Slide 1

Aspects of Structural Equation Modeling

- **Substantive Theory**
  - Concepts
  - Constructs
  - Formalization
- **Basic Issues**
  - Causality
  - Model building: theory vs data driven
  - Exploratory vs confirmatory analysis
  - Units of measurement and standardization
  - Scale types

Slide 2

- **Statistical Issues**
  - Specification of the model
  - Identification of models and parameters
  - Estimation of models and parameters
  - Fitting and testing of models
    - Assessment of fit
    - Detection of lack of fit
  - Testing structural hypotheses
- **Computational Aspects**

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 3

Two Programs: PRELIS and LISREL

Typical Uses

Use PRELIS for

- Data screening
- Data exploration
- Data summarization
- Computation of
  - Covariance matrix
  - Correlation matrix
  - Moment matrix
  - Asymptotic covariance matrix
  - PRELIS system file (PSF file)

Slide 4

Use LISREL for

- Model building
- Precise estimation and assessment of fit
- Many standard and non-standard models can be fitted and tested

Slide 5

PRELIS

Variables

- Ordinal (up to 15 categories)
- Continuous – normal or non-normal
- Censored – above, below or both
- Mixtures
- Nominal
- Fixed or Covariates

Slide 6

Data

- Cross-sectional
- Longitudinal
- Single sample
- Multiple samples
- Case-weighted

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 7

Data transformations

- Linear
- Normal scores
- Power
- Logarithmic
- Recoding
- Selection of cases
- Selection of variables
- Deletion of variables

Slide 8

Missing Values

- Listwise deletion
- Pairwise deletion
- Imputation by matching
- Multiple imputation

Slide 9

Matrices

CM Covariance matrix

MM Moment matrix

AM Augmented moment matrix

KM Matrix of product moment (Pearson) correlations

OM Matrix of canonical correlations or correlations of optimal scores

PM Matrix of polychoric and polyserial correlations

RM Matrix of Spearman rank correlations

TM Matrix of Kendall Tau-c correlations

AC Asymptotic covariance matrix

Slide 10

Features

- Fast and efficient data screening and summarization
- Analysis of all variables jointly
- No limit on sample size
- Regression equations by OLS and TSLS
- Exploratory factor analysis with factor scores
- Principal components with component scores

Slide 11

- Treatment of ordinal variables
  - Estimation of thresholds
  - Test of equal thresholds
  - Test of underlying bivariate normality
- Univariate and multivariate probit regression
- Multilevel analysis
- Formal inference-based recursive modeling (FIRM)

Slide 12

LISREL

Models

- Measurement models
- Exploratory and confirmatory factor analysis
- Multitrait-multimethod matrices
- Path analysis
- Correlated errors
- Panel models
- Univariate and multivariate regression
- ANOVA, ANCOVA, MANOVA, and MANCOVA
- MIMIC models

Slide 13

- Growth curves
- Latent growth curves
- Multilevel SEM models
- Econometric models
- Seemingly unrelated relations
- Mean structures
- General covariance structures
- ANCOVA with fallible covariates

Slide 14

Methods

IV Instrumental variables

TSLS Two-stage least-squares

ULS Unweighted least-squares

GLS Generalized least-squares

ML Maximum likelihood

WLS Weighted least-squares

DWLS Diagonally weighted least-squares

RO Ridge option

Slide 15

Features

- Parameter estimates, standard errors, and t-values
- Residuals and standardized residuals
- Stemleaf plots of residuals and standardized residuals
- Q-plot of standardized residuals
- Modification indices and estimated parameter change
- Parameter plot of concentrated fit function
- Multiple goodness-of-fit measures
- Test of close fit
- Standardized and completely standardized solutions
- Indirect and total effects and their standard errors and t-values
- Admissibility test of the model
- Test of model identification
- Full information ML of model with data missing at random

Slide 16

Books on Structural Equation Modeling

Jöreskog, K.G., & Sörbom, D. (1979) *Advances in Factor Analysis and Structural Equation Models*. Abt Books.

Dwyer, J.H. (1983) *Statistical Models for the Social and Behavioral Sciences*. Oxford University Press.

Long, J.S. (1983a) *Confirmatory Factor Analysis: A Preface to* LISREL. Sage Publications.

Long, J.S. (1983b) *Covariance Structure Models: An Introduction to* LISREL. Sage Publications.

Saris, W.E & Stronkhorst, L.H. (1984) *Causal Modelling in Non-Experimental Research*. Sociometric Research Foundation.

Hayduk, L.A. (1987) *Structural Equation Modelling with* LISREL*: Essentials and Advances*. The Johns Hopkins University Press.

Slide 17

Jöreskog, K.G. & Sörbom, D. (1989) LISREL 7 - *A guide to the program and applications.* Second Edition. SPSS Publications.

Bollen, K.A. (1989) *Structural Equations with Latent Variables*. Wiley.

Bollen, K.A. & Long, J.S. (Editors) (1993) *Testing Structural Equation Models*. Sage Publications.

Marcoulides, G.A. & Schumacker, R.E. (Editors) (1996) *Advanced Structural Equation Modeling: Issues and Techniques*. Lawrence Erlbaum Publishers.

Mueller, R.D. (1996) *Basic Principles of Structural Equation Modeling: An Introduction to LISREL and EQS*. Springer Verlag.

Slide 18

Schumacker, R.E. & Lomax, R.G. (1996) *A Beginner's Guide to Structural Equation Modeling*. Lawrence Erlbaum Publishers.

Hayduk, L.A. (1996) *LISREL Issues, Debates, and Strategies*. The Johns Hopkins University Press.

Kline, R.B. (1998) *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.

Maruyama, G.M. (1998) *Basics of Structural Equation Modeling*. Thousand Oaks, CA: Sage Publications.

Schumacker, R.E. & Marcoulides, G.A. (Eds) (1998) *Interaction and Nonlinear Effects in Structural Equation Modeling*. Mahwah, N.J: Lawrence Erlbaum Associates, Publishers.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 19

Byrne, B.M. (1998) *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications and Programming*. Mahwah, N J: Lawrence Erlbaum Associates, Publishers.

Loehlin, J.C. (1998) *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*. Third Edition. Mahwah, N J: Lawrence Erlbaum Associates, Publishers.

Raykov, T. & Marcoulides, G.A. (2000) *A First Course in Structural Equation Modeling*. Mahwah, N J: Lawrence Erlbaum Associates, Publishers.

Kaplan, D. (2000) *Structural Equation Modeling: Foundations and Extensions* Thousand Oaks, CA: Sage Publications.

Marcoulides, G.A. & Schumacker, R.E. (Editors) (2001) *New Developments and Techniques in Structural Equation Modeling*. Lawrence Erlbaum Publishers.

Slide 20

Books on LISREL published by SSI

Jöreskog, K.G. & Sörbom, D. (1999a) PRELIS 2 *User's Reference Guide*. Chicago: Scientific Software International.

Jöreskog, K.G. & Sörbom, D. (1999b) LISREL 8*: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International.

Jöreskog, K.G. & Sörbom, D. (1999c) LISREL 8 *User's Reference Guide*. Chicago: Scientific Software International.

