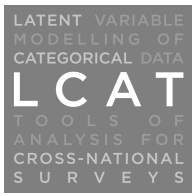
Latent Trait and Latent Class Analysis for Multiple Groups

Day 1: Single-group analysis

LCAT Training Workshop

2012



LCAT workshops

- Training component of the research project *Latent variable modelling of categorical data: Tools of analysis for cross-national surveys*, or **LCAT** for short
  - Funded by ESRC grant RES-239-25-0022, under the *Methods for Comparative Cross-National Research* initiative
  - See `http://stats.lse.ac.uk/lcat/` for more
- Three 2-day workshops in April-May 2012:
  - London (LSE)
  - Manchester (CCSR)
  - Edinburgh (AQMeN)
- Lecturers: Jouni Kuha, Irini Moustaki, Sally Stares, and Jonathan Jackson
  - All of the Methodology Institute and/or Department of Statistics, London School of Economics and Political Science

Outline of the workshop

- Day 1: Models for single groups
  - Session 1.1: Introduction and latent trait models
  - Session 1.2: Latent class models and model assessment
- Day 2: Models for multiple groups
  - Session 2.1: Cross-group comparisons of latent distributions
  - Session 2.2: Examining measurement equivalence and non-equivalence
- Each session consists of a lecture and a computer class

Session 1.1

1.1(a): Introduction to Latent Variable Models

Outline of Session 1.1

1.1(a): Introduction to latent variable models

1.1(b): Latent trait models for single groups

- Models with one trait
  - Specification: Measurement models and structural models
  - Fitting the model in Mplus
  - Interpretation: Item response probabilities
- Models with two traits
  - New issues in measurement and structural models

Example: Social life feelings study, Schuessler (1982)

Survey sample of 1490 Germans

Scale of "Economic self-determination": Yes or No responses to the following five questions:

1. Anyone can raise his standard of living if he is willing to work at it.
2. Our country has too many poor people who can do little to raise their standard of living.
3. Individuals are poor because of the lack of effort on their part.
4. Poor people could improve their lot if they tried.
5. Most people have a good deal of freedom in deciding how to live.

What is going on here?

Latent variables and measurement

Using statistical models to understand constructs better: a question of **measurement**

- Many theories in behavioral and social sciences are formulated in terms of theoretical constructs that are not directly observed

  attitudes, opinions, abilities, motivations, etc.

- The measurement of a construct is achieved through one or more observable **indicators** (questionnaire **items**).
- The purpose of a measurement model is to describe how well the observed indicators serve as a measurement instrument for the constructs, also known as **latent variables**.
- Measurement models often suggest ways in which the observed measurements can be improved.

Latent variables and substantive theories

Using statistical models to understand relationships between constructs and to test **theories** about those relationships.

- Often measurement by multiple indicators may involve more than one latent variable.
- Subject-matter theories and research questions usually concern relationships among the latent variables, and perhaps also observed explanatory variables.
- These are captured by statistical models for those variables: **structural models**.

Aims of latent variable modelling

- Measurement models:
  - Study the relationships among a set of observed indicators. Identify underlying constructs that explain the relationships among the indicators.
  - Derive measurement scales for the constructs.
  - Scale individuals on the identified latent dimensions.
  - Reduce dimensionality of the observed data.
- Structural models:
  - Study relationships among the constructs and explanatory variables, and test hypotheses about them.

Notation for variables

Consider the following variables for each subject (e.g. survey respondent):

- Observed indicators $\mathbf{y} = (y_1, \ldots, y_p)$
- Latent variables $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_q)$
  - We focus on cases with 1 or 2 latent variables, i.e. $\boldsymbol{\eta} = \eta_1 = \eta$ or $\boldsymbol{\eta} = (\eta_1, \eta_2)$.
- Explanatory variables $\mathbf{x}$, i.e. observed variables which are treated as predictors rather than measures of $\boldsymbol{\eta}$
  - These will be introduced tomorrow, but not included today.

Latent variable models

In general, a latent variable model (for one subject) is defined as

$$p(\mathbf{y}, \boldsymbol{\eta}|\mathbf{x}) = p(\mathbf{y}|\boldsymbol{\eta}, \mathbf{x})\, p(\boldsymbol{\eta}|\mathbf{x})$$

where $p(\cdot|\cdot)$ are (multivariate) conditional distributions.

- $p(\mathbf{y}|\boldsymbol{\eta}, \mathbf{x})$ is the **measurement model**
- $p(\boldsymbol{\eta}|\mathbf{x})$ is the **structural model**

Particular models are obtained with different choices of these distributions.

The first big choice is the *type* of the variables in this, i.e.

- **continuous** or
- **categorical** (i.e. nominal, ordinal, binary)

Latent variable models

		Observed indicators	
		Continuous	*Categorical*
Latent variables	*Continuous*	Factor analysis	Latent trait models
	Categorical	Latent profile analysis	Latent class models

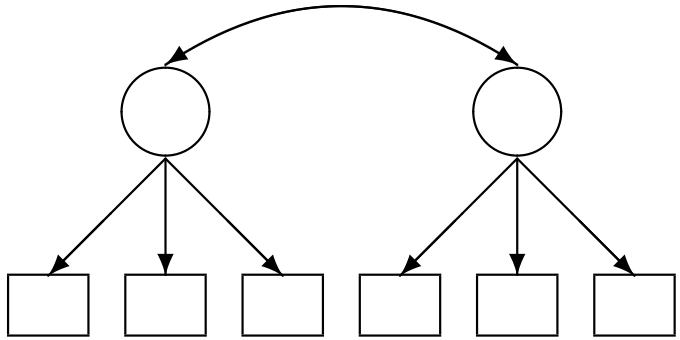
- We assume that you are somewhat familiar with linear factor analysis (including structural equation models).
- The topic of this workshop is models for categorical indicators, i.e. latent trait and latent class models.
  - Useful, because many items in surveys (and elsewhere) are categorical.

Path diagrams

Widely used to represent latent variable models graphically.

Basic elements:

- ◯ denotes latent variables
- □ denotes observed variables
- → represents a regression relationship (directed association)
- ↶ represents a correlation (undirected association)
- For example:



Readings

Theoretical:

- Bartholomew, D.J., Knott, M. and Moustaki, I. (2011). Latent Variable Models and Factor Analysis: A Unified Approach (3rd ed). Wiley.
- Skrondal, A. and Rabe-Hesketh, S. (2005). Generalized Latent Variable Models. Chapman and Hall/CRC.

Applied:

- Bartholomew, D.J., Steele, F., Moustaki, I. and Galbraith, J. (2008). The Analysis of Multivariate Social Science Data (2nd ed). Chapman and Hall/CRC.
(`http://www.cmm.bris.ac.uk/team/amssd.shtml`)

Software

In the computer classes of this workshop we will use

- **Mplus** for fitting the models themselves
  - Very general latent variable modelling software (`http://www.statmodel.com/`)
- LCAT functions in the general-purpose, free statistical package **R** (`http://cran.r-project.org/`) for post-processing and displaying the results

See instructions for the classes, and a computing manual at the LCAT website (`http://stats.lse.ac.uk/lcat/`) for more detailed instructions.

Session 1.1

1.1(b): Latent Trait Models for Single Groups

Example: Attitudes to abortion

From the 2004 British Social Attitudes Survey: *"Here are a number of circumstances in which a woman might consider an abortion. Please say whether or not you think the law should allow an abortion in each case."* (1=Yes, 2=No) :

1. The woman decides on her own that she does not wish to have the child. [WomanDecide]
2. The couple agree that they do not wish to have the child. [CoupleDecide]
3. The woman is not married and does not wish to marry the man. [NotMarried]
4. The couple cannot afford any more children. [CannotAfford]

(Bartholomew et al. (2008) analyse these same items for the 1986 BSA.)

Example: Attitudes to science and technology

From the Consumer Protection and Perceptions of Science and Technology section of the 1992 Eurobarometer Survey, GB respondents:

1. Science and technology are making our lives healthier, easier and more comfortable. [Comfort]
2. The application of science and new technology will make work more interesting. [Work]
3. Thanks to science and technology, there will be more opportunities for the future generations. [Future]
4. The benefits of science are greater than any harmful effects it may have. [Benefit]

Response alternatives: Strongly disagree (1), Disagree to some extent (2), Agree to some extent (3), Strongly agree (4).

(See Bartholomew et al. (2008) for more detailed analysis.)

