

Zadanie 1

Opis zmiennych:

nocars – ilość sprzedanych nowych samochodów w tysiącach sztuk

pop – ilość ludności w milionach

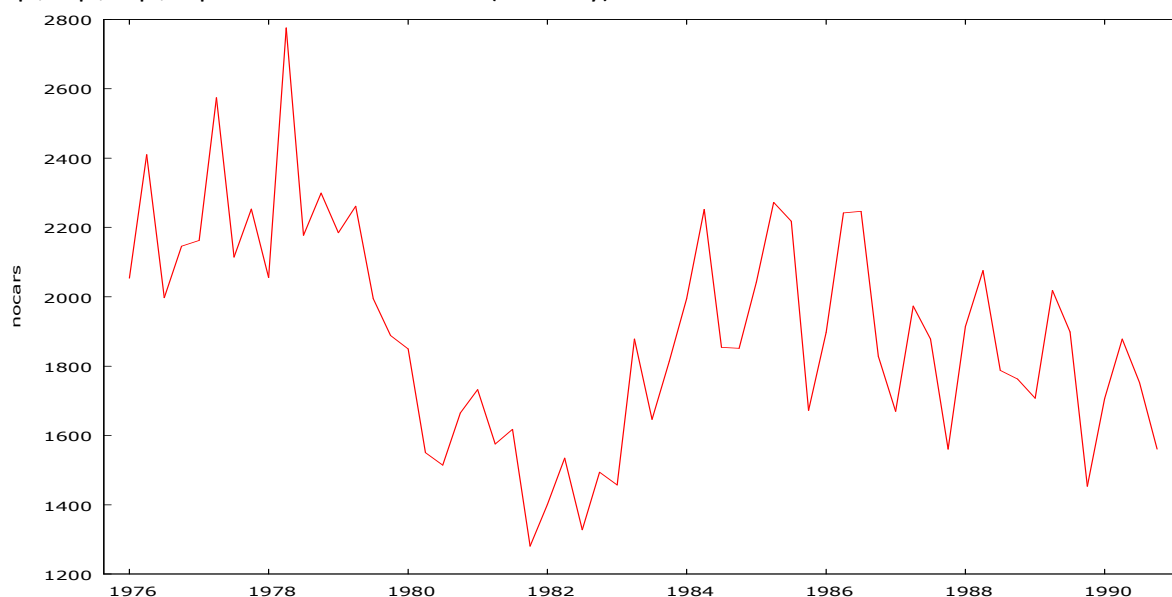
Y – dochód rozporządzalny per capita w tysiącach dolarów

price – indeks cen nowych samochodów

primert – stopa procentowa (%)

unemp – stopa bezrobocia (%)

dq1, dq2, dq3, dq4 – zmienne sezonowe (kwartały)



1. Chcemy oszacować regresję zmiennej nocars na pozostałe zmienne znajdujące się w bazie. Jakie znaki będą miały oszacowane parametry?
2. Następnie oszacowano tę regresję. Czy przypuszczenia się potwierdziły?

gretl: model 2

Plik Edycja Testy Zapisz Wykresy Analiza LaTeX

Model 2: Estymacja KMNK z wykorzystaniem 60 obserwacji 1976:1-1990:4
Zmienna zależna: nocars

	współczynnik	błąd standardowy	t-Student	wartość p	
const	8552,44	3494,61	2,447	0,0177	**
pop	-60,4498	23,4674	-2,576	0,0128	**
Y	757,040	136,762	5,535	9,82E-07	***
price	-5,50342	9,43847	-0,5831	0,5623	
dq2	247,554	74,3300	3,330	0,0016	***
dq3	26,1601	74,9925	0,3488	0,7286	
dq4	-39,8940	74,4281	-0,5360	0,5942	

Średnia arytmetyczna zmiennej zależnej = 1894,16
Odchylenie standardowe zmiennej zależnej = 311,346
Suma kwadratów reszt = 2,1825e+006
Błąd standardowy reszt = 202,927
Wsp. determinacji R-kwadrat = 0,61839
Skorygowany wsp. R-kwadrat = 0,57519
Statystyka F (6, 53) = 14,3144 (wartość p < 0,00001)
Statystyka testu Durbina-Watsona = 1,16128
Autokorelacja reszt rzędu pierwszego = 0,418676

2. Które zmienne są istotne?
3. Dokonać interpretacji oszacowanych parametrów.
4. Zaproponować przetestowanie hipotezy o łącznej nieistotności efektów sezonowych.

