Задача 6.15

Сборник задач к начальному курсу эконометрики_Катышев, Магнус и др_2007 -368с.

Подготовил: Гусынин В.И

Данные из газеты «Из рук в руки» за период с декабря 1996 г. по сентябрь 1997 г., касающиеся стоимости однокомнатных квартир в юго-западной части Москвы. Данные приведены в файле room1.xls

Сборник задач к начальному курсу эконометрики_Катышев, Магнус и др_2007 -368с.

Решение

- Итак, перечислим переменные, которые будут включены в нашу модель:

$$n \ - \text{номер по порядку}$$
$$distc \ - \text{удаленность от центра км}$$
$$distm \ - \text{удаленность от метро в мин.}$$
$$totsq \ - \text{общая площадь квартиры в кв. м.}$$
$$kitsq \ - \text{площадь кухни в кв. м.}$$
$$livsq \ - \text{площадь комнаты в кв. м.}$$
$$\mathrm{f}loor - \text{этаж}(1 - \text{первый или последний}, 0 - \text{остальные})$$
$$cat \ - \text{категория дома } (1 - \text{кирпичный}, 0 - \text{нет})$$
$$date \ - \text{дата рекламного объявления}$$
$$price \ - \text{цена квартиры}$$

Решение

Выведем описательные статистики. Для этого выбираем нужные нам переменные, далее Open as Group - > View -> Descriptive Stats ->Common Sample

EViews - [Group: UNTITLED Workfile: ROOM1::Room1\]

File Edit Object View Proc Quick Options Add-ins Window Help

View Proc Object Print Name Freeze Sample Sheet Stats Spec

	TOTSQ	PRICE	LIVSQ	KITSQ	FLOOR	DISTM	DISTC	CAT
Mean	35.39710	39.65217	19.61159	7.931884	0.797101	8.275362	10.55072	0.289855
Median	34.00000	37.00000	19.00000	8.000000	1.000000	7.000000	11.40000	0.000000
Maximum	54.00000	75.00000	34.00000	13.00000	1.000000	15.00000	14.80000	1.000000
Minimum	28.00000	28.00000	14.00000	5.200000	0.000000	1.000000	4.000000	0.000000
Std. Dev.	5.147043	10.04377	3.155972	1.722216	0.405104	4.158210	3.622342	0.457019
Skewness	1.913338	1.370128	2.871282	0.433033	-1.477537	0.315290	-0.308328	0.926371
Kurtosis	6.994038	4.803092	13.87159	2.770473	3.183117	2.037220	1.794097	1.858163
Jarque-Bera	87.96292	30.93540	434.6095	2.307919	25.20225	3.808158	5.274090	13.61728
Probability	0.000000	0.000000	0.000000	0.315386	0.000003	0.148960	0.071572	0.001104
Sum	2442.400	2736.000	1353.200	547.3000	55.00000	571.0000	728.0000	20.00000
Sum Sq. Dev.	1801.459	6859.652	677.2907	201.6899	11.15942	1175.768	892.2525	14.20290
Observations	69	69	69	69	69	69	69	69

Path = c:\users\влад\documents DB = none WF = room1

Заметки к слайду

Решение

Создадим модель зависимости price от totsq , livsq , kitsq, distc, distm. Как мы видим , качество модели, оставляет желать лучшего.

Equation: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: PRICE
Method: Least Squares
Date: 12/08/15 Time: 00:37
Sample: 1 69
Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTSQ	1.152291	0.259872	4.434070	0.0000
LIVSQ	-0.095778	0.355475	-0.269435	0.7885
KITSQ	1.036487	0.567535	1.826295	0.0725
DISTC	-1.020966	0.219282	-4.655941	0.0000
DISTM	-0.336667	0.178652	-1.884485	0.0641
C	6.079447	6.690856	0.908620	0.3670

R-squared	0.677517	Mean dependent var	39.65217
Adjusted R-squared	0.651924	S.D. dependent var	10.04377
S.E. of regression	5.925621	Akaike info criterion	6.479389
Sum squared resid	2212.118	Schwarz criterion	6.673660
Log likelihood	-217.5389	Hannan-Quinn criter.	6.556463
F-statistic	26.47188	Durbin-Watson stat	1.919389
Prob(F-statistic)	0.000000		

Решение

Попробуем убрать из модели, на предыдущем слайде наименее значимые переменные.