Latent trait models

By a **latent trait model** we mean a latent variable model where

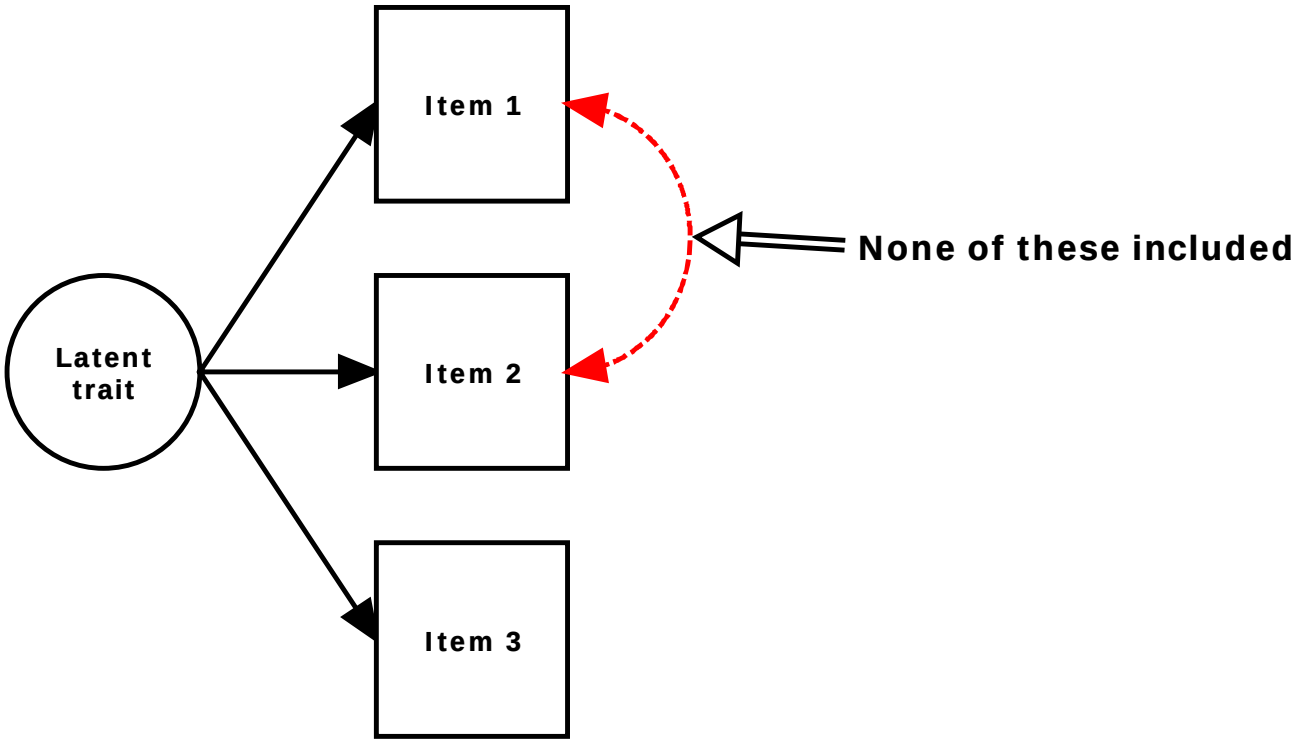
- latent variables $\boldsymbol{\eta}$ (**latent traits**) are continuous (like in factor analysis)
- observed indicators $\mathbf{y}$ are treated as categorical (unlike in factor analysis)

Such models are very commonly used also in educational and psychological testing, where they are known as Item Response Theory (IRT) models.

We begin with one-trait models, to introduce basic concepts.

- Here the focus is on the use of the model for measurement.

Assumption of conditional independence



Throughout this workshop (for both latent trait and latent class models) we consider only models where the items $\mathbf{y} = (y_1, \ldots, y_p)$ are conditionally independent of each other, given the latent variables.

The one-trait model

Under the assumption of conditional independence, a latent trait model with one trait η is given by

$$p(\mathbf{y}, \eta) = \left[\prod_{j=1}^{p} p(y_j|\eta) \right] p(\eta) = \left[p(y_1|\eta) \times \cdots \times p(y_p|\eta) \right] p(\eta).$$

We thus need to specify only

- distribution $p(\eta)$ of the latent trait (the structural model)
- models $p(y_j|\eta)$ for each indicator y_j given the trait (the measurement models)

The one-trait model: The structural model

Assume that latent trait η is normally distributed with mean κ and variance ϕ, i.e.

$$\eta \sim N(\kappa, \phi)$$

where we impose the constraints that $\kappa = 0$, $\phi = 1$.

Fixing (κ, ϕ) in this way is needed to identify the scale of the latent variable.

- This could also be achieved by freeing (κ, ϕ) but fixing parameters in one measurement model.
- However, a constraint on the distribution of η will be more convenient in multigroup analysis tomorrow, so we use it throughout.
- In multigroup analysis, (κ, ϕ) only needs to be fixed in one group.
- Fixing $(\kappa, \phi) = (0, 1)$ still leaves the *direction* of the trait undefined, so it may be reversed if convenient.

The one-trait model: The measurement models

Here each item y_j is categorical, so it has L_j possible levels (categories) $l = 1, \ldots, L_j$.

- Different items may have different values of L_j.
- If the item is ordinal, the numbering of the levels is in order and cannot be changed (except reversed).
- If the item is nominal, the numbering of the levels is arbitrary.
- If $L_j = 2$, the item is binary. This can be treated as either ordinal or nominal — the model is the same either way.

A measurement model for y_j is a regression model for the probabilities of the categories

$$\pi_{jl}(\eta) = P(y_j = l|\eta)$$

with the latent trait η as an explanatory variable.

The one-trait model: The measurement models

For the measurement models, Mplus uses standard types of regression models for categorical response variables:

- For a nominal item, a **multinomial logistic model**

$$\pi_{jl}(\eta) = \frac{\exp(\tau_{jl} + \lambda_{jl}\,\eta)}{\sum_{m=1}^{L_j} \exp(\tau_{jm} + \lambda_{jm}\,\eta)} \quad \text{for } l = 1, \ldots, L_j$$

  with the constraint $\tau_{jL_j} = \lambda_{jL_j} = 0$ — i.e. the *highest* category of the item is the baseline category.

The one-trait model: The measurement models

- For an ordinal item, an **ordinal logistic model**

$$\nu_{jl}(\eta) = P(y_j \leq l|\eta) = \frac{\exp(\tau_{jl} - \lambda_j \eta)}{1 + \exp(\tau_{jl} - \lambda_j \eta)} \quad \text{for } l = 1, \ldots, L_j - 1.$$

From this, the probabilities of individual levels of y_j are

$$\pi_{jl}(\eta) = \nu_{jl}(\eta) - \nu_{j,l-1}(\eta) \quad \text{for } l = 1, \ldots, L_j$$

where we take $\nu_{j0} = 0$ and $\nu_{jL_j} = 1$.

The one-trait model: The measurement models

- For a binary item, the multinomial model gives

$$\pi_{j1}(\eta) = \frac{\exp(\tau_{j1} + \lambda_{j1}\eta)}{1 + \exp(\tau_{j1} + \lambda_{j1}\eta)}$$

and the ordinal model

$$\nu_{j1}(\eta) = \pi_{j1}(\eta) = \frac{\exp(\tau_{j1} - \lambda_{j}\eta)}{1 + \exp(\tau_{j1} - \lambda_{j}\eta)}$$

which are the same, with $\lambda_{j1} = -\lambda_j$. Obviously $\pi_{j2}(\eta) = 1 - \pi_{j1}(\eta)$.

  - In the output of the `lcat` functions in R, we reverse the signs of the loadings λ_j from all ordinal models from Mplus, so that these two will agree.

Latent trait models in Mplus: Input

Types of indicator variables are declared by the `Variable` command, e.g.:

```
Data:
  File = bsa04ab.dat;
Variable:
  Names = item1 item2 item4 item4;
  Categorical = item1 item2;
  Nominal = item3 item4;
```

where `Categorical` means ordinal, and `Nominal` means nominal.

Latent trait(s) are declared and the model specified by the `Model` command, e.g.

```
Model:
  trait BY item1* item2 item3 item4;
  [trait@0]; trait@1;
```

Here `trait` is the name of the latent trait, `[trait@0]` fixes its mean (κ) at 0 and `trait@1` its variance (ϕ) at 1, and `item1*` causes the loading of the first item (`item1`) to be estimated (rather than fixed, as by default).
(More complete instructions in the computer class.)

Latent trait models in Mplus: Output

Suppose `trait` is the name of a latent trait, `ynom` an item declared to be nominal, and `yord` an item declared to be ordinal.

Mplus table of parameter estimates has following types of entries and headings for different types of parameters:

		Estimate	S.E.	Two-Tailed Est./S.E.	P-Value
	TRAIT BY				
λ_j:	YORD	-1.911	0.102	-18.786	0.000
λ_{jl}:	YNOM#1	2.985	0.209	14.265	0.000
	Thresholds				
τ_{jl}:	YORD$1	-0.042	0.059	-0.708	0.479
	Intercepts				
τ_{jl}:	YNOM#1	-1.154	0.100	-11.592	0.000
	Means				
κ:	TRAIT	0.000	0.000	999.000	999.000
	Variances				
ϕ:	TRAIT	1.000	0.000	999.000	999.000

(Note: `S.E. = 0.000` indicates a fixed parameter.)

Using the `lcat` R functions with Mplus

In the computer the classes, we will work as follows:

- Estimate a model in Mplus.
- In R, read in and post-process the results:

```
lt1.models <- lcat("ltmod1.out",path="c:/lcatworkshop")
```

- Display estimates and residuals, draw plots, etc. in R:

```
lt1.models
print(lt1.models,1)
reorder(lt1.models,1,traits=-1)
resid(lt1.models,1,sort=T)
plot(lt1.models,models=1,items=1:4,levels=1)
```

What all this means will be revealed in the classes.

Example: Attitudes to abortion

From the 2004 British Social Attitudes Survey: *"Here are a number of circumstances in which a woman might consider an abortion. Please say whether or not you think the law should allow an abortion in each case."* (1=Yes, 2=No) :

1. The woman decides on her own that she does not wish to have the child. [WomanDecide]
2. The couple agree that they do not wish to have the child. [CoupleDecide]
3. The woman is not married and does not wish to marry the man. [NotMarried]
4. The couple cannot afford any more children. [CannotAfford]

(Bartholomew et al. (2008) analyse these same items for the 1986 BSA.)

Example: Mplus input

```
Title: Attitudes to abortion, BSA04. 1-trait latent trait model.
Data:
  File = bsa04ab.dat;
Variable:
  Names = abort1 abort2 abort3 abort4;
  Missing = all (99) ;
  Categorical = abort1-abort4;
Analysis:
  Estimator=ML;
  Starts = 20 10;
Model:
  attitude BY abort1* abort2-abort4;
  [attitude@0];
  attitude@1;
Savedata:
  File="tmp.dat";
```

Example: Mplus output (parameter estimates)

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
ATTITUDE BY				
ABORT1	4.216	0.541	7.795	0.000
ABORT2	5.175	0.782	6.614	0.000
ABORT3	3.786	0.450	8.409	0.000
ABORT4	3.172	0.342	9.272	0.000
Means				
ATTITUDE	0.000	0.000	999.000	999.000
Thresholds				
ABORT1$1	1.462	0.258	5.664	0.000
ABORT2$1	3.111	0.477	6.525	0.000
ABORT3$1	0.997	0.213	4.678	0.000
ABORT4$1	1.011	0.184	5.499	0.000
Variances				
ATTITUDE	1.000	0.000	999.000	999.000

Example: Part of LCAT output

```
Trait  ATTITUDE :
      Mean sd
(All)    0  1

Parameters of the measurement model:
'$' indicates intercept of an ordinal logistic model,
and '#' of a multinomial logistic model.

Positive loading of a trait indicates that higher values of the trait
correspond to higher probabilities lower-numbered categories in ordinal model
and higher probability of a category relative to the highest-numbered category
in multinomial model.

         Constant ATTITUDE
ABORT1$1    1.462    4.216
         Constant ATTITUDE
ABORT2$1    3.111    5.175
         Constant ATTITUDE
ABORT3$1    0.997    3.786
         Constant ATTITUDE
ABORT4$1    1.011    3.172
```

(Here the trait itself has been reversed from the Mplus results.)

