The Single-Factor Model and True-Score Equivalent Models

Measurement Methods

Lecture 11

Today's Class

- More of the single factor model.
  - Remaining slides from last time.
  - How to fit the true-score model as a special case of the single factor model.

- Estimation of the single factor model in SAS.

The Single-Factor Model

The Spearman Single Factor Model

- In relating an examinee's level of the common factor to their performance on an item, we introduce the Spearman single factor model:

$$X_j = \mu_j + \lambda_j F + E_j$$

- X_j is the score for an examinee on the j^{th} item.

- F is the examinee's measure of the common attribute.

The Spearman Single Factor Model, Continued

$$X_j = \mu_j + \lambda_j F + E_j$$

- E_j is the examinee's measure of the unique or idiosyncratic property of item *j*.
  - The amount by which item *j* is shifted.
- μ_j is the overall mean for an item.
  - Allowing for differing item difficulties.
- λ_j is called the factor loading of item *j*.
  - We will come to discuss this factor loading quite a bit.
  - This is where the "magic" happens.

Model Parameters

- So, for our single factor model, we have a set of parameters we need to estimate:
  - *m* factor loadings – $\lambda_1, \lambda_2, \ldots, \lambda_m$
  - *m* unique variances – $\Psi_1^2, \Psi_2^2, \ldots, \Psi_m^2$
- Note that item means are not necessary to be estimated.
  - We have these already.
- How do we go about getting estimates of these parameters?
- How do we estimate things in general?

Model Estimation

- The key in model estimation is to find a set of parameters that minimize the discrepancy between the observed covariance matrix and the model-predicted covariance matrix.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \lambda_1^2+\psi_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 & \lambda_1\lambda_4 & \lambda_1\lambda_5 \\ \lambda_1\lambda_2 & \lambda_2^2+\psi_2^2 & \lambda_2\lambda_3 & \lambda_2\lambda_4 & \lambda_2\lambda_5 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2+\psi_3^2 & \lambda_3\lambda_4 & \lambda_3\lambda_5 \\ \lambda_1\lambda_4 & \lambda_2\lambda_4 & \lambda_3\lambda_4 & \lambda_4^2+\psi_4^2 & \lambda_4\lambda_5 \\ \lambda_1\lambda_5 & \lambda_2\lambda_5 & \lambda_3\lambda_5 & \lambda_4\lambda_5 & \lambda_5^2+\psi_5^2 \end{bmatrix}$$

- To demonstrate, lets look at an example from the SWLS.

- We essentially have to search for values of our model parameters to minimize the distance from Σ to S.

$$\boldsymbol{S} = \begin{bmatrix} 2.566 & 1.560 & 1.487 & 1.195 & 1.425 \\ 1.560 & 2.493 & 1.283 & 0.845 & 1.313 \\ 1.487 & 1.283 & 2.462 & 1.127 & 1.313 \\ 1.195 & 0.845 & 1.127 & 2.769 & 1.323 \\ 1.425 & 1.313 & 1.313 & 1.323 & 3.356 \end{bmatrix}$$

Model Estimation

- We could try out numerous values for our model parameters.

- For instance consider the following:

λ	Ψ
1.0	0.5
1.0	0.5
1.0	0.5
1.0	0.5
1.0	0.5

- These parameters give us a model predicted covariance matrix of:

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} 1.5 & 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 1.5 & 1.0 & 1.0 & 1.0 \\ 1.0 & 1.0 & 1.5 & 1.0 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.5 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.5 \end{bmatrix}$$

Realistically, Though…

- In reality, a computer will search through the parameter space and find a set that will minimize the discrepancy between Σ and S.
- There are many ways of defining a discrepancy, though.
- Our book introduces a very simple one: the unweighted least squares (ULS) function.

$$q_u = \frac{1}{m^2} \sum_j \sum_k \left(s_{jk} - \sigma_{jk}\right)^2$$

Example Results

TABLE 6.4
Satisfaction With Life Scale—Spearman Analysis

Model Estimates

	(a) λ	ψ^2
1	1.290	0.901
2	1.104	1.274
3	1.148	1.144
4	0.952	1.863
5	1.185	1.951

Model Predicted Covariance Matrix

	(b) 1	2	3	4	5
1	2.565	1.424	1.481	1.328	1.529
2	1.424	2.493	1.267	1.051	1.308
3	1.481	1.267	2.462	1.093	1.360
4	1.328	1.051	1.093	2.769	1.128
5	1.529	1.308	1.360	1.128	3.355

