实验七 回归分析消除异方差

2014211436 蔡梦璋

数据标准化

. use "C:\Users\orangeade\Desktop\数据分析\费用表.dta", clear

. egen x1=std(var4)

. egen x2=std(var1)

. egen x3=std(var2)

. egen x4=std(var3)

回归分析

. reg x1-x4

Source | SS df MS Number of obs

> = 39

-------------+------------------------------ F( 3, 35)

> = 8.76

Model | 16.2935992 3 5.43119974 Prob > F

> = 0.0002

Residual | 21.7064028 35 .620182937 R-squared

> = 0.4288

-------------+------------------------------ Adj R-squared

> = 0.3798

Total | 38.000002 38 1.00000005 Root MSE

> = .78752

--------------------------------------------------------------------

> ----------

x1 | Coef. Std. Err. t P>|t| [95% Conf.

> Interval]

-------------+------------------------------------------------------

> ----------

x2 | -.1642624 .1572676 -1.04 0.303 -.4835327

> .1550078

x3 | .1175943 .2168496 0.54 0.591 -.3226337

> .5578223

x4 | .5732259 .1915767 2.99 0.005 .1843046

> .9621472

\_cons | 2.59e-09 .1261036 0.00 1.000 -.256004

> .256004

--------------------------------------------------------------------

> ----------

Prob > F = 0.0002

F检验的P值为0.00，模型总体是显著的；

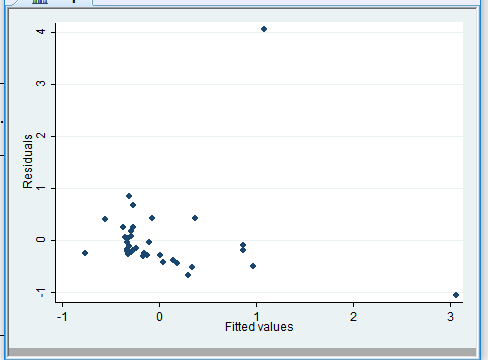
回归方程为：

lirun=-.1642624\*x1+ .1175943\*x2 .5732259\*x3+2.59e-09

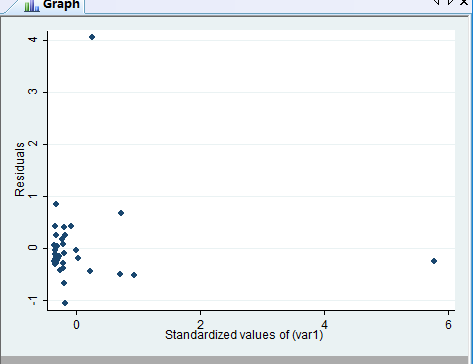
异方差检验

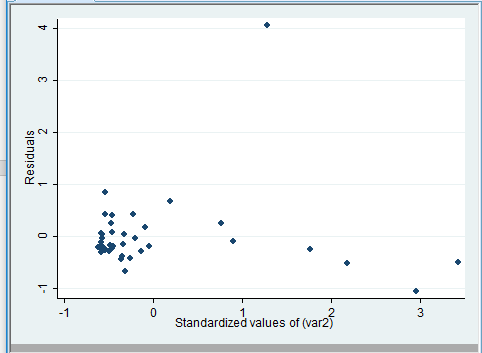
（1）作图

rvfplot



rvpplot x2





怀特检验

. estat imtest,white

White's test for Ho: homoskedasticity

against Ha: unrestricted heteroskedasticity

chi2(9) = 33.07

Prob > chi2 = 0.0001

Cameron & Trivedi's decomposition of IM-test

---------------------------------------------------

Source | chi2 df p

---------------------+-----------------------------

Heteroskedasticity | 33.07 9 0.0001

Skewness | 4.48 3 0.2145

Kurtosis | 1.20 1 0.2743

---------------------+-----------------------------

Total | 38.74 13 0.0002

---------------------------------------------------

.p=0.00002 所以 存在异方差

Bp检验

. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of x1

chi2(1) = 43.62

Prob > chi2 = 0.0000

P=0 存在异方差

改变异方差

（1）加权最小二乘法

. predict u, res

. gen u2=ln(u^2)

