UNIVERSITY OF CAMBRIDGE

The Psychometrics Centre

Introduction to Mplus: Latent variables, traits and classes

Peterhouse College, Cambridge

24th -25th January 2011

This course is prepared by

Anna Brown, PhD ab936@medschl.cam.ac.uk

Research Associate

Tim Croudace, PhD tjc39@cam.ac.uk

Senior Lecturer in Psychometric Epidemiology

UNIVERSITY OF CAMBRIDGE

School of Clinical Medicine > Department of Psychiatry

This course is funded by the ESRC RDI and hosted by The Psychometrics Centre

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

RDI Researcher Development Initiative

UNIVERSITY OF CAMBRIDGE The Psychometrics Centre

Day 2

- The theme of today will be models involving multiple groups.
- We start with logistic regression (extension of the multiple regression topic), using it for detecting DIF
- We explore tests of group invariance using latent trait models with continuous and categorical variables.
- We discuss the group-covariate approach and the multi-group approach with equivalence constraints.
- Finally, we introduce the latent class analysis (LCA) and show how to use Mplus to explore the presence of unobserved homogeneous groups in the data.

Regression with binary dependent variables

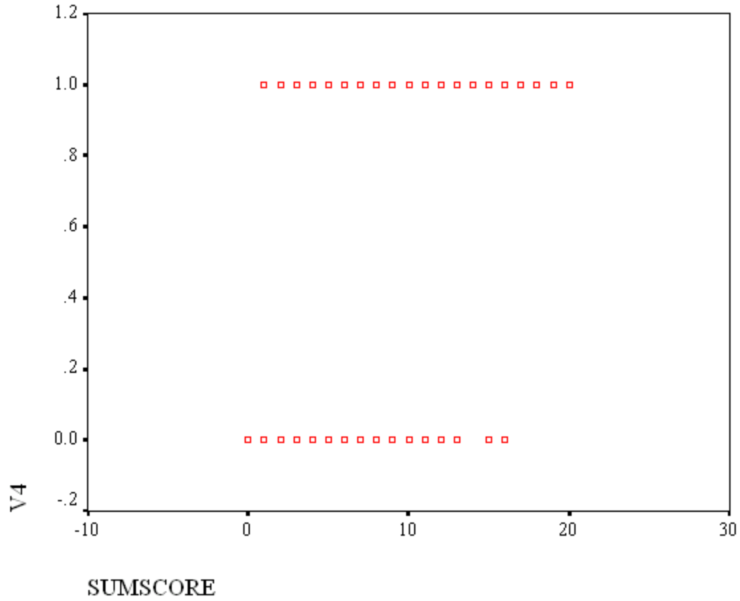
LOGISTIC REGRESSION

Binary variables: Example

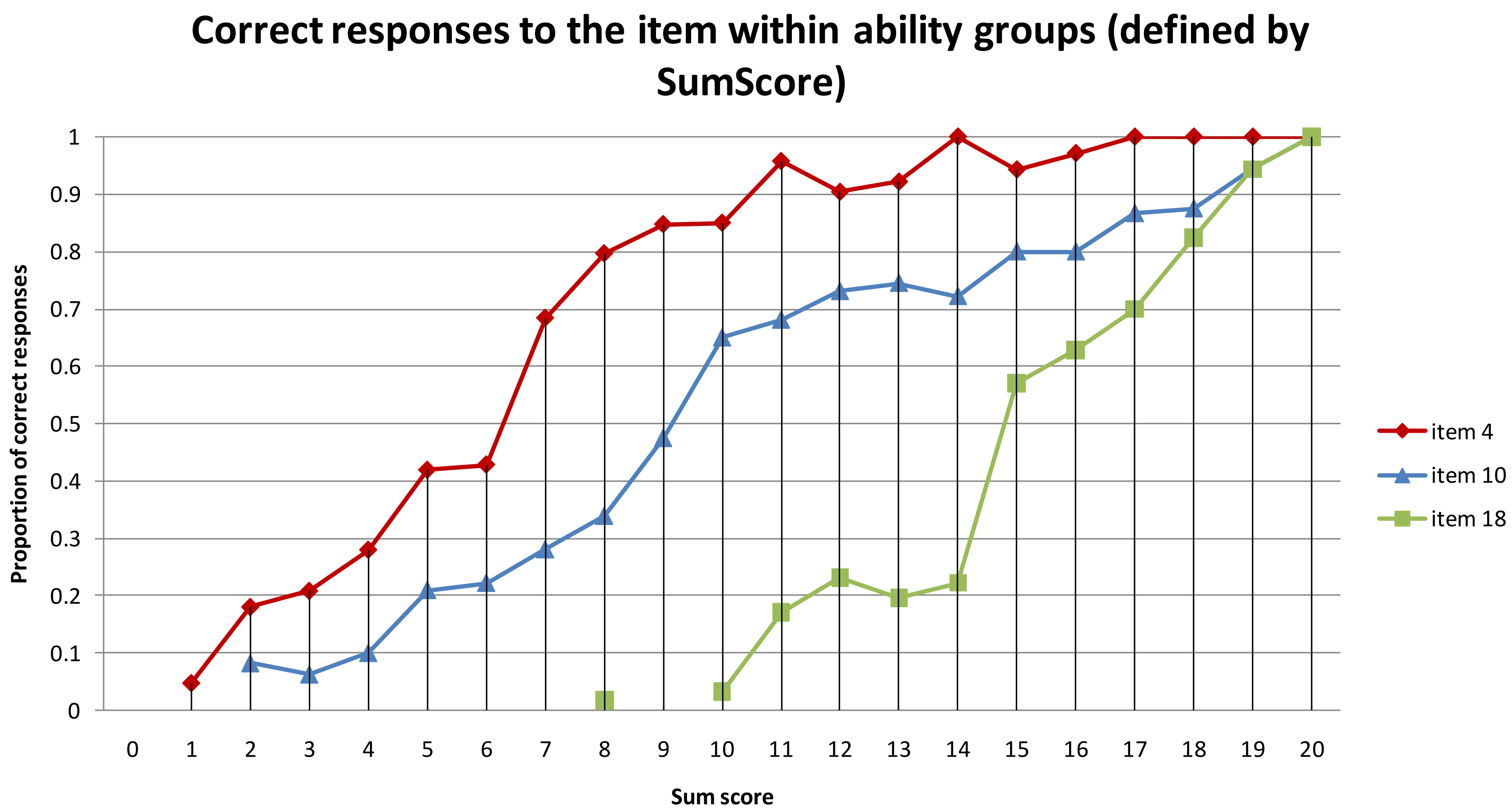
- Consider a test measuring *aptitude for mathematics* with 20 short tasks ("items").
- Each item is an experiment with 2 possible outcomes – correct or incorrect.
- Each item is assumed to 'sample' one underlying (latent) dimensions of 'ability'.
- Can we predict what the item response (binary outcome variable) will be, given the ability (continuous variable)?
  - We can count items that were answered correctly for each examinee (number correct), and use this score as "mathematical aptitude" score.

Linear regression is inappropriate

- Although we expect that ability should be quite a strong predictor of correct response, relationship is clearly not linear.
- We need another type of relationship between these variables
- We can look at proportions of correct responses on this item for each separate value of ability score



Likelihood of correct response as function of ability



UNIVERSITY OF CAMBRIDGE
The Psychometrics Centre

Log odds

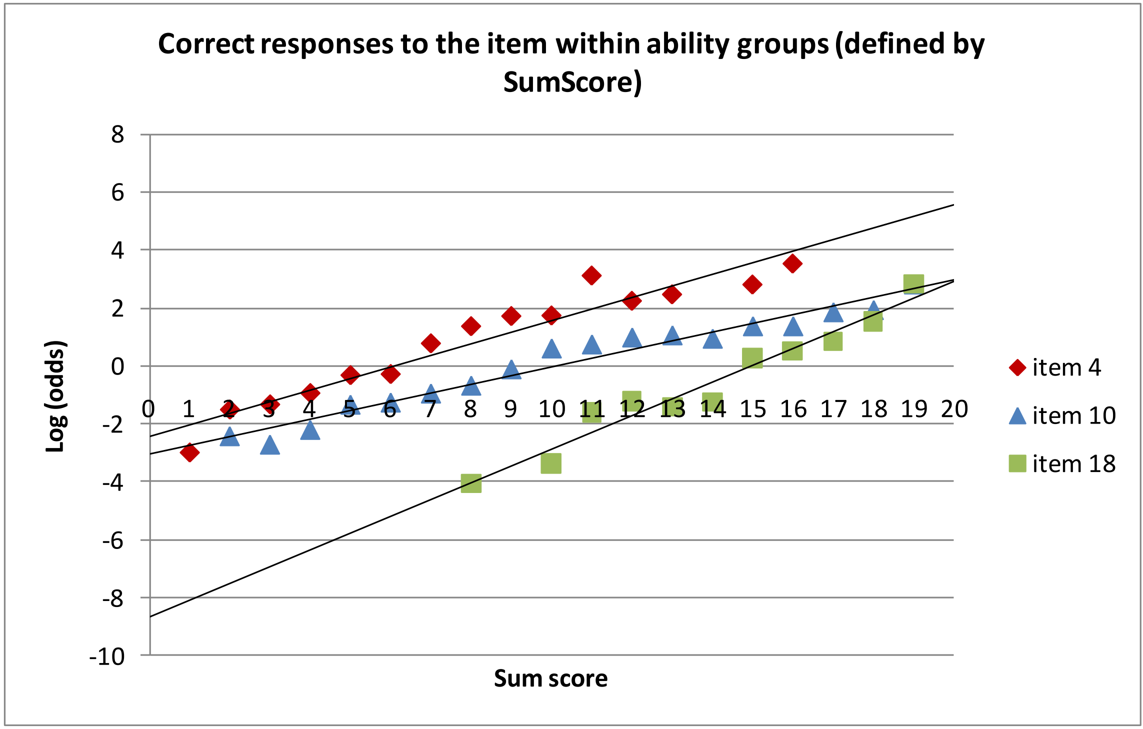
Odds = p/(1-p)