Jöreskog, K.G., Sörbom, D., Du Toit, S., & Du Toit M. (2000) LISREL 8*: New Statistical Features*. Second printing with revisions. Chicago: Scientific Software International.

Slide 21

Cudeck, R., Du Toit, S., & Sörbom, D. (Eds) (2001) *Structural Equation Modeling: Present and Future – A Festschrift in honor of Karl G. Jöreskog.* Chicago: Scientific Software International.

Du Toit, M. & Du Toit, S. (2001) *Interactive LISREL User's Guide.* Chicago: Scientific Software International.

Slide 22

Journal

Structural Equation Modeling: A Multidisciplinary Journal. Editor: George A.Marcoulides. Publisher: Lawrence Erlbaum. 1993 -

SEMNET

To subscribe, send an email message, containing the single line

"subscribe - your first name your last name"

to listserv@bama.ua.edu

To post a message, send your message to: semnet@bama.ua.edu

For latest information on LISREL, see `www.ssicentral.com`

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 23

Availability of LISREL

The current version of LISREL is 8.50

LISREL is available from

```
Scientific Software International
7383 North Lincoln Ave, Suite 100
LINCOLNWOOD, IL 60712-1704
USA

Phone: +1 847 675 0720 (sales)
Fax:   +1 847 675-2140

E-mail: info@ssicentral.com
URL: http://www.ssicentral.com
```

Slide 24

```
In Europe contact

ProGamma
P.O.Box 841
9700 AV Groningen
The Netherlands

Phone: +31 50 5771811
Fax:   +31 50 5778831

E-mail: gamma.post@gamma.rug.nl
URL: http://www.gamma.rug.nl
```

Slide 25

Ways of Running LISREL

- Point and click interface
- Syntax command files using
  - PRELIS syntax
  - LISREL syntax
  - SIMPLIS syntax

Slide 26

LISREL Syntax

- Uses Greek notation
- Uses matrices
- Uses mnemonic keywords with default values
- Requires some knowledge of matrix algebra

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 27

SIMPLIS Syntax

- SIMPLIS is a free form English command language
- Name all observed and latent (if any) variables
- Formulate the model to be estimated by paths or relationships (equations)
- It is not necessary to be familiar with the LISREL model or any of its submodels.
- No Greek or matrix notations are required.
- There are no complicated options to learn.
- Anyone who can formulate the model as a path diagram can use the SIMPLIS command language immediately.

Slide 28

Stages of Data Analysis

- Data Management
- Data Exploration
- Data Processing
- Data Modeling
  - Theory
  - Model
  - Estimation
  - Testing

Slide 29

Two Basic Problems

Broadly speaking, there are two basic problems that are important in the social and behavioral sciences. The first problem is concerned with the measurement properties—validities and reliabilities—of the measurement instruments. The second problem concerns the causal relationships among the variables and their relative explanatory power.

Slide 30

Measurement Error

Most measurements used in the behavioral and social sciences contain sizable measurement errors, which, if not taken into account, can cause severe bias in results. Adequate modelling should take measurement errors into account whenever possible.

Measurement errors occur because of imperfections in measurement instruments (questionnaires, interviews, tests, etc) and measuring procedures (recording, coding, scaling, grouping, aggregation, etc).

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 31

Measurement Models

Most theories and models in the social and behavioral sciences are formulated in terms of theoretical or hypothetical concepts, or constructs, or latent variables, which are not directly measurable or observable. Examples of such constructs are prejudice, radicalism, alienation, conservatism, trust, self-esteem, discrimination, motivation, ability, and anomia. The measurement of a hypothetical construct is accomplished indirectly through one or more observable indicators, such as responses to questionnaire items, that are *assumed* to represent the construct adequately.

Slide 32

The purpose of a measurement model is to describe how well the observed indicators serve as a measurement instrument for the latent variables. The key concepts here are measurement, reliability, and validity. Measurement models often suggest ways in which the observed measurements can be improved.

Measurement models are important in the social and behavioral sciences when one tries to measure such abstractions as people's behavior, attitudes, feelings and motivations. Most measures employed for such purposes contain sizable measurement errors and the measurement models allow us to take these errors into account.

Slide 33

Reliability and Validity

Reliability

The reliability of a measure measures how well the measure measures whatever it measures.

Validity

The validity of a measure measures how well the measure measures what it is supposed to measure.

Slide 34

Latent Variables and Indicators

Psychology

Intelligence	
Visual Perception	
Verbal Ability	Spelling Punctuation Vocabulary Word Fluency Reading Writing

Slide 35

Education

Academic Performance	
Science Achievement in Grade 9	
Pre-school Verbal Stimulation in the Home	Children Books Library Visits TV Watching Correcting Speach ... ...

Slide 36

Sociology

Socio-Economic Status	
Ambition	
Occupational Aspiration	
Attitudes towards Sex-typing	Shoveling Snow Cleaning the House Washing the Car Making Beds

Slide 37

Economics

Income	Register
	Questionnaire
	Interview by Phone
	Interview by Visit
Economic Expectation	
Economic Activity	
Permanent Income	Income
	Consumption

Slide 38

Political Science

Industrial Development	GNP
	Energy Consumption
	Labor Diversification
Political Development	Executive Functioning
	Legislative Functioning
	Party Organization
	Power Diversification
	Citizen Influence
Political Efficacy	NOSAY
	VOTING
	COMPLEX
	NOCARE

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

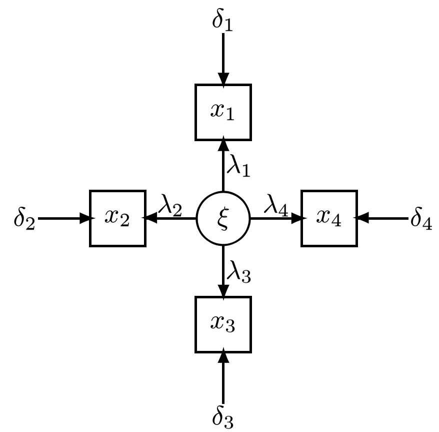
Slide 39

Other Latent Variables

Ability	Mood
Achievement	Motivation
Alienation	Perception
Ambition	Performance
Anomia	Prejudice
Aptitude	Radicalism
Aspiration	Satisfaction
Attitude	Self-esteem
Behavior	Trust
Conservatism	
Discrimination	
Feelings	
Health	

Slide 40

The Congeneric Measurement Model



Slide 41

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} \xi + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}$$

$$\mathbf{x} = \mathbf{\Lambda}_x \xi + \boldsymbol{\delta} \; .$$

$$\mathbf{\Sigma} = \mathbf{\Lambda}_x \mathbf{\Lambda}_x{}' + \mathbf{\Theta}_\delta = \begin{pmatrix} \lambda_1^2 + \theta_{11} & & & \\ \lambda_2\lambda_1 & \lambda_2^2 + \theta_{22} & & \\ \lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2 + \theta_{33} & \\ \lambda_4\lambda_1 & \lambda_4\lambda_2 & \lambda_4\lambda_3 & \lambda_4^2 + \theta_{44} \end{pmatrix}$$

Slide 42

Tau-equivalent measures

$$\lambda_1^2 = \cdots = \lambda_4^2 \; .$$

Parallel measures

$$\lambda_1^2 = \cdots = \lambda_4^2 \qquad \theta_{11} = \cdots = \theta_{44} \; .$$

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 43

Scaling the Latent Variables

Latent variables are unobservable and have no definite scales. Both the origin and the unit of measurement in each latent variable are arbitrary. To define the model properly, the origin and the unit of measurement of each latent variable must be defined.

