Interpretacja parametrów

Stanisław Cichocki

Natalia Nehrebecka

Wykład 5-6

Plan wykładu

- 1. Interpretacja parametrów przy zmiennych objaśniających ciągłych
    - Interpretacja współczynników regresji w modelu liniowym
    - Elastyczność
    - Semielastyczność

- 2. Zastosowanie modelu potęgowego
    - Przekształcenie Boxa-Coxa

Plan wykładu

- 1. Interpretacja parametrów przy zmiennych objaśniających ciągłych
  - Interpretacja współczynników regresji w modelu liniowym
  - Elastyczność
  - Semielastyczność

- 2. Zastosowanie modelu potęgowego
  - Przekształcenie Boxa-Coxa

Interpretacja współczynników regresji w modelu liniowym

- Model regresji liniowej dla i-tej obserwacji:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

- Wartość oczekiwana zmiennej objaśnianej przy danych wartościach zmiennych objaśniających wynosi:

$$E(Y_i) = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} \cdots + \quad \beta_K X_{Ki}$$

- Pochodną cząstkową warunkowej wartości oczekiwanej po x_{Ki}

$$\frac{\partial E(Y_i)}{\partial X_{ki}} = \beta_k$$

β_k mierzy oczekiwaną zmianę $\mathbf{Y}_i$ jako efekt zmiany X_{Ki} o jedną jednostkę, **gdy wartości innych zmiennych objaśniających modelu pozostają niezmienione** (ceteris paribus).

Przykład

$$Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

$$\hat{Y}_i = b_1 + b_2 X_{2i} + \ldots + b_K X_{Ki}$$

β_2 – współczynnik

INTERPRETACJA: jeżeli wartość zmiennej niezależnej X_{2i} wzrośnie o jednostkę, to wartość zmiennej zależnej y :

- wzrośnie (jeżeli $b_2>0$) o $|b_2|$ jednostek lub
- spadnie (jeżeli $b_2<0$) o $|b_2|$ jednostek

ceteris paribus.

β_1 – wyraz wolny

Uwaga ! Wyrazu wolnego nie interpretujemy.

Przykład

$$placa_i = \beta_1 + \beta_2 nauka_i + \beta_3 wiek_i + \varepsilon_i$$

$$\widehat{placa}_i = -993{,}26 + 73{,}59 \cdot nauka_i + 35{,}09 \cdot wiek_i$$

- Zmienna ***nauka**$_i$* - lata nauki i-tej osoby

- Interpretacja:

  Miesięczne wynagrodzenie wzrasta przeciętnie o 73,59 zł przy wzroście liczby lat nauki o jeden rok, przy założeniu pozostałych charakterystyk na niezmienionym poziomie.

Elastyczność

- Elastyczności mogą być wyznaczane z modelu - **LOGLINIOWYM**, w którym zarówno zmienna objaśniana jak i zmienne objaśniające są logarytmami zmiennych pierwotnych.

$$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \ldots + \beta_K \ln X_{Ki} + \varepsilon_i$$

$$\ln \mathrm{E}(Y_i) = \beta_1 + \beta_2 \ln X_{2i} + \ldots + \beta_K \ln X_{Ki}$$

$$\frac{\partial \ln E(Y_i)}{\partial \ln X_{ki}} = \beta_k$$

β_k mierzy o ile procent zmieni się zmienna objaśniana, gdy zmienna objaśniająca zmieni się o jeden procent, **gdy wartości innych zmiennych objaśniających modelu pozostają niezmienione** (ceteris paribus).

Przykład

$$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \ldots + \beta_K \ln X_{Ki} + \varepsilon_i$$

$$\ln \hat{Y}_i = b_1 + b_2 \ln X_{2i} + \ldots + b_K \ln X_{Ki}$$

β_2 – współczynnik

INTERPRETACJA: jeżeli wartość zmiennej niezależnej X_{2i} wzrośnie o 1% ,
to wartość zmiennej zależnej y :

- **wzrośnie (jeżeli $b_2>0$) o $|b_2|$ % lub**
- **spadnie (jeżeli $b_2<0$) o $|b_2|$ %.**

ceteris paribus.

