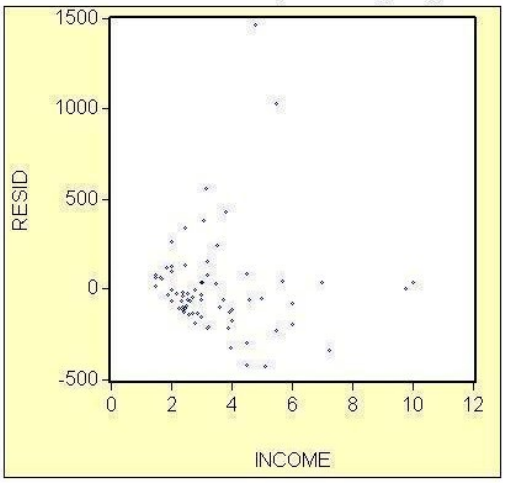
AVGEXP = average monthly credit card expenditures
AGE = age in years
OWNRENT = 1 if individual owns, 0 if rents, home
INCOME = income, divided by 10 000

Dependent Variable: AVGEXP
Dependent Variable: AVGEXP
Method: Least Squares
Date: 04/22/06 Time: 20:13
Sample: 29 100 ← uvažujeme iba dáta s kladnou hodnotou AVGEXP
Included observations: 72

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	-237.1465	199.3517	-1.189589	0.2384
AGE	-3.081814	5.514717	-0.558835	0.5781
OWNRENT	27.94091	82.92232	0.336953	0.7372
INCOME	234.3470	80.36595	2.915999	0.0048
INCOME^2	-14.99684	7.469337	-2.007788	0.0487

R-squared	0.243578	Mean dependent var	262.5321
Adjusted R-squared	0.198418	S.D. dependent var	318.0468
S.E. of regression	284.7508	Akaike info criterion	14.20802
Sum squared resid	5432562.	Schwarz criterion	14.36612
Log likelihood	-506.4888	F-statistic	5.393722
Durbin-Watson stat	0.485690	Prob(F-statistic)	0.000795

Graf INCOME - RESID (rezíduá z regresie)



Whitov odhad kovariančnej matice

Dependent Variable: AVGEXP
Method: Least Squares
Date: 04/22/06 Time: 20:17
Sample: 29 100
Included observations: 72
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	-237.1465	220.7950	-1.074058	0.2866
AGE	-3.081814	3.422641	-0.900420	0.3711
OWNRENT	27.94091	95.56573	0.292374	0.7709
INCOME	234.3470	92.12260	2.543860	0.0133
INCOME^2	-14.99684	7.199027	-2.083177	0.0411

R-squared	0.243578	Mean dependent var	262.5321
Adjusted R-squared	0.198418	S.D. dependent var	318.0468
S.E. of regression	284.7508	Akaike info criterion	14.20802
Sum squared resid	5432562.	Schwarz criterion	14.36612
Log likelihood	-506.4888	F-statistic	5.393722
Durbin-Watson stat	0.485690	Prob(F-statistic)	0.000795

Whitov test heteroskedasticity - verzia 1

White Heteroskedasticity Test:

F-statistic	1.339024	Probability	0.252849
Obs*R-squared	7.920384	Probability	0.243994

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/22/06 Time: 20:19
Sample: 29 100
Included observations: 72

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-488575.1	588227.2	-0.830589	0.4092
AGE	21993.12	34461.58	0.638192	0.5256
AGE^2	-276.0592	498.9109	-0.553324	0.5819
OWNRENT	-85501.14	81196.10	-1.053020	0.2962
INCOME	39116.36	206785.6	0.189164	0.8506
INCOME^2	7785.868	32477.55	0.239731	0.8113
(INCOME^2)^2	-108.0669	160.9298	-0.671516	0.5043

R-squared	0.110005	Mean dependent var	75452.25
Adjusted R-squared	0.027852	S.D. dependent var	279705.5
S.E. of regression	275782.8	Akaike info criterion	27.98478
Sum squared resid	4.94E+12	Schwarz criterion	28.20612
Log likelihood	-1000.452	F-statistic	1.339024
Durbin-Watson stat	0.500140	Prob(F-statistic)	0.252849

Whitov test heteroskedasticity - verzia 2

White Heteroskedasticity Test:

F-statistic	1.221595	Probability	0.290513
Obs*R-squared	14.32895	Probability	0.280197

Goldfeld - Quandtov test: premenná INCOME

Dáta sú zoradené podľa premennej INCOME.

