

- 1、计算多相序列相关矩阵（这里不能用 Pearson 相关，因为不是连续变量）

```
poly_res <- polychoric(item_df)
R <- poly_res$rho # 多项序列相关矩阵（用于后续 EFA）
```

- 2、KMO 检验与 Bartlett 检验

KMO 检验：MSA = 0.78，还算挺不错的

Bartlett 检验：p < .01

Kaiser-Meyer-Olkin factor adequacy

Call: KMO(r = R)

Overall MSA = 0.78

MSA for each item =

Item	IHAS1	IHAS2	IHAS3	IHAS4	IHAS5	IHAS6	IHAS7	IHAS8
IHAS1	0.87	0.75	0.89	0.67	0.72	0.86	0.93	0.88
IHAS9	0.75	0.78	0.82	0.83	0.66	0.76	0.88	0.83
IHAS17	0.75	0.84	0.83	0.79	0.66	0.87	0.33	0.78
IHAS25	0.84	0.85	0.67	0.82	0.80	0.83	0.90	0.72
IHAS33	0.78	0.44	0.78	0.82	0.58	0.85	0.79	0.91
IHAS41	0.66	0.76	0.58	0.71	0.82	0.59	0.88	0.77
IHAS49	0.84	0.76						

\$schisq

[1] 9589.03

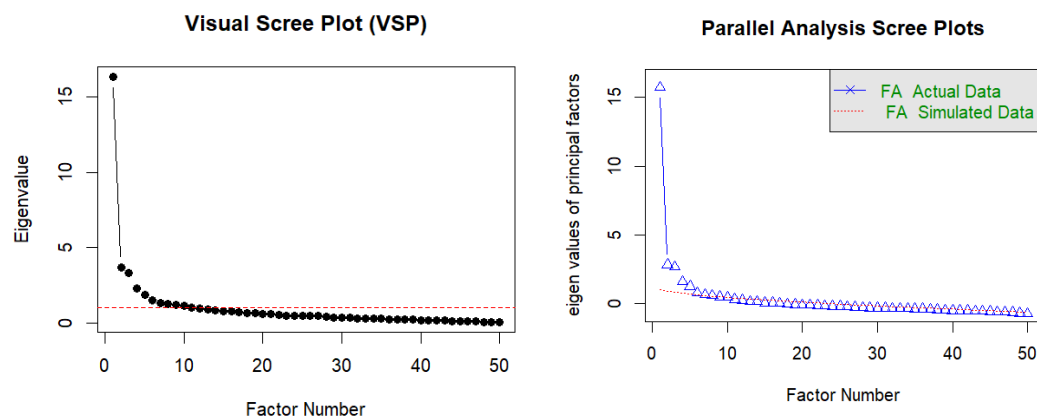
\$p.value

[1] 0

\$df

[1] 1225

- 3、决定因子数量（Visual Scree Plot (VSP) + Parallel Analysis)



根据碎石图的拐点和平行分析的结果，选择因子数量 $n = 10$ 。这里不能简单地取大于 1 的本征值数作为因子数，这一点在很多文献中提到¹⁻³。

- 4、运行第一次 EFA

使用的方法：主轴分析法和斜交旋转。关于用主成分分析法 vs 主轴分析法，正交旋转（如 varimax）vs 斜交旋转存在一些 debate，后者比较明确，因为心理学调查中 item 之间几乎不可能独立，所以应该使用斜交旋转²。前者争议比较大，但近期论文认为 PCA 用于 EFA 是错误的，在文献里之所以大家都用 PCA 是因为 SPSS 的 default 就是 PCA，而正确的方法应该是 Principal Axis Factoring 或 Maximum Likelihood³。

筛选标准：（1）共同度大于 0.4 的，保留²；（2）load onto primary factor > 0.4, load onto alternative factor < 0.3, and load_pri - load_alter > 0.2，满足上述三个条件才保留³

	PA1	PA3	PA5	PA7	PA2	PA9	PA8	PA6	PA10	PA4
SS loadings	4.52	4.31	3.65	3.20	3.07	2.96	2.66	2.29	2.17	1.23
Proportion Var	0.09	0.09	0.07	0.06	0.06	0.06	0.05	0.05	0.04	0.02
Cumulative Var	0.09	0.18	0.25	0.31	0.37	0.43	0.49	0.53	0.58	0.60
Proportion Explained	0.15	0.14	0.12	0.11	0.10	0.10	0.09	0.08	0.07	0.04
Cumulative Proportion	0.15	0.29	0.41	0.52	0.62	0.72	0.81	0.89	0.96	1.00

Loadings:

	PA1	PA3	PA5	PA7	PA2	PA9	PA8	PA6	PA10	PA4
IHAS1	0.627									
IHAS2	0.828									
IHAS3	0.632									
IHAS4								0.610		
IHAS5	0.624									
IHAS6	0.629									
IHAS7	0.676									
IHAS8	0.542									
IHAS9	0.340							0.482		
IHAS10	0.373							0.470		
IHAS11		0.545								
IHAS12		0.690								
IHAS13		0.878								
IHAS14		0.854								
IHAS15		0.585								
IHAS16								0.437		
IHAS17							0.428			
IHAS18						0.522				
IHAS19						0.673				
IHAS20						0.693				
IHAS21						0.417				0.448
IHAS22				0.511						
IHAS23						0.329				
IHAS24										
IHAS25				0.828						
IHAS26				0.694						
IHAS27				0.331						
IHAS28				0.326						
IHAS29		0.301					0.317			
IHAS30							0.346			
IHAS31							0.560			
IHAS32					0.654					
IHAS33							0.566			
IHAS34					0.736					
IHAS35					0.626					
IHAS36					0.488					
IHAS37					0.872					
IHAS38										
IHAS39			0.723							
IHAS40			0.686							
IHAS41							0.313			
IHAS42			0.586							
IHAS43								0.316		
IHAS44								0.403		
IHAS45							0.349			
IHAS46			0.464							
IHAS47			0.643							
IHAS48										
IHAS49								0.595		
IHAS50								0.572		

5、这样子删掉了 24 道题，具体删除原因可见 excel 表。剩下 6 个因子，26 道题（参加 final_round1 这一列）。

- 6、接下来再做一次 EFA 验证这样的因子结构是否合适，还有无修建的可能。同上述流程操作一遍，再删去一道题（第 36 题，因为它在 primary factor 上的 loading < 0.4）。最后总共剩下 25 题，6 个因子。
- 7、再做最后一次 EFA，不再有要删的题，最后方差解释率达到 65%

	PA1	PA3	PA4	PA2	PA6	PA5
SS loadings	4.44	2.87	2.79	2.36	2.06	1.73
Proportion Var	0.18	0.11	0.11	0.09	0.08	0.07
Cumulative Var	0.18	0.29	0.40	0.50	0.58	0.65
Proportion Explained	0.27	0.18	0.17	0.15	0.13	0.11
Cumulative Proportion	0.27	0.45	0.62	0.77	0.89	1.00

References

1. Flora DB, Flake JK. The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Can J Behav Sci Rev Can Sci Comport.* 2017;49(2):78-88. doi:10.1037/cbs0000069
2. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. doi:10.7275/JYJ1-4868
3. Howard MC. A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *Int J Human-Computer Interact.* 2016;32(1):51-62. doi:10.1080/10447318.2015.1087664