Przeprowadzono następujące testy:

a) punktu zwrotnego Chow'a.

gretl: wyniki testu Chowa

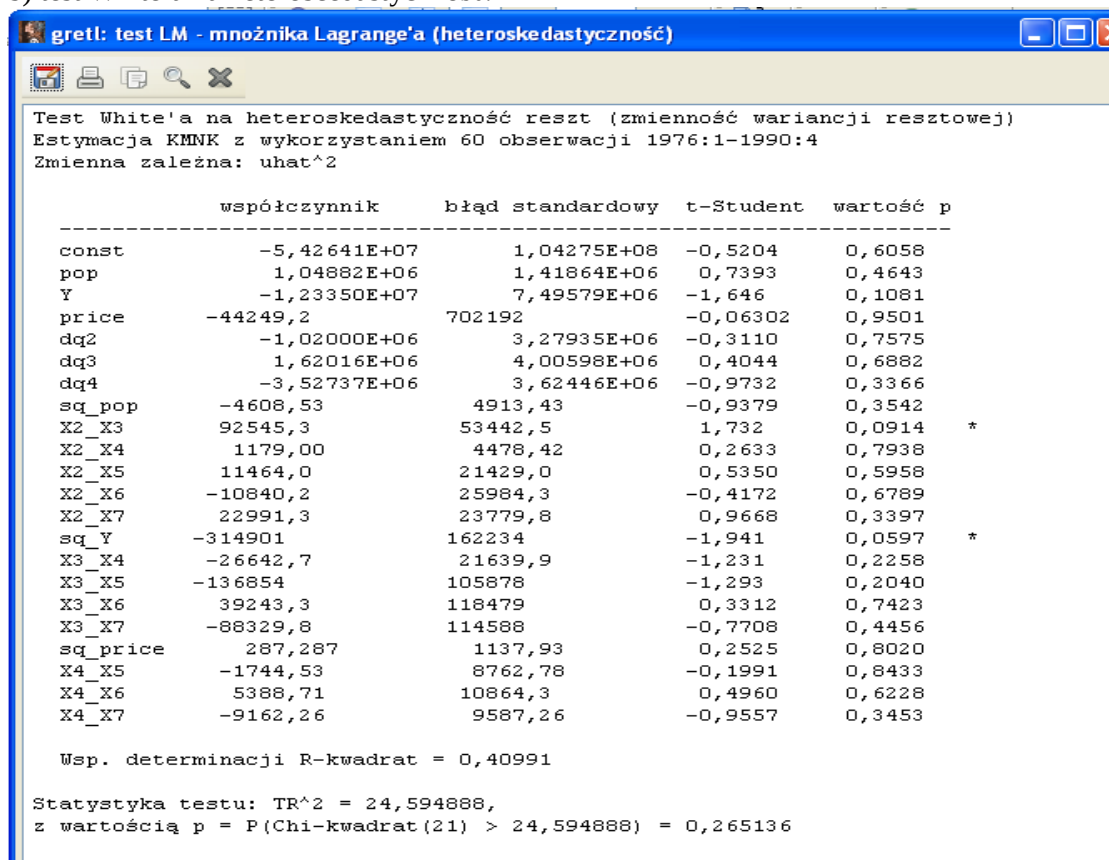
Pomocnicze równanie regresji dla testu Chowa
Estymacja KMNK z wykorzystaniem 60 obserwacji 1976:1-1990:4
Zmienna zależna: nocars

	współczynnik	błąd standardowy	t-Student	wartość p	
const	-23587,8	11241,1	-2,098	0,0414	**
pop	114,607	61,0894	1,876	0,0670	*
Y	716,858	156,855	4,570	3,66E-05	***
price	-88,5232	25,6099	-3,457	0,0012	***
dq2	195,545	69,8986	2,798	0,0075	***
dq3	-119,729	71,8302	-1,667	0,1023	
dq4	18,1140	69,8454	0,2593	0,7965	
splitdum	37447,1	11605,5	3,227	0,0023	***
sd_pop	-192,970	64,2770	-3,002	0,0043	***
sd_Y	-431,400	233,876	-1,845	0,0715	*
sd_price	121,795	32,8787	3,704	0,0006	***
sd_dq2	74,3916	100,012	0,7438	0,4608	
sd_dq3	248,299	103,714	2,394	0,0208	**
sd_dq4	-140,627	99,9511	-1,407	0,1662	