Discrepancies: From S - Σ

	(c) 1	2	3	4	5
1	.0	.135	.006	−.033	−.104
2	.135	.0	.015	−.206	.004
3	.006	.015	.0	.035	−.048
4	−.033	−.206	.035	.0	.195
5	−.104	.004	−.048	.195	.0

- To provide an example, consider the following estimates provided in our textbook (we will learn how to get estimates next week).

Great, But Does The Model "Fit?"

- Once you get model parameter estimates, you must check to see if the model "fits."
  - Measures of model fit are plentiful.
  - For now, we will consider one measure.

- Basically, we want to say a model "fits" when the discrepancy between the observed and model-predicted covariance matrix is small.

- So, we need a function that will quantify the size of the model discrepancy.

The Goodness-of-Fit Index

- Introducing the GFI – an index based on our ULS discrepancy function:

$$q_u = \frac{1}{m^2} \sum_j \sum_k \left(s_{jk} - \sigma_{jk}\right)^2$$

- The GFI is then:

$$GFI = 1 - \frac{q_u}{c}$$

- Where:

$$c = \frac{1}{m^2} \sum_j \sum_k s_{jk}^2$$

Properties of the GFI

- The GFI weights the size of the discrepancies by the size of the covariances in the matrix.
  - That is why we divide by c.
  - Actually, q_u is the average discrepancy.
- If the model fits perfectly, the GFI will be 1.0.
- The GFI will approach 0.0 as the fit gets worse.
- The computer package you use to estimate your model will tell you the goodness of fit.

What is the GFI of the SWLS?

- The GFI of the SWLS was 0.9968.
- From practical experience, we come to learn this is very good fit.
  - Rod says that 0.90 is acceptable and the fit is “good” when the GFI is above 0.95.

What About the Residuals?

- It is also good practice to look at the discrepancy matrix to see the fit.
  - This will tell you a better story of why a model does not fit.
  - It will also tell you if some pairs of items are not well predicted by the model.
- The GFI and other goodness of fit indices are single number summaries of the elements in this matrix.

TABLE 6.4
Satisfaction With Life Scale—Spearman Analysis

	(a)		(b)				
	λ	ψ^2	*1*	*2*	*3*	*4*	*5*
1	1.290	0.901	2.565	1.424	1.481	1.328	1.529
2	1.104	1.274	1.424	2.493	1.267	1.051	1.308
3	1.148	1.144	1.481	1.267	2.462	1.093	1.360
4	0.952	1.863	1.328	1.051	1.093	2.769	1.128
5	1.185	1.951	1.529	1.308	1.360	1.128	3.355

	(c)				
	1	*2*	*3*	*4*	*5*
1	.0	.135	.006	−.033	−.104
2	.135	.0	.015	−.206	.004
3	.006	.015	.0	.035	−.048
4	−.033	−.206	.035	.0	.195
5	−.104	.004	−.048	.195	.0

Back To Test Homogeneity

- By applying the single factor model to the SWLS data, we were in essence trying to determine if the test was homogeneous.
  - If the model fit, we say the test measured a single factor – and was homogeneous.
  - If the model didn't fit, we say the test measured more than a single factor – and was not homogeneous.
- By our standards, we would say our model fit, and that the SWLS was a homogeneous test.

Variations of the Single Factor Model

Variations of the Single Factor Model

- Leaving the generality of the single factor model, we will now introduce two versions of the model with fewer parameters to be estimated:
  - True-score equivalent items model.
  - Parallel items model.

- For both models:
  - F and E will still be uncorrelated for every item j.
  - E_j and E_k are uncorrelated for all distinct pairs of items.

True-Score Equivalent Items

- For true-score equivalent items, we make the constraint that all *j* factor loading parameters are equal:

$$\lambda_1 = \lambda_2 = \ldots = \lambda_j$$

- This model is also called the ***essentially tau-equivalent items model***.

