

社会统计学及SPSS软件应用

STATISTICS WITH SPSS

Instructor: 王荣欣

Email: rxwang@qq.com

周一3-4节、单周周四3-4节, 3A106-2

2020年11月30日

CONTENTS

1 因子分析

- 因子分析例子
- 统计软件操作

因子分析的步骤

- 1 计算所有变量的相关矩阵。
 - (1) 大部分相关系数大于0.3。
 - (2) KMO测度。0.9以上为非常好，0.8以上为好，0.7为一般，0.5以下不能接受。
- 2 提取因子（求解初始因子）。
- 3 进行因子旋转（通过坐标变换使因子解的实际意义更容易解释）。
- 4 因子的解释与命名。
- 5 计算因子值。

- 1 1990年30个省、自治区、直辖市的数据
- 2 变量包括：初中以上文化程度的人口比例（x3）、人均国民收入（x4）、城镇人口比例（x5）、多孩率（x1）、综合节育率（x2）

```
. corr x1 x2 x3 x4 x5  
(obs=30)
```

	x1	x2	x3	x4	x5
x1	1.0000				
x2	-0.7610	1.0000			
x3	-0.5418	0.2929	1.0000		
x4	-0.4528	0.2528	0.7712	1.0000	
x5	-0.4534	0.2447	0.8488	0.8777	1.0000

多孩率 (x1) 和综合节育率 (x2) 之间存在较强的相关关系，另外三个社会经济变量之间也存在较强的相关关系。

```
. estat kmo
```

Kaiser-Meyer-Olkin measure of sampling adequacy

Variable	kmo
x1	0.6577
x2	0.5770
x3	0.7989
x4	0.7831
x5	0.7000
Overall	0.7132

KMO值在0.7以上，说明可以对该数据进行因子分析。

选择任意一种求解方法：

- 1 factor x1-x5,pf
- 2 factor x1-x5,pcf
- 3 factor x1-x5,ipf
- 4 factor x1-x5,ml

以主成分分析法为例

- 1 factor x1-x5,pcf /*主成分分析法*/
- 2 greigen, yline(1) /*特征值*/
- 3 screeplot /*碎石图*/
- 4 rotate,varimax /*因子旋转（最大方差法）*/


```
. factor x1-x5,pcf
(obs=30)
```

```
Factor analysis/correlation
Method: principal-component factors
Rotation: (unrotated)
```

```
Number of obs   =      30
Retained factors =       2
Number of params =       9
```

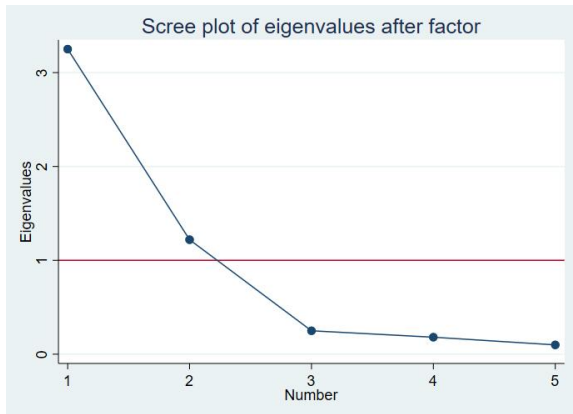
Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	3.25028	2.03049	0.6501	0.6501
Factor2	1.21979	0.97016	0.2440	0.8940
Factor3	0.24963	0.06862	0.0499	0.9439
Factor4	0.18102	0.08174	0.0362	0.9801
Factor5	0.09928	.	0.0199	1.0000

```
LR test: independent vs. saturated:  chi2(10) = 110.81 Prob>chi2 = 0.0000
```

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Uniqueness
x1	-0.7616	0.5543	0.1127
x2	0.5684	-0.7684	0.0865
x3	0.8918	0.2550	0.1396
x4	0.8704	0.3466	0.1223
x5	0.8912	0.3701	0.0689

第一个因子与x3、x4、x5更紧密，第二个因子与x1、x2更紧密。



曲线开始变平的前一个点被认为是提取的最大因子数。后面的这些散点就像山脚下的碎石，舍去这些“碎石”，并不会损失很多信息。

```
. rotate, varimax
```

```
Factor analysis/correlation      Number of obs   =      30
Method: principal-component factors  Retained factors =       2
Rotation: orthogonal varimax (Kaiser off)  Number of params =       9
```

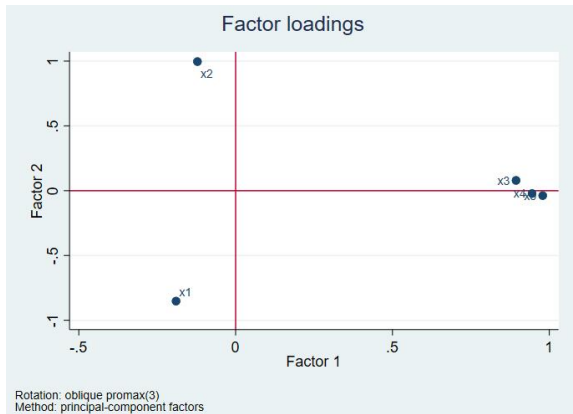
Factor	Variance	Difference	Proportion	Cumulative
Factor1	2.68177	0.89348	0.5364	0.5364
Factor2	1.78829	.	0.3577	0.8940

```
LR test: independent vs. saturated:  chi2(10) = 110.81 Prob>chi2 = 0.0000
```

```
Rotated factor loadings (pattern matrix) and unique variances
```

Variable	Factor1	Factor2	Uniqueness
x1	-0.3530	-0.8733	0.1127
x2	0.0757	0.9528	0.0865
x3	0.8917	0.2555	0.1396
x4	0.9220	0.1664	0.1223
x5	0.9520	0.1575	0.0689

通过因子旋转后，使每个因子上的载荷尽可能地拉开距离。



因子旋转后，x1和x2更接近十字线的纵轴（因子2），x3、x4、x5更接近横轴（因子1）

参考文献

- 1 郭志刚，2015，《社会统计分析方法——SPSS软件应用（第二版）》，北京：中国人民大学出版社