Example: Part of LCAT output

Models for the the latent traits:

Trait ATTITUDE :

	Mean	sd
(All)	0	1

Measurement probabilities
conditional on each latent trait at m+(-2,-1,0,1,2)*sd
where m and sd are the mean and standard deviation of the latent trait

Given trait ATTITUDE :

	m-2sd	m-1sd	mean	m+1sd	m+2sd
ABORT1#1	0.001	0.060	0.812	0.997	1.000
ABORT1#2	0.999	0.940	0.188	0.003	0.000
ABORT2#1	0.001	0.113	0.957	1.000	1.000
ABORT2#2	0.999	0.887	0.043	0.000	0.000
ABORT3#1	0.001	0.058	0.730	0.992	1.000
ABORT3#2	0.999	0.942	0.270	0.008	0.000
ABORT4#1	0.005	0.103	0.733	0.985	0.999
ABORT4#2	0.995	0.897	0.267	0.015	0.001

Example: Estimates of the measurement model

Item j	$\hat{\tau}_{j1}$	(s.e.)	$\hat{\lambda}_j$	(s.e.)	$\hat{\pi}_{j1}(0)$
WomanDecide	1.46	(0.26)	4.22	(0.54)	0.81
CoupleDecide	3.11	(0.48)	5.18	(0.78)	0.96
NotMarried	1.00	(0.21)	3.79	(0.45)	0.73
CannotAfford	1.01	(0.18)	3.17	(0.34)	0.73

Here $\hat{\pi}_{j1}(0)$ is the probability of 1=Yes (should be legal) when $\eta = 0$.

Parameters of the measurement model: Interpretation

For a binary item y_j with values $l = 1, 2$, we are using the model

$$\pi_{j1}(\eta) = P(y_j = 1|\eta) = \exp(\tau_{j1} + \lambda_j\eta)/\left[1 + \exp(\tau_{j1} + \lambda_j\eta)\right].$$

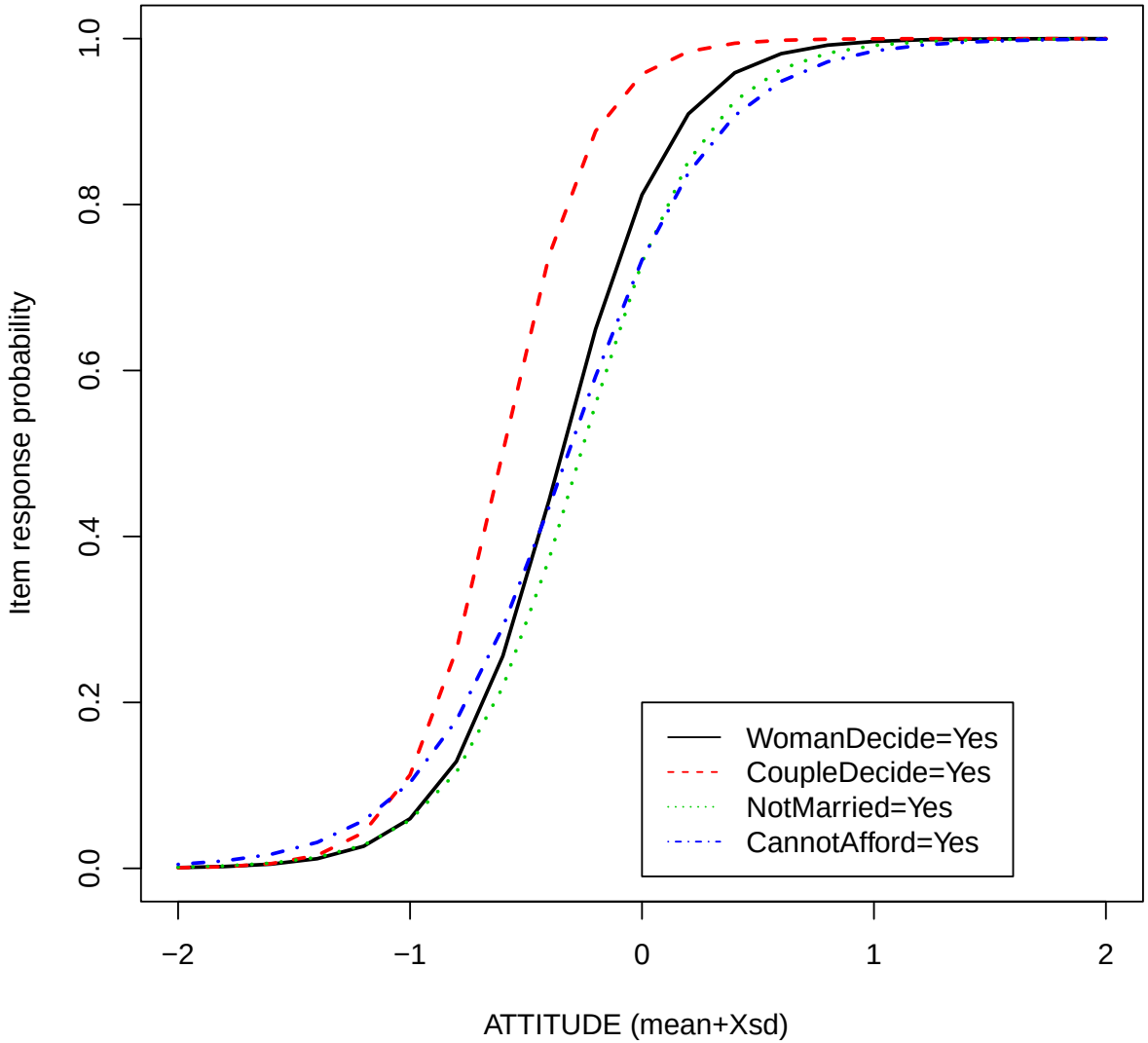
In educational testing, the **intercept** τ_{j1} is called the **difficulty** parameter, because it is related to the overall magnitude of $\pi_{j1}(\eta)$ across η. In particular, for the average individual ($\eta = 0$),

$$\pi_{j1}(0) = \exp(\tau_{j1})/\left[1 + \exp(\tau_{j1})\right].$$

The coefficient (**loading**) λ_j is also called the **discrimination** parameter, because it shows how fast $\pi_{j1}(\eta)$ varies as η varies, i.e. how well y_j discriminates between individuals with different values of η.

It is easiest to see these by drawing curves of $\pi_{jl}(\eta)$ as functions of η (**item response curves**).

Abortion example: Item response probabilities



Example: Attitudes to science and technology

From the Consumer Protection and Perceptions of Science and Technology section of the 1992 Eurobarometer Survey, GB respondents:

1. Science and technology are making our lives healthier, easier and more comfortable. [Comfort]
2. The application of science and new technology will make work more interesting. [Work]
3. Thanks to science and technology, there will be more opportunities for the future generations. [Future]
4. The benefits of science are greater than any harmful effects it may have. [Benefit]

Response alternatives: Strongly disagree (1), Disagree to some extent (2), Agree to some extent (3), Strongly agree (4).

(See Bartholomew et al. (2008) for more detailed analysis.)

