Multivariate Statistical Techniques
CLPS2908, Spring 2019
Prof. Bertram Malle

Homework 8:

Factor Analysis (FA)

The files you need for this assignment can be found on Canvas. The data consist of a 20-item Self-Esteem (SE) questionnaire. Complete it first for yourself before moving on.

1. The SE questionnaire was designed to have three subscales: performance self-esteem, social self-esteem, and appearance self-esteem.
   (a) To compute these subscales you will need to recode some items. Specifically, reverse-score the following items:
   `SE2 SE4 SE5 SE7 SE8 SE10 SE13 SE15 SE16 SE17 SE18 SE19 SE20`

   (b) Now you can compute the subscale means, following the information below. To make your interpretation of later results easier, first rename the items SE1 to SE20 such that they reflect their subscale meaning, so SE1 → PERF1, SE4 → PERF2, SE2 → SOC1, etc.

   Performance self-esteem:
   Average of `SE1, SE4, SE5, SE9, SE14, SE18, SE19`
   Social self-esteem:
   Average of `SE2, SE8, SE10, SE13, SE15, SE17, SE20`
   Appearance self-esteem:
   Average of `SE3, SE6, SE7, SE11, SE12, SE16`
2. (a) Run the RELIABILITY procedure on the entire set of 20 items *and* on each the three subscales. For each run, report (or copy the output for) the average inter-item correlation and Cronbach's α.
   (b) Comment on the degree of reliability of the total scale and each subscale.
   (c) Now compute the intercorrelations among the three subscales.
   (d) From the information gleaned so far, would you say that the whole 20-item measure is unidimensional? Or does it consist of independent dimensions?

3. Run a *PCA* on the 20 items.
   (a) Settle, after some exploration, on an optimal number of factors and briefly discuss your solution. To aid your discussion, include the scree plot and the table of eigenvalues (and percent variance explained) from your output.
   Extra credit: Run a parallel analysis to inform your decision.
   (b) Apply a VARIMAX rotation. How does the rotated solution differ from the unrotated one? How do you interpret the components of the rotated solution? To facilitate interpretation, use the SPSS subcommand that blanks out loadings below a certain value (e.g., .30) and orders variables within the factors on which they most strongly load:
   `/FORMAT = SORT BLANK (.30).`
   (Add the loading matrix to your write-up.)

4. Now run a *factor analysis*, using principal axis factoring.
   (a) Settle on an optimal number of factors, using whatever data and arguments you have available (if you run the analysis a few times, don't add all the output).
   (b) Apply a VARIMAX rotation. Comment on the meaning of the rotated factors. Show the loading matrix, with the loadings ordered within each factor and with very small loadings deleted.
   (c) In a few sentences, compare the results and interpretation of this FA with those of the PCA analysis you conducted earlier.
5. Now apply an oblique rotation, using OBLIMIN. Are the results different from the orthogonal rotation? To answer this question,
   (a) examine the FACTOR INTERCORRELATION MATRIX to determine how much obliqueness among factors the rotation produced;
   (b) compare the FACTOR MATRIX (from the VARIMAX rotation) to the STRUCTURE MATRIX (from the OBLIMIN rotation); and
   (c) do you notice differences between loadings in the PATTERN matrix and the STRUCTURE matrix? If so, explain those differences.
6. Are there any items in the self-esteem questionnaire that you would exclude because they consistently perform badly in your analyses (i.e., the reliability analysis, the PCA, and the FA)?
7. Spend considerable time on your summary, which should contain decisions and justifications regarding
   (a) whether one should favor the results from the PCA or the factor analysis,
   (b) how many dimensions the measure contains, and
   (c) whether these dimensions are independent (orthogonal) or should be treated as oblique.