β_1 **– wyraz wolny**

Uwaga ! Wyrazu wolnego nie interpretujemy.

Przykład

$$\ln(wydatki_i) = \beta_1 + \beta_2 \ln(dochód_i) + \varepsilon_i$$

$$\widehat{\ln(wydatki_i)} = 2 + 0{,}25 \cdot \ln(dochód_i)$$

- Interpretacja:

Miesięczne wydatki wzrastają przeciętnie o 0,25% przy wzroście miesięcznego wynagrodzenia o 1%, przy założeniu pozostałych charakterystyk na niezmienionym poziomie.

Semielastyczność

- Semielastyczności mogą być wyznaczane z modelu , w którym zmienna objaśniana jest zlogarytmowana a zmienne objaśniające nie są logarytmami zmiennych pierwotnych.

$$\ln Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

$$\ln \mathrm{E}(Y_i) = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki}$$

$$\frac{\partial \ln E(Y_i)}{\partial X_{ki}} = \beta_k$$

β_k ***100%** mierzy o ile procent zmieni się zmienna objaśniana, gdy zmienna objaśniająca zmieni się o jedną jednostkę, **gdy wartości innych zmiennych objaśniających modelu pozostają niezmienione** (ceteris paribus).

Przykład

$$\ln Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

$$\ln \hat{Y}_i = b_1 + b_2 X_{2i} + \ldots + b_K X_{Ki}$$

β_2 – współczynnik

INTERPRETACJA: jeżeli wartość zmiennej niezależnej X_{2i} wzrośnie o 1 jednostkę, to wartość zmiennej zależnej y :

- wzrośnie (jeżeli $b_2>0$) o $|b_2|$ *100% lub
- spadnie (jeżeli $b_2<0$) o $|b_2|$ *100%.

ceteris paribus.

β_1 – wyraz wolny

Uwaga ! Wyrazu wolnego nie interpretujemy.

Przykład

$$\ln(placa_i) = \beta_1 + \beta_2 wiek_i + \varepsilon_i$$

$$\ln(\widehat{placa_i}) = 2{,}34 + 0{,}04 \cdot wiek_i$$

- Interpretacja:

  Płaca wzrasta przeciętnie o 4% przy wzroście wieku o 1 rok, przy założeniu pozostałych charakterystyk na niezmienionym poziomie.

Zadanie

$$\widehat{\ln(wydatki_i)} = 3{,}6 + 0{,}35 \cdot \ln(dochód)_i + 0{,}11 \cdot dzieci_i$$

- Interpretacja:

  Elastyczność: wzrost dochodu o 1% powoduje wzrost wydatków o 0,35% przy założeniu pozostałych charakterystyk na niezmienionym poziomie.

  Semielastyczność: wzrost liczby dzieci o 1 powoduje wzrost wydatków o 11%=0,11*100% przy założeniu pozostałych charakterystyk na niezmienionym poziomie.

