



# The Reliability of a Homogeneous Test

Measurement Methods

Lecture 12



# Today's Class

- Estimating the reliability of a homogeneous test.
  - McDonald's Omega.
  - Guttman-Chronbach Alpha.
  - Spearman-Brown.
- What to do with binary items.
  - Wait until we talk about more appropriate methods for binary items.
- Brief notes about generalizability.



# The Reliability of a Homogeneous Test



# The Single-Factor Model for Test Scores

- Recall from the previous chapter our discussion was framed about how to estimate the precision of a test.
- The classical true-score model was introduced as a method to get at reliability and the standard error of measurement.
- We will now see what happens when we combine a total test score with the use of the single-factor model.



# Total Test Score

- Taking what we know about our single-factor model, let's see if we can express the total test score in terms of the model:
- The test score is: 
$$Y = \sum_j X_j$$
- The single factor model states: 
$$X_j = \mu_j + \lambda_j F + E_j$$



# Putting the Two Together

- So plugging in for  $X$  we get:

$$Y = \sum_j X_j = \sum_j (\mu_j + \lambda_j F + E_j)$$
$$= \sum_j \mu_j + \sum_j (\lambda_j) F + \sum_j E_j$$

- We can now re-write into:

$$Y = C + U + \mu_y$$

- $C$  is the common part of  $Y$  due to the common factor.
- $U$  is the unique part due to the idiosyncratic properties of the test.
- $\mu_Y$  is the mean of the test  $Y$ .



# Doesn't That Look Familiar?

- So, we are now back to looking at a test as a decomposition into a “true score” and an “error.”
  - The previous slide looked similar to  $Y = T + E$
- The factor model gives us an interpretation of the total true score,  $T_Y$  as the common part  $C$  of  $Y$ , and of the error  $E$  as the unique part  $U$ .
  - So we can rewrite our model as:

$$Y = T_Y + E_Y + \mu_Y$$



# Where Do We Go From Here?

- You can guess what is coming now, can't you?
  - In Chapter 5, we defined the classical true-score model, then went and showed the variance of  $Y$  as a decomposition into true and error variances.
- Under the single factor model, we can also show the variance of  $Y$ .
  - Only now, our decomposition will be having parameters of our single-factor model.





# Variance of Y

- Using the algebra of expectations, we have:

$$\begin{aligned}\text{Var}(Y) &= \text{Var}[\mu_Y + \Sigma (\lambda_j) F + \Sigma(E_j)] \\ &= \text{Var}[\Sigma (\lambda_j) F + \Sigma(E_j)] \\ &= [\Sigma (\lambda_j)]^2 \text{Var}[F] + \text{Var}[\Sigma(E_j)] + \\ &\quad 2\text{Cov}[F, \Sigma(E_j)] \\ &= [\Sigma (\lambda_j)]^2 + \Sigma \Psi_j^2 \\ &= \sigma_{Ty}^2 + \sigma_{Ey}^2\end{aligned}$$



# Reliability Coefficient

- So, we have  $\sigma_Y^2 = \sigma_{Ty}^2 + \sigma_{Ey}^2$
- Recall the reliability coefficient defined in Chapter 5 was:

$$\rho_r = \sigma_{Ty}^2 / (\sigma_{Ty}^2 + \sigma_{Ey}^2) = \sigma_{Ty}^2 / \sigma_Y^2$$

- We now define a new version of the reliability coefficient, constructed with the single-factor model in mind.



# McDonald's Omega

- McDonald's Omega is a measure of reliability of a homogeneous test based on the single-factor model:

$$\omega = \sigma_{T_y}^2 / \sigma_Y^2 = [\sum (\lambda_j)]^2 / (\sigma_Y^2)$$

-or-

$$\omega = \sigma_{T_y}^2 / \sigma_Y^2 = [\sum (\lambda_j)]^2 / ([\sum (\lambda_j)]^2 + \sum \Psi_j^2)$$

- Omega is a ratio of the true-score variance of Y to the total variance of Y.



# Estimation of McDonald's Omega

- To get an estimate of McDonald's Omega, we must first fit our single-factor model.
- If the fit is adequate – and hence, we have a homogeneous test – we then take our factor loading parameters:
  - sum them
  - then square the sum
  - then divide by the variance of Y (variance of a total test score).



# Properties of McDonald's Omega

- Omega is the ratio of the variance due to the common factor to the total variance of  $Y$ .
- Omega is the *square* of the correlation between  $Y$  and the common factor  $F$ .