Dependent Variable: AVGEXP
Method: Least Squares
Date: 04/22/06 Time: 20:23
Sample: 29 64
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	153.1301	501.5152	0.305335	0.7622
AGE	-4.137398	2.670104	-1.549527	0.1314
OWNRENT	108.8721	45.24256	2.406408	0.0223
INCOME	16.88575	475.5545	0.035507	0.9719
INCOME^2	3.693360	107.1857	0.034458	0.9727

R-squared	0.183179	Mean dependent var	116.6133
Adjusted R-squared	0.077782	S.D. dependent var	106.8257
S.E. of regression	102.5871	Akaike info criterion	12.22755
Sum squared resid	326247.3	Schwarz criterion	12.44748
Log likelihood	-215.0958	F-statistic	1.737999
Durbin-Watson stat	2.035947	Prob(F-statistic)	0.166809

Dependent Variable: AVGEXP
Method: Least Squares
Date: 04/22/06 Time: 20:24
Sample: 65 100
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	-259.1083	637.7350	-0.406295	0.6873
AGE	-1.940400	11.89225	-0.163165	0.8714
OWNRENT	-52.82816	163.0564	-0.323987	0.7481
INCOME	250.1354	226.8140	1.102822	0.2786
INCOME^2	-16.11412	17.82176	-0.904183	0.3729

R-squared	0.067690	Mean dependent var	408.4508
Adjusted R-squared	-0.052609	S.D. dependent var	387.2786
S.E. of regression	397.3351	Akaike info criterion	14.93568
Sum squared resid	4894130.	Schwarz criterion	15.15562
Log likelihood	-263.8423	F-statistic	0.562681
Durbin-Watson stat	2.247802	Prob(F-statistic)	0.691465

Breusch - Paganov test: premenné INCOME, INCOME^2

Dependent Variable: G
Method: Least Squares
Date: 04/22/06 Time: 20:34
Sample: 29 100
Included observations: 72

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	-4.020459	2.053893	-1.957482	0.0543
INCOME	2.261001	0.953436	2.371424	0.0205
INCOME^2	-0.187619	0.091829	-2.043134	0.0449

R-squared	0.085929	Mean dependent var	1.000000
Adjusted R-squared	0.059434	S.D. dependent var	3.707054
S.E. of regression	3.595204	Akaike info criterion	5.437852
Sum squared resid	891.8588	Schwarz criterion	5.532713
Log likelihood	-192.7627	F-statistic	3.243227
Durbin-Watson stat	2.208335	Prob(F-statistic)	0.045062

PRÍKLAD 2

Snow geeese

Aerial survey methods are regularly used to estimate the number of snow geese in their summer range areas west of Hudson Bay in Canada. To obtain estimates, small aircraft fly over the range and, when a flock of geese is spotted, an experienced person estimates the number of geese in the stock. To investigate the reliability of this method of counting, an experiment was conducted in which an airplane carrying two observers flew over $n = 45$ flocks, and each observer made an independent estimate of the of birds in each flock. Also, a photograph of the flock was taken so that an exact count of the number of birds in the flock could be made. The resulting data are in given in geese.txt.

As a result of this experiment, the practice of using visual counts of flock size to determite population estimates was discontinued in favor of using photographs.

S. Weisberg: Applied Linear Regression. John Wiley & Sons, 1985.
R. D. Cook, J. O. Jacobsen: Analysis of 1977 West Hudson Bay snow goose surveys. Unpublished report, Canadian Wildlife Services, 1977.

Dependent Variable: PHOTO
Method: Least Squares
Date: 05/01/06 Time: 16:52
Sample: 1 45
Included observations: 45
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	26.64957	10.27434	2.593798	0.0129
OBS1	0.882557	0.199470	4.424504	0.0001

R-squared	0.750294	Mean dependent var	89.31111
Adjusted R-squared	0.744487	S.D. dependent var	87.84796
S.E. of regression	44.40568	Akaike info criterion	10.46804
Sum squared resid	84790.18	Schwarz criterion	10.54833
Log likelihood	-233.5309	F-statistic	129.2023
Durbin-Watson stat	2.107238	Prob(F-statistic)	0.000000

Wald Test:
Equation: EQ02

Null Hypothesis:	C(1)=0 C(2)=1		
F-statistic	11.52196	Probability	0.000098
Chi-square	23.04392	Probability	0.000010

DO PROJEKTU:

- Overenie predpokladov modelu: testovanie heteroskedasticity, normality rezíduí.
- Ak zistíte, že niektorý predpoklad regresného modelu nie je splnený: Ak treba, korekcia predchádzajúcich testov signifikancie. Vysvetlite, čo spravíte pre zabezbečenie korektnosti ďalšej analýzy. Inou možnosťou je preformulovať model tak, aby v novom modeli boli predpoklady splnené.
- Samostatne vymyslite, čo by bolo zaujímavé s vašim modelom robiť. Napríklad:
  - Ak zistíte heteroskedasticitu, skúste nájsť premennú, ktorá ju spôsobuje.