Measurement probabilities for non-binary items

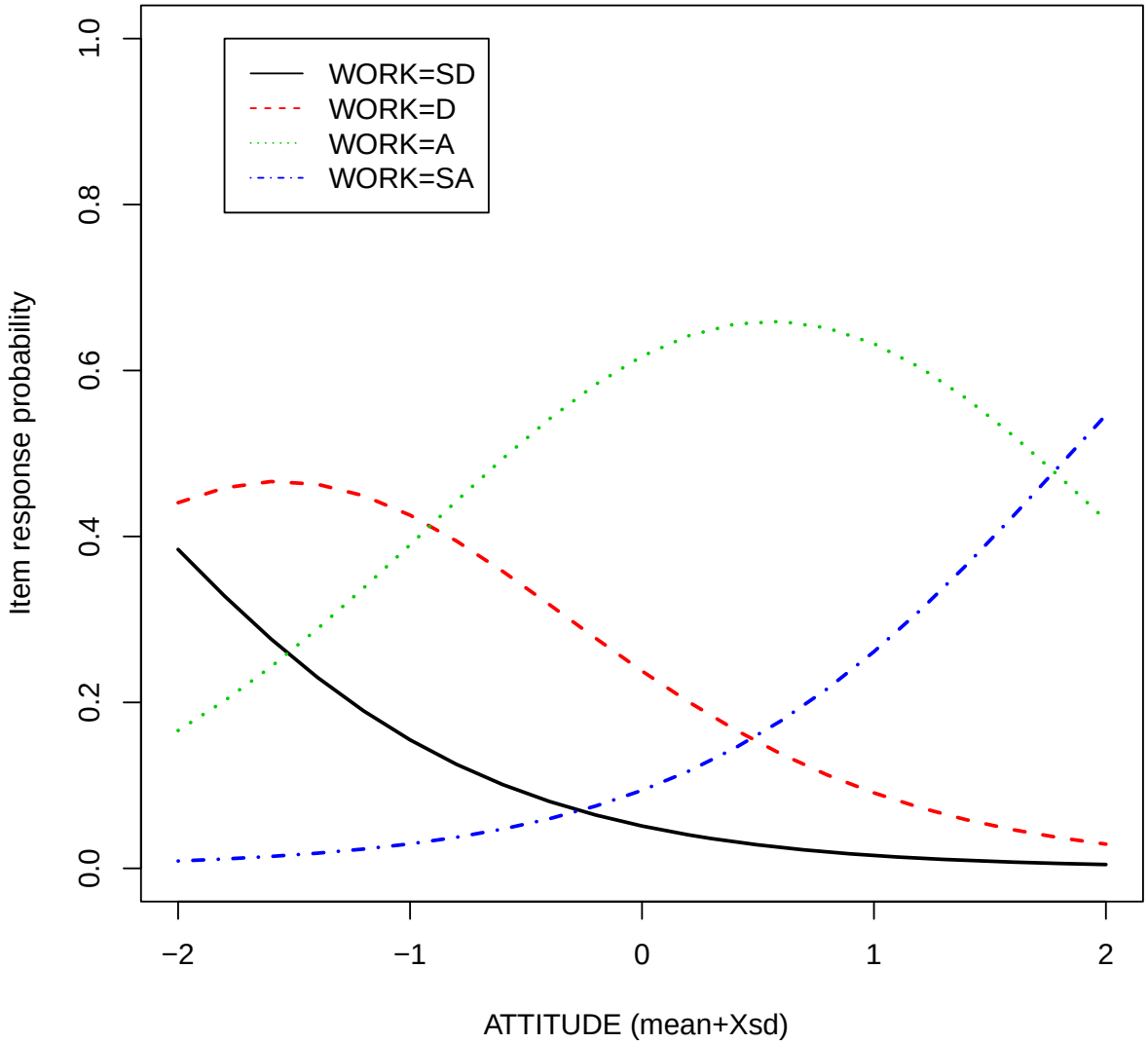
The intercepts and loadings of ordinal and multinomial logistic measurement models can also be interpreted as "difficulty" and "discrimination" parameters.

However, this can get complicated. It is much easier to interpret the measurement model by drawing item response curves again.

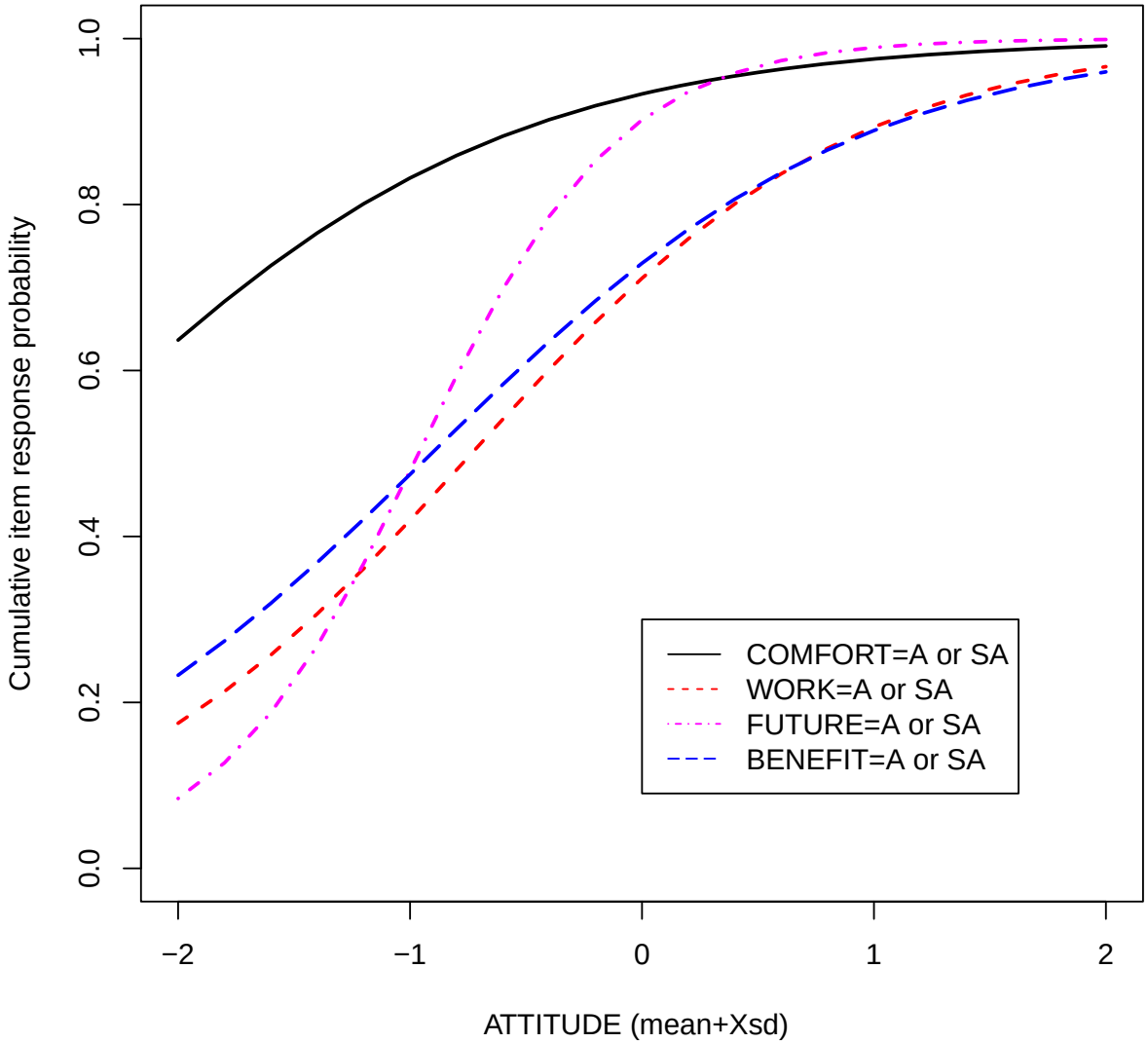
On the next slides, some ICCs for the science and technology example, where the items have been modelled as ordinal.

- Clearly here higher values of the latent trait indicate higher levels of support for science and technology.

Science example: Item response probabilities



Science example: Cumulative response probabilities



Trait scores

One use of a latent variable model is to derive predicted values (**scores**) of the latent variables for individuals, given their values of the items **y**.

For a latent trait model, we use the conditional ("posterior") means

$$\mathrm{E}(\eta|\mathbf{y}) = \frac{\int \eta\, p(\mathbf{y}|\eta) p(\eta)\, d\eta}{\int p(\mathbf{y}|\eta) p(\eta)\, d\eta}.$$

In Mplus, use the SAVEDATA command, as in:

```
Variable:
   Idvariable = idno;
Savedata:
   File = outfile.dat;
   Save = fscores;
```

(Here `idno` is an ID variable in the input data set which will also be included in the output data set `outfile.dat`, to allow merging back into a data set in other software.)

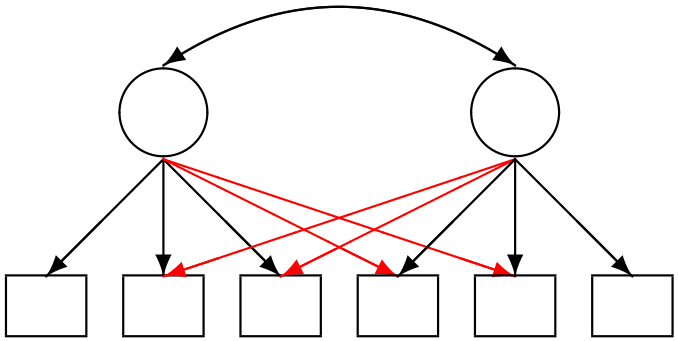
Models with more than one trait

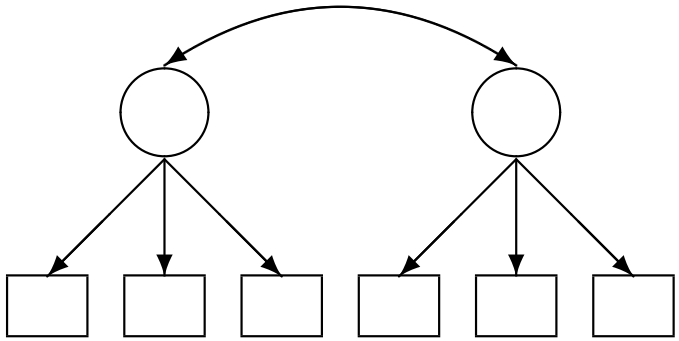
(Here we focus on the 2 traits $\boldsymbol{\eta} = (\eta_1, \eta_2)$, but the same ideas apply more generally.)

When there are more than one trait, new questions arise for both measurement and structural models:

- Measurement models: Cross-loadings, i.e. items which measure more than one trait.
- Structural models: Associations/regression models among the latent traits.

2 Traits: Measurement models





On the left is the largest possible measurement model

- For identifiability, each trait must have 1 item which measures only that trait
- This is analogous to Exploratory factor analysis with "oblique rotation"

On the right is smallest sensible model: Each trait measures only one trait.

Everything in between is also possible.