Typically, there are two ways in which this is done. The most useful and convenient way of assigning the units of measurement of the latent variables is to assume that they are standardized so that they have zero means and unit variances in the population.

Slide 44

Another way to assign a unit of measurement for a latent variable, is to fix a non-zero coefficient (usually one) in the relationship for one of its observed indicators. This defines the unit for each latent variable in relation to one of the observed variables, a so-called *reference variable*. In practice, one chooses as reference variable the observed variable, which, in some sense, best represents the latent variable.

Slide 45

LISREL 8 defines the units of latent variables as follows

- If a reference variable is assigned to a latent variable by the specification of a fixed non-zero coefficient in the relationship between the reference variable and the latent variable, then this defines the scale for that latent variable.
- If no reference variable is specified for a latent variable, by assigning a fixed non-zero coefficient for an observed variable, the program will standardize this latent variable.

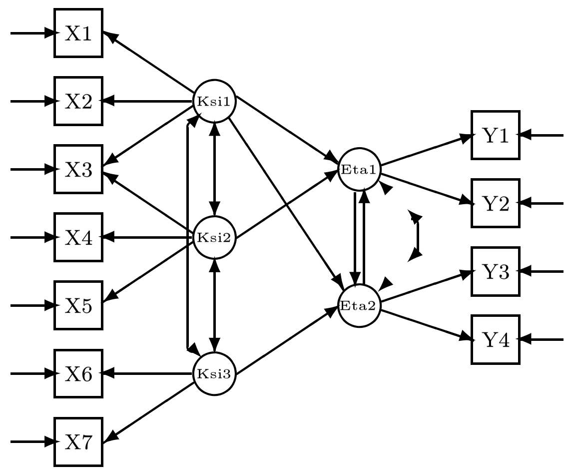
Slide 46

Further considerations are necessary when the observed variables in the model are ordinal, because then, even the observed variables do not have any units of measurement. Other considerations are necessary in longitudinal and multiple group studies in order to put the variables on the same scale at different occasions and in different groups. In such studies one can also relax the assumption of zero means for the latent variables.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

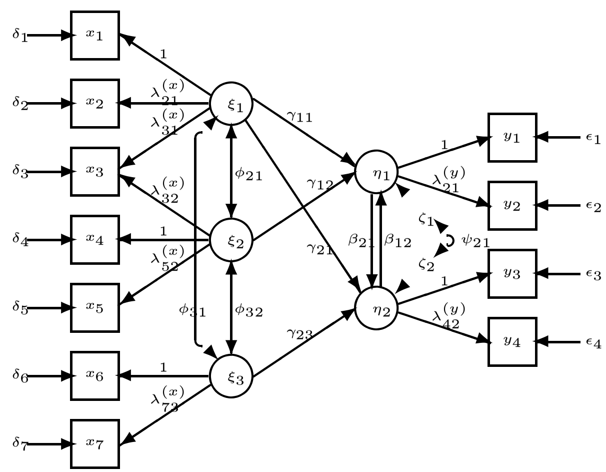
Slide 47

The LISREL Model in SIMPLIS Notation



Slide 48

The LISREL Model in LISREL Notation



Slide 49

The *structural equations* are

$$\begin{aligned}\eta_1 &= \beta_{12}\eta_2 + \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \zeta_1 \\ \eta_2 &= \beta_{21}\eta_1 + \gamma_{21}\xi_1 + \gamma_{23}\xi_3 + \zeta_2\end{aligned}$$

$$\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} = \begin{pmatrix}0 & \beta_{12}\\ \beta_{21} & 0\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} + \begin{pmatrix}\gamma_{11} & \gamma_{12} & 0\\ \gamma_{21} & 0 & \gamma_{23}\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix} + \begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix}$$

The *measurement model* equations for y-variables are

$$\begin{aligned}y_1 &= \eta_1 + \epsilon_1 \\ y_2 &= \lambda_{21}^{(y)}\eta_1 + \epsilon_2 \\ y_3 &= \eta_2 + \epsilon_3 \\ y_4 &= \lambda_{42}^{(y)}\eta_2 + \epsilon_4\end{aligned}$$

Slide 50

$$\begin{pmatrix}y_1\\ y_2\\ y_3\\ y_4\end{pmatrix} = \begin{pmatrix}1 & 0\\ \lambda_{21}^{(y)} & 0\\ 0 & 1\\ 0 & \lambda_{42}^{(y)}\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} + \begin{pmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\\ \epsilon_4\end{pmatrix} .$$

The *measurement model* equations for x-variables are

$$\begin{aligned}x_1 &= \xi_1 + \delta_1 \\ x_2 &= \lambda_{21}^{(x)}\xi_1 + \delta_2 \\ x_3 &= \lambda_{31}^{(x)}\xi_1 + \lambda_{32}^{(x)}\xi_2 + \delta_3 \\ x_4 &= \xi_2 + \delta_4 \\ x_5 &= \lambda_{52}^{(x)}\xi_2 + \delta_5 \\ x_6 &= \xi_3 + \delta_6 \\ x_7 &= \lambda_{73}^{(x)}\xi_3 + \delta_7\end{aligned}$$

Slide 51

$$\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ \lambda_{21}^{(x)} & 0 & 0\\ \lambda_{31}^{(x)} & \lambda_{32}^{(x)} & 0\\ 0 & 1 & 0\\ 0 & \lambda_{52}^{(x)} & 0\\ 0 & 0 & 1\\ 0 & 0 & \lambda_{73}^{(x)}\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix} + \begin{pmatrix}\delta_1\\ \delta_2\\ \delta_3\\ \delta_4\\ \delta_5\\ \delta_6\\ \delta_7\end{pmatrix} .$$

All these equations correspond to the path diagram in Slide 48 One λ in each column of $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ has been set equal to 1 to fix the scales of measurement in the latent variables.