Equation: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: PRICE
Method: Least Squares
Date: 12/08/15 Time: 00:44
Sample: 1 69
Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTSQ	1.091617	0.188356	5.795495	0.0000
KITSQ	1.085216	0.551773	1.966781	0.0535
DISTC	-0.929074	0.210249	-4.418922	0.0000
C	2.206677	5.993476	0.368180	0.7139

R-squared	0.659339	Mean dependent var	39.65217
Adjusted R-squared	0.643616	S.D. dependent var	10.04377
S.E. of regression	5.995919	Akaike info criterion	6.476258
Sum squared resid	2336.818	Schwarz criterion	6.605772
Log likelihood	-219.4309	Hannan-Quinn criter.	6.527640
F-statistic	41.93511	Durbin-Watson stat	1.839770
Prob(F-statistic)	0.000000		

Решение

- Попробуем добавить в преыдущую модель переменную cat. Качество модели улучшилось. Но цена отрицательно коррелирована с distc. Это можно исправить.

Equation: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

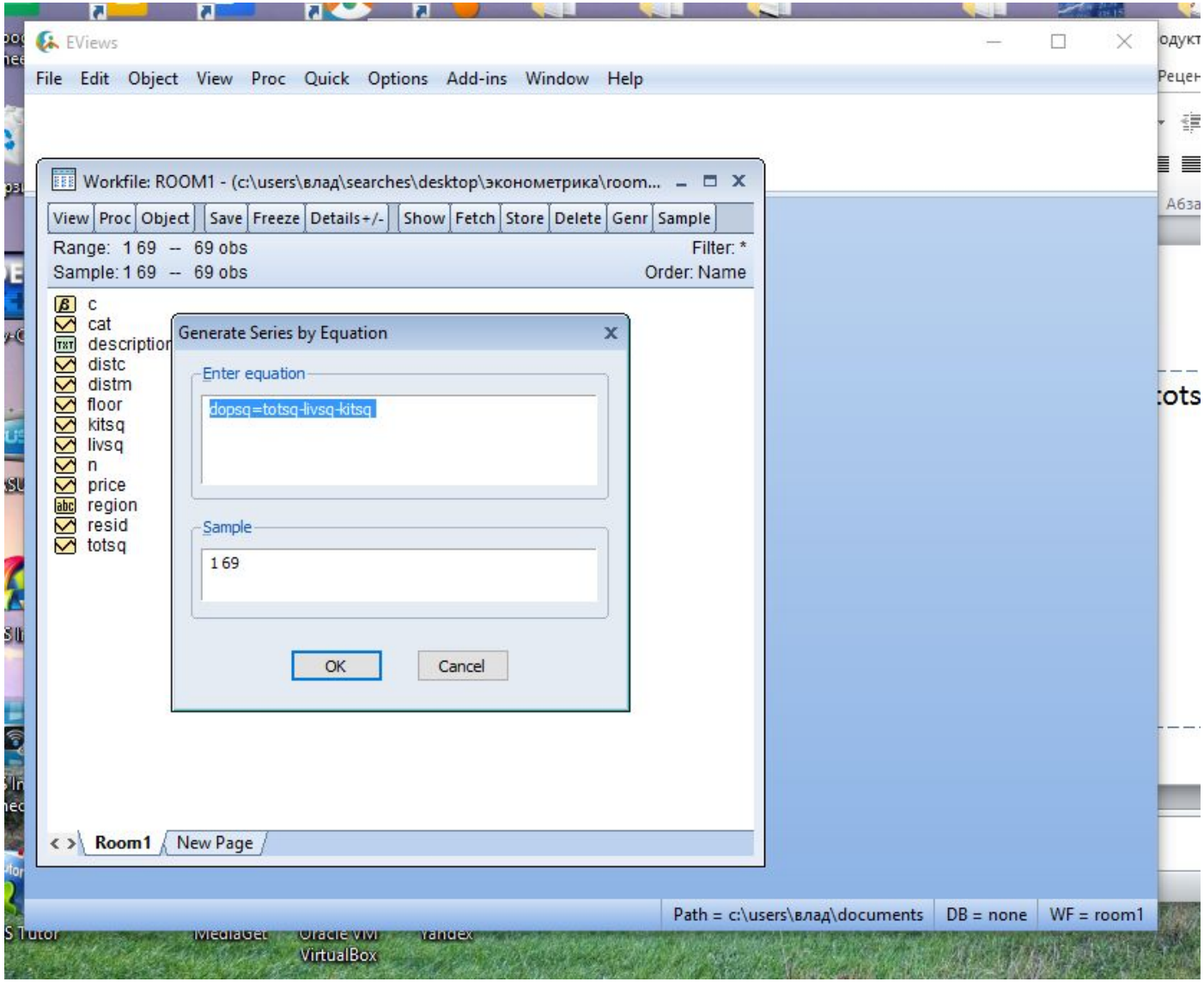
Dependent Variable: PRICE
Method: Least Squares
Date: 12/08/15 Time: 00:47
Sample: 1 69
Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTSQ	0.883679	0.173538	5.092135	0.0000
KITSQ	1.395404	0.493838	2.825632	0.0063
DISTC	-0.255618	0.242269	-1.055097	0.2953
CAT	8.663741	1.994026	4.344848	0.0001
C	-2.509968	5.417700	-0.463290	0.6447

