

Zadanie 1

OPIS ZMIENNYCH:

1) **klm** – klasa miejscowości

4 = >100 tys.

3 = 20 tys. – 100 tys.

2 = miasto do 20 tys.

1 = wieś

2) **pl** – płeć respondenta

0 = kobieta

1 = mężczyzna

4) **wyksz** – wykształcenie

1 = podstawowe

2 = średnie

3 = wyższe

5) **dochód** – miesięczne zarobki netto

6) **wiek** – wiek respondenta w latach

gretl: model 1

Plik Edycja Testy Zapisz Wykresy Analiza

Model 1: Estymacja KMNK z wykorzystaniem 1118 obserwacji 1-1118
Zmienna zależna: dochod

Zmienna	Współczynnik	Błąd stand.	Statystyka t	Wartość p
const	300,633	114,768	2,619	0,00893 ***
wiek	6,12471	2,10372	2,911	0,00367 ***
pl	294,443	45,1271	6,525	<0,00001 ***
sre	213,453	69,3871	3,076	0,00215 ***
wyzsze	1047,81	84,6752	12,374	<0,00001 ***
duze_m	491,413	54,7537	8,975	<0,00001 ***
srednie_m	-49,5269	70,1756	-0,706	0,48049
male_m	98,0211	69,2835	1,415	0,15741

Srednia arytmetyczna zmiennej zależnej = 1224,23
Odchylenie standardowe zmiennej zależnej = 864,584
Suma kwadratów reszt = 6,07111e+008
Błąd standardowy reszt = 739,558
Wsp. determinacji R-kwadrat = 0,272889
Skorygowany wsp. R-kwadrat = 0,268303
Statystyka F (7, 1110) = 59,5127 (wartość p < 0,00001)
Logarytm wiarygodności = -8967,93
Kryterium informacyjne Akaika (AIC) = 17951,9
Kryterium bayesowskie Schwarz (BIC) = 17992
Kryterium infor.Hannana-Quinna (HQC) = 17967

Wyłączając stałą, największa wartość p jest dla zmiennej 9 (srednie_m)

Zamknij

a) Przetestować istotność zmiennych.

b) Przetestować hipotezę o łącznej istotności równania regresji.

c) Dokonać interpretacji współczynników przy zmiennych objaśniających.