Slide 52

The covariance matrix of $\boldsymbol{\xi}$,

$$\mathbf{\Phi} = \begin{pmatrix}\phi_{11} & & \\ \phi_{21} & \phi_{22} & \\ \phi_{31} & \phi_{32} & \phi_{33}\end{pmatrix} ,$$

The covariance matrix of $\boldsymbol{\zeta}$,

$$\mathbf{\Psi} = \begin{pmatrix}\psi_{11} & \\ \psi_{21} & \psi_{22}\end{pmatrix} ,$$

The covariance matrix of $\boldsymbol{\epsilon}$, a diagonal matrix,

$$\mathbf{\Theta}_\epsilon = \mathsf{diag}(\theta_{11}^{(\epsilon)}, \theta_{22}^{(\epsilon)}, \ldots, \theta_{44}^{(\epsilon)}) ,$$

The covariance matrix of $\boldsymbol{\delta}$, a diagonal matrix,

$$\mathbf{\Theta}_\delta = \mathsf{diag}(\theta_{11}^{(\delta)}, \theta_{22}^{(\delta)}, \ldots, \theta_{77}^{(\delta)}) .$$

Slide 53

The LISREL Model with Means

$$\mathbf{y} = \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon}$$

$$\mathbf{x} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

$\mathbf{y}, \mathbf{x}$ = Observed Variables $\boldsymbol{\eta}, \boldsymbol{\xi}$ = Latent Variables

$\boldsymbol{\epsilon}, \boldsymbol{\delta}$ = Measurement Errors $\boldsymbol{\zeta}$ = Structural Errors

$\boldsymbol{\tau}_y, \boldsymbol{\tau}_x, \boldsymbol{\alpha}$ = Intercept Terms

$\boldsymbol{\Lambda}_y, \boldsymbol{\Lambda}_x, \mathbf{B}, \boldsymbol{\Gamma}$ = Parameter Matrices

Slide 54

Assumptions

- $\boldsymbol{\epsilon}$ is uncorrelated with $\boldsymbol{\eta}$
- $\boldsymbol{\delta}$ is uncorrelated with $\boldsymbol{\xi}$
- $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\xi}$
- $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\epsilon}$ and $\boldsymbol{\delta}$
- $\mathbf{I} - \mathbf{B}$ is non-singular

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 55

LISREL as a Mean and Covariance Structure

Let

$$\boldsymbol{\kappa} = E(\boldsymbol{\xi}) \quad \boldsymbol{\Phi} = Cov(\boldsymbol{\xi})$$

$$\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\Theta}_\epsilon & \boldsymbol{\Theta}'_{\delta\epsilon} \\ \boldsymbol{\Theta}_{\delta\epsilon} & \boldsymbol{\Theta}_\delta \end{pmatrix} = Cov \begin{pmatrix} \boldsymbol{\epsilon} \\ \boldsymbol{\delta} \end{pmatrix}$$

Then it follows that the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{z} = (\mathbf{y}', \mathbf{x}')'$ are

Slide 56

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\kappa}) \\ \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x \boldsymbol{\kappa} \end{pmatrix} ,$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Lambda}_y \mathbf{A} (\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}' + \boldsymbol{\Psi}) \mathbf{A}' \boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_\epsilon & \boldsymbol{\Lambda}_y \mathbf{A} \boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Lambda}'_x + \boldsymbol{\Theta}'_{\delta\epsilon} \\ \boldsymbol{\Lambda}_x \boldsymbol{\Phi}\boldsymbol{\Gamma}' \mathbf{A}' \boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_{\delta\epsilon} & \boldsymbol{\Lambda}_x \boldsymbol{\Phi} \boldsymbol{\Lambda}'_x + \boldsymbol{\Theta}_\delta \end{pmatrix} ,$$

where $\mathbf{A} = (\mathbf{I} - \mathbf{B})^{-1}$.

The elements of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are functions of the elements of $\boldsymbol{\kappa}$, $\boldsymbol{\alpha}$, $\boldsymbol{\tau}_y$, $\boldsymbol{\tau}_x$, $\boldsymbol{\Lambda}_y$, $\boldsymbol{\Lambda}_x$, $\mathbf{B}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Theta}_\epsilon$, $\boldsymbol{\Theta}_\delta$, and $\boldsymbol{\Theta}_{\delta\epsilon}$ which are of three kinds:

- *fixed parameters* that have been assigned specified values,
- *constrained parameters* that are unknown but linear or non-linear functions of one or more other parameters, and
- *free parameters* that are unknown and not constrained.

Slide 57

Statistical Principles of Estimation

Maximum Likelihood

Given a specified functional form $p(x, \theta)$ of the distribution of observations x, choose parameters θ which maximizes the likelihood of the actual observations x.

Example:

$$x_1, x_2, \ldots, x_n \sim iid\ N(\mu, 1)$$

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\mu)^2}$$

$$\ln L = const - \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2$$

Slide 58

Maximizing $\ln L$ is equivalent to minimizing

$$\sum_{i=1}^{n} (x_i - \mu)^2$$

which gives

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}$$

Least Squares

Choose parameters which minimizes the (weighted) sum of squares between observed and expected (predicted) quantities.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 59

Fitting $\Sigma(\boldsymbol{\theta})$ to S

Sample Covariance Matrix: $\mathbf{S}$
Model: $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ as functions of parameters $\boldsymbol{\theta}$

How do we fit $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ to $\mathbf{S}$? Choose $\boldsymbol{\theta}$ so that $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is as close as possible to $\mathbf{S}$, in some sense, i.e., choose a fit function $F = F(\mathbf{S}, \boldsymbol{\Sigma})$ to be minimized with respect to $\boldsymbol{\theta}$

Slide 60

Three Specific Fit Functions

Unweighted Least Squares(ULS)

$$F = \frac{1}{2} tr[(\mathbf{S} - \boldsymbol{\Sigma})^2]$$

Generalized Least Squares(GLS)

$$F = \frac{1}{2} tr[(\mathbf{I} - \mathbf{S}^{-1}\boldsymbol{\Sigma})^2]$$

Maximum Likelihood(ML)

$$F = \log \|\boldsymbol{\Sigma}\| + tr(\mathbf{S}\boldsymbol{\Sigma}^{-1}) - \log \|\mathbf{S}\| - (p + q)$$

Slide 61

Estimation Methods

A LISREL model may be estimated by seven different methods:

IV Instrumental Variables

TSLS two-stage least squares

ULS unweighted least squares

GLS generalized least squares

ML maximum likelihood

WLS weighted least squares

DWLS diagonally weighted least squares

Under general assumptions, all methods give consistent estimates of parameters.

Slide 62

- TSLS and IV are non-iterative and very fast. They are used to compute starting values for the other methods but can also be requested as final estimates.
- ULS, GLS, ML, WLS, and DWLS estimates are obtained by an iterative procedure that minimizes a particular fit function.
- WLS requires an estimate of the asymptotic covariance matrix of the sample variances and covariances or correlations being analyzed.
- Similarly, DWLS requires an estimate of the asymptotic variances of the sample variances and covariances or correlations being analyzed.
- These asymptotic variances and covariances are obtained by PRELIS which saves them in a binary file to be read by LISREL.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 63

General Fit Function

Let $\mathbf{z} = (\mathbf{y}', \mathbf{x}')'$ be a vector of the observed variables in the model and let $\boldsymbol{\theta}$ be a vector of all free and independent parameters of the model. Then the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{z}$ are functions of $\boldsymbol{\theta}$, see Slide 56. Let $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N$ be N independent observations of the vector $\mathbf{z}$ and let $\bar{\mathbf{z}}$ and $\mathbf{S}$ be the sample mean vector and covariance matrix. Parameter estimates are obtained by minimizing some fit function

$$\mathsf{F}(\boldsymbol{\theta}) = \mathsf{F}(\bar{\mathbf{z}}, \mathbf{S}, \boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

of $\boldsymbol{\mu}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. Only the most important ones are considered here.

