第十章 典型相关分析

一、某地区的矿物的相关分析

1、矿区下部矿Pt、Pd与Cu、Ni的数据

| **Obs** | **x1** | **x2** | **y1** | **y2** |
| --- | --- | --- | --- | --- |
| **1** | 0.14 | 0.30 | 0.03 | 0.14 |
| **2** | 0.20 | 0.50 | 0.14 | 0.22 |
| **3** | 0.06 | 0.11 | 0.03 | 0.02 |
| **4** | 0.07 | 0.11 | 0.04 | 0.13 |
| **5** | 0.12 | 0.22 | 0.06 | 0.12 |
| **6** | 0.52 | 0.87 | 0.19 | 0.20 |
| **7** | 0.23 | 0.47 | 0.14 | 0.10 |
| **8** | 1.19 | 0.38 | 0.09 | 0.11 |
| **9** | 0.37 | 0.66 | 0.14 | 0.15 |
| **10** | 0.36 | 0.60 | 0.12 | 0.14 |
| **11** | 0.42 | 0.77 | 0.17 | 0.10 |
| **12** | 0.35 | 0.85 | 0.30 | 0.19 |
| **13** | 0.50 | 0.87 | 0.23 | 0.22 |
| **14** | 0.56 | 1.15 | 0.29 | 0.28 |
| **15** | 0.43 | 0.90 | 0.13 | 0.22 |
| **16** | 0.47 | 0.97 | 0.26 | 0.22 |
| **17** | 0.49 | 0.79 | 0.21 | 0.20 |
| **18** | 0.47 | 0.77 | 0.51 | 0.22 |
| **19** | 0.40 | 0.88 | 0.33 | 0.19 |
| **20** | 0.66 | 1.30 | 0.21 | 0.30 |
| **21** | 0.52 | 1.43 | 0.31 | 0.23 |
| **22** | 0.44 | 0.87 | 0.17 | 0.25 |
| **23** | 0.03 | 0.07 | 0.05 | 0.08 |
| **24** | 0.20 | 0.28 | 0.04 | 0.08 |
| **25** | 0.04 | 0.10 | 0.11 | 0.07 |
| **26** | 0.17 | 0.28 | 0.15 | 0.09 |

2、SAS程序

data chapter10;

input x1 x2 y1 y2;

cards;

0.14 0.30 0.03 0.14

0.20 0.50 0.14 0.22

0.06 0.11 0.03 0.02

0.07 0.11 0.04 0.13

0.12 0.22 0.06 0.12

0.52 0.87 0.19 0.20

0.23 0.47 0.14 0.10

1.19 0.38 0.09 0.11

0.37 0.66 0.14 0.15

0.36 0.60 0.12 0.14

0.42 0.77 0.17 0.10

0.35 0.85 0.30 0.19

0.50 0.87 0.23 0.22

0.56 1.15 0.29 0.28

0.43 0.90 0.13 0.22

0.47 0.97 0.26 0.22

0.49 0.79 0.21 0.20

0.47 0.77 0.51 0.22

0.40 0.88 0.33 0.19

0.66 1.30 0.21 0.30

0.52 1.43 0.31 0.23

0.44 0.87 0.17 0.25

0.03 0.07 0.05 0.08

0.20 0.28 0.04 0.08

0.04 0.10 0.11 0.07

0.17 0.28 0.15 0.09

;

proc print;

run;

proc cancorr all;

var x1 x2;

with y1 y2;

run;

3、结果分析

| **均值和标准偏差** | | |
| --- | --- | --- |
| **变量** | **均值** | **标准差** |
| **x1** | 0.361923 | 0.248435 |
| **x2** | 0.634615 | 0.382614 |
| **y1** | 0.171154 | 0.113853 |
| **y2** | 0.164231 | 0.071564 |