i.e. Probability of event occurring ÷ Probability of event not occurring

Log odds = ln(p/(1-p))

Happens to be a linear function of ability

ln(p/(1-p)) = a + b*X



Parameters in logistic regression

- Probability of keyed response on the item

$$P(u_i = 1 \mid x) = \frac{e^{(a_i + b_i x)}}{1 + e^{(a_i + b_i x)}}$$

- Slope parameter *b*
- Intercept parameter *a*
- Attention! Mplus prints threshold τ, which equals –*a*

$$P(u_i = 1 \mid x) = \frac{e^{(b_i x - \tau_i)}}{1 + e^{(b_i x - \tau_i)}}$$

Logistic regression example

- A 20-item ability test, N=1000 examinees
  - 717 majority group, 283 minority group.
- Each item is coded 1=correct or 0=incorrect.
- The number of items answered correctly for each examinee (number correct) is used as “mathematical aptitude” score.
- Predict the probability of correctly answering a particular item given the ability score
  - Then see if the group membership adds to this prediction

Ability test: logistic regression syntax

```
VARIABLE: NAMES ARE i1-i20 group;
    USEVARIABLES ARE i10 group ability;
    CATEGORICAL ARE i10;
DEFINE:
   ability = SUM(i1-i9 i11-i20); !sum score excluding item 10
ANALYSIS:
    ESTIMATOR=ML;
MODEL:
i10  ON  ability group@0;     !fix in the first run and then release
```

Regression on the ability score

i10 ON ability group@0;

- Log likelihood = -409.147 (2 parameters)
- R-square = 0.577 (se=0.031)
- Estimates

I10 ON ABILITY	0.366	(0.022)	p=0.000	*this is **b***
I10$1	3.728	(0.231)	p=0.000	*this is -**a***

LOGISTIC REGRESSION ODDS RATIO RESULTS

I10 ON ABILITY 1.44

exp(*0.366*)=1.442

Interpretation: as ability increases by 1 point, the odds of getting item 10 right increases by 1.44

Adding the grouping variable

i10 ON ability group;

- Log likelihood = -386.723 (3 parameters)
- R-square = 0.625 (se=0.030)
- Estimates

I10	ON	ABILITY	0.381	(0.023)	p=0.000	*this is **b1***
		GROUP	-1.391	(0.218)	p=0.000	*this is **b2***
I10$1			3.513	(0.236)	p=0.000	*this is **-a***

LOGISTIC REGRESSION ODDS RATIO RESULTS

I10	ON	ABILITY	1.464	*exp(0.381)*=1.464
		GROUP	0.249	*exp(-1.391)*=0.249

Interpretation: as ability increases by 1 point, the odds of getting item 10 right increases by 1.464; for group 1 (minority) the odds of getting item 10 right

Differential Item Functioning

- In fact, what we have just done is tested for uniform DIF
- DIF is present when there is lower (or higher) chance for members of a certain group to get the item correct, *given the same level of ability*
- Logistic regression is a popular method of testing for DIF
- How do we know DIF was present?
  - Group variable improved the prediction
    - Log likelihood improved (test difference *2, as chi-square with 1 degree of freedom)
    - R-square improved (large effect size > 0.07, medium > 0.035)

Calculating probabilities

- Calculating the probability of getting item right

  L = 0.381*x – 1.391*g – 3.513

  $$P(u_i = 1 \mid x) = \frac{e^L}{1 + e^L}$$

  Note the reversed threshold to make the intercept parameter

- For an individual with test score x=10
  - If from the majority group (g=0)

    L=0.381*10 – 1.391*0 -3.513 =0.297

    P=exp(0.297)/(1+exp(0.297))=0.574
  - If from the minority group (g=1)