Slide 64

Maximum Likelihood

The maximum likelihood (ML) approach will estimate $\boldsymbol{\theta}$ by minimizing the fit function

$$\mathsf{F}(\boldsymbol{\theta}) = \log||\boldsymbol{\Sigma}|| + \mathsf{tr}(\mathbf{S}\boldsymbol{\Sigma}^{-1}) - \log||\mathbf{S}|| - k + (\bar{\mathbf{z}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{z}} - \boldsymbol{\mu}) \, .$$

where k is the number of variables in $\mathbf{z}$.

This fit function assumes that the observed variables $\mathbf{z}$ have a multinormal distribution.

Slide 65

Weighted Least Squares

The weighted least squares (WLS) approach will estimate $\boldsymbol{\theta}$ by minimizing the fit function

$$\mathsf{F}(\boldsymbol{\theta}) = (\mathbf{s} - \boldsymbol{\sigma})'\mathbf{W}^{-1}(\mathbf{s} - \boldsymbol{\sigma}) + (\bar{\mathbf{z}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{z}} - \boldsymbol{\mu}) \ ,$$

where $\mathbf{s}' = (s_{11}, s_{21}, s_{22}, s_{31}, \ldots, s_{kk})$, $\boldsymbol{\sigma}' = (\sigma_{11}, \sigma_{21}, \sigma_{22}, \sigma_{31}, \ldots, \sigma_{kk})$, and $\mathbf{W}$ is a symmetric positive definite matrix. The usual way of choosing $\mathbf{W}$ in weighted least squares is to let $\mathbf{W}$ be a consistent estimate of the asymptotic covariance matrix of $\mathbf{s}$.

WLS does not assume multivariate normality of $\mathbf{z}$ but it does assume that $\bar{\mathbf{z}}$ and $\mathbf{S}$ are asymptotically independent.

Slide 66

Weighted Least Squares based on Augmented Moment Matrix

To avoid the mentioned problems associated with ML and WLS, one can use the augmented moment matrix

$$\mathbf{A} = (1/N)\sum_{c=1}^{N} \begin{pmatrix} \mathbf{z}_c \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}'_c & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{S} + \bar{\mathbf{z}}\bar{\mathbf{z}}' & \\ \bar{\mathbf{z}}' & 1 \end{pmatrix} .$$

This is the matrix of sample moments about zero for the vector $\mathbf{z}$ augmented with a variable which is constant equal to one for every case.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 67

The corresponding population matrix is

$$(\alpha_{ij}) = \mathsf{E} \begin{pmatrix} \mathbf{z} \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}' & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}' & \\ \boldsymbol{\mu}' & 1 \end{pmatrix} .$$

Note that the last element in these matrices is a fixed constant equal to 1.

Slide 68

Let

$$\begin{aligned} \mathbf{a}' &= (a_{11}, a_{21}, a_{22}, a_{31}, \ldots, a_{k+1,k}, 1) \\ \boldsymbol{\alpha}' &= (\alpha_{11}, \alpha_{21}, \alpha_{22}, \alpha_{31}, \ldots, \alpha_{k+1,k}, 1) \ . \end{aligned}$$

Applying weighted least squares to $\mathbf{A}$ instead of $\mathbf{S}$ gives another fit function (WLSA):

$$\mathsf{F}(\boldsymbol{\theta}) = (\mathbf{a} - \boldsymbol{\alpha})'\mathbf{W}_a^{-}(\mathbf{a} - \boldsymbol{\alpha}) \ ,$$

where $\mathbf{W}_a$ is a consistent estimate of the covariance matrix of $\mathbf{a}$ and $\mathbf{W}_a^{-}$ is a Moore-Penrose generalized inverse of $\mathbf{W}_a$. Note that since the last element in $\mathbf{a}$ is a fixed constant, the last row of $\mathbf{W}_a$ is zero. Hence, $\mathbf{W}_a$ is singular.

Slide 69

Nature of Inference

- Once a model has been carefully formulated, it can be confronted with empirical data and, if all assumptions hold, LISREL can be used to test if the model is consistent with the data.
- The inference problem is of the following general form. Given assumptions A, B, C, . . ., test model M against a more general model M_A. Most researchers will take assumptions A, B, C, . . ., for granted and proceed to test M formally. But the assumptions should be checked before the test is made, whenever this is possible.

Slide 70

- If the model is rejected by the data, the problem is to determine what is wrong with the model and how the model should be modified to fit the data better.
- If the model *fits* the data, it does not mean that it is *the "correct" model or even the "best" model*. In fact, there can be many *equivalent* models, all of which will fit the data equally well as judged by any goodness-of-fit measure. The direction of causation and the causal ordering of the constructs cannot be determined by the data[a]. To conclude that the fitted model is the "best," one must be able to exclude all models equivalent to it on logical or substantive grounds.

[a]Different equivalent models will give different parameter estimates, and some may give estimates that are not meaningful. This fact may be used to distinguish some equivalent models from others.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 71

- Even outside the class of models equivalent to the given model, there may be many models that fit the data almost as well as the given model. SEM techniques can be used to distinguish sharply between models that fit the data *very badly* and those that fit the data *reasonably well.* To discriminate further between models that fit the data almost equally well requires a careful power study.

Slide 72

Three Situations

`SC` In a *strictly confirmatory* situation the researcher has formulated one single model and has obtained empirical data to test it. The model should be accepted or rejected.

`AM` The researcher has specified several *alternative models* or *competing models* and on the basis of an analysis of a single set of empirical data, one of the models should be selected.

Slide 73

`MG` The researcher has specified a tentative initial model. If the initial model does not fit the given data, the model should be modified and tested again using the same data. Several models may be tested in this process. The goal may be to find a model which not only fits the data well from a statistical point of view, but also has the property that every parameter of the model can be given a substantively meaningful interpretation. The re-specification of each model may be theory-driven or data-driven. Although a model may be tested in each round, the whole approach is *model generating* rather than model testing.

Slide 74

In practice, the `MG` situation is by far the most common. The `SC` situation is very rare because few researchers are content with just rejecting a given model without suggesting an alternative model. The `AM` situation is also rare because researchers seldom specify the alternative models a priori.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 75

Testing the Model

- In the `SC` situation one can test the model by using a chi-square statistic which is $N-1$ times the minimum value of the fit function $F(\boldsymbol{\theta})$ used to estimate the model, where N is the sample.
- If the model holds, this is approximately distributed in large samples as χ^2 with $d = s - t$ degrees of freedom, where $s = k(k+1)/2$ and t is the number of independent parameters estimated. k is the number of observed variables in the model.

Slide 76

- To use this test formally, one chooses a significance level α and rejects the model if chi-square exceeds the $(1-\alpha)$ percentile of the χ^2 distribution with d degrees of freedom. For further information about this and other chi-square statistics, see *Standard Errors and Chi-squares* in the **Help** documentation.
- If the model does not hold, c is distributed as non-central χ^2 with $s-t$ degrees of freedom and non-centrality parameter λ which may be estimated as $\hat{\lambda} = \max\{(c-d), 0\}$. One can also set up a confidence interval for non-centrality parameter.