R-squared	0.736934	Mean dependent var	39.65217
Adjusted R-squared	0.720492	S.D. dependent var	10.04377
S.E. of regression	5.309989	Akaike info criterion	6.246761
Sum squared resid	1804.543	Schwarz criterion	6.408652
Log likelihood	-210.5132	Hannan-Quinn criter.	6.310988
F-statistic	44.82119	Durbin-Watson stat	1.849360
Prob(F-statistic)	0.000000		

Решение

Создадим переменную Object -> Generate Series , затем допишем в открывшемся окне dopsq=totsq-livsq-kitsq



Решение

Теперь посмотрим, как коррелируют между собой параметры. Как мы можем увидеть цена положительно коррелирована со всеми переменными кроме расстояния до центра и расстояния до метро. Это вполне очевидно.

\users\влад\searches\desktop\эконометрика\room...

Group: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

Correlation

	PRICE	TOTSQ	LIVSQ	KITSQ	FLOOR	DOPSQ	DISTM	DISTC	CAT
PRICE	1.000000	0.742732	0.562053	0.517725	0.141430	0.404517	-0.012462	-0.440216	0.598964
TOTSQ	0.742732	1.000000	0.711689	0.619779	-0.063762	0.640836	0.080499	-0.202908	0.311698
LIVSQ	0.562053	0.711689	1.000000	0.264705	0.106540	0.017909	0.012080	-0.343631	0.282100
KITSQ	0.517725	0.619779	0.264705	1.000000	0.049458	0.221027	0.010461	0.044973	0.004901
FLOOR	0.141430	-0.063762	0.106540	0.049458	1.000000	-0.261985	-0.280628	0.108335	-0.074826
DOPSQ	0.404517	0.640836	0.017909	0.221027	-0.261985	1.000000	0.125187	-0.013050	0.246601
DISTM	-0.012462	0.080499	0.012080	0.010461	-0.280628	0.125187	1.000000	-0.211535	0.104412
DISTC	-0.440216	-0.202908	-0.343631	0.044973	0.108335	-0.013050	-0.211535	1.000000	-0.677026
CAT	0.598964	0.311698	0.282100	0.004901	-0.074826	0.246601	0.104412	-0.677026	1.000000

Решение

Equation: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: PRICE
Method: Least Squares
Date: 11/20/15 Time: 10:05
Sample: 1 69
Included observations: 69

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LIVSQ	0.871627	0.207280	4.205072	0.0001
KITSQ	2.214960	0.365270	6.063897	0.0000
FLOOR	4.840677	1.505149	3.216079	0.0021
DOPSQ	0.923971	0.225195	4.102975	0.0001
1/DISTC	38.92379	15.45162	2.519076	0.0144
CAT	7.350850	1.774277	4.143011	0.0001
C	-12.58312	4.541122	-2.770927	0.0074

R-squared	0.791983	Mean dependent var	39.65217
Adjusted R-squared	0.771852	S.D. dependent var	10.04377
S.E. of regression	4.797389	Akaike info criterion	6.069947
Sum squared resid	1426.926	Schwarz criterion	6.296596
Log likelihood	-202.4132	Hannan-Quinn criter.	6.159866
F-statistic	39.34202	Durbin-Watson stat	1.780046
Prob(F-statistic)	0.000000		

Построим линейную спецификацию следующим образом. Модель достаточно простая, коэффициенты имеют явную интерпретацию, также она является наиболее подходящей для прогнозирования, также удалось увидеть зав-ть цены квартиры от расстояния до центра .

Решение

Протестируем модель на гетероскедостичность с помощью теста Уайта. Значение статистики равно 48, p-значение равно 0,0037. Гипотеза о гомоскедостичности отвергается. Т. о. цена зависит от следующих величин: livsq, kitsq, 1/distc,dopsq,cat, floor.

Equation: UNTITLED Workfile: ROOM1::Room1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Heteroskedasticity Test: White