    L=0.381*10 – 1.391*1 -3.513 =-1.094

    P=exp(-1.094)/(1+exp(-1.094))=0.251

Observed grouping

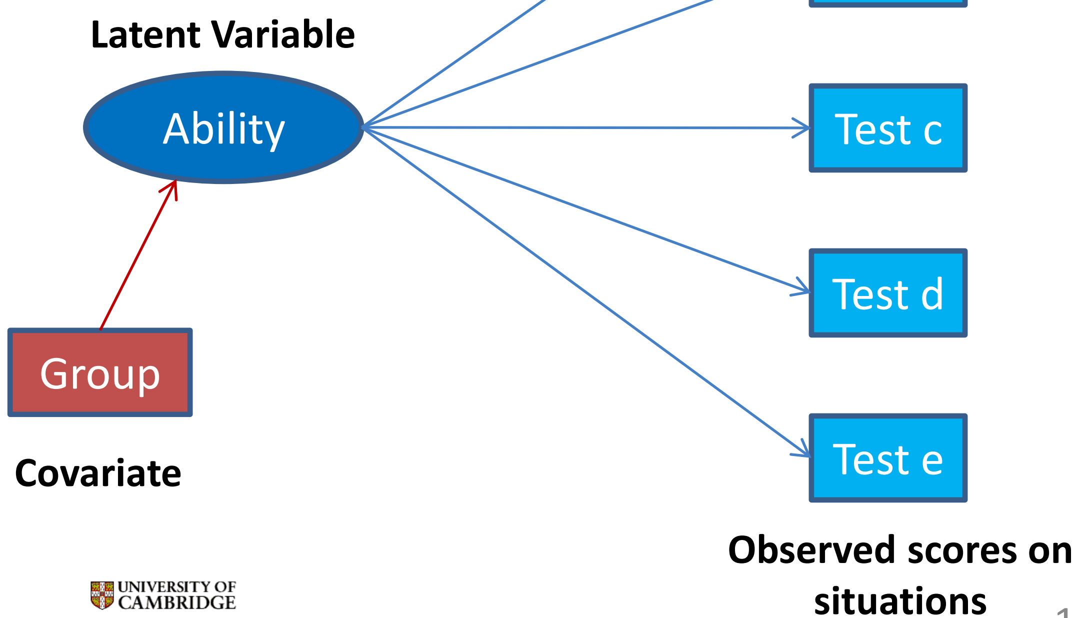
GROUPING AS COVARIATE

Inductive reasoning test

- Fragment of a paper & pencil test assessing aptitude for finding patterns and rules and applying them
- Consists of cards describing different problems ("situations") – we will consider 5 here:
    A. *"Frequent flyer" scheme rules*
    B. *Figures on employment of graduates*
    C. *Rules for video conference booking*
    D. *Tax duties on goods at an airport*
    E. *Stock records on books*
- There are 3 problems to solve about each "situation"
- We consider data from $n=451$ student volunteers, out of which 356 were native English speakers, 96 non-native

The common factor model

- We can use the observed "*nat_eng*" variable as a covariate in the model
- To test if the inductive reasoning ability (as measured by this test) varies for native and non-native speakers



CFA with covariate syntax

```
TITLE:  CFA with covariate on Inductive Reasoning test
DATA:  FILE IS IndReasoning.dat;
VARIABLE:  NAMES ARE a b c d e nat_eng;
  ! 1=native english speaker; 2=non-native speaker
   USEVARIABLES ARE ALL;
   MISSING ARE  .;
ANALYSIS:  ESTIMATOR IS ML;
MODEL:
   Ind_R  BY  a b c d e;  !first loading is fixed to 1 by default
   Ind_R  ON  nat_eng d@0;  !we will release this later
OUTPUT: MODINDICES (ALL); STAND;
```

CFA with covariate - Results

- Regression path estimation significant (standardized estimate)

  IND_R ON NAT_ENG -0.262 (SE=0.063; p=0.000)

- Model fits reasonably well

  Chi-Square 15.352 (df = 9; P = 0.082)