Example: Attitudes to science and technology

Same example as before, but now with these 6 items:

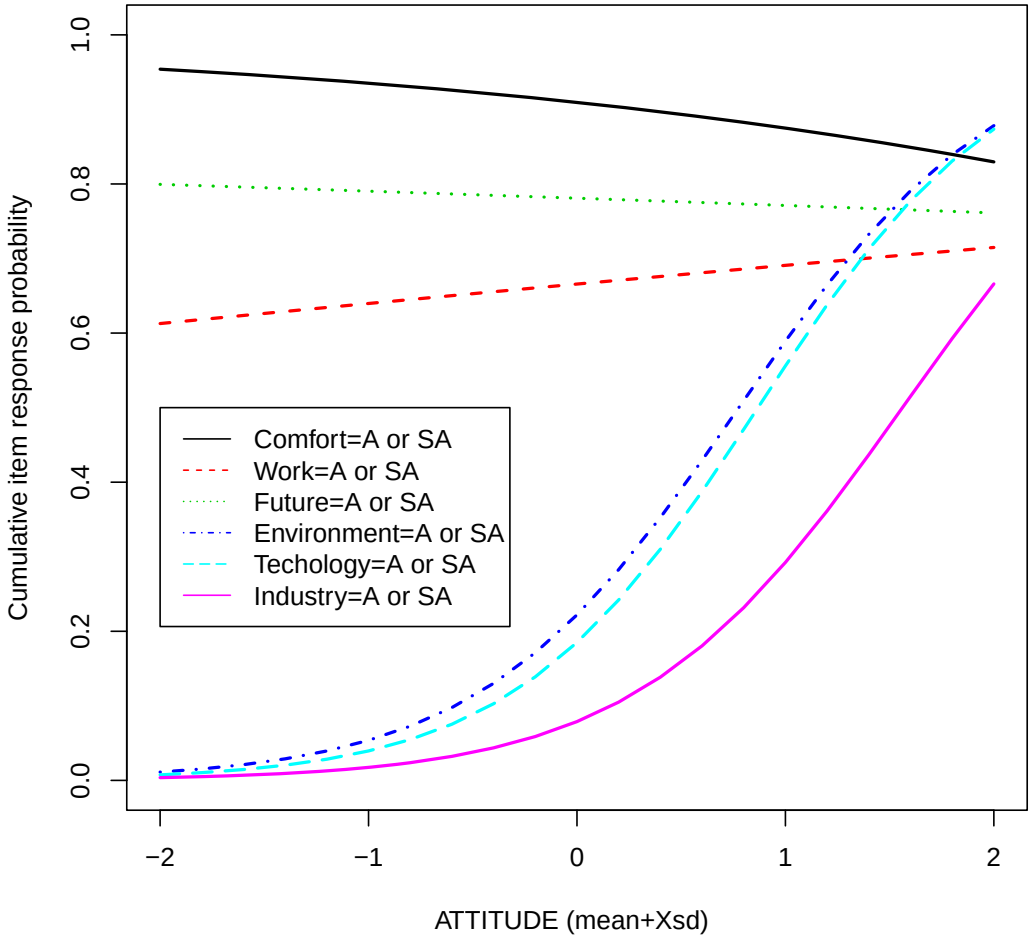
1. Science and technology are making our lives healthier, easier and more comfortable. [Comfort]
2. The application of science and new technology will make work more interesting. [Work]
3. Thanks to science and technology, there will be more opportunities for the future generations. [Future]
4. Scientific and technological research cannot play an important role in protecting the environment and repairing it. [Environment]
5. New technology does not depend on basic scientific research. [Technology]
6. Scientific and technological research do not play an important role in industrial development. [Industry]

Response alternatives: Strongly disagree (1), Disagree to some extent (2), Agree to some extent (3), Strongly agree (4).

(See Bartholomew et al. (2008) for more detailed analysis.)

Example: Attitudes to science and technology

Item response curves for a 1-trait model:



The trait seems to be more strongly associated with 3 of the items.

Example: A 2-trait model

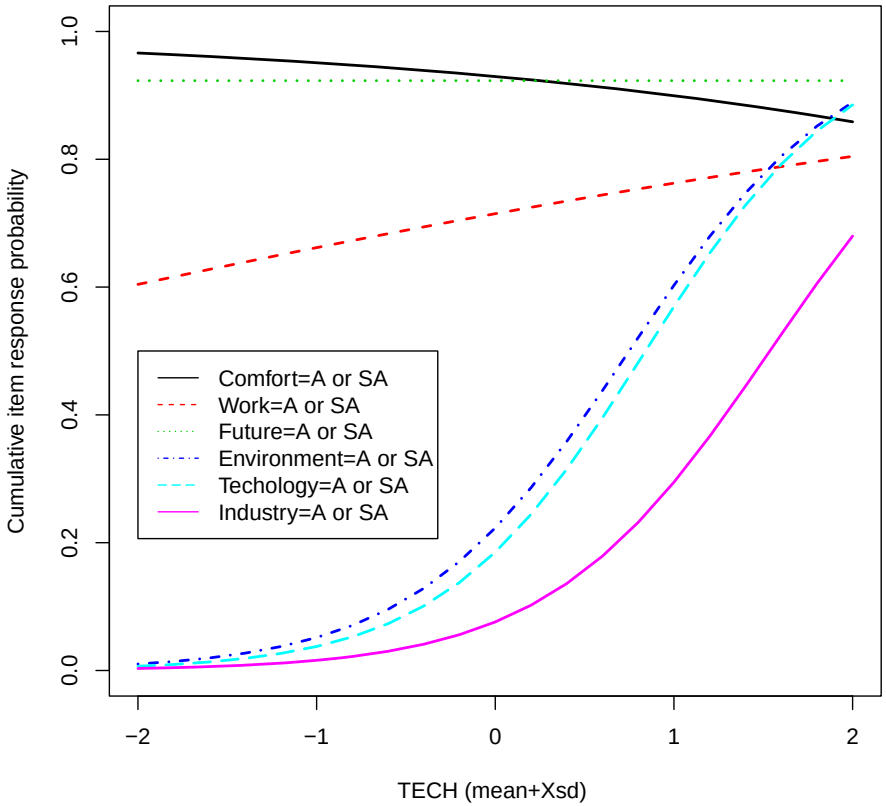
In Mplus - A full measurement model ("Model 1" below):

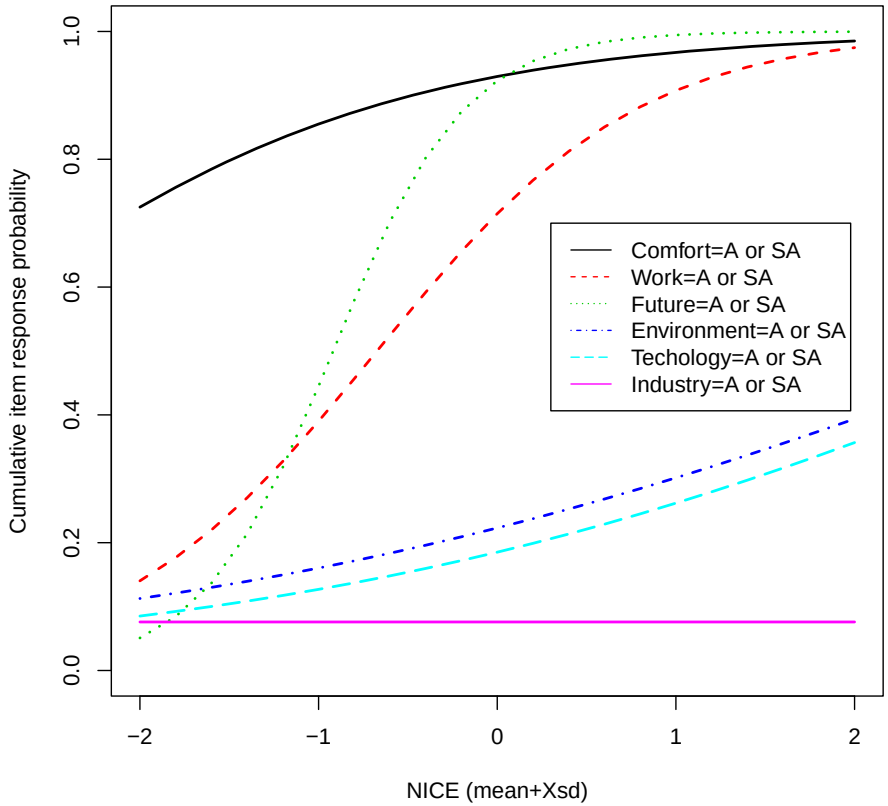
```
Model:
  tech BY comfort* environ work future@0 technol industry;
  nice BY comfort* environ work future technol industry@0;
  [tech@0];  tech@1;
  [nice@0];  nice@1;
```

and one restricted model ("Model 2"):

```
Model:
  tech BY environ* technol industry;
  nice BY comfort* work future;
  [tech@0];  tech@1;
  [nice@0];  nice@1;
```

Example: A 2-trait model (Model 1)





Example: A 2-trait model (Model 1)

Estimated loadings $\hat{\lambda}_j$ and standard errors for Model 1:

Item j	$\hat{\lambda}_{j\text{TECH}}$	(s.e.)	$\hat{\lambda}_{j\text{NICE}}$	(s.e.)
Comfort	0.39	(0.15)	-0.81	(0.17)
Work	-0.25	(0.16)	-1.34	(0.26)
Future	0.00		-2.71	(0.94)
Environment	-1.67	(0.25)	-0.41	(0.34)
Technology	-1.76	(0.27)	-0.45	(0.21)
Industry	-1.63	(0.25)	0.00	

The parameters in grey are set to 0 in Model 2.

Selecting the measurement model

Nested models, Like Models 1 and 2 here, can be compared using the likelihood ratio test:

```
> lcat.lrtest(workshop.scien6i,2,3)

Likelihood ratio test:
H0: scien6i_2lt2          H1: scien6i_2lt1
LR = 16.626      df = 4   P-value = 0.002
```

Here the conclusion is that at least some of the cross-loadings in Model 2 are significant.

- However, we might still decide to omit them, for simplicity.

We will discuss model assessment and model selection in more detail in the afternoon.

2 Traits: Structural models

When there are more than one trait, we can start considering models for relationships between the traits.

These relationships can be specified as undirected (correlation, on the left) or directed (regression, on the right).

The two formulations are equivalent, so the choice depends on what best matches theory and research questions.

Structural correlation vs. regression

In the correlation formulation,

$$\eta_1 \sim N(\kappa_1, \phi_{11}), \quad \eta_2 \sim N(\kappa_1, \phi_{22}), \quad \text{cov}(\eta_1, \eta_2) = \phi_{12}$$

with identifiability constraints $(\kappa_1, \phi_{11}) = (\kappa_2, \phi_{22}) = (0, 1)$.

