**JOTor Thogersen**

**DA 410 -- Multivariate Analysis -- Winter 2018**

**Project 8**

**Part 1**: Read Example 14.3.3, and use SAS or R (I would recommend SAS) to complete Problem 14.4. part (a). Make sure you include the commands and outputs, as well as the interpretations of the outputs.

**# Load Data**

DATA probe;  
INFILE "/folders/myfolders/probe.txt";  
INPUT observation $ 1-3 y1 4-6 y2 7-9 y3 10-12 y4 13-15 y5 16-17;  
run;  
  
**/\* ML factor analysis with a varimax rotation \*/**  
PROC FACTOR DATA=PROBE HEY   
 METHOD=ML **/\* Change this to METHOD=PRIN for a principal factor** solution \*/  
 PRIORS=MAX   
 ROTATE=VARIMAX **/\* Change this to ROTATE=NONE for an unrotated** solution \*/  
 NFACT=2 **/\* Specifies the number of factors desired \*/** RESIDUALS OUTSTAT=factout;  
 VAR y1 y2 y3 y4 y5;  
RUN;  
  
**/\* \*\*\*\*\*\*\* Some diagnostic plots \*\*\*\*\*\*\*\* \*/  
/\* Just copy all this into SAS \*/**  
DATA tempc;  
 SET factout;  
 vtemp=\_NAME\_;  
 KEEP vtemp \_NUMERIC\_;  
 WHERE \_TYPE\_="CORR";  
RUN;  
  
DATA tempr;  
 SET factout;  
 vtemp=\_NAME\_;  
 KEEP vtemp \_NUMERIC\_;  
 WHERE \_TYPE\_="RESIDUAL";  
RUN;  
  
PROC TRANSPOSE DATA=tempc OUT=tempc2;  
 VAR \_NUMERIC\_;  
 BY vtemp;  
RUN;  
  
PROC TRANSPOSE DATA=tempr OUT=tempr2;  
 VAR \_NUMERIC\_;  
 BY vtemp;  
RUN;  
  
DATA tempc3;  
 SET tempc2;  
 pair=trim(vtemp)||trim(\_NAME\_);  
 original=COL1;  
 KEEP pair original;  
 WHERE vtemp>\_NAME\_;  
RUN;  
  
DATA tempr3;  
 SET tempr2;  
 pair=trim(vtemp)||trim(\_NAME\_);  
 residual=COL1;  
 KEEP pair residual;  
 WHERE vtemp>\_NAME\_;  
RUN;  
  
DATA fitdata;  
 MERGE tempc3 tempr3;  
 BY pair;  
 predicted=original-residual;  
RUN;  
  
**/\* Predicted correlations vs. original correlations \*/**  
PROC GPLOT DATA=fitdata;  
 PLOT predicted\*original;  
 run;  
  
 **/\* Residual correlations vs. Predicted correlations \*/**  
PROC GPLOT DATA=fitdata;  
 PLOT residual\*predicted / vref=0;  
 run;  
 **/\* Histogram, Stem-and-leaf plot, and summary statistics for the residuals \*/**  
PROC UNIVARIATE DATA=fitdata PLOT;  
 VAR residual;  
 HISTOGRAM residual;  
run;

**/\* ML factor analysis with a varimax rotation \*/**PROC FACTOR DATA=PROBE

HEY   
 METHOD=ML /**\* Change this to METHOD=PRIN for a principal factor solution \*/**  
 PRIORS=MAX   
 ROTATE=VARIMAX **/\* Change this to ROTATE=NONE for an unrotated solution \*/**  
 NFACT=2 **/\* Specifies the number of factors desired \*/**  
 RESIDUALS OUTSTAT=factout;  
 VAR y1 y2 y3 y4 y5;  
RUN;

**# Goodness of fit:**

**Overall model range for TLI and CFI are in the good range**

|  |  |
| --- | --- |
| **Chi-Square without Bartlett's Correction** | 1.5376366 |
| **Akaike's Information Criterion** | -0.4623634 |
| **Schwarz's Bayesian Criterion** | -0.8602586 |
| **Tucker and Lewis's Reliability Coefficient** | 1.0370694 |

**# Overall the models Root Mean Square Off-Diagonal Residuals: Overall = 0.05032111 < .05 was in the good range.**

**# Thus two factor model does fit the data very well.**