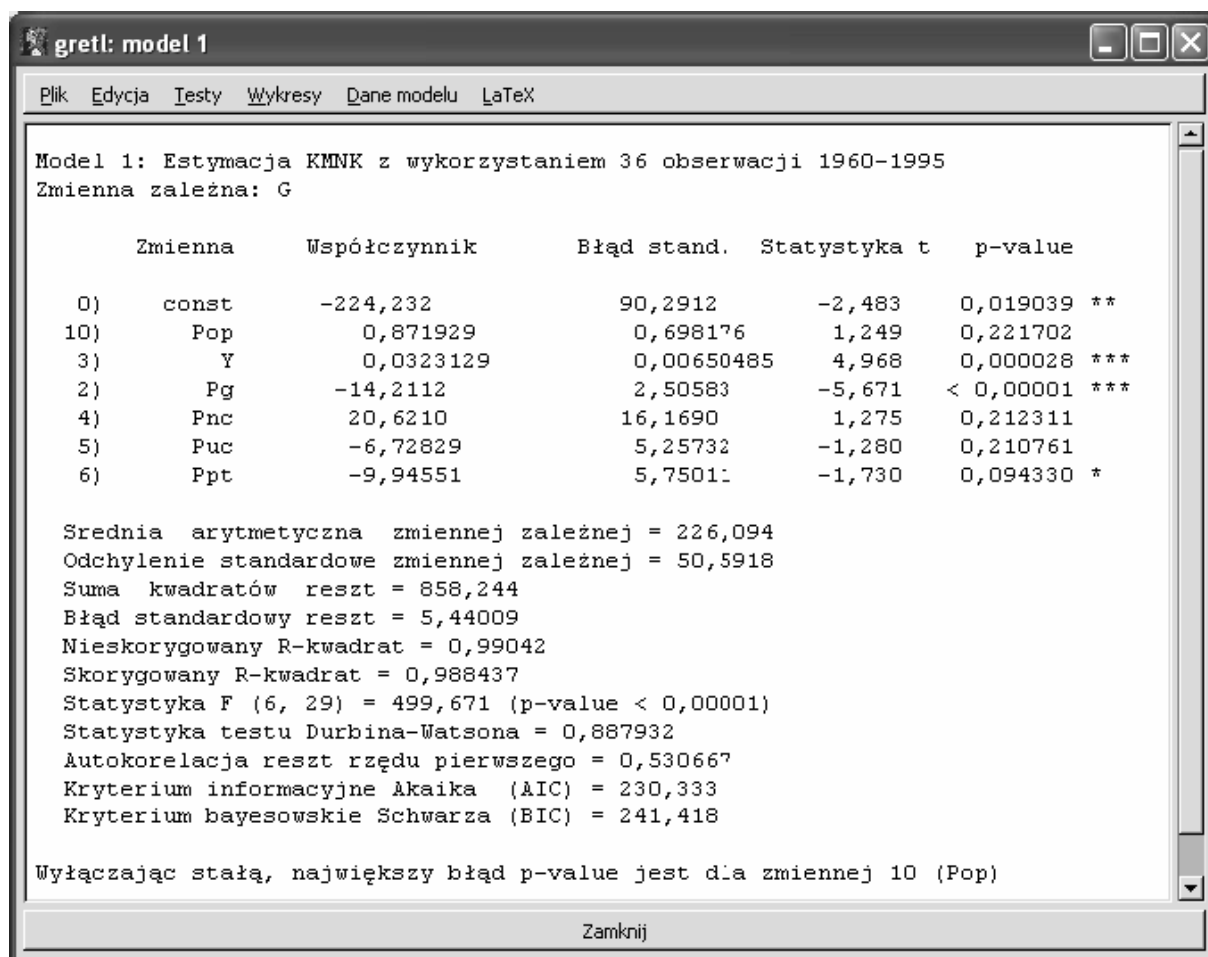
Zadanie 2

G – całkowita konsumpcja benzyny w USA (w mld galonów)

Pop – populacja (w mln), Y – dochód per capita (w 100 \$), Pg – cena benzyny (w \$),

Pnc – średnia cena nowych samochodów (w 1000\$), Puc – średnia cena używanych samochodów (w 1000\$),

Ppt – ceny transportu publicznego (w 10\$)



Model 1: Estymacja KMNK z wykorzystaniem 36 obserwacji 1960-1995
Zmienna zależna: G