In the regression formulation,

$$\eta_1 \sim N(\kappa_1, \phi_{11}) \quad \text{and} \quad \eta_2 = \gamma_0 + \gamma_1 \eta_1 + \zeta, \quad \text{with } \zeta \sim N(0, \psi)$$

and identifiability constraints $(\kappa_1, \phi_{11}) = (\gamma_0, \psi) = (0, 1)$.

Structural correlation vs. regression

In the science and technology example above:

In Mplus, the correlation formulation is given by

```
Model:
  nice WITH tech;
```

(which is the default)

and the regression formulation (if we decide that NICE will be a response variable to TECH) by

```
Model:
  nice ON tech;
```

Inference for the structural model

The only estimable parameters in these structural models are the association parameters between η_1 and η_2:

Correlation formulation: $\phi_{12} = \text{cov}(\eta_1, \eta_2)$

- Mplus output in the example:

		Estimate	S.E.	Est./S.E.	P-Value
NICE	WITH				
TECH		0.014	0.078	0.176	0.860

Regression formulation: Regression coefficient γ_1

-

		Estimate	S.E.	Est./S.E.	P-Value
NICE	ON				
TECH		0.015	0.078	0.193	0.847

Mplus output contains the Wald test of the parameter, or we could also use the likelihood ratio test.

Here the association between the two traits is not actually significant.

Session 1.2

1.2(a): Latent Class Models for Single Groups

Outline of Session 1.2

1.2(a): Latent class models for single groups

- Definition
- Methods of estimation (also apply to latent trait models)
- Fitting in Mplus
- Interpretation: Estimated class and measurement probabilities
- Class allocation

1.2(b): Model assessment for latent trait and latent class models

- Likelihood ratio tests
- AIC and BIC
- Measures based on bivariate marginal residuals

Example: Public engagement with science and technology

Based on Mejlgaard and Stares (*Public Understanding of Science*, 2010).

Sample of 1,307 UK respondents from Eurobarometer survey 63.1 on Europeans, Science and Technology, fielded in 2005.

Questions asking respondents if they ever...

Item	Description	% 'Yes'
read	Read articles on science in newspapers, magazines or on the internet	80
talk	Talk with your friends about science and technology	74
meet	Attend public meetings or debates about science and technology	22
protest	Sign petitions or join street demonstrations about nuclear power, biotechnology or the environment	25

Latent class models

A **latent class model** is a latent variable model where the latent variables $\boldsymbol{\eta}$ as well as the observed items are categorical.

- Here we consider only the case of a single latent variable η.
- The items may be nominal, ordinal and/or binary, as before.

The latent variable η then has C levels (**latent classes**) $c = 1, \ldots, C$.

Latent class models

Two basic elements of a (single-group) latent class models are

- Measurement model: The item response probabilities

$$\pi_{jl(c)} = P(y_j = l|\eta = c)$$

for items $j = 1, \ldots, p$, item levels $l = 1, \ldots, L_j$ and latent classes $c = 1, \ldots, C$.

- Structural model: the latent class probabilities

$$\alpha_c = P(\eta = c) \qquad \text{for } c = 1, \ldots, C.$$

Practical purposes of latent class models

- A formal statistical model for classifying ("segmenting") respondents.
- Measurement model: The patterns of item response probabilities within each class.
  - This also gives an interpretation of the 'contents' of the classes.
- Data reduction technique: aim to classify a large set of response profiles into a smaller number of classes.
  - Can construct (in 'posterior' analysis) a nominal variable grouping cases into classes, for use in subsequent analyses.
- Structural model: Estimate probabilities of the latent classes.

ML estimation of latent variable models

Before proceeding with the latent class model, a brief discussion of how latent variable models are estimated.

- This applies to both latent trait and latent class models.

We consider only **maximum likelihood** (ML) estimation.

ML estimates of the model parameters are the values of the parameters which yield a maximum value of the likelihood function

$$L = \prod_{i=1}^{n} p(\mathbf{y}_i|\mathbf{x}_i)$$

for the observed data $(\mathbf{y}_i, \mathbf{x}_i)$ for units (e.g. survey respondents) $i = 1, \ldots, n$.

- Here we include covariates $\mathbf{x}_i$, which will be used tomorrow.

Likelihood function for latent variable models

The contribution of a single unit i to the likelihood is

$$\begin{aligned} L_i = p(\mathbf{y}_i|\mathbf{x}_i) &= \int p(\mathbf{y}_i|\boldsymbol{\eta}_i, \mathbf{x}_i)\, p(\boldsymbol{\eta}_i|\mathbf{x}_i)\, d\boldsymbol{\eta}_i \\ &= \int \left[\prod_{j\in\mathcal{O}_i} p(y_{ij}|\boldsymbol{\eta}_i, \mathbf{x}_i)\right] p(\boldsymbol{\eta}_i|\mathbf{x}_i)\, d\boldsymbol{\eta}_i \end{aligned}$$

where $\mathcal{O}_i$ is the set of items y_{ij} that are observed for unit i.

- This shows that estimation can easily accommodate data where some items are missing for some units.
- If all items are observed for unit i, $\mathcal{O}_i = \{1, 2, \ldots, p\}$.

Likelihood function for latent class models

For a latent class model with single latent variable η with classes $c = 1, \ldots, L$, the likelihood contribution of a unit i is

$$L_i = \sum_{c=1}^{C} \left\{ \left[\prod_{j \in \mathcal{O}_i} p(y_{ij} | \eta_i = c, \mathbf{x}_i) \right] P(\eta_i = c | \mathbf{x}_i) \right\}$$

i.e. the integral in the likelihood is a sum over the possible values of η.

For a latent trait model, the integral does not reduce to a simple sum, so it needs to be approximated using numerical integration.

ML estimation: Numerical challenges

ML estimation of latent variable models for categorical items is a non-trivial task:

- It requires an iterative algorithm, of course.
  - Mplus uses (by default) the EM algorithm, with occasional Quasi-Newton and Fisher scoring steps
- For latent trait models, numerical integration is needed.
- The likelihood is often multimodal, and algorithms are not guaranteed to converge to a global maximum (i.e. the ML estimate).
  - It is always advisable to run the algorithm with multiple starting points. In Mplus, this is set by the `Starts` option of the `Model` command.

Specifying and fitting latent class models

A latent class model is specified by the following choices:

- The number C of latent classes.
  - The classes are taken to be unordered, and there are usually no constraints on their probabilities α_c.
- Measurement models for the items y_j
  - These are effectively standard regression models for categorical responses y_j, with dummy variables for the levels of η as explanatory variables.
  - In a single-group analysis, Mplus always uses the multinomial logistic model, i.e. items are treated as nominal even when they are specified as ordinal ("`categorical`").
  - Instead of the parameters (intercepts and loadings) of these models, we usually examine the probabilities $\pi_{jl(c)} = P(y_j = l|\eta = c)$ implied by them.

Latent class models in Mplus: Input

- The latent class variable is declared under the `VARIABLE` command:

```
Variable:
  Classes = class(3);
```

— here called `class`, with $C = 3$ latent classes.

- A latent class model is requested by the `Type=Mixture` option of the `ANALYSIS` command:

```
Analysis:
  Type=Mixture;
  Estimator=ML; ! Requests ML estimation; we always use this.
  Starts=20 10; ! Number of starts for estimation algorithm
```

- The measurement model is by default a multinomial logistic model for each item, and does not need to be specified at all
  - ...unless further constraints, starting values etc. are wanted

Latent class models in Mplus: Output

- Mplus output contains estimates both for the parameters (intercepts and loadings) of the structural and measurement models, and for corresponding probabilities
  - ...except that the item response probabilities are not shown if the items are specified as `Nominal`.
- Below and in the computer classes we will instead show the same results as presented by the `lcat` functions in R.

Engagement example: Mplus input

```
Title:
    LCAT workshop examples.
    Engagement with science and technology (EB data).
    Latent class model, 3 classes.
Data:
    File = engagement.dat;
Variable:
    Names = read talk meet protest interest informed knowledg;
    Missing = all(5 9);
    Usevariables = read-protest;
    Categorical = read-protest;
    Classes = class(3);
Analysis:
    Type=Mixture;
    Estimator=ML;
    Starts=20 10;
Savedata:
    File="tmp.dat";
    Save=Cprobabilities;
```

Identification of the latent class model

A latent variable model is statistically **identified** if different values of its parameters imply different fitted values for the data

- ...and not identified if exact same fit is produced by different parameter values.

For a latent class model, main issue of identifiability is the number C of classes. The model is *not* identified if

$$df = \{L_1 \times \cdots \times L_p - 1\} - \{(C-1) + C \times [(L_1 - 1) + \cdots + (L_p - 1)]\} < 0$$

- In our example $p = 4$, $L_1 = \cdots = L_4 = 2$ and $C = 3$, so $df = 1$. Thus the 3-class model for 4 binary items is identified, but provides only a minimally more parsimonious representation of the data than the original 2^4 table.