Slide 77

- Chi-square can be regarded as a chi-square statistic for testing the model against the alternative that the covariance matrix of the observed variables is unconstrained. This is valid if all assumptions are satisfied, if the model holds and the sample size is sufficiently large.
- In practice it is more useful to regard chi-square as a *measure* of fit rather than as a *test statistic*. In this view, chi-square is a measure of overall fit of the model to the data.
- Even if all the assumptions of the chi-square test hold, it may not be realistic to assume that the model holds exactly in the population. In this case, chi-square should be compared with a non-central rather than a central chi-square distribution.

Slide 78

Four Different Chi-squares

The four different chi-squares are denoted C1, C2, C3, C4. Which ones are obtained in different situations is seen in the following table, where • means obtained and ○ means not obtained (⋆ means that C2 will be obtained only if the asymptotic variances are provided).

Slide 79

Asymptotic Covariance Matrix not Provided

	ULS	GLS	ML	WLS	DWLS
C1	○	•	•	○	○
C2	•	•	•	○	⋆
C3	○	○	○	○	○
C4	○	○	○	○	○

Asymptotic Covariance Matrix Provided

	ULS	GLS	ML	WLS	DWLS
C1	○	•	•	•	○
C2	•	•	•	○	•
C3	•	•	•	○	•
C4	•	•	•	○	•

Slide 80

- C1 is $N-1$ times the minimum value of the fit function.
- C2 is $N-1$ times the minimum of the WLS fit function using a weight matrix estimated under multivariate normality.
- C3 is the Satorra-Bentler scaled chi-square statistic or its generalization to mean and covariance structures and multiple groups.
- C4 is $N-1$ times the minimum of the WLS fit function using a weight matrix estimated under multivariate non-normality (as provided by PRELIS).

Slide 81

Under multivariate normality of the observed variables, C1 and C2, whenever provided, are asymptotically equivalent and have an asymptotic chi-square distribution if the model holds exactly and an asymptotic non-central chi-square distribution if the model holds approximately. The same holds for C4 under the more general assumption that the observed variables have a multivariate distribution with finite moments up to order four. C3 is a correction to C2 which makes C3 have the correct asymptotic mean even under non-normality. For more information, see Chapter 4 and Appendix A of SSI's book *New Statistical Features*, see Slide 20.

Slide 82

The Four Chi-Squares in the Output

```
C1          Minimum Fit Function Chi-Square
C2  Normal Theory Weighted Least Squares Chi-Square
C3          Satorra-Bentler Scaled Chi-Square
C4        Chi-Square Corrected for Non-Normality
```

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 83

Some Formulas

Definitions

$$
\begin{aligned}
k &= \text{number of observed variables} \\
s &= \frac{1}{2}k(k+1) \\
t &= \text{number of independent parameters} < s \\
d &= s - t \\
\underset{k^2\times s}{\mathbf{K}} &= \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\text{, where } \underset{k^2\times s}{\mathbf{D}} \text{ is the duplication matrix} \\
\underset{s\times 1}{\mathbf{s}} &= \mathbf{K}'\mathsf{vec}(\mathbf{S}) \qquad \mathsf{vec}(\mathbf{S}) = \mathbf{D}\mathbf{s} \\
\underset{s\times s}{\boldsymbol{\Omega}} &= n\mathsf{ACov}(\mathbf{s}) \qquad n = N - 1
\end{aligned}
$$

Slide 84

$$
\begin{aligned}
\underset{s\times s}{\mathbf{W}} &= n\,\mathsf{Est}[\mathsf{ACov}(\mathbf{s})] \\
\mathbf{W}_{\mathsf{NT}} &= 2\mathbf{K}'(\hat{\boldsymbol{\Sigma}}\otimes\hat{\boldsymbol{\Sigma}})\mathbf{K} \text{ under } \mathsf{NT}, \\
\mathbf{W}_{\mathsf{NNT}} &= \text{computed by PRELIS under } \mathsf{NNT}, \\
\boldsymbol{\sigma} &= \underset{s\times 1}{\boldsymbol{\sigma}}(\underset{t\times 1}{\boldsymbol{\theta}}) \\
\underset{s\times t}{\boldsymbol{\Delta}} &= \frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{\theta}'} \quad \text{evaluated at } \hat{\boldsymbol{\theta}} \\
\underset{s\times d}{\boldsymbol{\Delta}_c} &= \text{orthogonal complement to } \boldsymbol{\Delta} \\
& \quad \boldsymbol{\Delta}_c'\boldsymbol{\Delta} = \mathbf{0} \qquad [\boldsymbol{\Delta}|\boldsymbol{\Delta}_c] \quad \text{non-singular}
\end{aligned}
$$

Slide 85

Fit Functions

$$\mathsf{F} = (\mathbf{s} - \boldsymbol{\sigma})' \underset{s\times s}{\mathbf{V}} (\mathbf{s} - \boldsymbol{\sigma}) \,,$$

where the weight matrix $\mathbf{V}$ is defined differently for different fit functions:

$$\begin{array}{lll}
\mathsf{ULS:} & \mathbf{V} = & \mathbf{I}^* = \mathsf{diag}\ (1,2,1,2,2,1,\ldots) \\
\mathsf{GLS:} & \mathbf{V} = & \mathbf{D}'(\mathbf{S}^{-1} \otimes \mathbf{S}^{-1})\mathbf{D} \\
\mathsf{ML:} & \mathbf{V} = & \mathbf{D}'(\hat{\mathbf{\Sigma}}^{-1} \otimes \hat{\mathbf{\Sigma}}^{-1})\mathbf{D} \\
\mathsf{WLS:} & \mathbf{V} = & \mathbf{W}_{\mathsf{NNT}}^{-1} \quad \text{or} \quad \mathbf{W}_{\mathsf{NNT}}^{-} \quad \text{if} \quad \mathbf{W}_{\mathsf{NNT}} \quad \text{singular} \\
\mathsf{DWLS:} & \mathbf{V} = & \mathbf{D}_{\mathsf{W}}^{-1} = [\mathsf{diag}\mathbf{W}]^{-1} \\
\text{with} & & \mathbf{W} = \mathbf{W}_{\mathsf{NT}} \quad \text{if AC not read} \\
 & & \mathbf{W} = \mathbf{W}_{\mathsf{NNT}} \quad \text{if AC read}
\end{array}$$