这个表求出了x1x2y1y2的均值标准差。

|  | **典型 相关** | **调整 典型 相关** | **近似 标准 误差** | **典型 相关 平方** | **特征值: Inv(E)\*H = CanRsq/(1-CanRsq)** | | | | **H0 检验: 当前行和之后的所有行的典型相关都是零** | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **特征值** | **差分** | **比例** | **累积** | **似然 比** | **近似 F 值** | **分子自由度** | **分母自由度** | **Pr > F** |
| **1** | 0.881573 | 0.875390 | 0.044566 | 0.777171 | 3.4877 | 3.4875 | 0.9999 | 0.9999 | 0.22278617 | 12.30 | 4 | 44 | <.0001 |
| **2** | 0.013935 | . | 0.199961 | 0.000194 | 0.0002 |  | 0.0001 | 1.0000 | 0.99980583 | 0.00 | 1 | 23 | 0.9473 |

| **Multivariate Statistics and F Approximations** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **S=2 M=-0.5 N=10** | | | | | |
| **统计量** | **值** | **F 值** | **分子自由度** | **分母自由度** | **Pr > F** |
| **Wilks' Lambda** | 0.22278617 | 12.30 | 4 | 44 | <.0001 |
| **Pillai's Trace** | 0.77736474 | 7.31 | 4 | 46 | 0.0001 |
| **Hotelling-Lawley Trace** | 3.48793166 | 18.93 | 4 | 25.396 | <.0001 |
| **Roy's Greatest Root** | 3.48773745 | 40.11 | 2 | 23 | <.0001 |

输入用似然比检验典型相关系数与0是否有显著性差异。

典型相关分析

| **VAR 变量 的原始典型系数** | | |
| --- | --- | --- |
|  | **V1** | **V2** |
| **x1** | -0.117500292 | 5.0220949278 |
| **x2** | 2.6585280081 | -1.889840001 |

| **WITH 变量 的原始典型系数** | | |
| --- | --- | --- |
|  | **W1** | **W2** |
| **y1** | 2.7717034134 | -11.10287639 |
| **y2** | 10.730817444 | 14.707471072 |

分别给出了var变量与with变量的两个典型变量的非标准和标准的组合系数估计值。

V1=-0.117500292x1+2.6585280081

V2=5.0220949278x1+14.707471072x2

W1=2.7717034134y1+10.730817444

W2=-11.10287639y1+14.707471072

典型结构

| **VAR 变量 及其典型变量之间的相关性** | | |
| --- | --- | --- |
|  | **V1** | **V2** |
| **x1** | 0.5794 | 0.8151 |
| **x2** | 0.9997 | 0.0234 |

| **WITH 变量 及其典型变量之间的相关性** | | |
| --- | --- | --- |
|  | **W1** | **W2** |
| **y1** | 0.8078 | -0.5894 |
| **y2** | 0.9702 | 0.2422 |

| **VAR 变量 和 WITH 变量 的典型变量之间的相关性** | | |
| --- | --- | --- |
|  | **W1** | **W2** |
| **x1** | 0.5108 | 0.0114 |
| **x2** | 0.8813 | 0.0003 |

| **WITH 变量 和 VAR 变量 的典型变量之间的相关性** | | |
| --- | --- | --- |
|  | **V1** | **V2** |
| **y1** | 0.7122 | -0.0082 |
| **y2** | 0.8553 | 0.0034 |

二、康复部俱乐部20名成员测试数据的典型相关分析

1、康复俱乐部20名尘成员的测试数据

| **Obs** | **x1** | **x2** | **x3** | **y1** | **y2** | **y3** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | 191 | 36 | 50 | 5 | 162 | 60 |
| **2** | 193 | 38 | 58 | 12 | 101 | 101 |
| **3** | 189 | 35 | 46 | 13 | 155 | 58 |
| **4** | 211 | 38 | 56 | 8 | 101 | 38 |
| **5** | 176 | 31 | 74 | 15 | 200 | 40 |
| **6** | 169 | 34 | 50 | 17 | 120 | 38 |
| **7** | 154 | 34 | 64 | 14 | 215 | 105 |
| **8** | 193 | 36 | 46 | 6 | 70 | 31 |
| **9** | 176 | 37 | 54 | 4 | 60 | 25 |
| **10** | 156 | 33 | 54 | 15 | 255 | 73 |
| **11** | 189 | 37 | 52 | 2 | 110 | 60 |
| **12** | 162 | 35 | 62 | 12 | 105 | 37 |
| **13** | 182 | 36 | 56 | 4 | 101 | 42 |
| **14** | 167 | 34 | 60 | 6 | 125 | 40 |
| **15** | 154 | 33 | 56 | 17 | 251 | 250 |
| **16** | 166 | 33 | 52 | 13 | 210 | 115 |
| **17** | 247 | 46 | 50 | 1 | 50 | 50 |
| **18** | 202 | 32 | 52 | 11 | 230 | 80 |
| **19** | 138 | 33 | 68 | 2 | 110 | 43 |