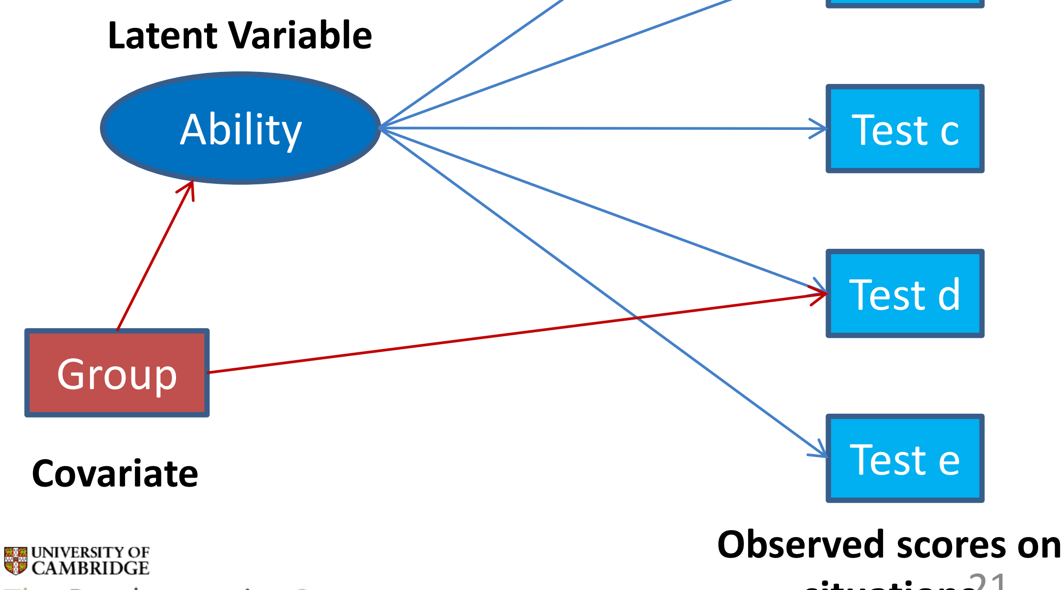
  RMSEA = 0.040 90 Percent C.I. (0.000 0.073)

  CFI = 0.946

- Explanation for the result? Can we conclude from this data that the non-native speakers' have lower inductive reasoning ability?

Direct effect of grouping on item response

- In a fair test, all differences in performance on subtests should be explained by the difference in inductive reasoning ability
- If this is not the case, and a direct path exists between the grouping and the subtest variable, we observe Differential Item Functioning (DIF)



UNIVERSITY OF CAMBRIDGE
The Psychometrics Centre

Direct effect of grouping variable

- Direct regression path just significant (standardized estimate)

  IND_R ON NAT_ENG -0.307 (SE=0.067; p=0.000)

  D ON NAT_ENG **0.112** (SE=0.057; p=0.049)

- Model fits better

  Chi-Square 11.206 (df = 8; P = 0.190)

  RMSEA = 0.030 90 Percent C.I. (0.000 0.067)

  CFI = 0.973

- Explanation for the result?

Observed grouping

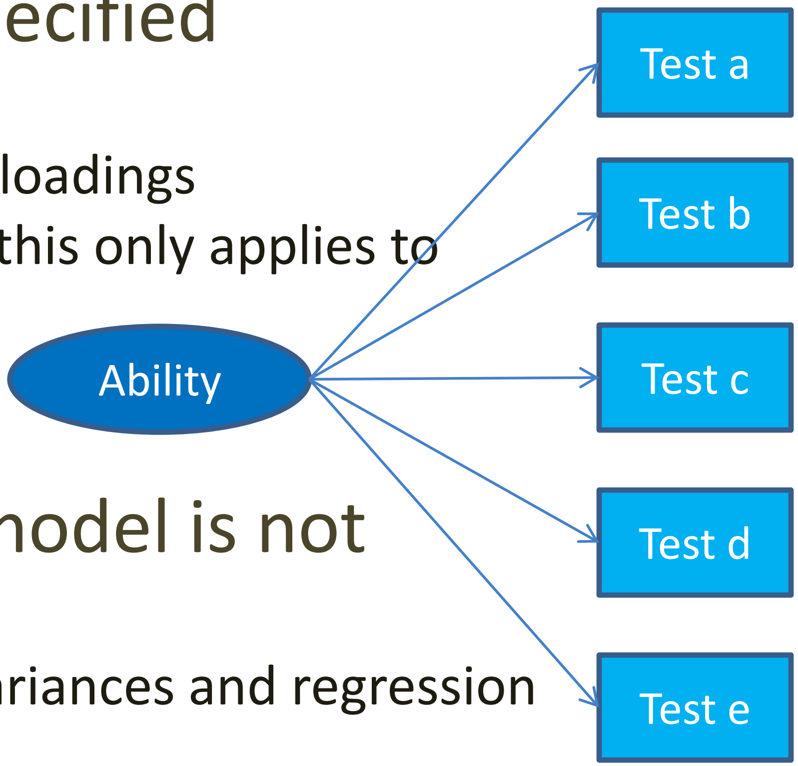
MULTI-GROUP ANALYSIS

CFA – multigroup approach

- Approach with covariates was only able to detect differences in means (intercepts), or uniform DIF
- Confirmatory approach with multiple groups can be used to test for any combinations of the following
  - Measurement parameters (measurement invariance)
    - Intercepts (*item difficulty – uniform DIF*)
    - Factor loadings paths (*item discrimination – non-uniform DIF*)
    - Residual variances
  - Structural parameters (population heterogeneity)
    - Latent means
    - Latent variances/covariances/regression paths
- One of the most attractive features is that more than 2 groups can be tested