Slide 86

Results

$$\begin{array}{rcl}
\underset{s\times s}{\mathbf{E}} & = & \mathbf{\Delta}'\mathbf{V}\mathbf{\Delta} \\
n\mathsf{ACov}(\hat{\boldsymbol{\theta}}) & = & \mathbf{E}^{-1}\mathbf{\Delta}'\mathbf{V}\mathbf{W}\mathbf{V}\mathbf{\Delta}\mathbf{E}^{-1} \\
n\mathsf{ACov}(\mathbf{s} - \hat{\boldsymbol{\sigma}}) & = & \mathbf{W} - \mathbf{\Delta}\mathbf{E}^{-1}\mathbf{\Delta}' \\
\text{with } \mathbf{W} & = & \mathbf{W}_{\mathsf{NT}} \quad \text{if AC not read} \\
\mathbf{W} & = & \mathbf{W}_{\mathsf{NNT}} \quad \text{if AC read} \\
c_1 & = & n \ \mathsf{min}\ \mathsf{F} \quad \text{for ULS, GLS, WLS, DWLS} \\
 & = & n\ (\log||\hat{\mathbf{\Sigma}}|| + \mathsf{tr}(\mathbf{S}\hat{\mathbf{\Sigma}}^{-1}) - \log||\mathbf{S}|| - k) \quad \text{for ML} \\
c_2 & = & n(\mathbf{s} - \hat{\boldsymbol{\sigma}})'\mathbf{\Delta}_c(\mathbf{\Delta}_c'\mathbf{W}_{\mathsf{NT}}\mathbf{\Delta}_c)^{-1}\mathbf{\Delta}_c'(\mathbf{s} - \hat{\boldsymbol{\sigma}}) \\
h_1 & = & \mathsf{tr}[(\mathbf{\Delta}_c'\mathbf{W}_{\mathsf{NT}}\mathbf{\Delta}_c)^{-1}(\mathbf{\Delta}_c'\mathbf{W}_{\mathsf{NNT}}\mathbf{\Delta}_c)] \\
c_3 & = & (d/h_1)c_2 \\
c_4 & = & n(\mathbf{s} - \hat{\boldsymbol{\sigma}})'\mathbf{\Delta}_c(\mathbf{\Delta}_c'\mathbf{W}_{\mathsf{NNT}}\mathbf{\Delta}_c)^{-1}\mathbf{\Delta}_c'(\mathbf{s} - \hat{\boldsymbol{\sigma}})
\end{array}$$

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 87

Does Non-Normality Matter?

HOSWGW: Original Scores

	C_1	C_2	C_3	C_4
Model 1	51.19	48.62	47.22	57.92
Model 2	28.10	27.71	27.64	30.13

HOSWGW: Normalized Scores

	C_1	C_2	C_3	C_4
Model 1	48.74	46.87	46.66	56.30
Model 2	25.75	25.71	26.17	27.94

Slide 88

Model Selection

In the AM situation, three strongly related criteria can be used. These are simple functions of chi-square and the degrees of freedom:

$$\mathbf{AIC} = c + 2t$$

$$\mathbf{CAIC} = c + (1 + \ln N)t$$

$$\mathbf{ECVI} = (c/n) + 2(t/n)$$

To apply these measures to the decision problem, one estimates each model, ranks them according to one of these criteria and chooses the model with the smallest value.

Slide 89

Model Selection

HOSWGW: Original Scores

	Inde..dence	Model 1	Model 2	Saturated
df	36	24	23	0
χ^2	502.29	48.62	27.71	.00
RMSEA	0.300	0.084	0.038	?
AIC	520.29	90.62	71.71	90.00
CAIC	556.08	174.13	159.20	268.95
ECVI	3.613	0.629	0.498	0.625

Slide 90

Population Error of Approximation

Models don't have to be true to be useful.

The use of chi-square as a central χ^2-statistic is based on the assumption that the model holds exactly in the population. As already pointed out, this may be an unreasonable assumption in most empirical research. A consequence of this assumption is that models which hold approximately in the population will be rejected in large samples. A fit measures which take particular account of the error of approximation in the population is *population discrepancy function*:

$$\hat{F}_0 = \mathsf{Max}\{\hat{F} - (d/n), 0\} \ , \qquad (1)$$

where $\hat{F}$ is the minimum value of the fit function, $n = N - 1$ and d is the degrees of freedom.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 91

Since $\hat{F}_0$ generally decreases when parameters are added in the model Steiger (1990) suggested using a Root Mean Square Error of Approximation (RMSEA):

$$\epsilon = \sqrt{\hat{F}_0/d} \qquad (2)$$

as a measure of *population discrepancy per degree of freedom*. The LISREL output gives the point estimate of ϵ called RMSEA, a 90% confidence limit for ϵ and the P-value for the test of $\epsilon < 0.05$, called P-value for test of *close fit*

Slide 92

"Practical experience has made us feel that a value of RMSEA of about 0.05 or less would indicate a close fit of the model in relation to the degrees of freedom. This figure is based on subjective judgment. It cannot be regarded as infallible or correct, but it is more reasonable than the requirement of exact fit with RMSEA = 0.0. We are also of the opinion that a value of about 0.08 or less for the RMSEA would indicate a reasonable error of approximation and would not want to employ a model with a RMSEA greater than 0.1." (Browne & Cudeck, 1993)

Slide 93

Other Fit Measures

Since chi-square is $N-1$ times the minimum value of the fit function, chi-square tends to be large in large samples if the model does not hold. A number of goodness-of-fit measures have been proposed to eliminate or reduce its dependence on sample size.

Goodness-of-Fit Indices

$$\mathrm{GFI} = 1 - \frac{F[\mathbf{S}, \mathbf{\Sigma}(\hat{\theta})]}{F[\mathbf{S}, \mathbf{\Sigma}(\mathbf{0})]} .$$

$$\mathrm{AGFI} = 1 - \frac{k(k+1)}{2d}(1 - \mathrm{GFI}) ,$$

Slide 94

Other Fit Indices

Let F be the minimum value of the fit function for the estimated model, let F_i be the minimum value of the fit function for the independence model, and let d and d_i be the corresponding degrees of freedom. Furthermore, let $f = nF/d$, $f_i = nF_i/d_i$, $\tau = \mathsf{max}(nF - d, 0)$, and $\tau_i = \mathsf{max}(nF_i - d_i, nF - d, 0)$.

Then

$$\mathrm{NFI} = 1 - F/F_i$$

$$\mathrm{PNFI} = (d/d_i)(1 - F/F_i)$$

$$\mathrm{NNFI} = \frac{f_i - f}{f_i - 1}$$

$$\mathrm{CFI} = 1 - \tau/\tau_i$$

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 95

$$\mathrm{IFI} = \frac{nF_i - nF}{nF_i - d}$$

$$\mathrm{RFI} = \frac{f_i - f}{f_i}$$

$$\mathrm{PGFI} = (2d/k(k+1))\mathrm{GFI}$$

Hoelter (1983) proposed a critical N (CN) statistic:

$$\mathrm{CN} = \frac{\chi^2_{1-\alpha}}{F} + 1 ,$$

where $\chi^2_{1-\alpha}$ is the $1-\alpha$ percentile of the chi-square distribution. This is the sample size that would make the obtained chi-square just significant at the significance level α.

Slide 96

"We want a model with relatively few parameters in which each is significant and meaningful" (Jöreskog, 1981)

Model Assessment and Modification

Consider the `MG` situation. The output from LISREL provides much information useful for model evaluation and assessment of fit. It is helpful to classify this information into the three groups

- Examination of the solution
- Measures of overall fit
- Detailed assessment of fit

and to proceed as follows.

