Exploratory Factor Analysis

PSY 600K
Dr. Alyssa Gibbons
February 27, 2018

Agenda

- When would you use an EFA?
- The logic of EFA
- Extracting factors
  - The eigenvalue > 1 rule
  - Scree plots
  - Parallel analysis
- Communality & uniqueness

Exploratory Factor Analysis

- When you have no hypotheses about how the items relate to the common factors.
  - Or you simply aren't willing to state these hypotheses.
- Unrestricted factor model – we don't specify:
  - How many factors there are.
  - Which items load on which factors.
  - How many factors a single item can load on.
    - At least at first.

Arguments for EFA

- When we are just starting to measure a construct, it is presumptuous to claim we "know" what the structure should be.
- We might have missed something – what we think is a unitary construct might be more complex.
- Less restrictive model is better because it allows the data to "speak."
  - Instead of "forcing" it to fit a particular pattern.
- Sometimes CFA models don't fit, and then we need to make sense of the data.

The Logic of EFA

- In EFA, we can have (up to) as many factors as we have items.
  - $X_1 = a_{11}f_1 + a_{12}f_2 + \cdots + a_{1m}f_m + u_1$, up to $m = p$.
  - Items can load on one, some, or all of the factors.
- Why doesn't each item just load onto its own factor?
  - Because the items are correlated.
  - Items 1, 2, and 3 may all be highly influenced by Factor 1, etc.
- When we run an EFA, our software will estimate a set of factor loadings λ_1 through λ_m that recreate our observed covariance matrix as closely as possible.

Extracting Factors

- Having as many factors as we have items doesn't really help us much.
- We want to find the smallest number of factors that:
  - Can be interpreted.
  - Account for the majority of variance in our data (not overlooking anything important).
- This process is called *extracting factors*.
- How do we do it?
  - We use a matrix algebra concept called *eigenvalues*…

Eigenvalues

- An **eigenvalue** of a matrix is a scalar (single value) that essentially sums up all of the information in a matrix.
  - For a given square matrix **A:**
  - **λX = AX**
  - Where λ is the eigenvalue and X is the corresponding *eigenvector*.
  - Multiplying λ by X gives the same result as multiplying **A** by X; therefore, λ is an efficient way to summarize **A.**
- Matrices can have multiple sets of eigenvalues/ eigenvectors.
  - In fact, we can find as many eigenvalues as we have items.

Size of Eigenvalues

- Eigenvalues can be thought of as explaining variance within a correlation matrix.
  - A large eigenvalue suggests a substantial amount of shared variance from a common source.
- In EFA, we look to eigenvalues to tell us how many factors we need to explain most of the variance in our data.

The Kaiser / Eigenvalue > 1 Rule

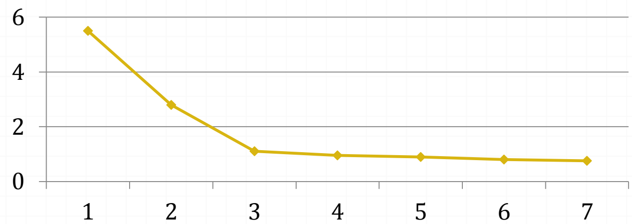
- When we factor analyze a correlation matrix, the variance of each item = 1.
  - Because correlations are a *standardized* covariance – we transformed the item to have variance = 1.
- Any eigenvalue > 1, then, explains more variance than a single item does.
  - This means that it is actually more efficient to combine items than to use them individually.
  - This is, of course, a big part of our goal.
- Guttman (1954) noted that the minimum number of dimensions for a correlation matrix is the number of eigenvalues that are > 1.
- Kaiser (1960) argued that any factor with an eigenvalue < 1 would have negative reliability.

The Kaiser / Eigenvalue > 1 Rule

- Somehow, this was transformed into a prescriptive rule that the number of eigenvalues > 1 = the number of meaningful factors in a set of items.
  - But this is not what Guttman or Kaiser said!
- Early work on this criterion was based on principal components analysis, which is **not** FA.
- In general, the eigenvalues > 1 rule tends to extract **too many** factors – more than are really meaningful.
  - What are the practical implications of this?
- So what should we use instead?

Scree Plot

- The first eigenvalue is always the largest. They go in descending order by size, and the relationship isn't linear.
- In fact, we can plot them like this:

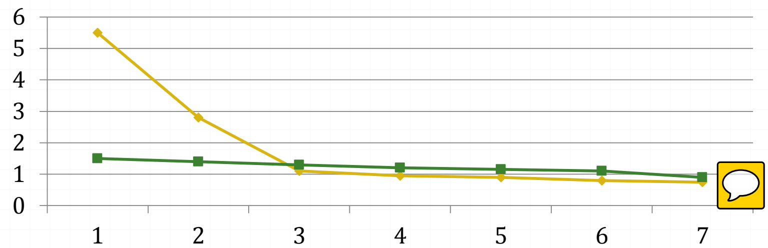


- There is a natural break point where the plot appears to level off. Many argue that this suggests the best number of factors to retain.
  - After this point, the added explanatory power of additional factors is minimal.

Parallel Analysis

- Often recommended as a less subjective alternative to the scree plot.
- For PA:
  - Simulate a matrix of **random** data that is the same size as your actual data matrix.
  - Perform an EFA on this random matrix as well as on your real matrix.
  - Only retain factors with eigenvalues larger than those you get from the random data.

Example of Parallel Analysis



- Simulation studies suggest that PA is **the most accurate** strategy.
- Although it may not work so well with correlated factors…

Parting Thoughts on Extracting Factors

- "Although the default criterion in both SPSS and SAS, K1 has consistently been found to be inaccurate, and review articles are unanimous in recommending against its use."
  - \- Bandalos & Boehm-Kaufman (2008).

- "In the end, researchers should retain a factor only if they can interpret it in a meaningful way no matter how solid the evidence for its retention based on the empirical criteria."
  - \- Worthington & Whittaker (2006), quoted in Bandalos & Boehm-Kaufman (2008).

Communality

- EFA results allow us to calculate a statistic called the **communality** of each item:
  - The proportion of **item** variance that is due to the (whole set of) common factors.
- $Var(X_j) = Var(a_{j1}f_1 + a_{j2}f_2 + \cdots + a_{jm}f_m) + Var(u_j)$
  - Does this look familiar?
- All we've done is separate the variance into a common part and a unique part.
- If the factors are uncorrelated:
  - $h_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jm}^2$
- If the factors are correlated:
  - $h_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jm}^2 + 2(a_{j1}a_{j2}\varphi_{12} + \cdots + a_{j,m-1}a_{jm}\varphi_{m,m-1})$

Communality & Uniqueness

- If we use correlations, the variance of each item = 1.
  - $1 = h_j^2 + \theta_j^2$
  - 1 = communality + uniqueness.
- Uniqueness = sources of variance that are not accounted for by any common factors.
  - Includes error that is specific to this item **and** random measurement error.
- Theoretically, we can decompose this variance further:
  - $1 = h_j^2 + \delta_j^2 + \omega_j^2$
  - 1 = communality + specificity + random error.
- High communalities on all items indicate that our solution is pretty good.
  - Our factors are accounting for most of the variance in our items.

To be continued..

- Orthogonal vs. oblique factors
- Rotation
- Factor analysis vs. principal components analysis

Questions?

For next time:
Issues in EFA
Read: DeVellis pp. 132-151 AND R & M 3.5, 4.1 – 4.2
Bandalos & Boehm-Kaufman (2008)
7th Reading Response