1

Tests of Unidimensionality

This primer describes how to conduct tests of unidimensionality using SEM. I review Mplus syntax and output for such tests. I assume you are familiar with SEM. If not, I cover the basics of it in Chapter 7 of my book. Read it first. The test requires that you fit a confirmatory factor analysis (CFA) model consisting of a single latent variable to the items comprising your scale. I consider first the case where the response metric for the items is interval-level in character or reasonably close to it. I then consider the case of ordinal response metrics.

CASE I: INTERVAL LEVEL RESPONSE METRICS

The relevant Mplus syntax for items whose responses approximate interval level properties is in Table 1. I number the lines for reference, but Mplus syntax excludes the numbers. The syntax could be more efficient but I sacrifice efficiency in the interest of pedagogy. I assume you have reviewed the basics of Mplus syntax on the two syntax tabs on my website.

Table 1: Mplus Syntax for One Factor CFA: Interval Level Item Responses

```
1.  TITLE: TEST OF UNIDIMENSIONALITY ;
2.  DATA: FILE IS c:\ret\temp3b.dat ;
3.  VARIABLE:
```

```
4.    NAMES ARE v1 v2 v3 v4 v5 v6 id rep ;
5.    USEVARIABLES ARE v1 v2 v3 v4 v5 v6;
6.    MISSING ARE ALL (-9999) ;
7.  ANALYSIS:
8.    ESTIMATOR = MLR ;
9.  MODEL:
10.   LFACTOR BY v1* v2* v3* v4* v5* v6* ;
11.   LFACTOR@1 ;
12. OUTPUT: SAMP RESIDUAL CINTERVAL TECH4 STAND(STDYX) MOD(ALL 4) ;
```

Line 1 is the title line. I can provide any title I want. Line 2 tells Mplus where to find the data file. Each line in the data file contains 8 values, space delimited, providing the scores for a given individual on the 8 variables. Line 3 tells Mplus I am going to provide information about the variables that are in the data set. Line 4 provides the names I want to assign to the variables in the order they are encountered in the data file. There are 8 names because there are 8 variables. Line 5 specifies the subset of variables I want to use in the model. Line 6 tells Mplus that if it encounters the value -9999 for any of the variables, it should treat it as missing data. By default, Mplus uses full information maximum likelihood (FIML) for missing data for a single factor CFA. Line 7 tells Mplus I am going to provide information about the type of analysis I want. Line 8 specifies the estimator for the analysis to be robust maximum likelihood, to help deal with non-normality. Line 9 tells Mplus I am going to provide information about the model I want Mplus to fit to the data. Line 10 specifies a model with a latent variable called `LFACTOR` (I can name this factor anything, but I cannot have a name that exceeds 8 characters) as reflected BY 6 observed indicators, variables v1 through v6. These are the variables that comprise the items on the scale. The * after each variable name tells Mplus to estimate the factor loading for the indicator. Line 10 tells Mplus to fix the variance of the latent variable to 1.0. The @ sign is read as "fix the referenced parameter to a value of….", followed by the value you want to fix the parameter to. In Mplus, listing a variable by name refers to the variance of the variable (or the error variance, if the variable is endogenous). This is why line 10 fixes the factor variance to 1.0. It can be read as "for the variance of `LFACTOR`, fix it to a value of 1.0." This makes the metric of the latent variable be standardized in form, i.e., it has a mean of 0 (the Mplus default) and a variance of 1.0. This is an alternative strategy to the more common reference indicator approach for assigning a metric to a latent variable. Line 12 tells Mplus what I want to see on the output. I discuss the different options on the first syntax tab of my webpage.