Slide 97

- Examine the parameter estimates to see if there are any unreasonable values or other anomalies. Parameter estimates should have the right sign and size according to theory or a priori specifications. Examine the squared multiple correlation R^2 for each relationship in the model. The R^2 is a measure of the strength of linear relationship. A small R^2 indicates a weak relationship and suggests that the model is not effective.
- Examine the measures of overall fit of the model, particularly the chi-square. A number of other measures of overall fit, all of which are functions of chi-square, may also be used. If any of these quantities indicate a poor fit of the data, proceed with the detailed assessment of fit in the next step.

Slide 98

- The tools for examining the fit in detail are the residuals and standardized residuals, the modification indices and the corresponding expected change. Each of these quantities may be used to locate the source of misspecification and to suggest how the model should be modified to fit the data better.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 99

Detailed Assessment of Fit

If, on the basis of overall measures of fit or other considerations, it is concluded that the model does not fit sufficiently well, one can examine the fit more closely to determine possible sources of the lack of fit. For this purpose, fitted and standardized residuals and modification indices are useful.

Slide 100

Fitted and Standardized Residuals

A residual is an observed minus a fitted covariance (variance). A standardized residual is a residual divided by its estimated standard error. There are such residuals for every pair of observed variables. Fitted residuals depend on the unit of measurement of the observed variables. If the variances of the variables vary considerably from one variable to another, it is rather difficult to know whether a fitted residual should be considered large or small. Standardized residuals, on the other hand, are independent of the units of measurement of the variables. In particular, standardized residuals provide a "statistical" metric for judging the size of a residual.

Slide 101

A large positive residual indicates that the model underestimates the covariance between the two variables. On the other hand, a large negative residual indicates that the model overestimates the covariance between the variables. In the first case, one should modify the model by adding paths which could account for the covariance between the two variables better. In the second case, one should modify the model by eliminating paths that are associated with the particular covariance.

Slide 102

All the standardized residuals may be examined collectively in two plots: a stemleaf plot and a Q-plot. A good model is characterized by a stemleaf plot in which the residuals are symmetrical around zero, with most in the middle and fewer in the tails. An excess of residuals on the positive or negative side indicates that residuals may be systematically under- or overestimated in the above sense. In the Q-plot, a good model is characterized by points falling approximately on a 45° line. Deviations from this pattern are indicative of specification errors in the model, non-normality in the variables or of nonlinear relationships among the variables. In particular, standardized residuals that appear as outliers in the Q-plot are indicative of a specification error in the model.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 103

Modification Index

A modification index may be computed for each fixed and constrained parameter in the model. Each such modification index measures how much chi-square is expected to decrease if this particular parameter is set free and the model is reestimated. Thus, the modification index is approximately equal to the difference in chi-square between two models in which one parameter is fixed or constrained in one model and free in the other, all other parameters being estimated in both models. The largest modification index shows the parameter that improves the fit most when set free.

Slide 104

Associated with each modification index, there is an expected parameter change (EPC). This measures how much the parameter is expected to change, in the positive or negative direction, if it is set free. If the units of measurement in observed and/or latent variables are of no particular interest, one can use a scale free variant SEPC of EPC.

Slide 105

Modification indices are used in the process of model evaluation and modification in the following way. If chi-square is large relative to the degrees of freedom, one examines the modification indices and relaxes the parameter with the largest modification index *if this parameter can be interpreted substantively*. If it does not make sense to relax the parameter with the largest modification index, one considers the second largest modification index, etc. If the signs of certain parameters are specified a priori, positive or negative, the expected parameter changes associated with the modification indices for these parameters can be used to exclude models with parameters having the wrong sign.

Slide 106

Testing Models and Hypotheses: Summary

Exact Fit = Model fits exactly in the population ($\epsilon_a = 0$)

Close Fit = Model fits approximately in the population ($\epsilon_a \leq 0.05$)

Test of Exact Fit (SC situation)

Use Chi-square and reject the model if P-value is less than 0.05

Test of Close Fit (SC situation)

Use RMSEA and accept the model if RMSEA is 0.05 or smaller and the upper confidence limit for RMSEA is 0.08 or smaller or if the P-value for test of close fit is larger than 0.50

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 107

Finding the Best Fitting Model (MG situation)

- Make sure the Independence model is rejected
- If the initial model fits well, eliminate all parameters (except error variances) with non-significant t-values and make sure the resulting model makes sense and all its parameters can be interpreted
- If the initial model does not fit, use modification indices to determine the parameters to be added. Add only parameters that make sense and can be interpreted
- Among all models generated, including the Independence model and the Saturated model, choose the model with the smallest ECVI (or AIC or CAIC)

Slide 108

Testing a Structural Hypothesis

First establish a valid base model that fits the data and note its chi-square c_1 and degrees of freedom d_1. Then estimate this model under the hypothesis and note its chi-square c_0 and degrees of freedom d_0. To test the hypothesis use $c = c_0 - c_1$ as chi-square with $d = d_0 - d_1$ degrees of freedom.

Slide 109

Strategy of Analysis

A suitable strategy for data analysis in the `MG` case may be the following.

- Specify an initial model on the basis of substantive theory, stated hypotheses, or at least some tentative ideas of what a suitable model should be.
- Estimate the measurement model for each construct separately, then for each pair of constructs, combining them two by two. Then estimate the measurement model for all the constructs without constraining the covariance matrix of the constructs. Finally, estimate the structural equation model for the constructs jointly with the measurement models.

Slide 110

- For each model estimated in Step 2, evaluate the fit. In particular, pay attention to chi-square, standard errors, t-values, standardized residuals, and modification indices. If chi-square is large relative to the degrees of freedom, the model must be modified to fit the data better. For model modification use the modification indices. If chi-square is small relative to the degrees of freedom, the model is overfitted and parameters with very small t-values could possibly be eliminated. If chi-square is in the vicinity of the degrees of freedom, the model may be acceptable, but examine the estimated solution to see if there are any unreasonable values or other anomalies.

  For each model estimated, if this step leads to a modified model, repeat this step on each modified model.

Karl G Jöreskog LISREL Workshop A, Jena, Spring 2003

Slide 111

- Hopefully, the last model estimated in the previous step is one which fits the data of the sample reasonably well and in which all parameters are meaningful and substantively interpretable. However, this does not necessarily mean that it is the “best” model, because its results may have been obtained to some extent by “capitalizing on chance.” The model modification process may have generated several “reasonable models” which should be cross-validated on independent data. If no independent sample is available but the initial sample is large, one may consider splitting the sample into two subsamples and use one (the calibration sample) for exploration and the other (the validation sample) for cross-validation.

Slide 112

- The cross-validation is done by computing a cross-validation index CVI (Cudeck & Browne, 1983) for each model. The CVI index is a measure of the distance (difference, discrepancy, deviance) between the fitted covariance matrix in the calibration sample and the sample covariance matrix of the validation sample. The model with the smallest validation index is the one which is expected to be most stable in repeated samples. If the smallest validation index occurs for the best fitted model in the calibration sample it is good. If the smallest validation index occurs for some of the other models that have been fitted, one must make a decision on substantive grounds, which of the two models to retain.