- Note that we reduce the number of parameters from 2m down to m+1.
  - More parsimonious model – but is it always realistic?

Features of the Model

- The true-score equivalent items model gets its name from the following features of the parameterization:

$$X_j = \lambda F + E_j + \mu_j$$

  – Here there is no subscript on the factor loading.
    - They are all equal to the same value for each item.

You look familiar, have we met before?

- If we rewrite the model in a slightly different fashion, we get something that may be familiar to you:

$$X_j = T_x + E_j + \mu_j$$

- Here we write T_x for λF.
  - We are "rescaling" the common attribute/factor so it is measured in the same units as the items.
- The "item true-score" is the common factor multiplied by the factor loading.

More Features of the Model

- In the true-score equivalence model, all the item true scores are said to have variance λ^2.

- Each item is said to measure the attribute equally well.
  - Each has the same level of discrimination.

Model Predicted Variances/Covariances

- The true-score equivalence model posits the covariances for pairs of items as:

$$\sigma_{jk} = \lambda^2 = \sigma^2_{Tx}$$

- The variances for each item is then:

$$\sigma_{jj} = \lambda^2 + \Psi^2_j = \sigma^2_{Tx} + \sigma^2_j$$

The Parallel Items Model

- The parallel-items model takes the true-score equivalence model to a greater length:
  - the factor loadings constrained to be equal

$$\lambda_1 = \lambda_2 = \ldots = \lambda_j$$

  - the unique variances are constrained to be equal:

$$\Psi_1^2 = \Psi_2^2 = \ldots = \Psi_j^2$$

- Now only two parameters need to be estimated, λ and Ψ^2

Model Predicted Variances/Covariances

- The parallel-items model posits the covariances for pairs of items as:

$$\sigma_{jk} = \lambda^2 = \sigma^2_{Tx}$$

  – Same as the true-score equivalence model.

- The variances for each item is then:

$$\sigma_{jj} = \lambda^2 + \Psi^2 = \sigma^2_{Tx} + \Psi^2$$

Model Features

- The parallel-items model posits that:
  - All item variances are equal to each other
    - The diagonal of the covariance matrix
  - All item covariances are equal to each other
    - The off-diagonal of the covariance matrix

- Some writers use the distinction parallel-items to have the additional constraint of equality of means.
  - Rod calls this the strictly parallel items model.

- Item parameters for both models can be estimated by hand.
  - But who wants to do that?

What About Fit for the SWLS?

- Applying both models to the SWLS yielded the discrepancy matrices to the right.
- The GFI for the true-score equivalence model was 0.991
- The GFI for the parallel-items model was 0.949.

TABLE 6.5
Discrepancy Matrices—Restricted Models

(a) True-Score Equivalence Model

	1	*2*	*3*	*4*	*5*
1	.0	.273	.200	−.092	.138
2	.273	.0	−.004	−.442	.026
3	.200	−.004	.0	−.160	.026
4	−.092	−.442	−.160	.0	.036
5	.138	.026	.026	.036	.0

(b) Parallel Items Model

	1	*2*	*3*	*4*	*5*
1	−.163	.273	.200	−.092	.138
2	.273	−.236	−.004	−.442	.026
3	.200	−.004	−.268	−.160	.026
4	−.092	−.442	−.160	.040	.036
5	.138	.026	.026	.036	.627

Estimation of the One-Factor Model

Estimation of the One-Factor Model

- To estimate the one-factor model, we will use the Calis procedure in SAS.
- WARNING: Proc Calis can be a complicated procedure that leaves you very frustrated.
- Used properly, Calis will give you what you need for the one-factor model.

Proc Calis Example Files

- Example files for Calis are available on blackboard – please look at these.

Wrapping Up

- The single-factor model can be a general way to check the homogeneity of a test.
- The true-score equivalence model is a more restricted model.
  - It is similar to what is used in the classical true-score model.
  - We will see it is what is hypothesized when the most common coefficient of reliability (alpha) is computed.
- The parallel-items model takes constraints to the extreme.
  - Sometimes that is what is assumed.

Next Time

- Test reliability.
  - McDonald's Ω – a measure of the reliability of a test.
  - The Guttman-Chronbach α – the standard measure of the reliability of a test.

- The Spearman-Brown prophecy formula.