Mplus Output for Test of Unidimensionality

Negative items are reverse scored. An omnibus test of unidimensionality uses global fit indices for the one factor SEM model. Here is selected output for the indices:

```
MODEL FIT INFORMATION

Chi-Square Test of Model Fit

          Value                              12.710*
          Degrees of Freedom                      9
          P-Value                            0.1762
RMSEA (Root Mean Square Error Of Approximation)

          Estimate                            0.029
          90 Percent C.I.                   0.000  0.062
          Probability RMSEA <= .05            0.830

CFI/TLI

          CFI                                 0.996
          TLI                                 0.994
SRMR (Standardized Root Mean Square Residual)

          Value                               0.016
```

The chi square for model fit was 12.71 (df=9), which yielded a statistically non-significant p value (0.18). This is consistent with a reasonable fitting model of unidimensionality. The RMSEA was 0.029 with an upper 90% confidence interval of 0.062, both of which are consistent with a reasonable model fit. The CFI was 1.00 and the standardized RMR was 0.02. Overall, the global fit indices suggest the data are consistent with unidimensionality.

Here are the z tests for the difference between the predicted and observed covariances on a cell-by-cell basis for the unidimensional model:

```
          Standardized Residuals (z-scores) for Covariances
              V1            V2            V3            V4            V5
              ________      ________      ________      ________      ________
 V1          999.000
 V2           -0.065       999.000
 V3            0.366        -0.436       999.000
 V4           -1.803        -0.913         1.026       999.000
 V5            1.963         1.735        -0.485        -0.813       999.000
 V6            0.059        -0.315        -0.423         2.876        -1.658

           Standardized Residuals (z-scores) for Covariances
              V6
              ________
 V6          999.000
```

A value of 999 means the test could not be computed. We are interested in any absolute value greater than 1.96 (except 999). There is one; the predicted and observed covariance between variables 4 and 6 was statistically significantly different from 0. I return to this result shortly.

Here are the relevant modification indices for the model:

```
Minimum M.I. value for printing the modification index     4.000

                                       M.I.     E.P.C.    Std E.P.C.  StdYX E.P.C.
WITH Statements

V6       WITH V4                       5.454     0.072      0.072         0.130
```

Of interest are modification indices (in the column labeled `M.I.`) greater than 4 because they suggest parameters that would yield statistical significance if added to the model. The `WITH` statements refer to correlated errors because both v4 and v6 are endogenous. The value of 0.13 (last column) indicates what the correlation likely would be if added to the model.

Both the modification indices and the z tests of predicted and observed covariances suggest a point of stress in the model for items 4 and 6. Because these tests are performed in the context of a large number of significance tests, they could be chance based. In fact, I created simulated data for this example in which the population model conformed to unidimensionality, so I know the result is chance based. However, I would not know this in my typical role as a researcher.

There are 21 tests of significance for the z tests. For a 0.05 alpha level, I would expect at least one of them to be statistically significant. There were 13 contrasts for the correlated errors. When I examined the content of the two offending items, I saw no substantive reason why I would expect their errors to be correlated. I decide to inspect other features of the data before making a final conclusion about the correlated error.

Here are the unstandardized factor loadings for the one factor model:

```
MODEL RESULTS

                                                       Two-Tailed
                    Estimate       S.E.    Est./S.E.    P-Value

 LFACTOR  BY
    V1                 0.713      0.042     16.921      0.000
    V2                 0.688      0.042     16.386      0.000
    V3                 0.726      0.042     17.360      0.000
    V4                 0.699      0.043     16.180      0.000
    V5                 0.628      0.040     15.738      0.000
    V6                 0.690      0.043     15.940      0.000
```

The standard deviations of v1 through v6 (not shown here) are all close to 1.0, so these loadings are similar to standardized loadings. The main result I look for is that the loadings in the `Estimate` column are similar to one another, which then justifies my plan to use uniform weighting when forming a composite. All seems to be in order in this regard.[1]

Here are the standardized loadings:

```
STDYX Standardization
                                                        Two-Tailed
                     Estimate       S.E.    Est./S.E.    P-Value

LFACTOR   BY
   V1                   0.686      0.028     24.499      0.000
   V2                   0.670      0.029     23.428      0.000
   V3                   0.741      0.025     29.480      0.000
   V4                   0.682      0.029     23.726      0.000
   V5                   0.657      0.029     22.816      0.000
   V6                   0.685      0.029     23.225      0.000
```

