch-Godfrey LM test for autocorrelation

| lags(p) | chi2 | df | Prob > chi2 |
|---------|-------|----|-------------|
| 1 | 0.472 | 1 | 0.4920 |
| 2 | 3.000 | 2 | 0.2231 |
| 3 | 3.161 | 3 | 0.3674 |

H0: no serial correlation

```
. estat bgodfrey, lags(1 2 3) small
```

Breusch-Godfrey LM test for autocorrelation

| lags(p) | F | df | Prob > F |
|---------|-------|-----------|----------|
| 1 | 0.472 | (1, 12) | 0.5050 |
| 2 | 1.500 | (2, 11) | 0.2654 |
| 3 | 1.054 | (3, 10) | 0.4112 |

Test RESET (1969)

- James B. Ramsey (1969). "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," Journal of the Royal Statistical Society, series B: 31:350–371.
- Test poprawności specyfikacji formy funkcyjnej modelu
- **Regression Equation Specification Error Test**
- Do modelu regresji liniowej

$$y = X\beta + \varepsilon$$

- Dodawana jest macierz dodatkowych regresorów Z

$$y = X\beta + Z\gamma + \varepsilon$$

- Weryfikowana jest hipoteza

$$H_0 : \gamma = 0$$

- przeciw alternatywie

$$H_0 : \gamma \neq 0$$

Test RESET

- Procedura testowa jest analogiczna do testu łącznej istotności
- Statystyka testowa ma rozkład $F(r(Z), N - k)$
- Alternatywna postać testu wykorzystuje rozwinięcie w szereg Taylora

$$y = \gamma_0 + \gamma_1 X\beta + \gamma_2 (X\beta)^2 + \dots + \gamma_p (X\beta)^p + \varepsilon$$

- Podstawiając wartość dopasowaną uzyskujemy

$$y = \gamma_0 + \gamma_1 \hat{y} + \gamma_2 \hat{y}^2 + \dots + \gamma_p \hat{y}^p + \varepsilon$$

- Test łącznej istotności

$$LM = nR^2 \sim_a \chi^2(p)$$

Test RESET - próba o dużej liczbie obserwacji

```
. estat ovtest, rhs  
(note: wiek2 dropped because of collinearity)  
(note: wiek2^2 dropped because of collinearity)
```

Ramsey RESET test using powers of the independent variables

Ho: model has no omitted variables

```
F(5, 16141) = 30.55  
Prob > F = 0.0000
```

```
. estat ovtest
```

Ramsey RESET test using powers of the fitted values of lzarobki

Ho: model has no omitted variables

```
F(3, 16142) = 12.36  
Prob > F = 0.0000
```

Test RESET - próba o małej liczbie obserwacji

```
. estat ovtest, rhs  
(note: wiek2 dropped because of collinearity)  
(note: wiek2^2 dropped because of collinearity)
```

Ramsey RESET test using powers of the independent variables

Ho: model has no omitted variables

F(5, 158) = 0.87
Prob > F = 0.4994

```
. estat ovtest
```

Ramsey RESET test using powers of the fitted values of lzarobki

Ho: model has no omitted variables

F(3, 159) = 0.25
Prob > F = 0.8606

Przekształcenie Boxa-Coxa

- George E. P. Box, David R. Cox (1964). "An analysis of transformations". Journal of the Royal Statistical Society, Series B. 26 (2): 211–252.
- Forma przekształcenia

$$g(x, \lambda) = \frac{x^\lambda - 1}{\lambda}$$

Przekształcenie Boxa-Coxa - przykład

```

Log likelihood = -100329.78
Number of obs   =    16162
LR chi2(16)    =    5185.61
Prob > chi2    =    0.000
-----+-----
      zarobki |      Coef.  Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      /theta |  -0.2597717  .0121652   -21.35  0.000   -0.283615   -0.2359284
  
```

Estimates of scale-variant parameters

```

-----+-----
              |      Coef.
-----+-----
Notrans      |
  _Iplec_2    |  -0.0603042
    wiek     |   0.0063774
    wiek2    |  -0.0000647
  _Iwyksztal^2 | -0.0502765
  _Iwyksztal^3 | -0.0522569
  _Iwyksztal^4 | -0.0468753
  _Iwyksztal^5 | -0.0779558
  _Iwyksztal^6 | -0.1028573
  _Iwyksztal^7 | -0.1412295
  _Iklm_12_1  |  -0.0071674
  _Iklm_12_2  |  -0.0150532
  _Iklm_12_3  |  -0.014285
  _Iklm_12_4  |  -0.0247102
  _Iklm_12_5  |  -0.0238557
  _Iklm_12_6  |  -0.0206737
  _Iklm_12_9  |  -0.0286794
  _cons      |    2.97119
  
```

Przekształcenie Boxa-Coxa - przykład

| Test | Restricted | LR statistic | P-value |
|------------|----------------|--------------|-------------|
| H0: | log likelihood | chi2 | Prob > chi2 |
| theta = -1 | -102742.96 | 4826.35 | 0.000 |
| theta = 0 | -100547.59 | 435.62 | 0.000 |
| theta = 1 | -105237.15 | 9814.73 | 0.000 |

Rozszerzenia regresji

- 1 modele wielomianowe
- 2 modele schodkowe
- 3 modele krzywej łamanej

Krzywa łamana - przykład

- Model uzależniający wydatki na żywność na osobę w gospodarstwie domowym od dochodu na osobę
- Dopuszczamy by wraz ze wzrostem dochodu zmieniała się elastyczność dochodowa wydatków na żywność
- Progi: 1000 zł oraz 4000 zł

| Source | SS | df | MS | Number of obs | = | 31,679 |
|----------|------------|--------|------------|---------------|---|---------|
| ----- | | | | | | |
| | | | | F(3, 31675) | = | 3536.24 |
| Model | 78921608.7 | 3 | 26307202.9 | Prob > F | = | 0.0000 |
| Residual | 235640314 | 31,675 | 7439.31536 | R-squared | = | 0.2509 |
| ----- | | | | | | |
| Total | 314561923 | 31,678 | 9929.98052 | Adj R-squared | = | 0.2508 |
| | | | | Root MSE | = | 86.251 |

| wz_os | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] | |
|-----------|-----------|-----------|--------|-------|----------------------|-----------|
| ----- | | | | | | |
| doch_os | .1372183 | .0022743 | 60.33 | 0.000 | .1327606 | .141676 |
| doch_os_1 | -.0274148 | .0016915 | -16.21 | 0.000 | -.0307303 | -.0240994 |
| doch_os_2 | -.0734132 | .0016275 | -45.11 | 0.000 | -.0766032 | -.0702232 |
| _cons | 115.9215 | 1.354548 | 85.58 | 0.000 | 113.2665 | 118.5764 |

Krzywa łamana - przykład