Defaults for multi-group setup

- The measurement part of the model is assumed invariant if not specified otherwise
  - Intercepts, thresholds, factor loadings
  - (except error variances – but this only applies to continuous indicators)



- The structural part of the model is not assumed invariant
  - Factor means, variances, covariances and regression coefficients

Syntax for multi-group analysis

- Testing for ***measurement*** invariance using default settings:

```
VARIABLE:   <all commands as before>
GROUPING IS nat_eng (1=native, 2=non-native);
ANALYSIS:  ESTIMATOR  IS  ML;
MODEL:    Ind_R BY a b c d e;  !overall part
OUTPUT:  MODINDICES (ALL 3.84);
```

- Examine the output – which parameters does Mplus constrain to be equal?

Testing for measurement invariance

- The default model (measurement model constrained and structural model free) does not quite fit the data:

  Chi-Square 29.638 (df = 18, P-Value = 0.0411)

  RMSEA = 0.054 90 Percent C.I. 0.011 0.087

  CFI = 0.884

- Examining the modification indices:
  - Factor loading to test d needs freeing

  MODEL non-native: Ind_R BY d*;

    - Loading estimated 2.199 for native group and 0.581 (n/s) for non-native
    - Now the model fits: chi-square 21.980 (df=17, p=0.1855)

Measurement invariance model parameters

- Measurement part - Factor loadings and intercepts are the same across groups
- Factor means and variances
  - Native speakers mean= 0 (fixed), var=0.090
  - Non-native speakers mean = -0.239, var = 0.116
- Looks like the non-native group is different in terms of both their mean and variance

Testing for equality of means and variances

- Imposing parameter constraints (one by one)

```
MODEL:
  Ind_R BY a b c d e;     !overall part
  Ind_R (1);
  ![Ind_R] @0;   !this will imply equality of means
MODEL non-native: Ind_R BY d*; !freeing factor loading
```

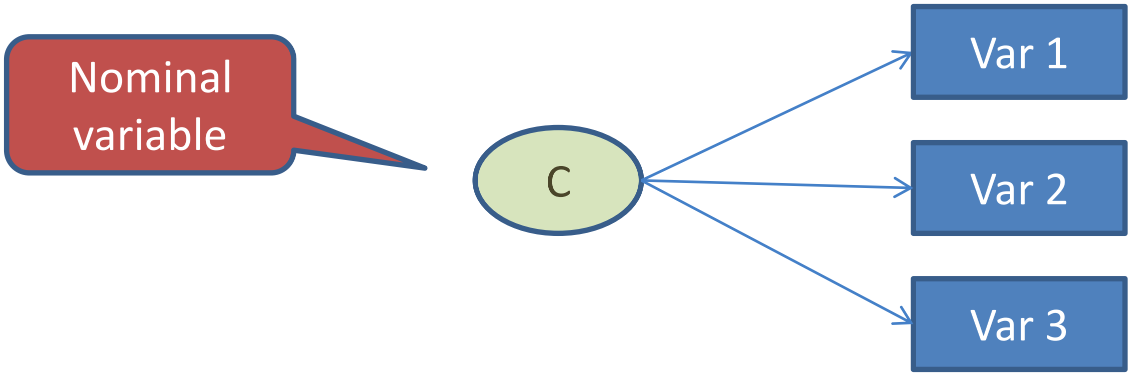
- The variances are not significantly different
  - Chi-square 22.343 (df=18, p=0.217)
- The means are different
  - chi-square 39.996 (df=19, p=0.0033)

Unobserved grouping

LATENT CLASS ANALYSIS

Aims of Latent Class Analysis

- The aim of LCA is to reduce the complexity of data by explaining the associations between the observed variables in terms of membership of a small number of unobserved (latent) classes
- Typical applications: learning theory, psychiatric diagnosis, medical diagnosis.
- Latent class analysis is available for continuous, ordinal, nominal and count observed variables.