Plan wykładu

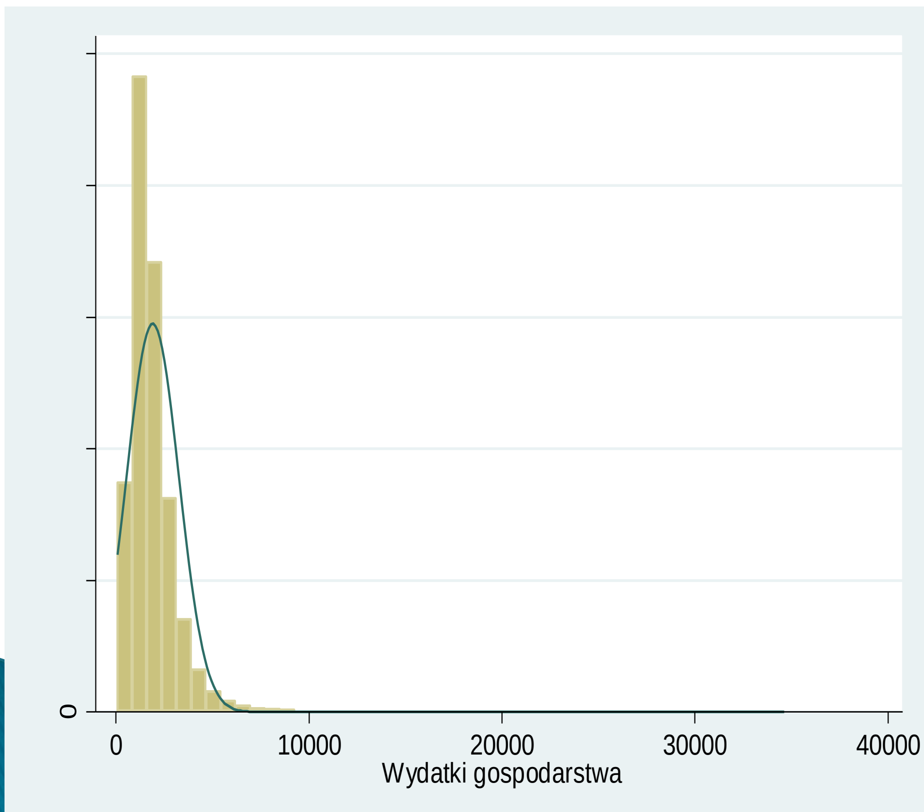
- 1. Interpretacja parametrów przy zmiennych objaśniających ciągłych
  - Interpretacja współczynników regresji w modelu liniowym
  - Elastyczność
  - Semielastyczność

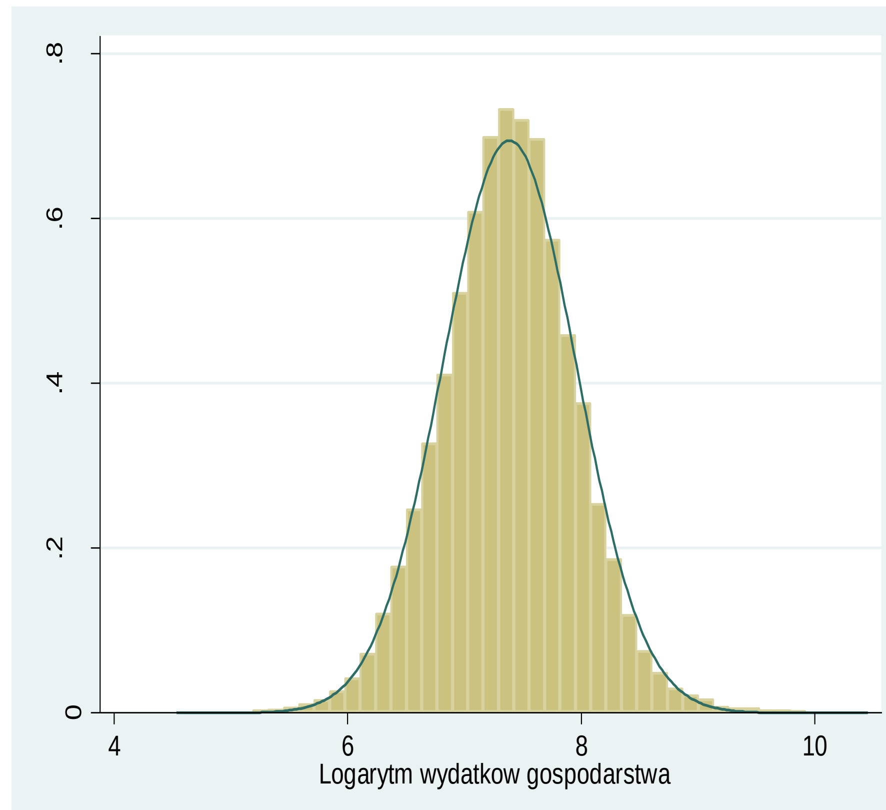
- 2. Zastosowanie modelu potęgowego
  - Przekształcenie Boxa-Coxa

Zastosowanie modelu potęgowego - przykład

Modelujemy wydatki gospodarstw domowych za pomocą dochodu tych gospodarstw.