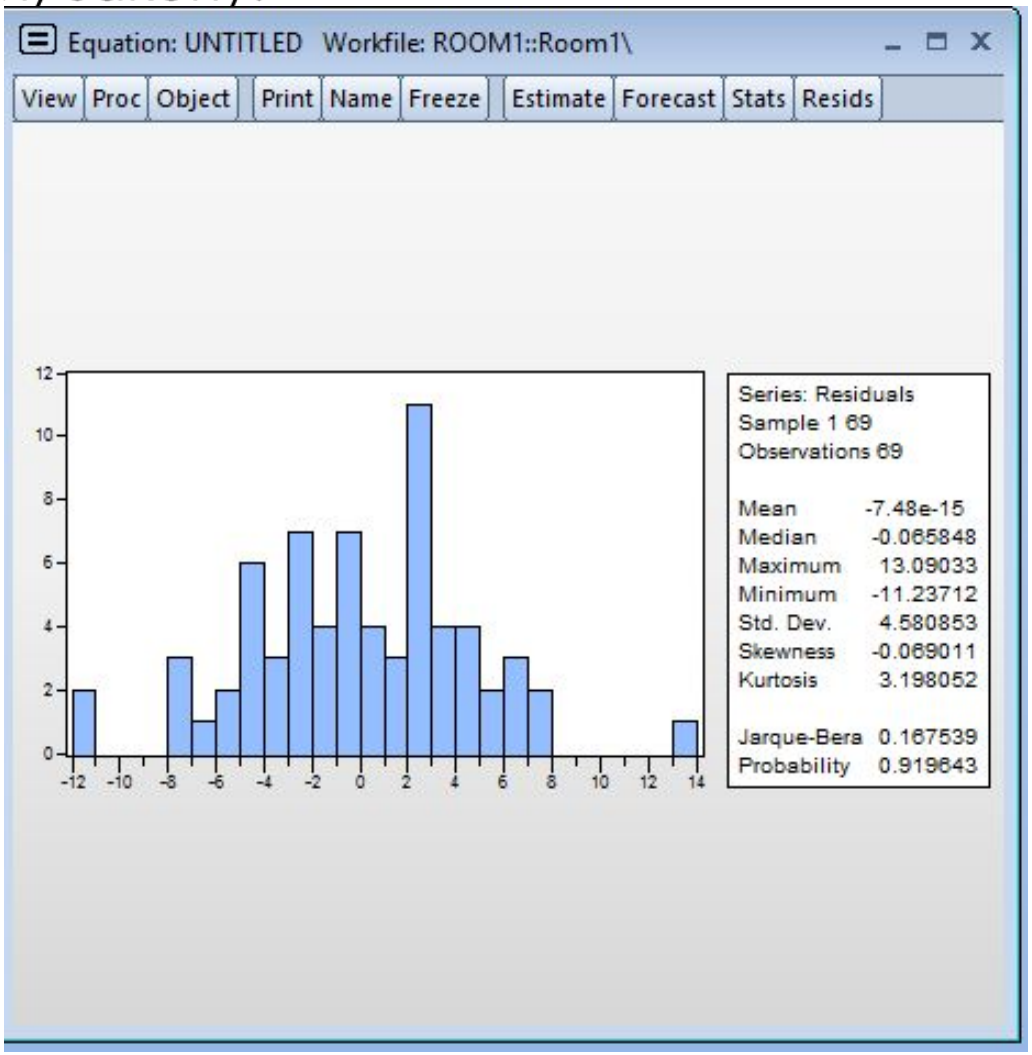
F-statistic	3.931983	Prob. F(25,43)	0.0000
Obs*R-squared	48.00206	Prob. Chi-Square(25)	0.0037
Scaled explained SS	42.59445	Prob. Chi-Square(25)	0.0155

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/20/15 Time: 10:17
Sample: 1 69
Included observations: 69
Collinear test regressors dropped from specification

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	449.8440	231.9325	1.939547	0.0590
LIVSQ^2	0.270368	0.412304	0.655749	0.5155
LIVSQ*KITSQ	0.429635	1.213732	0.353978	0.7251
LIVSQ*FLOOR	2.231887	4.425867	0.504282	0.6166
LIVSQ*DOPSQ	2.664313	1.027909	2.591973	0.0130
LIVSQ*1/DISTC	-64.58902	80.95480	-0.797841	0.4293
LIVSQ*CAT	-5.252790	7.042722	-0.745847	0.4598
LIVSQ	-29.09134	15.04060	-1.934188	0.0597
KITSQ^2	-0.246636	1.352709	-0.182327	0.8562
KITSQ*FLOOR	4.327285	6.173445	0.700951	0.4871
KITSQ*DOPSQ	-0.661617	1.179457	-0.560951	0.5777

Решение

- Протестируем модель на нормальность остатков. $Prob = 0{,}919643 > 0{,}05$, откуда следует, что остатки распределены по нормальному закону.



Решение

- Коэффициент при livsq равен 0,872, что означает, что увеличение площади комнаты на 1 кв.м. при прочих равных приводит к увеличению цены в среднем на 872 $
- Коэффициент 38,924 при 1/distc означает, что при увеличении расстояния до центра на b км при прочих равных приводит к уменьшению цены в среднем на ($38{,}924/distc^2$)р $, учитывая значимость коэффициента можно утверждать, что цена квартиры зависит от расстояния до центра.
- Коэффициент 7,351 при cat означает, что при прочих равных квартира в кирпичном доме стоит в среднем дороже на 7,351тыс. $ дороже, чем в панельном