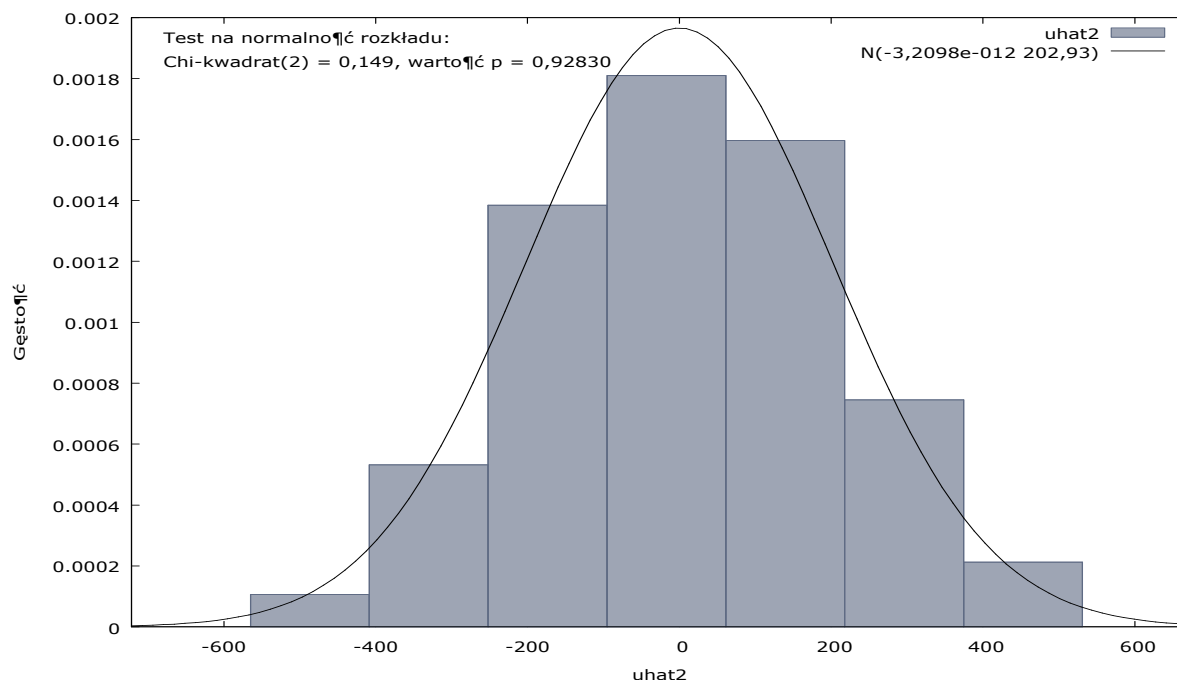
Średnia arytmetyczna zmiennej zależnej = 1894,16
Odchylenie standardowe zmiennej zależnej = 311,346
Suma kwadratów reszt = 834735
Błąd standardowy reszt = 134,709
Wsp. determinacji R-kwadrat = 0,85405
Skorygowany wsp. R-kwadrat = 0,81280
Statystyka F (13, 46) = 20,7055 (wartość p < 0,00001)
Statystyka testu Durbina-Watsona = 2,26166
Autokorelacja reszt rzędu pierwszego = -0,153308
Logarytm wiarygodności = -371,352
Kryterium informacyjne Akaike'a (AIC) = 770,704
Kryterium bayesowskie Schwarza (BIC) = 800,025
Kryterium infor. Hannana-Quinna (HQC) = 782,173

Test Chowa na zmiany strukturalne przy podziale próby w obserwacji 1983:2
F(7, 46) = 10,6103 z wartością p 0,0000