The square of each loading is an estimate of the reliability of the respective item. For example, v1 has an estimated reliability of 0.686*0.686 = 0.47; v2 has an estimated reliability of 0.670*0.670 = 0.45; and so on. The product of any two loadings represents the predicted correlation between the items. For example, the predicted correlation between v1 and v2 is 0.686*0.670 = 0.46; the predicted correlation between v1 and v3 is 0.686*0.741 = 0.50. A rough estimate of the margin of error (MOE) for a loading is double its standard error (in the column labeled `S.E.`). The standardized factor loading for v1 is 0.686 ± 0.056. The standardized factor loading for v2 is 0.670 ± 0.058. Given the small MOEs across the loadings, the estimates seem reasonably trustworthy relative to the referent population.

The magnitude of the standardized loadings, the uniformity of the unstandardized loadings, the global fit statistics, and the localized tests of fit, overall, suggest that the unidimensional model is viable. I can further gain perspectives on this by contrasting the observed and predicted correlations between the items. Here are the observed correlations:

```
          Correlations
             V1            V2            V3            V4            V5
          ________      ________      ________      ________      ________
V1           1.000
V2           0.458         1.000
V3           0.514         0.490         1.000
V4           0.436         0.439         0.522         1.000
V5           0.478         0.478         0.477         0.433         1.000
V6           0.471         0.454         0.502         0.510         0.416
```

[1] I can conduct formal significance tests of loading equivalence by fitting a model with all the loadings constrained to be equal, but I do not do so here.

and here are the predicted correlations:

Model Estimated Correlations

	V1	V2	V3	V4	V5
V1	1.000				
V2	0.459	1.000			
V3	0.508	0.497	1.000		
V4	0.468	0.457	0.506	1.000	
V5	0.451	0.441	0.487	0.449	1.000
V6	0.470	0.459	0.508	0.468	0.450

They are close in value. Recall that the correlation between v4 and v6 was "problematic" based on model diagnostics. The predicted correlation between these two items based on a unidimensional model was 0.47; the observed correlation was 0.51. This discrepancy does not seem bothersome in the current context.[2] Even if I grant the presence of a "minor factor" between v4 and v6 (see Chapter 3), it is unlikely to be an issue when I form my composite because the items appear to be, for the most part, *functionally* unidimensional. When I estimate the composite reliability using omega-hierarchical per my discussion in Chapter 3, it will take into account the presence of minor factors.

CASE II: ORDINAL LEVEL RESPONSE METRICS

CFA as applied to item ordinal response metrics traditionally focus on the analysis of polychoric correlations. These are correlations based on the assumption that each item response is a crude indicator of a continuous item response that is normally distributed. For example, v1 might be a binary response (0 = disagree, 1 = agree) to an item that represents a crude measure of a continuous agree-disagree dimension for that item. Polychoric correlations estimate the item correlations for the continuous agreement constructs and these correlations are subjected to the one factor unidimensional model. For more information about polychoric correlations, see my book.

The relevant Mplus syntax for the modeling is in Table 2. There are 5 items each responded to on a binary response metric (0 = disagree, 1 = agree). All items are scored in the same direction, i.e., negative items are reverse scored. All line parallel with syntax for our previous example in Table 1, except lines 6 and 8. The `CATECORICAL` command tells Mplus to treat the variables listed on the line as ordinal. Line 8 tells Mplus to use the WLSMV estimator, which for the CFA will invoke a focus on polychoric correlations and a weight least

[2] When I estimated a model to include the correlated error between V4 and V6 per the modification index (b adding a model command `V4 WITH V6 ;`), the predicted correlation between V4 and V6 became 0.51

squares estimator for it (see my book for details).