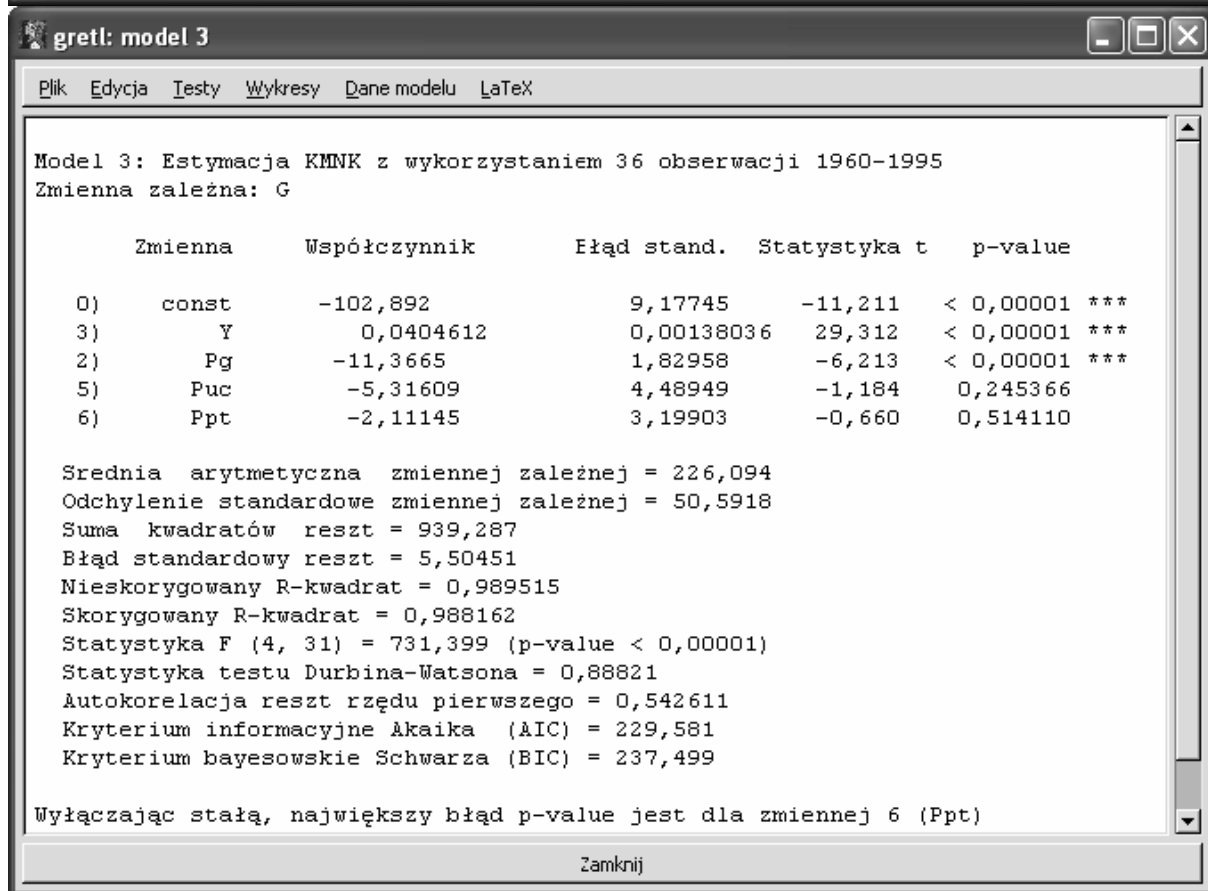
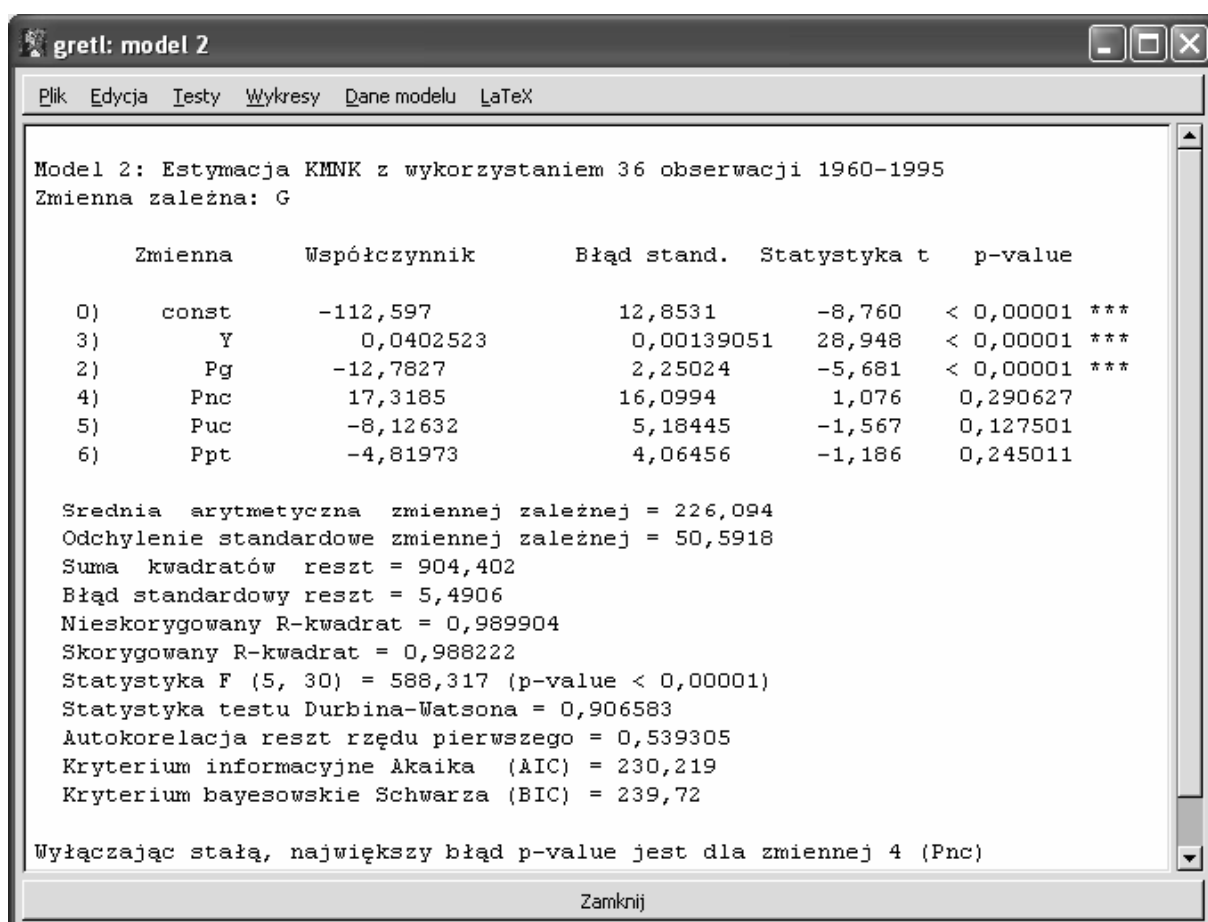
	Zmienna	Współczynnik	Błąd stand.	Statystyka t	p-value
0)	const	-224,232	90,2912	-2,483	0,019039 **
10)	Pop	0,871929	0,698176	1,249	0,221702
3)	Y	0,0323129	0,00650485	4,968	0,000028 ***
2)	Pg	-14,2112	2,50583	-5,671	< 0,00001 ***
4)	Pnc	20,6210	16,1690	1,275	0,212311
5)	Puc	-6,72829	5,25732	-1,280	0,210761
6)	Ppt	-9,94551	5,75011	-1,730	0,094330 *

