实验5

蔡梦璋 2014211436

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

. egen x1=std(var1)

.

. egen x2=std(var2)

.

. egen x3=std(var3)

.

. egen x4=std(var4)

. describe

Contains data from C:\Users\orangeade\Desktop\数据分析\费用表.dta

obs: 39

vars: 8 8 May 2017 21:36

size: 1,248

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

storage display value

variable name type format label variable label

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

var1 float %8.0g 销售费用

var2 float %8.0g 管理费用

var3 float %8.0g 财务费用

var4 float %8.0g 利润总额

x1 float %9.0g Standardized values of (var1)

x2 float %9.0g Standardized values of (var2)

x3 float %9.0g Standardized values of (var3)

x4 float %9.0g Standardized values of (var4)

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

Sorted by:

Note: dataset has changed since last saved

. \*主成分估计\*/

. pca x1-x4

Principal components/correlation Number of obs = 39

Number of comp. = 4

Trace = 4

Rotation: (unrotated = principal) Rho = 1.0000

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

Component | Eigenvalue Difference Proportion Cumulative

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

Comp1 | 2.25456 1.08938 0.5636 0.5636

Comp2 | 1.16518 .779036 0.2913 0.8549

Comp3 | .386141 .192015 0.0965 0.9515

Comp4 | .194126 . 0.0485 1.0000

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

Principal components (eigenvectors)

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

Variable | Comp1 Comp2 Comp3 Comp4 | Unexplained

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

x1 | 0.2370 0.8242 0.3992 0.3244 | 0

x2 | 0.5885 0.2667 -0.3479 -0.6793 | 0

x3 | 0.5930 -0.2182 -0.4271 0.6468 | 0

x4 | 0.4958 -0.4494 0.7329 -0.1223 | 0

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

前2个特征值累计贡献率已达85.49%，说明前2个主成分基本包含了全部指标具有的信息，我们取前2个特征值

. predict f1 f2, score

(2 components skipped)

Scoring coefficients

sum of squares(column-loading) = 1

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

Variable | Comp1 Comp2 Comp3 Comp4

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

x1 | 0.2370 0.8242 0.3992 0.3244

x2 | 0.5885 0.2667 -0.3479 -0.6793

x3 | 0.5930 -0.2182 -0.4271 0.6468

x4 | 0.4958 -0.4494 0.7329 -0.1223

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

. gen f=0.5636\*f1+0.2913\*f2

. reg var4 f

Source | SS df MS Number of obs = 39

-------------+------------------------------ F( 1, 37) = 14.37

Model | 29502081.9 1 29502081.9 Prob > F = 0.0005

Residual | 75948295.2 37 2052656.63 R-squared = 0.2798

-------------+------------------------------ Adj R-squared = 0.2603

Total | 105450377 38 2775009.93 Root MSE = 1432.7

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

var4 | Coef. Std. Err. t P>|t| [95% Conf. Interval]

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

f | 976.0012 257.4436 3.79 0.001 454.3708 1497.632

\_cons | 912.7118 229.4171 3.98 0.000 447.8685 1377.555

Prob > F = 0.0005 所以是符合的

f | 976.0012

所以在其他条件不变情况下，费用每单位的增长会增长976单位的利润。