Identification of the latent class model

- Even when the model is identified in having not too many classes, it has an inherent but trivial non-identifiability: The labelling of the classes.
- Which class is numbered "1", which one "2" etc. is arbitrary, and all permutations of the labels give the same model.
- We should choose an ordering which is convenient for presentation.
- The `lcat` function `reorder.lcat.list` can be used (among other things) to reorder the classes:

```
workshop.eng4 <- lcat("engage_3cl.out",path="c:/lcatworkshop")
reorder(workshop.eng4,1,classes=c(3,1,2))
```

Engagement example: lcat output

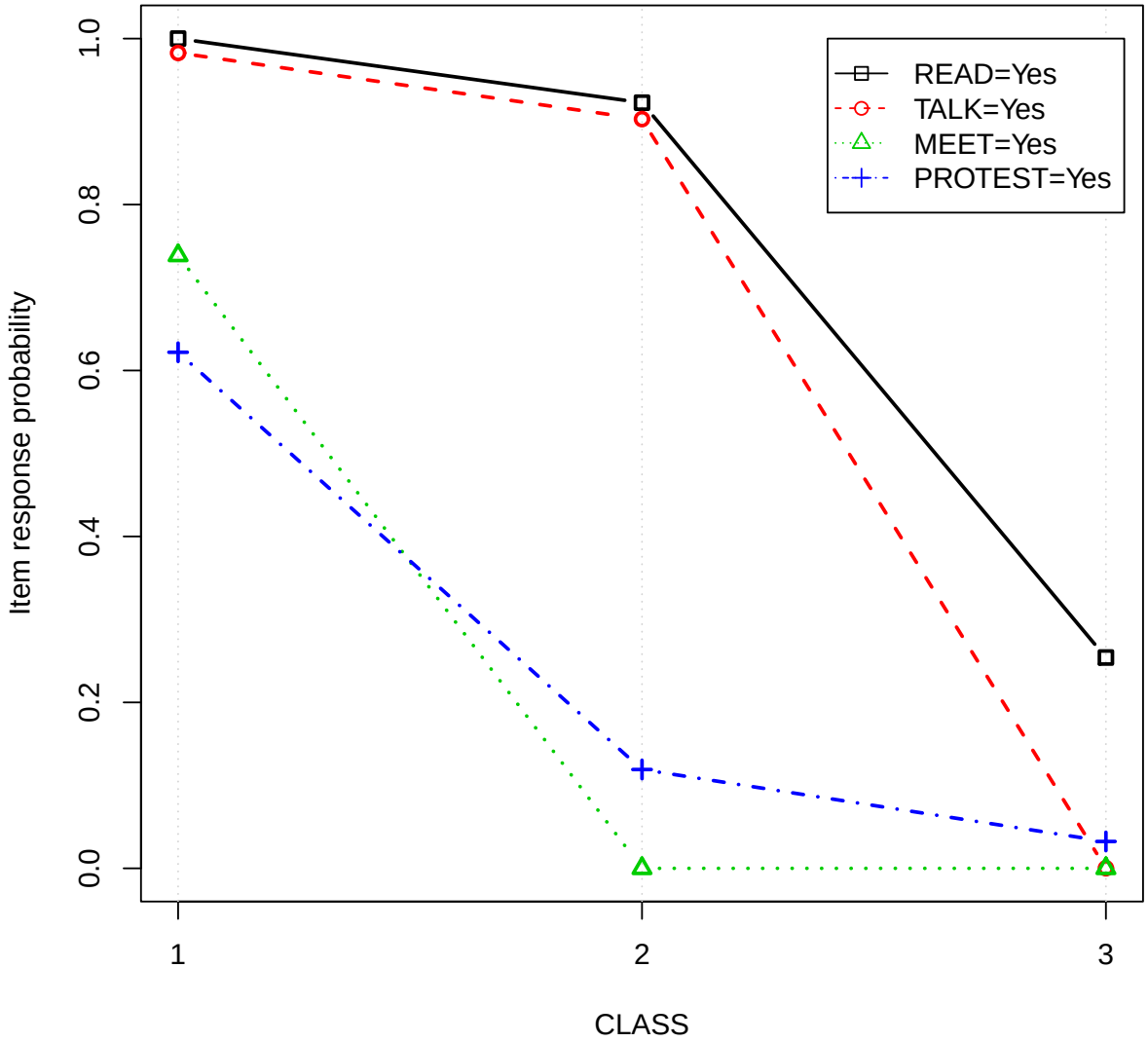
```
------------------------------------------------------------
LCAT output
Mplus file:  engage_3cl
Latent class model, latent class variable CLASS with 3 classes

Probabilities of latent classes:
                        CLASS#1 CLASS#2 CLASS#3
(All)                   0.302   0.461   0.236

Measurement probabilities:
                        CLASS#1 CLASS#2 CLASS#3
        READ#1          0.000   0.077   0.746
        READ#2          1.000   0.923   0.254

        TALK#1          0.017   0.097   1.000
        TALK#2          0.983   0.903   0.000

        MEET#1          0.261   1.000   1.000
        MEET#2          0.739   0.000   0.000

        PROTEST#1       0.378   0.881   0.968
        PROTEST#2       0.622   0.119   0.032
------------------------------------------------------------
```

Engagement example: Estimated probabilities

	Probability of 'Yes' response:		
	Class 1 ('Everything')	Class 2 ('Non-political')	Class 3 ('Nothing')
Item	$\hat{\pi}_{11(1)}$	$\hat{\pi}_{21(1)}$	$\hat{\pi}_{31(1)}$
read	$> .99$	.92	.25
talk	.98	.90	$< .01$
meet	.74	$< .01$	$< .01$
protest	.62	.12	.03
Estimated proportion ($\hat{\alpha}_c$):	.30	.46	.24

The labelling of each class is up to us, and meant to be descriptive of the profile of the response probabilities within the class.

Engagement example: Item probabilities



Allocating cases to classes

- We can assign individuals to classes based on the fitted model.
  - This is analogous to calculating trait scores for a latent trait model.
- Use the estimated conditional probabilities for membership in each class, given response profiles $\mathbf{y}$:

$$P(\eta = c|\mathbf{y}) = \frac{P(\mathbf{y}|\eta = c)P(\eta = c)}{\sum_{c'=1}^{C} P(\mathbf{y}|\eta = c')P(\eta = c')}$$

  - These are often termed *posterior probabilities* of the classes.
- Each response profile is allocated to the class for which its posterior probability is highest.

Allocating cases to classes

Uses of the class allocation:

- For data reduction: generate a new summary variable categorising each individual to a class.
  - This may be used as a derived variable in subsequent analyses.
- For data analysis: inspect the posterior probabilities to see how 'clean' the class allocation is
  - If for each profile there is one very high probability, this suggests strong clustering.
  - If not, this suggests weaker clustering. It might be interesting to see where the grey areas occur.

Class allocations in the engagement example

Item response profiles				Obs	Modal	*'Everything'*	*'Non-pol'*	*'Nothing'*
read	talk	meet	protest	freq	class	Class 1	Class 2	Class 3
yes	yes	yes	yes	177	1	.999	.001	.000
yes	yes	yes	no	109	1	.991	.009	.000
yes	no	yes	yes	4	1	.996	.004	.000
yes	no	yes	no	2	1	.954	.036	.010
yes	yes	no	no	480	2	.054	.934	.012
yes	yes	no	yes	122	2	.346	.653	.002
no	yes	no	no	37	2	.000	.589	.412
no	yes	no	yes	5	2	.002	.879	.119
yes	no	no	yes	11	2	.082	.645	.273
no	no	no	no	227	3	.000	.008	.991
yes	no	no	no	125	3	.004	.314	.682
no	no	no	yes	8	3	.000	.043	.957

Session 1.2

1.2(b): Model assessment for latent class and latent trait models

Model assessment

By **model assessment** (or "model selection") we mean the process of choosing the model(s) that we use for presentation and interpretation.

For latent trait and latent class models, this involves various choices:

- Latent trait vs. latent class
- The number of traits or number of latent classes
- Treating an item as ordinal or nominal (if it can be ordinal)
- If multiple traits, which items measure which traits.
- If multiple traits, are the traits associated.
- Tomorrow, whether parameters are equal across groups.

Fitted and observed values

When items $y_1, \ldots, y_p$ are categorical, with $L_1, \ldots, L_p$ levels, the observed data in a single-group analysis are a p-way contingency table with $K = L_1 \times \cdots \times L_p$ cells.

- We can report the data through the **observed frequencies** $\mathbf{O} = (O_1, \ldots, O_K)$
- ...and a fitted model produces **expected frequencies** (fitted values) $\mathbf{E} = (E_1, \ldots, E_K)$.

Model assessment is, one way or another, based on the comparison of **O** to **E**.

- If they are similar (in some sense), the model fits well; if not, the model does not fit well.

Challenges in model assessment

Model assessment for latent trait and latent class models is not easy:

- When the number of items p is large relative to the sample size, the contingency table can be very sparse.
  - Thus some conventional model selection statistics do not work.
- Formal theory of model selection for these models is not (and perhaps cannot be) complete, and properties of model assessment statistics are not fully understood.
  - So often use the statistics fairly informally and with rules of thumb, to guide but not entirely determine model selection.
- Often a good fit according to strict criteria is obtained only for unhelpfully complex models.
  - A need to balance fit and parsimony, to obtain an interpretable model which fits well enough
  - ...without being completely subjective about what "well enough" means.

Methods of model assessment

We will mention the following approaches:

- Likelihood ratio tests for nested models
- Global goodness of fit tests
- AIC and BIC
- Statistics based on bivariate marginal residuals

Likelihood ratio test of nested models

This is a general test which you have probably seen in other contexts.

Suppose we have fitted two models M_0 and M_1 with r_0 and r_1 parameters, where M_0 is nested within M_1.

The LR test statistic of the null hypothesis that M_0 holds is

$$G^2_{01} = 2(\log L_1 - \log L_0)$$

where L_j denotes the likelihood of model M_j.

This test statistic is referred to the χ^2 distribution with $r_1 - r_0$ degrees of freedom. Small p-value indicates that M_0 is rejected in favour of M_1.

Likelihood ratio test of nested models

Nested hypotheses which occur in latent variable modelling:

1. In a multitrait model, some cross-loadings are 0 (vs. not).
2. In a multitrait model, association between traits is 0 (vs. not).
3. In a multigroup model (tomorrow), some parameters are the same across groups (vs. not).

Likelihood ratio test of nested models

Example: In the latent trait section we considered a 2-trait model with traits TECH and NICE (slides 44–54).