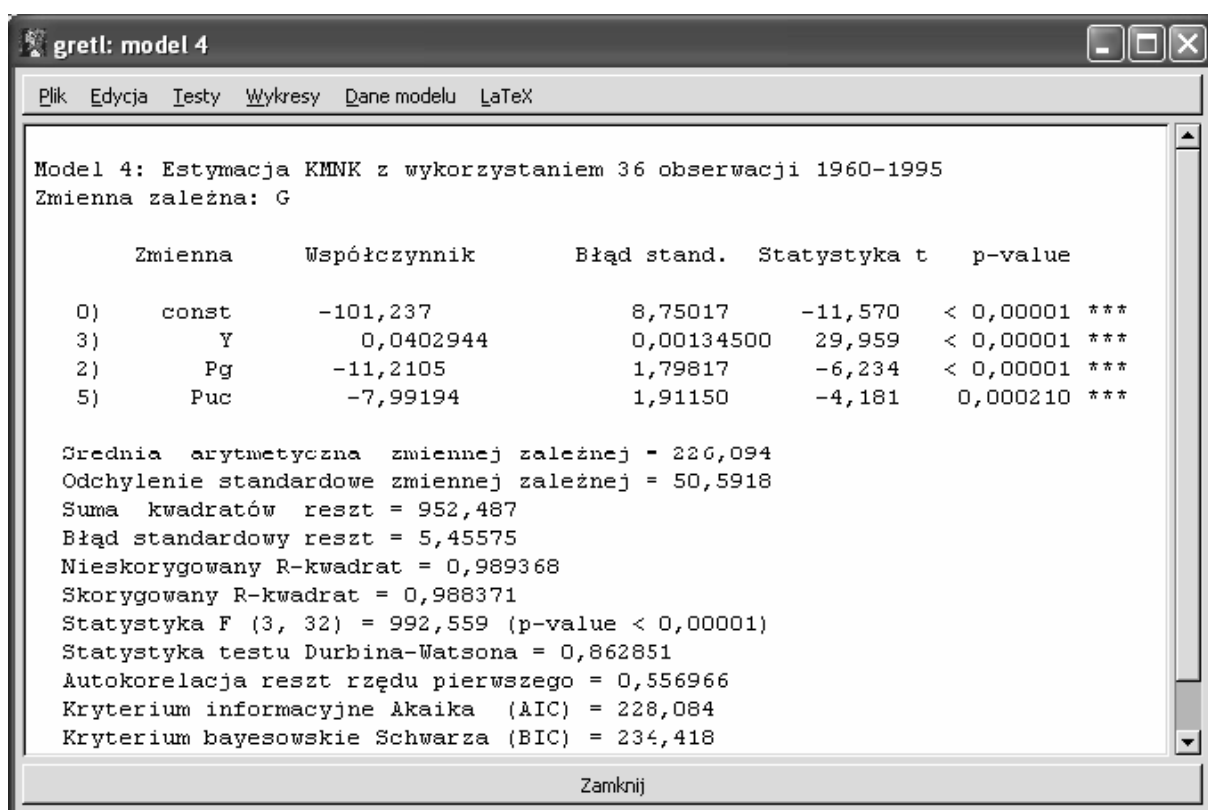
Srednia arytmetyczna zmiennej zależnej = 226,094
Odchylenie standardowe zmiennej zależnej = 50,5918
Suma kwadratów reszt = 858,244
Błąd standardowy reszt = 5,44009
Nieskorygowany R-kwadrat = 0,99042
Skorygowany R-kwadrat = 0,988437
Statystyka F (6, 29) = 499,671 (p-value < 0,00001)
Statystyka testu Durbina-Watsona = 0,887932
Autokorelacja reszt rzędu pierwszego = 0,530667
Kryterium informacyjne Akaika (AIC) = 230,333
Kryterium bayesowskie Schwarza (BIC) = 241,418