. reg u2 x2 x3 x4

Source | SS df MS Number of obs

> = 39

-------------+------------------------------ F( 3, 35)

> = 3.55

Model | 32.2948486 3 10.7649495 Prob > F

> = 0.0242

Residual | 106.248742 35 3.03567833 R-squared

> = 0.2331

-------------+------------------------------ Adj R-squared

> = 0.1674

Total | 138.54359 38 3.64588396 Root MSE

> = 1.7423

---------------------------------------------------------------------

> ---------

u2 | Coef. Std. Err. t P>|t| [95% Conf.

> Interval]

-------------+-------------------------------------------------------

> ---------

x2 | .0871954 .3479423 0.25 0.804 -.619165

> .7935558

x3 | .2100982 .4797627 0.44 0.664 -.7638719

> 1.184068

x4 | .7442828 .4238484 1.76 0.088 -.1161752

> 1.604741

\_cons | -2.85184 .2789945 -10.22 0.000 -3.418228

> -2.285451

---------------------------------------------------------------------

> ---------

. predict g,xb

.

. gen h=exp(g)

.

. gen invar=1/h

. reg x1 x2 x3 x4 [aw=invar]

(sum of wgt is 8.4999e+02)

Source | SS df MS Number of obs

> = 39

-------------+------------------------------ F( 3, 35)

> = 4.84

Model | 2.65623871 3 .885412905 Prob > F

> = 0.0064

Residual | 6.40866457 35 .183104702 R-squared

> = 0.2930

-------------+------------------------------ Adj R-squared

> = 0.2324

Total | 9.06490328 38 .238550086 Root MSE

> = .42791

---------------------------------------------------------------------

> ---------

x1 | Coef. Std. Err. t P>|t| [95% Conf.

> Interval]

-------------+-------------------------------------------------------

> ---------

x2 | -.2202606 .1406305 -1.57 0.126 -.5057557

> .0652346

x3 | .2918215 .180042 1.62 0.114 -.0736831

> .6573262

x4 | .4661208 .176934 2.63 0.012 .1069257

> .825316

\_cons | -.0086559 .093969 -0.09 0.927 -.1994231

> .1821114

---------------------------------------------------------------------

> ---------

成功消除异方差

X1=-.2202606\*x2+ .2918215\*x3+.4661208\*x4--.0086559

（2）非量化修正

取对数

. gen y1=ln(var4)

(2 missing values generated)

. gen y2=ln(var1)

. gen y3=ln(var2)

. gen y4=ln(var3)

(11 missing values generated)

. reg y1 y2 y3 y4

Source | SS df MS Number of obs

> = 27

-------------+------------------------------ F( 3, 23)

> = 10.37

Model | 39.0726629 3 13.024221 Prob > F

> = 0.0002

Residual | 28.8739427 23 1.25538881 R-squared

> = 0.5750

-------------+------------------------------ Adj R-squared

> = 0.5196

Total | 67.9466056 26 2.61333098 Root MSE

> = 1.1204

---------------------------------------------------------------------

> ---------

y1 | Coef. Std. Err. t P>|t| [95% Conf.

> Interval]

-------------+-------------------------------------------------------

> ---------

y2 | -.0149636 .175612 -0.09 0.933 -.3782448

> .3483175

y3 | -.0149636 .3466301 1.28 0.215 -.2747459

> 1.159372

y4 | -.0149636 .153125 2.28 0.032 .0328569

> .6663834

\_cons | 2.048428 1.150793 1.78 0.088 -.3321694

> 4.429026

---------------------------------------------------------------------

> ---------

利润= -.0149636\*销售费用+ -.0149636管理费用+ -.0149636财务费用+2.048428

BP检验

. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of y1

chi2(1) = 0.00

Prob > chi2 = 0.615

P值为0.615远远大于0.1，故不存在异方差