Suppose we fit two models, M0 where the two traits are not associated and M1 where they are. Then in R we run the test as follows:

```
twomodels <- lcat("M0.out","c:/lcatworkshops")
twomodels <- lcat("M1.out","c:/lcatworkshops",addto=twomodels)
lcat.lrtest(twomodels,1,2)
```

```
Likelihood ratio test:
H0: M0          H1: M1
LR = 0.036        df = 1  P-value = 0.85
```

Here the hypothesis of no association between TECH and NICE is not rejected. This agrees with the Wald test shown on slide 54.

Where likelihood ratio test cannot be used

Many interesting hypotheses cannot be formulated as pairs of nested models for which the standard LR test is applicable:

- Latent class vs. latent trait model
- q vs. $q+1$ latent traits
- C vs. $C+1$ latent classes
- Ordinal vs. Multinomial logistic model for the same item.

So something else is needed.

Overall goodness-of-fit tests

Compare the observed frequencies $\mathbf{O}$ of the $K = L_1 \times \cdots \times L_p$ response patterns to the expected (fitted) frequencies $\mathbf{E}$ from a model by means of a X^2 Pearson goodness-of-fit or a likelihood ratio test G^2:

$$X^2 = \sum_{i=1}^{K} (O_i - E_i)^2 / E_i \quad \text{and} \quad G^2 = 2 \sum_{i=1}^{K} O_i \log O_i / E_i.$$

When n is large and K small and the model is true, these statistics follow approximately a χ^2 distribution with degrees of freedom equal to $(K-1) - r$, where r is the number of parameters in the fitted model. They are then test statistics of the overall goodness of fit of the model.

However, in latent variable models K is typically *not* small enough relative to n, so basing the test on the χ^2 sampling distribution is not valid.

More accurate p-value can be obtained using bootstrapping, but we do not discuss that here.

AIC and BIC

Two so-called "information criteria":

$$\begin{aligned} AIC_j &= -2 \log L_j + 2\, r_j \\ BIC_j &= -2 \log L_j + (\log n)\, r_j \end{aligned}$$

where L_j and r_j are the likelihood and number of parameters of a model M_j.

These are used to compare models (which need not be nested): The model with the *smallest* value of the statistic is preferred.

AIC and BIC do not always agree. If not, BIC prefers smaller (more parsimonious) models.

Bivariate marginal residuals

Consider the **two-way** tables of each pair of items y_i and y_j, and denote their observed frequencies $O_{rs}^{(ij)}$ for $r = 1, \ldots, L_i$ and $s = 1, \ldots, L_j$.

The corresponding expected frequencies are

$$\begin{aligned} E_{rs}^{(ij)} &= \hat{P}(y_i = r, y_j = s) \\ &= n\,\hat{P}(y_i = \text{obs}, y_j = obs) \int \hat{P}(y_i = r|\boldsymbol{\eta})\hat{P}(y_j = s|\boldsymbol{\eta})\, p(\boldsymbol{\eta})\, d\boldsymbol{\eta} \end{aligned}$$

- Here $\hat{P}(y_i = \text{obs}, y_j = \text{obs})$ is the observed proportion of observations where both y_i and y_j are observed.
  - (which is the expected proportion under MCAR nonresponse)
- For a latent class model, the integral is a sum over $\eta = 1, \ldots, C$.
- For a latent trait model, the `lcat` functions use brute-force Monte Carlo integration to approximate the integral.

Bivariate marginal residuals

Define the bivariate marginal residuals as $(O_{rs}^{(ij)} - E_{rs}^{(ij)})^2 / E_{rs}^{(ij)}$ and

$$S^{(ij)} = \sum_{r=1}^{L_i} \sum_{s=1}^{L_j} \frac{(O_{rs}^{(ij)} - E_{rs}^{(ij)})^2}{E_{rs}^{(ij)}}$$

their sums, for each pair of items i, j. We discuss the use of these to assess model fit for

- individual cells in the two-way marginal tables
- each pair of items overall
- the model overall

These assessments are exploratory, in the spirit of Bartholomew et al. (2008). Valid distributional results for bivariate residuals have been developed by Maydeu-Olivares and Joe (2006) and others. Tests based on these will be added to the `lcat` functions in the future.

Individual residuals

- We can examine individual residuals $(O_{rs}^{(ij)} - E_{rs}^{(ij)})^2/E_{rs}^{(ij)}$ to see in detail which cells in the bivariate tables are not well fitted by the data.
- Residual greater than 4 is suggestive of poor fit
  - This is loosely motivated by an analogue to the χ^2 goodness of fit test, but not formally justified. So it — like the other suggestions below — is just an informal rule of thumb.
- Below, illustrations for examples considered before.

Individual residuals

Attitudes to abortion, 1-trait model:

```
> resid(workshop.ab04,1)
    item1  item2 value1 value2 Observed Expected Residual Std.residual
1  ABORT1 ABORT2      1      1      451      452     -1.4          0.0
2  ABORT1 ABORT2      1      2       24       28     -4.3         -0.6
3  ABORT1 ABORT3      1      1      384      393     -9.1         -0.2
4  ABORT1 ABORT3      1      2       81       79      2.2          0.1
5  ABORT1 ABORT4      1      1      385      394     -8.6         -0.2
6  ABORT1 ABORT4      1      2       84       78      6.3          0.5
7  ABORT1 ABORT2      2      1       95       97     -2.3         -0.1
8  ABORT1 ABORT2      2      2      200      192      8.0          0.3
9  ABORT1 ABORT3      2      1       57       55      1.6          0.0
10 ABORT1 ABORT3      2      2      234      229      5.3          0.1
... etc.
```

All are very small, largest is 0.8.

("Std.residual" is the residual discussed above, with a sign added for convenience.)

Individual residuals

Engagement with science and technology, 2-class model:

```
> resid(workshop.eng4,2,over4=T,sort=T)
  item1   item2 value1 value2 Observed Expected Residual Std.residual
1  MEET PROTEST      2      2      181      103     78.5         60.1
2  MEET PROTEST      2      1      110      189    -79.3        -33.2
3  MEET PROTEST      1      2      146      225    -78.5        -27.5
4  MEET PROTEST      1      1      869      790     79.3          8.0
```

The only ones greater than 4 involve items **meet** and **protest**.

For the 3-class model, all the residuals are very small.

Sums of residuals for pairs of items

- The sum $S^{(ij)}$ of the bivariate residuals can be used as a quick summary of how well a model fits the observed joint distributions of pairs of items y_i, y_j
- A rough yardtick is to compare $S^{(ij)}$ to the χ^2 distribution with $L_i L_j - 1$ degrees of freedom.
  - For example, $S^{(ij)}$ being larger than the 95% quantile of this distribution is suggestive of poor fit.
- Example below: 2-trait model for 6 items on attitudes to science and technology, with no cross-loadings (see slides 45–49)

Sums of residuals for pairs of items

```
> resid(workshop.scien6i,4,sumitem2way=T)

          ENVIRON WORK   FUTURE TECHNOL INDUSTRY
COMFORT   14.2    22.4   11.5   19.7    16.7
ENVIRON           25.9*  23.1   20.5    31.2*
WORK                     9.6    19.7    29.8*
FUTURE                          22.3    28.9*
TECHNOL                                 33.6*
```

("*" indicates a sum which is greater than the 95% quantile of the χ^2 distribution with $L_i L_j - 1$ degree of freedom.)

Most large values involve item INDUSTRY. These are not improved by cross-loadings, or by modelling INDUSTRY as nominal.

Overall model fit

- We may also consider all the bivariate residuals together, to get an impression of the fit of the model overall.
- For example, examine what % of all residuals are greater than 4
  - A rough rule of thumb, at least for single-group models: Less than 10% suggests a reasonable fit.

Example: Fear of crime

- Consider four questions on fear of crime in Round 5 of the European Social Survey (2011)
  - Frequency of worry: "How often, if at all, do you worry about [crime]?"
  - Effect of worry: "Does this worry about [crime] have a [response option] [effect on the quality of your life]?"

  where [crime] was "becoming a victim of violent crime" or "your home being burgled", thus defining 4 questions in all
- 4 response options for the frequency questions, 3 for the effect questions
- Consider data on British (GB) respondents, with $n = 2421$.
- Consider latent class models with 1–7 classes.

Example: Fear of crime

Classes	AIC	BIC	%
1	16332	16390	85
2	14451	14572	41
3	14115	14301	34
4	13955	14204	16
5	13847	14160	11
6	13770	14146	1
7	13751	14191	1
%: of bivariate residuals > 4			

6 classes give a good fit to the data.

Example: Fear of crime

		Latent class					
		'Unworried'	'Occasional ineffective worry'	'Frequent ineffective worry'	'Burglary only'	'Effective worry'	'Persistent worry'
		1	2	3	4	5	6
	Probability of latent class:	0.44	0.24	0.11	0.10	0.08	0.03
Question	Response						
Violent crime: Frequency of worry	*Never*	1.00	0.00	0.00	0.47	0.00	0.00
	Just occasionally	0.00	0.99	0.21	0.37	0.43	0.08
	Some of the time	0.00	0.01	0.71	0.16	0.54	0.47
	All or most of the time	0.00	0.00	0.08	0.01	0.03	0.45
Effect of worry on quality of life	*No real effect*	1.00	0.94	0.58	1.00	0.00	0.03
	Some effect	0.00	0.06	0.37	0.00	0.99	0.40
	Serious effect	0.00	0.00	0.05	0.00	0.01	0.58
Burglary: Frequency of worry	*Never*	0.61	0.23	0.30	0.00	0.00	0.00
	Just occasionally	0.32	0.58	0.22	0.33	0.36	0.10
	Some of the time	0.07	0.18	0.39	0.45	0.51	0.30
	All or most of the time	0.00	0.00	0.08	0.22	0.13	0.60
Effect of worry on quality of life	*No real effect*	1.00	1.00	1.00	0.41	0.00	0.03
	Some effect	0.00	0.00	0.00	0.56	0.99	0.41
	Serious effect	0.00	0.00	0.00	0.03	0.01	0.55

That is all for day 1. See you tomorrow for more.

`stats.lse.ac.uk/lcat/`