LCA with binary variables

- The latent class model for p binary variables with C latent classes makes the following assumptions:
  - i) The n cases are a random sample from some population and every case in that population belongs to just one of the C latent classes
  - ii) The probability of giving a positive response to a particular item is the same for all cases in the same class but may be different for cases in different classes
  - iii) Once it is known to which latent class a case belongs, then the responses to different items are conditionally independent (no remaining within class association)

Example: Diagnosis of myocardial infarction

- Rindskopf and Rindskopf (1986) – data from a coronary care unit where patients were admitted to rule out "heart attack"
- Each of n=94 patients were assessed on four test criteria with 1= test result positive and 0= test negative
  - [Q-wave] – q-wave in ECG
  - [History] – classical clinical history
  - [LDH] – having a flipped LDH
  - [CPK] – CPK-MB
- We explore 2 classes (with and without MI) = "latent/true diagnoses"

CPK	LDH	History	Q-wave	count
1	1	1	1	**24**
1	1	1	0	**5**
1	1	0	1	**4**
1	1	0	0	**3**
1	0	1	1	**3**
1	0	1	0	**5**
1	0	0	1	**2**
1	0	0	0	**7**
0	1	0	0	**1**
0	0	1	0	**7**
0	0	0	0	**33**

What is estimated?

- In a simple LCA model with p categorical variables and C classes (like the MI example), we estimate two types of probabilities:

1. Probabilities of correct responses to each item p, given the latent class (these are called conditional probabilities)
2. Probability of belonging to class c (unconditional probability/class membership)

In clinical and epidemiological research 2) are prevalence of classes in the population.

LCA model, exact fit

- With p items, there are 2^p possible response patterns
- Observed (O) and expected (E) frequencies of each response pattern can be computed
- Pearson chi-square $$\chi_p^2 = \sum_r \frac{(O_r - E_r)^2}{E_r}$$
- Likelihood ratio test $$G = 2 \sum_r O_r \ln \left(\frac{O_r}{E_r} \right)$$
- For large n and small p, these statistics follow a chi-square distribution (BUT n is often small and p large! – sparse tables)
- The degrees of freedom are equal to the number of response patterns minus model parameters minus one.

  $df = 2^p - [pC - (C - 1)] - 1$

Mplus syntax for LCA

```
TITLE: Rindskopf & Rindskopf MI data
DATA: FILE IS MIdata.dat;
VARIABLE: NAMES ARE qwave history ldh cpk;
   CATEGORICAL ARE ALL;   ! binary indicators
   CLASSES = c (2); !two latent diagnosis classes
ANALYSIS: TYPE = MIXTURE;
OUTPUT: TECH 10;
```

The TECH10 option is used to request univariate, bivariate, and response pattern model fit information for the categorical dependent variables in the model.

MI data – model fit

- Degrees of Freedom $2^4-(2*4+1)-1 = 6$
- Pearson Chi-Square 4.223 ($p=0.647$)
- Likelihood Ratio Chi-Square 4.293 ($p=0.637$)
- The model fits well
  - but often we cannot interpret these Chi-square tests; particularly if they diverge a lot.
  - What to do instead?

MI data - Observed and expected counts

Response Pattern	Frequency Obs	Frequency Est	Stand. Residual	Chi-square Pearson	Chi-square Loglike.
1	24.00	21.62	0.58	0.26	5.01
2	5.00	6.63	-0.66	0.40	-2.82
3	4.00	5.70	-0.73	0.51	-2.83
4	3.00	1.95	0.76	0.57	2.59
5	3.00	4.49	-0.72	0.50	-2.43
6	5.00	3.26	0.98	0.93	4.28
7	2.00	1.18	0.75	0.56	2.10
8	7.00	8.17	-0.43	0.17	-2.16
9	1.00	0.89	0.12	0.01	0.24
10	7.00	7.78	-0.29	0.08	-1.48
11	33.00	32.11	0.19	0.02	1.80

MI model results - probabilities

Latent Class 1 — No MI

QWAVE	Estimate	S.E.
Category 1	1.000	0.000*
Category 2	0.000	0.000*
HISTORY		
Category 1	0.805	0.063
Category 2	0.195	0.063
LDH		
Category 1	0.973	0.027
Category 2	0.027	0.027
CPK		
Category 1	0.804	0.068
Category 2	0.196	0.068

Latent Class 2 — MI

QWAVE	Estimate	S.E.
Category 1	0.233	0.078
Category 2	0.767	0.078
HISTORY		
Category 1	0.209	0.065
Category 2	0.791	0.065
LDH		
Category 1	0.172	0.070
Category 2	0.828	0.070
CPK		
Category 1	0.000	0.000*
Category 2	1.000	0.000*

Specificity = conditional probability of having this symptom

MI model results - thresholds

Latent Class 1 — No MI

Thresholds	Estimate	S.E.
QWAVE$1	15.000	0.000
HISTORY$1	1.417	0.400
LDH$1	3.588	1.015
CPK$1	1.414	0.429