## More Properties of McDonald's Omega

- Omega is the correlation between two test scores  $Y$  and  $Y'$  that have the same sum of their loadings and the same sum of their unique variances.
- Omega is the square of the correlation between the total (or mean) score on the given  $m$  items and the means score on an infinite set of items from a homogeneous domain of items of which the  $m$  items in the test are a subset.
  - In this respect, it is a measure of the generalizability of the test items.



# Reliability of SWLS Data

- Recall our estimates of the SWLS data parameters.
  - From these we can compute omega.
- Sum of the loadings = 5.69
  - Square of 5.69 = 32.25
- Sum of the unique variances = 7.13
- Omega =  $32.25 / (32.25 + 7.13) = 0.82$

TABLE 6.4  
Satisfaction With Life Scale—Spearman Analysis

	(a)		(b)				
	$\lambda$	$\psi^2$	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



# Coefficient Alpha





# Another Measure of Reliability

- A much more frequently used measure of reliability is the Guttman-Chronbach Alpha.
  - Note that this is commonly called Chronbach's Alpha.
    - Rod (correctly) cites the first reporting of this coefficient in 1945 by Guttman.
- The Omega coefficient of reliability was based on the general single-factor model.
  - Alpha is based on the true-score equivalence model.



# How Does This Work?

- Consider the computation of coefficient omega under the true-score equivalence model:

$$\omega = \frac{m^2 \lambda^2}{\sigma_Y^2}$$

- For the true-score equivalence model in the SWLS example,  $\lambda = 1.134$ .
  - So omega for this case was:
    - $5^2 (1.134)^2 / 39.382 = 0.8170$
- Note that the G-C alpha is slightly lower than our estimate of omega.



# G-C Alpha

- It can be shown (see p. 91), that the classical definition of the G-C alpha is equal to what omega is under the true-score equivalence model.
- The classical definition of the G-C alpha is:

$$\hat{\alpha} = \left( \frac{m}{m-1} \right) \left( 1 - \frac{\sum_j s_{jj}}{s_Y^2} \right)$$



# Calculating Alpha for the SWLS

- To do this, we must sum the diagonal elements of the covariance matrix:
  - $2.566 + 2.493 + 2.462 + 2.769 + 3.356 = 13.646$
- Then we must sum every element in the covariance matrix to get  $s_Y^2$ .
  - $s_Y^2 = 39.382$
- Alpha is then  $(6/(6-1))*(1-(13.646/39.382)) = 0.8170$

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



# Omega and Alpha in Practice

- It is difficult to invent a homogeneous test where Omega differs greatly from Alpha.
- But part of the motivation for introducing Omega is the motivation behind the measure.
  - The factor model should be used to check for homogeneity of a test.
    - If a test is roughly homogeneous, then the sum score can be used.



# Other Meaning of Alpha

- Cronbach showed that Alpha gives the average correlation between all possible splits of the items into halves.
  - The split-half method is a classical method for estimating reliability – it yields  $Y$  and  $Y'$ .
- Alpha is also referred to as the “internal consistency” reliability of a test.
  - Rod recommends against this.



# The Spearman-Brown Prophecy Formula



# Covering All Our Models

- We showed that Omega is a reliability coefficient for the single factor model.
- When the true-score equivalence model is estimated, Omega is equal to the G-C Alpha.
- What about the parallel items model?





# Notes About Parallel Items

- Recall that Under the parallel-items model:
  - the factor loadings constrained to be equal

$$\lambda_1 = \lambda_2 = \dots = \lambda_j$$

- the unique variances are constrained to be equal:

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

- In this case, omega for a test is the same as a multiple for omega for a single item.
  - This is because the covariance for each pair of items is the same across all pairs



# The Spearman-Brown Prophecy Formula

- Specifically, the reliability for a single item is:

$$p_1 = \frac{\sigma_{T_x}^2}{\sigma_{T_x}^2 + \sigma_{E_x}^2}$$

- The reliability for the whole test is then:

$$p_m = \frac{m\sigma_{T_x}^2}{m\sigma_{T_x}^2 + \Psi^2}$$

- The final formula is the Spearman-Brown Prophecy formula.
  - It was designed to prophecy the increase in reliability by doubling the length of a test.