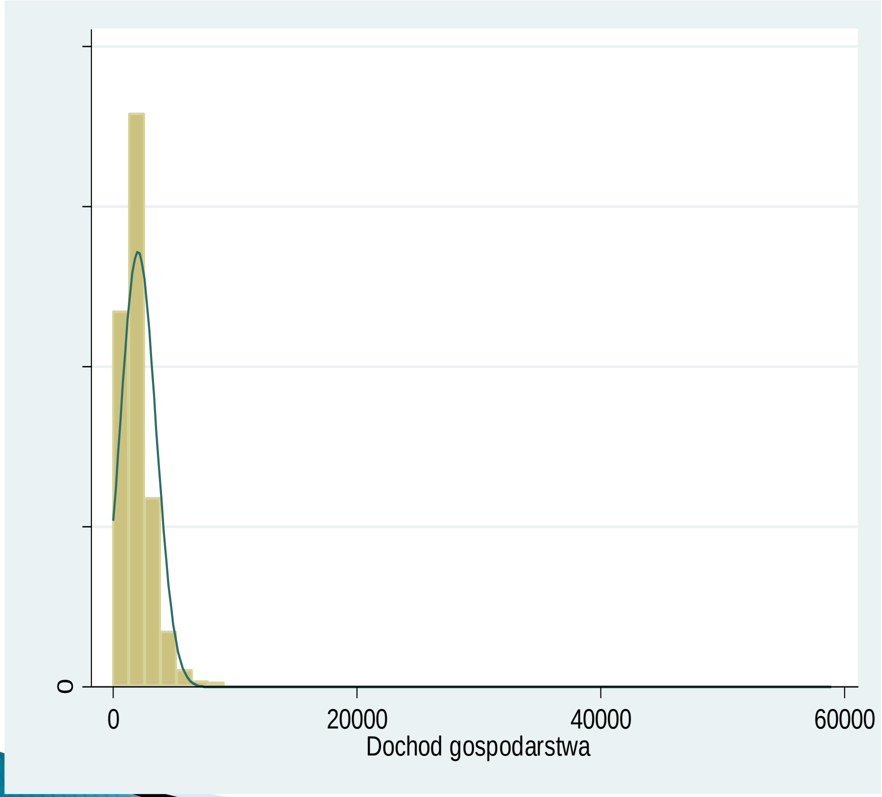
Histogram wydatków /logarytmu wydatków gospodarstw domowych:

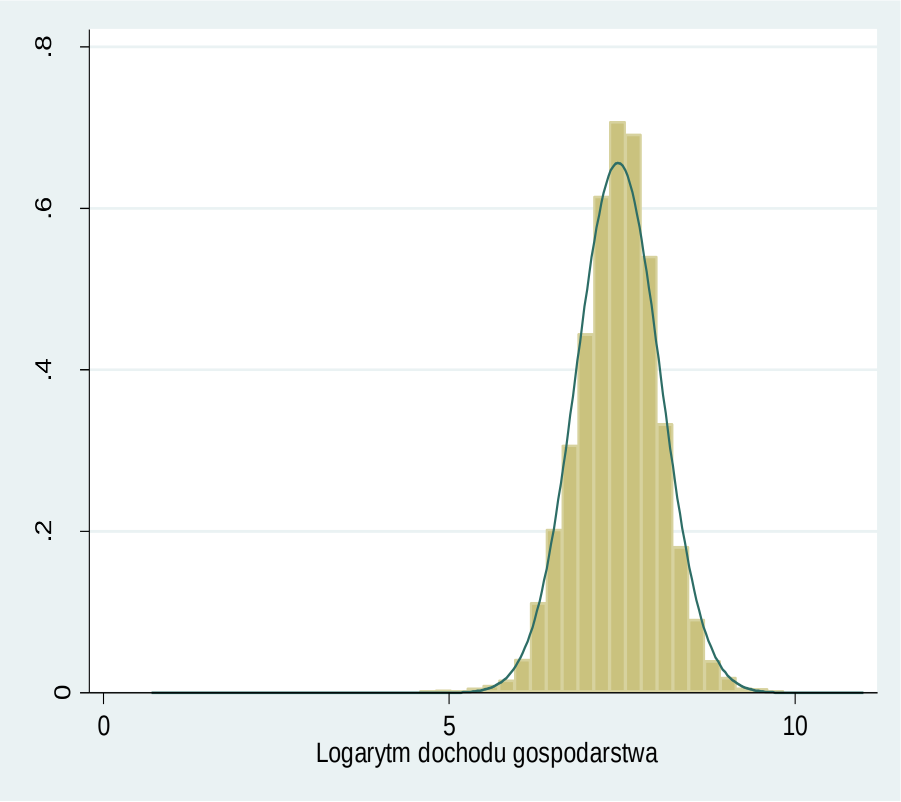




Zastosowanie modelu potęgowego - przykład

Histogram dochodów/logarytmu dochodów gospodarstw:





Zastosowanie modelu potęgowego - przykład

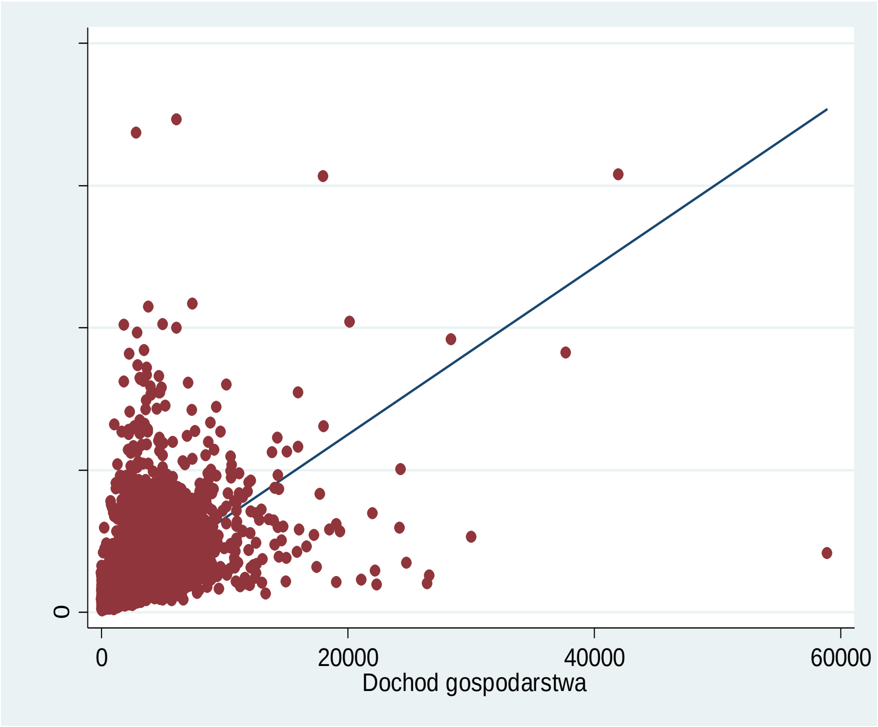
Wyniki regresji:

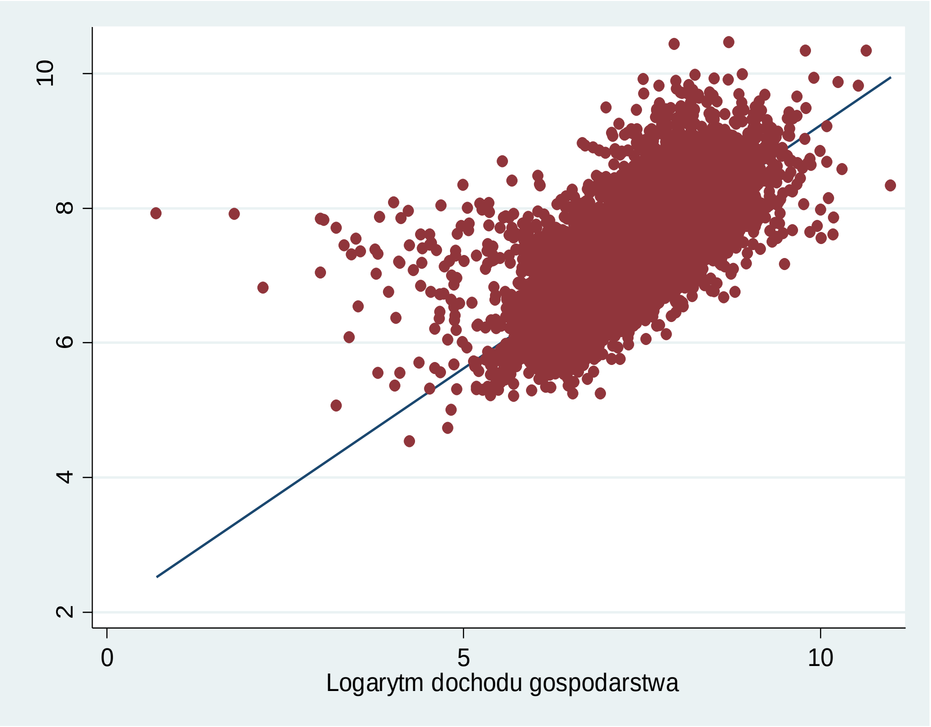
$\ln(Wydatki) = 2{,}02+0{,}72*\ln(Dochod)$ $R^2=0{,}58$

$Wydatki=712{,}81+0{,}58*Dochod$ $R^2=0{,}41$

Zastosowanie modelu potęgowego - przykład

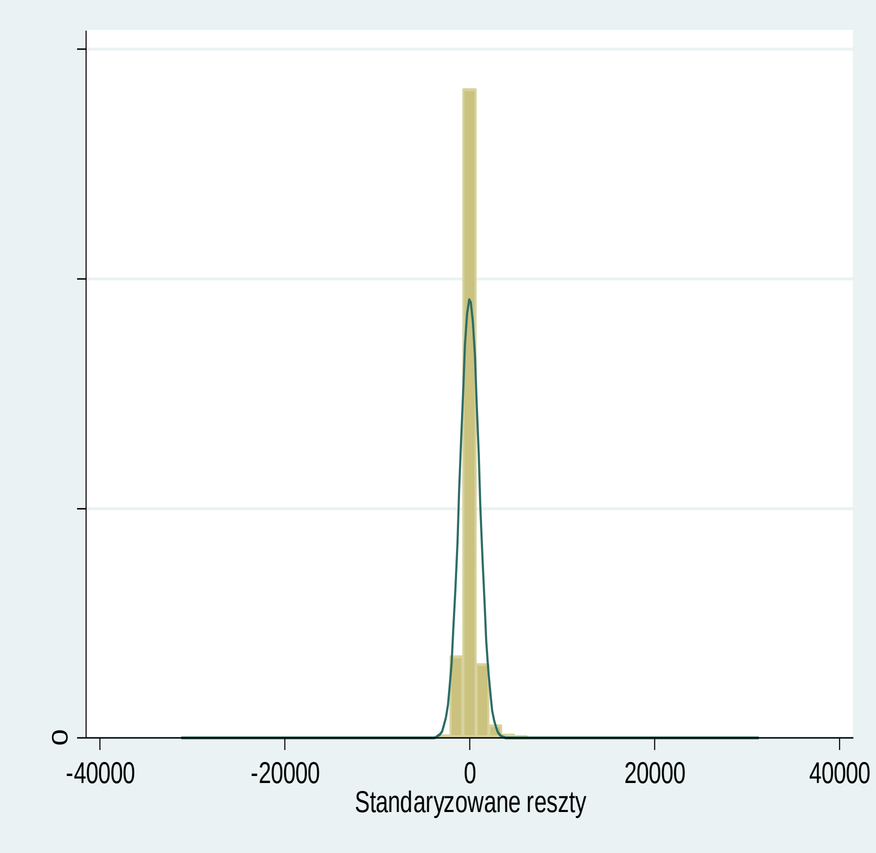
Regresja na poziomach i logarytmach:

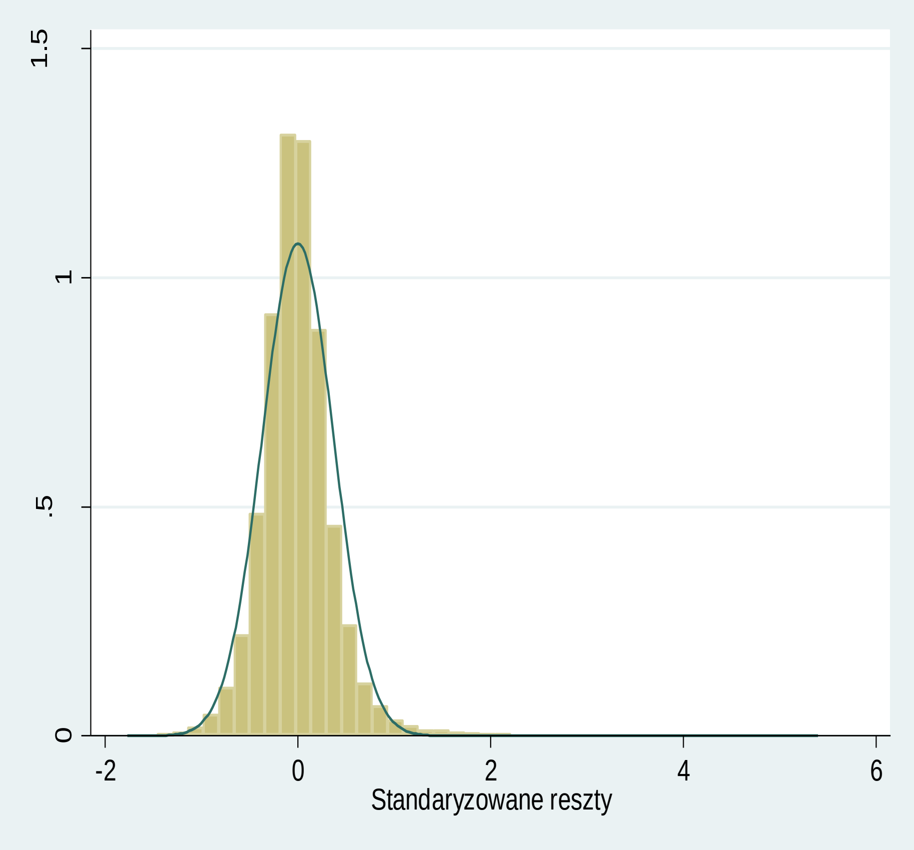




Zastosowanie modelu potęgowego - przykład

Reszty z regresji:





Przekształcenie Boxa-Coxa

- Pozwala na przeprowadzenie sformalizowanej procedury wyboru między modelem liniowym i potęgowym

- Postać przekształcenia:

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

Pytanie: Co otrzymujemy dla $\lambda = 1, \lambda = -1$, $\lambda = 0$?

Przekształcenie Boxa-Coxa

- Stosując to przekształcenie do zmiennej zależnej i zmiennych niezależnych:

$$y_i^{(\lambda)} = \beta_1 + x_{2i}^{(\lambda)}\beta_2 + \ldots + x_{Ki}^{(\lambda)}\beta_K + \varepsilon_i$$

- Dla $\lambda = 1$:

$$y_i = \beta^*{}_1 + x_{2i}\beta_2 + \ldots + x_{Ki}\beta_K + \varepsilon_i$$

Gdzie: $\beta_1^* = 1 + \beta_1 - \beta_2 - \ldots\beta_K$

Przekształcenie Boxa-Coxa

- Dla $\lambda = -1$:

$$\frac{1}{y_i} = \beta^*_1 + \frac{\beta_2}{x_{2i}} + \ldots + \frac{\beta_K}{x_{Ki}} + \varepsilon_i$$

Gdzie: $\beta_1^* = 1 - \beta_1 - \beta_2 - \ldots \beta_K$

Dla $\lambda = 0$:

$$\ln y_i = \beta_1 + \ln x_{2i}\beta_2 + \ldots + \ln x_{Ki}\beta_K + \varepsilon_i$$

Dziękuję za uwagę