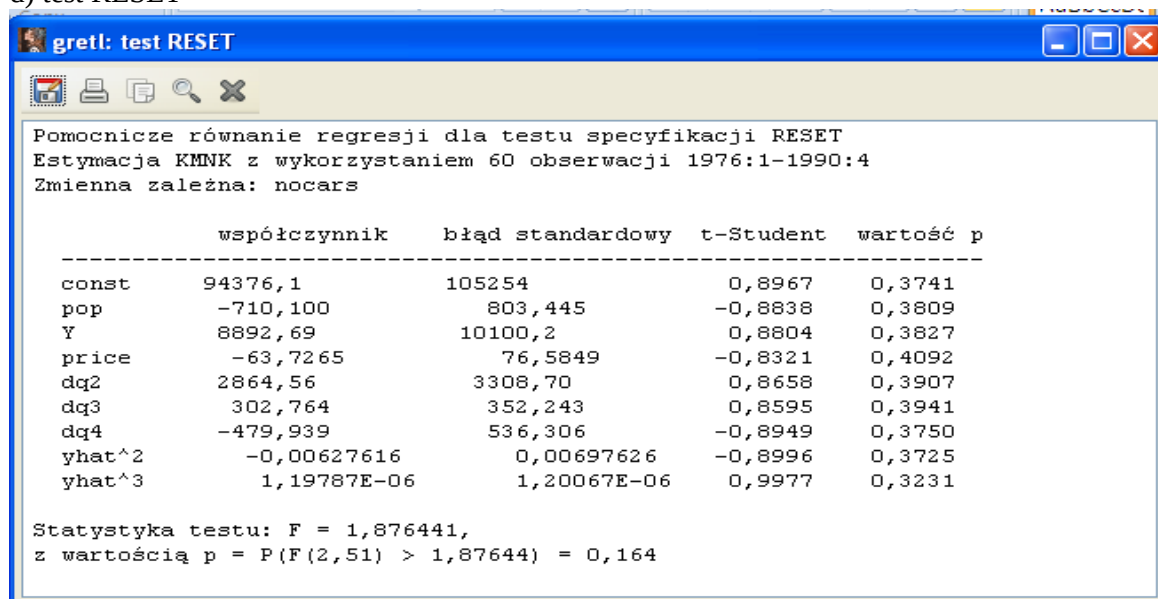
b) test White'a na heteroscedastyczność.



c) test na normalność zaburzeń losowych.



d) test RESET



Które z założeń Klasycznego modelu regresji liniowej są niespełnione? Jakie to ma konsekwencje dla wnioskowania statystycznego? Jakie są metody radzenia sobie z tym problemem?

Zadanie 2

ID #	Nazwa zmiennej	Pełny opis zmiennej
0	const	stała - automatycznie generowana
1	produkt	Produkcja przemysłu chemicznego w Polsce
2	zatrud	Zatrudnienie w przemyśle chemicznym w Polsce
3	inwest	Inwestycje w przemyśle chemicznym w Polsce
Roczne: Pełny zakres 1962 - 1985		

Oszacowanie tego modelu wygląda następująco:

Model 2: Estymacja KMNK z wykorzystaniem 24 obserwacji 1962-1985
Zmienna zależna: produk

	Zmienna	Współczynnik	Błąd stand.	Statystyka t	p-value
0)	const	102,251	5,76519	17,736	< 0,00001 ***
2)	zatrud	0,193058	0,00883051	21,863	< 0,00001 ***
3)	inwest	0,411536	0,103342	3,982	0,000678 ***

Srednia arytmetyczna zmiennej zależnej = 214,458
Odchylenie standardowe zmiennej zależnej = 15,8772
Suma kwadratów reszt = 160,175
Błąd standardowy reszt = 2,76177
Nieskorygowany R-kwadrat = 0,972374
Skorygowany R-kwadrat = 0,969743
Statystyka F (2, 21) = 369,576 (p-value < 0,00001)
Statystyka testu Durbina-Watsona = 1,11728
Autokorelacja reszt rzędu pierwszego = 0,336047

RESET:	Statystyka testu: F = 2,018,	p-value = 0,1135
Jarque-Bera:	Statystyka testu: chi2=2,18,	p-value = 0,3333
Chow:	Statystyka testu: F=3,87,	p-value=0,0087
White:	Statystyka testu: F= 0,351685,	p-value=0,8777
Breusch-Pagan:	Statystyka testu: F= 0,416328,	p-value=0,8153

Na podstawie powyższych wyników odpowiedzieć na następujące pytania. Każdą odpowiedź uzasadnić wielkościami odpowiednich statystyk testowych. Testy przeprowadzić na poziomie istotności $\alpha = 0,05$.

1. Zinterpretować parametry modelu.
2. Sprawdzić istotność oraz łączną istotność uzyskanych wyników.
3. Określić, czy model jest dobrze dopasowany.

4. Sprawdzić, czy parametry w modelu są stabilne.
5. Czy spełnione jest założenie o homoscedastyczności składnika losowego?
6. Czy spełnione jest założenie o normalności składnika losowego?
7. Proszę przeprowadzić test Reset.
8. Jeśli model nie spełnia założeń KMRL określić:
 - (a) które założenie nie jest spełnione?
 - (b) jakie ma to konsekwencje dla wnioskowania statystycznego?
 - (c) jakie są metody radzenia sobie z tym problemem?