# S-B and Omega

- We can show that Omega happens to be equal to the S-B reliability under the parallel-items model.
  - The S-B formula reported from time to time in the context of building test with split-half methods.
- Really, there is not a good reason for using the S-B formula over Alpha or Omega (p. 95).
  - It is presented here for historical reasons.



# S-B from the SWLS

- To demonstrate the S-B from the SWLS, recall the estimates of our parameters:

- $\lambda = 1.287$

- $\Psi^2 = 1.442$

- The formula for coefficient Omega:

$$[\Sigma (\lambda_j)]^2 / ([\Sigma (\lambda_j)]^2 + \Sigma \Psi_j^2) =$$

$$m^2 \lambda^2 / (m^2 \lambda^2 + \Sigma \Psi_j^2) = 0.8169$$



## Another Example – The Hope Scale



# The Hope Scale

- Supplied by Hal Shorey, the Hope Scale consists of 18 items measuring three attributes.
    - We will focus on the Pathways Attribute, as measured by the items to the right.
1. I'm not good at planning how to get things done.\*
  2. I can think of many ways to get out of a jam.
  3. I have difficulty finding ways to solve problems.\*
  4. I'm not good at coming up with solutions.\*
  5. I'm good at coming up with new ways to solve problems.
  6. I create alternate plans when blocked.



# Is The Test Homogenous?

```
PROC IMPORT OUT= WORK.hope
            DATAFILE= "C:\Documents and Settings\Jonat
            DBMS=EXCEL2000 REPLACE;
            GETNAMES=YES;
RUN;

*single factor for pathways;
proc calis cov data=work.hope residual method=uls;
lineqs
    h4=lambda4 f1 + e4,
    h5=lambda5 f1 + e5,
    h9=lambda9 f1 + e9,
    h12=lambda12 f1 + e12,
    h13=lambda13 f1 + e13,
    h15=lambda15 f1 + e15;
std
    f1=1,
    e4=psi4,
    e5=psi5,
    e9=psi9,
    e12=psi12,
    e13=psi13,
    e15=psi15;
run;
```

The SAS System

21:09 Monday,

## The CALIS Procedure Covariance Structure Analysis: Least-Squares Estimation

Fit Function	0.1651
Goodness of Fit Index (GFI)	0.9919
GFI Adjusted for Degrees of Freedom (AGFI)	0.9810
Root Mean Square Residual (RMR)	0.0887
Parsimonious GFI (Mulaik, 1989)	0.5951



# What is Omega?

## The CALIS Procedure Covariance Structure Analysis: Least-Squares Estimation

### Manifest Variable Equations with Estimates

H4	=	0.9171*f1	+	1.0000 e4
		lambda4		
H5	=	0.6234*f1	+	1.0000 e5
		lambda5		
H9	=	0.9728*f1	+	1.0000 e9
		lambda9		
H12	=	0.9897*f1	+	1.0000 e12
		lambda12		
H13	=	0.7723*f1	+	1.0000 e13
		lambda13		
H15	=	0.7178*f1	+	1.0000 e15
		lambda15		

### Variances of Exogenous Variables

Variable	Parameter	Estimate
f1		1.00000
e4	psi4	2.21130
e5	psi5	1.45741
e9	psi9	1.15605
e12	psi12	0.75474
e13	psi13	1.26163
e15	psi15	0.78997





# What is Alpha?

A	B	C	D	E	F	G	H	
_TYPE_	_NAME_	H4	H5	H9	H12	H13	H15	
COV	H4	3.052461	0.343256	0.939549	0.861857	0.770751	0.788631	
COV	H5	0.343256	1.846086	0.654012	0.65766	0.594421	0.497503	
COV	H9	0.939549	0.654012	2.102341	1.0889	0.623994	0.559452	
COV	H12	0.861857	0.65766	1.0889	1.734173	0.701174	0.630569	
COV	H13	0.770751	0.594421	0.623994	0.701174	1.858153	0.63627	
COV	H15	0.788631	0.497503	0.559452	0.630569	0.63627	1.305166	

Sum of diagonal elements: 11.898

Sum of matrix: 32.594



# Wrapping Up

- Reliability of a test can be estimated from a test itself.
  - The single-factor model has Omega as its estimate of reliability.
  - Omega equals Alpha under the true-score equivalence Model.
- The S-B formula is Omega under the parallel-items model
  - Because of the strict assumptions of the model, we cannot think this will be a good estimate of the reliability.
- Binary items will be dealt with in Chapter 11.



# Next Time

- Midterm review.
  - Bring your questions.