2、SAS程序

data chapter10;

input x1 x2 x3 y1 y2 y3;

cards;

191 36 50 5 162 60

193 38 58 12 101 101

189 35 46 13 155 58

211 38 56 8 101 38

176 31 74 15 200 40

169 34 50 17 120 38

154 34 64 14 215 105

193 36 46 6 70 31

176 37 54 4 60 25

156 33 54 15 255 73

189 37 52 2 110 60

162 35 62 12 105 37

182 36 56 4 101 42

167 34 60 6 125 40

154 33 56 17 251 250

166 33 52 13 210 115

247 46 50 1 50 50

202 32 52 11 230 80

138 33 68 2 110 43

;

proc print;

run;

proc concorr all;

var x1 x2 x3;

with y1 y2 y3;

run;

3、结果分析

|  | **典型 相关** | **调整 典型 相关** | **近似 标准 误差** | **典型 相关 平方** | **特征值: Inv(E)\*H = CanRsq/(1-CanRsq)** | | | | **H0 检验: 当前行和之后的所有行的典型相关都是零** | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **特征值** | **差分** | **比例** | **累积** | **似然 比** | **近似 F 值** | **分子自由度** | **分母自由度** | **Pr > F** |
| **1** | 0.792515 | 0.747174 | 0.087662 | 0.628080 | 1.6888 | 1.6036 | 0.9520 | 0.9520 | 0.34271649 | 1.95 | 9 | 31.789 | 0.0798 |
| **2** | 0.280189 | 0.114321 | 0.217198 | 0.078506 | 0.0852 | 0.0852 | 0.0480 | 1.0000 | 0.92147912 | 0.29 | 4 | 28 | 0.8806 |
| **3** | 0.004023 | . | 0.235698 | 0.000016 | 0.0000 |  | 0.0000 | 1.0000 | 0.99998382 | 0.00 | 1 | 15 | 0.9878 |

| **Multivariate Statistics and F Approximations** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **S=3 M=-0.5 N=5.5** | | | | | |
| **统计量** | **值** | **F 值** | **分子自由度** | **分母自由度** | **Pr > F** |
| **Wilks' Lambda** | 0.34271649 | 1.95 | 9 | 31.789 | 0.0798 |
| **Pillai's Trace** | 0.70660218 | 1.54 | 9 | 45 | 0.1632 |
| **Hotelling-Lawley Trace** | 1.77396085 | 2.45 | 9 | 17.5 | 0.0520 |
| **Roy's Greatest Root** | 1.68875044 | 8.44 | 3 | 15 | 0.0016 |

输入用似然法比检验典型相关系数与0，由于p并不是都很小，所以所做的典型相关分析不一定有统计意义。

| **VAR 变量 的原始典型系数** | | | |
| --- | --- | --- | --- |
|  | **V1** | **V2** | **V3** |
| **x1** | -0.019713515 | -0.05975416 | -0.019631999 |
| **x2** | 0.4046168947 | 0.1895808331 | 0.16610938 |
| **x3** | -0.001141988 | -0.094106135 | 0.1253144287 |

| **WITH 变量 的原始典型系数** | | | |
| --- | --- | --- | --- |
|  | **W1** | **W2** | **W3** |
| **y1** | -0.047337426 | 0.0631065821 | 0.2458238702 |
| **y2** | -0.016467784 | -0.008368616 | -0.015457876 |
| **y3** | 0.0113569816 | 0.0216644917 | -0.005774493 |

