Byrne – Chapter 4

1. Byrne lives in a magical statistical world where CFA is only used for confirming results from an exploratory analysis.
   1. This fact is true to the ideal of statistical steps – EFA then CFA
   2. However, I’ve had several reviews that suggest skipping the EFA altogether if you have clear theoretical reasons why questions should load onto separate latent factors.
   3. Often the best solution is to include both in a paper with a *new* scale (obliviously with old scales you might just do the CFA).
2. New Stuff – How to test assumptions with SEM
   1. Normality
      1. You really do not want kurtotic distributions – where the distribution is fatter or skinner than normal
      2. Skewness affects means, kurtosis affects tests using variance covariance matrices
      3. You can view this by checking boxes in the analysis properties > output > tests for normality and outliers
      4. Do not want kurtosis values over 7 (remember this is z-score standardized basically)
      5. Critical ratio for multivariate normality – nothing greater than 5
         1. If so: Asymptotic distribution free estimation, but only with very large sample sizes (10 times the number of estimated parameters – 1000 people is best)
         2. Use EQS with the Satorra Bentler correction on X2
   2. Outliers
      1. Use mahalanobis much in the same way as you would using SPSS, but here you don’t seem to have a clear critical value, just looking for a big jump in D2.
3. Correlated error variances
   1. May be appropriate when looking at questionnaires with very similar questions
      1. However you have think about also eliminating those questions…why measure the same idea twice?

Brown – Chapter 5

1. Reasons your CFA will suck:
2. Not the right number of factors
   1. This reason probably should not happen if you’ve done the EFA first
   2. Method effects – the way you asked the questions may influence the number of factors, when really you should be using correlated errors (this seems messy to me…)
      1. See examples on page 160, careful using positive and negatively worded items
   3. When you get correlated errors, consider is it the wording of the question OR are you missing a factor (since AMOS has no way of telling you that problem)
   4. Too many factors will be seen with covariance estimates between latent factors close to 1 (i.e. their correlation is high, meaning you are measuring the same thing)
3. The nested model idea
   1. Mostly nested models describe using multiple groups or layers in a model. Here Brown describes the different steps of model testing.
      1. Remember the RCADS example: We had 6-factor, 2-factor and 1-factor models.
      2. In theory, these are nested because the 1 factor model has the most DF, while getting more constricting (using more DFs to estimate parameters) moving from 2 to 6 factor models.
      3. You can use the X2 difference test to examine the differences between those models and figure out which one is better (discussed this last week).
      4. You want to use the difference in DF to look up the critical value, and use the difference in X2 to see if it exceeds this critical value
         1. If so 🡪 the one with the smaller chi square is the better model (remember you always want chi-square to be small)
         2. If not 🡪 models are equal, so you usually go with the parsimonious (simpler) or more theoretical model
4. Not nested model test
   1. AIC and ECVI 🡪 use the one with the lowest values
   2. Problem: no clear way to say statistically that model is better (probably want a big difference between these values if you are going to use them)
5. The measured variables are not put in the right places
   1. Double loadings ☹
   2. Should load on a different factor
   3. Shouldn’t load on anything
   4. Try dropping it and seeing what happens
   5. (Might also consider asking for correlations when checking this stuff – is it correlated to anything? Is the estimate significant?)
6. Correlated error variables
   1. May be because of the way the questions are worded, reverse coded, etc.
   2. Be sure you have a good reason
7. Positive definite matrices
   1. Basically you are finding the inverse of the covariance matrix
   2. You cannot find the inverse of a matrix where two columns are exactly (or correlated enough) the same
      1. It goes kablooey! Or in math terms, you are dividing by zero
      2. Neat trick – use PCA and look at the eigenvalues! If they are all over 0, then you’re good!
   3. Solutions:
      1. Try correlations and figure out which variables are offenders. Combine them or drop one of them.
      2. Lots of missing data filled in will cause this problem (so don’t do it)
      3. If everything is highly correlated, one relationship cannot be lowly correlated
      4. More people!
      5. Eliminate outliers
8. Equal models
   1. Mostly you reject based on logic or theory (should discuss but this doesn’t happen because you want a nice neat paper).

Example Data Depression Anxiety and Stress Scale

Depression: Questions 3, 5, 10, 13, 16, 17, 21

Anxiety: Questions 2, 4, 7, 9, 15, 19, 20

Stress: 1, 6, 8, 11, 12, 14, 18

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | X2 | DF | X2/df | RMSEA | SRMR | CFI | NFI | TLI |
| 3-factor model |  |  |  |  |  |  |  |  |
| 1-factor model |  |  |  |  |  |  |  |  |
| CORRELATED error |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

Start by programming the model as normal.

Then eliminate outliers and discuss if normality is an issue.

Is this a better one factor model? Use a nested chi-square test to examine if it is better.

Using modification indices, what might you change about this model?