Wyłączając stałą, największy błąd p-value jest dla zmiennej 10 (Pop)

Zamknij

- Ocen istotność zmiennych objaśniających w modelu.
- Jaka jest interpretacja oszacowań parametrów modelu? Zinterpretuj je.
- Czy regresja jest istotna na poziomie istotności 0,05?





Zadanie 3

Dla 35 obserwacji oszacowano następujący model: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$ i otrzymano:

$$\hat{y}_i = \underset{(20,95)}{-56,73} + \underset{(23,11)}{34,22}x_{2i} - \underset{(0,0024)}{0,03}x_{3i} - \underset{(6111)}{11089}x_{4i}$$

- Które zmienne są w nim istotne dla $\alpha = 0,05$?
- Jakie są przedziały ufności dla estymatorów?

Zadanie 4

Na próbie liczącej 25 obserwacji metodą najmniejszych kwadratów oszacowano model i otrzymano następujące wyniki:

$$\hat{y}_i = 1,5 - 2,5x_{2i} + 3x_{3i}$$

(0,5) (1,5) (0,75)

- Które zmienne są istotne?
- Wyznacz przedziały ufności.

Zadanie 5

Na podstawie danych:

y	x2	x3
10	0	0
11	4	2
12	1	1
10	0	1
8	0	1

oszacowano model: $\hat{y}_i = 10,45 + 0,75x_{2i} - x_{3i}$

Oblicz i zinterpretuj R^2 oraz oceń czy równanie regresji jest istotne (czy zmienne objaśniające są łącznie istotne) dla $\alpha = 0,05$.

Zadanie 6

Dysponujemy następującymi danymi: $n = 20$, $\hat{y}_i = 2 + 0,5x_{2i} + 0,05x_{3i}$, $e'e = 60$, $\sum (y_i - \bar{y})^2 = 100$.
Przetestuj hipotezę o łącznej istotności równania regresji.