| **VAR 变量 的标准化典型系数** | | | |
| --- | --- | --- | --- |
|  | **V1** | **V2** | **V3** |
| **x1** | -0.4894 | -1.4833 | -0.4873 |
| **x2** | 1.3218 | 0.6193 | 0.5427 |
| **x3** | -0.0083 | -0.6841 | 0.9109 |

| **WITH 变量 的标准化典型系数** | | | |
| --- | --- | --- | --- |
|  | **W1** | **W2** | **W3** |
| **y1** | -0.2554 | 0.3405 | 1.3265 |
| **y2** | -1.0688 | -0.5432 | -1.0033 |
| **y3** | 0.5825 | 1.1112 | -0.2962 |

由表可知var变量与with变量的两个典型相关的原始型及标准化后的典型系数的近似值，

如w1=-0.2554y1-1.0688y2+0.5825y3；

典型冗余分析结果如下

| **型变量号** | **它们自己的 典型变量** | | **典型 R 方** | **对立面 典型变量** | |
| --- | --- | --- | --- | --- | --- |
| **比例** | **累积 比例** | **比例** | **累积 比例** |
| **1** | 0.2727 | 0.2727 | 0.6281 | 0.1713 | 0.1713 |
| **2** | 0.4257 | 0.6985 | 0.0785 | 0.0334 | 0.2047 |
| **3** | 0.3015 | 1.0000 | 0.0000 | 0.0000 | 0.2047 |

|  | | | | | |
| --- | --- | --- | --- | --- | --- |
| **典型变量号** | **它们自己的 典型变量** | | **典型 R 方** | **对立面 典型变量** | |
| **比例** | **累积 比例** | **比例** | **累积 比例** |
| **1** | 0.4946 | 0.4946 | 0.6281 | 0.3107 | 0.3107 |
| **2** | 0.4296 | 0.9242 | 0.0785 | 0.0337 | 0.3444 |
| **3** | 0.0758 | 1.0000 | 0.0000 | 0.0000 | 0.3444 |

| **典型变量号** | **它们自己的 典型变量** | | **典型 R 方** | **对立面 典型变量** | |
| --- | --- | --- | --- | --- | --- |
| **比例** | **累积 比例** | **比例** | **累积 比例** |
| **1** | 0.4205 | 0.4205 | 0.6281 | 0.2641 | 0.2641 |
| **2** | 0.1915 | 0.6120 | 0.0785 | 0.0150 | 0.2791 |
| **3** | 0.3880 | 1.0000 | 0.0000 | 0.0000 | 0.2791 |

|  | | | | | |
| --- | --- | --- | --- | --- | --- |
| **典型变量号** | **它们自己的 典型变量** | | **典型 R 方** | **对立面 典型变量** | |
| **比例** | **累积 比例** | **比例** | **累积 比例** |
| **1** | 0.4430 | 0.4430 | 0.6281 | 0.2783 | 0.2783 |
| **2** | 0.4273 | 0.8703 | 0.0785 | 0.0335 | 0.3118 |
| **3** | 0.1297 | 1.0000 | 0.0000 | 0.0000 | 0.3118 |

| **VAR 变量 和 WITH 变量 前 M 个典型变量之间的多重相关系数平方** | | | |
| --- | --- | --- | --- |
| **M** | **1** | **2** | **3** |
| **x1** | 0.1748 | 0.2113 | 0.2113 |
| **x2** | 0.5669 | 0.5719 | 0.5719 |
| **x3** | 0.0506 | 0.0541 | 0.0542 |

| **WITH 变量 和 VAR 变量 前 M 个典型变量之间的多重相关系数平方** | | | |
| --- | --- | --- | --- |
| **M** | **1** | **2** | **3** |
| **y1** | 0.3196 | 0.3395 | 0.3395 |
| **y2** | 0.4870 | 0.4989 | 0.4989 |
| **y3** | 0.0282 | 0.0970 | 0.0970 |

由最后一个表可以看出var变量对with变量的预测能力，由表可知，x1x2对y1由很强的相关性，所以有很强的预测能力。

同理x1x2跟y2有很强的相关性，对y2有很强的预测能力.