Table 2: Mplus Syntax for One Factor CFA for Ordinal Level Item Response

```
TITLE: TEST OF UNIDIMENSIONALITY ;
DATA: FILE IS c:\ret\temp2b.dat ;
VARIABLE:
 NAMES ARE v1 v2 v3 v4 v5 ;
 MISSING ARE ALL (-9999) ;
 CATEGORICAL ARE v1 v2 v3 v4 v5 ;
ANALYSIS:
 ESTIMATOR = WLSMV ;
MODEL:
LFACTOR BY v1* v2* v3* v4* v5* ;
LFACTOR@1 ;
OUTPUT: SAMP RESIDUAL CINTERVAL TECH4 STAND(STDYX) MOD(ALL 4) ;
```

Mplus Output for Test of Unidimensionality

An omnibus test of unidimensionality focuses on the global fit indices for the one factor SEM model. Here is the key output:

```
Chi-Square Test of Model Fit

          Value                               6.065*
          Degrees of Freedom                      5
          P-Value                            0.2999

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                            0.021
          90 Percent C.I.                   0.000  0.068
          Probability RMSEA <= .05            0.806

CFI/TLI

          CFI                                 0.998
          TLI                                 0.997

SRMR (Standardized Root Mean Square Residual)

          Value                               0.024
```

All indices are interpreted as before and all suggest a reasonable model fit.

Mplus does not provide z tests for the difference between the predicted and observed covariances on a cell-by-cell basis because they are not tractable when working with polychoric correlations.[3] However, it does provide modification indices for the correlated errors for the continuous response model underlying the ordinal responses. Here are the results:

```
MODEL MODIFICATION INDICES

Minimum M.I. value for printing the modification index     4.000

                                   M.I.    E.P.C.    Std E.P.C.   StdYX E.P.C.
WITH Statements

V3        WITH V2                  4.903    -0.141     -0.141       -0.294
```

There is one value greater than 4, for items v2 and v3. The expected correlated error if I were to add the parameter is -0.29. As with the first example, when I generated the population data for this example, I created the data to conform to unidimensionality with no correlated errors, so I know the above result is chance based. However, in practice, I would not know this.

Mplus provides on its output the polychoric correlations from the observed data as well as the predicted continuous correlations based on the fitted model. It is instructive to examine these two matrices. The observed correlation matrix is taken from the section of the output called `ESTIMATED SAMPLE STATISTICS` in the initial portion of the output; the predicted correlations are taken from the `TECHNICAL 4 OUTPUT` section under the label `ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES`. Here are the observed correlations:

CORRELATION MATRIX

	V1	V2	V3	V4
V2	0.509			
V3	0.541	0.437		
V4	0.444	0.484	0.561	
V5	0.514	0.498	0.573	0.465

and here are the predicted correlations based on the model:

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	V1	V2	V3	V4
V2	0.473			
V3	0.540	0.511		
V4	0.484	0.458	0.523	
V5	0.515	0.487	0.556	0.499

[3] Mplus provides predicted versus observed estimates of the bivariate contingency tables between all pairs of items, but I do not delve into that output here.

The observed and predicted correlations seem reasonably close.

The unstandardized factor loadings for the one factor model are the same as the standardized factor loadings because I am analyzing a correlation matrix rather than a covariance matrix: Here are the results

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E. Est./S.E.     P-Value

 LFACTOR  BY
    V1                 0.707      0.049     14.562      0.000
    V2                 0.669      0.050     13.337      0.000
    V3                 0.763      0.046     16.703      0.000
    V4                 0.685      0.049     13.915      0.000
    V5                 0.728      0.048     15.151      0.000
```

I note that the loadings, in the `Estimate` column are similar to one another, which justifies my intent to use uniform weighting when forming a composite by summing or averaging the items. The loadings have the same properties as discussed for the first example, e.g., the square of each loading is an estimate of the reliability of the respective item, and so on.

The magnitude of the standardized loadings, the uniformity of the unstandardized loadings, the global fit statistics, and the localized tests of fit, overall, suggest that the unidimensional model is viable. Even if I grant the presence of a "minor factor" between v2 and v3, it is unlikely to be an issue when I form my composite because the items appear to be *functionally* unidimensional. When I estimate the composite reliability using omega-categorical per my discussion in Chapter 3, it will take into account the presence of minor factors.