Latent Class 2 — MI

Thresholds	Estimate	S.E.
QWAVE$1	-1.191	0.436
HISTORY$1	-1.333	0.391
LDH$1	-1.571	0.492
CPK$1	-15.000	0.000

MI data – prevalence

- Unconditional probability of having MI

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON THE ESTIMATED MODEL

Latent Classes		
1	50.96639	0.54220
2	43.03361	0.45780

Prevalence of MI is 46%

Plot results and save class memberships

- To plot conditional probabilities

PLOT: TYPE IS PLOT3;

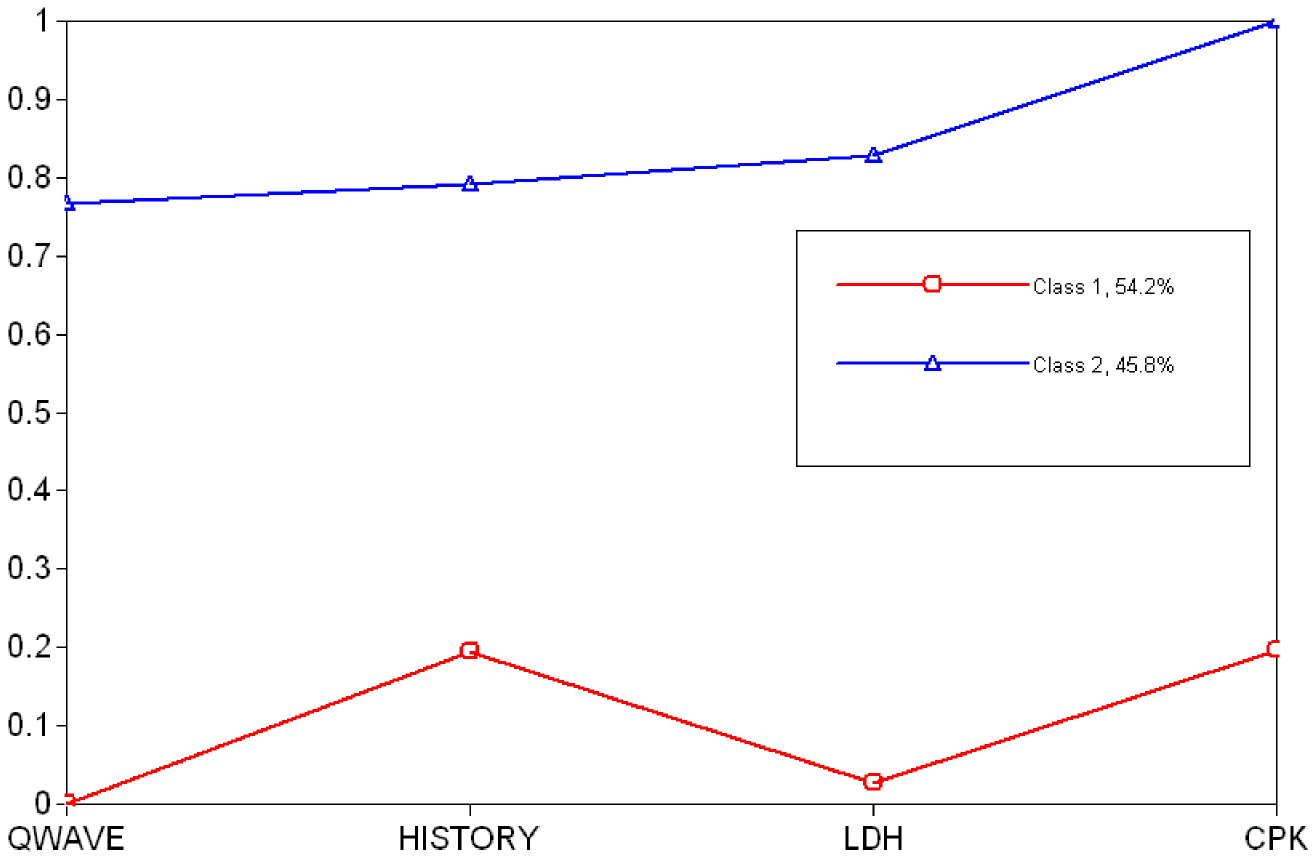
SERIES ARE qwave(1) history(2) ldh(3) cpk(4);

- To save class memberships (probabilities of belonging to class 1 and 2, and the most likely class)

SAVE: FILE IS ResultsMIdata.dat;

SAVE=CPROBABILITIES;

Estimated conditional probabilities



UNIVERSITY OF CAMBRIDGE
The Psychometrics Centre

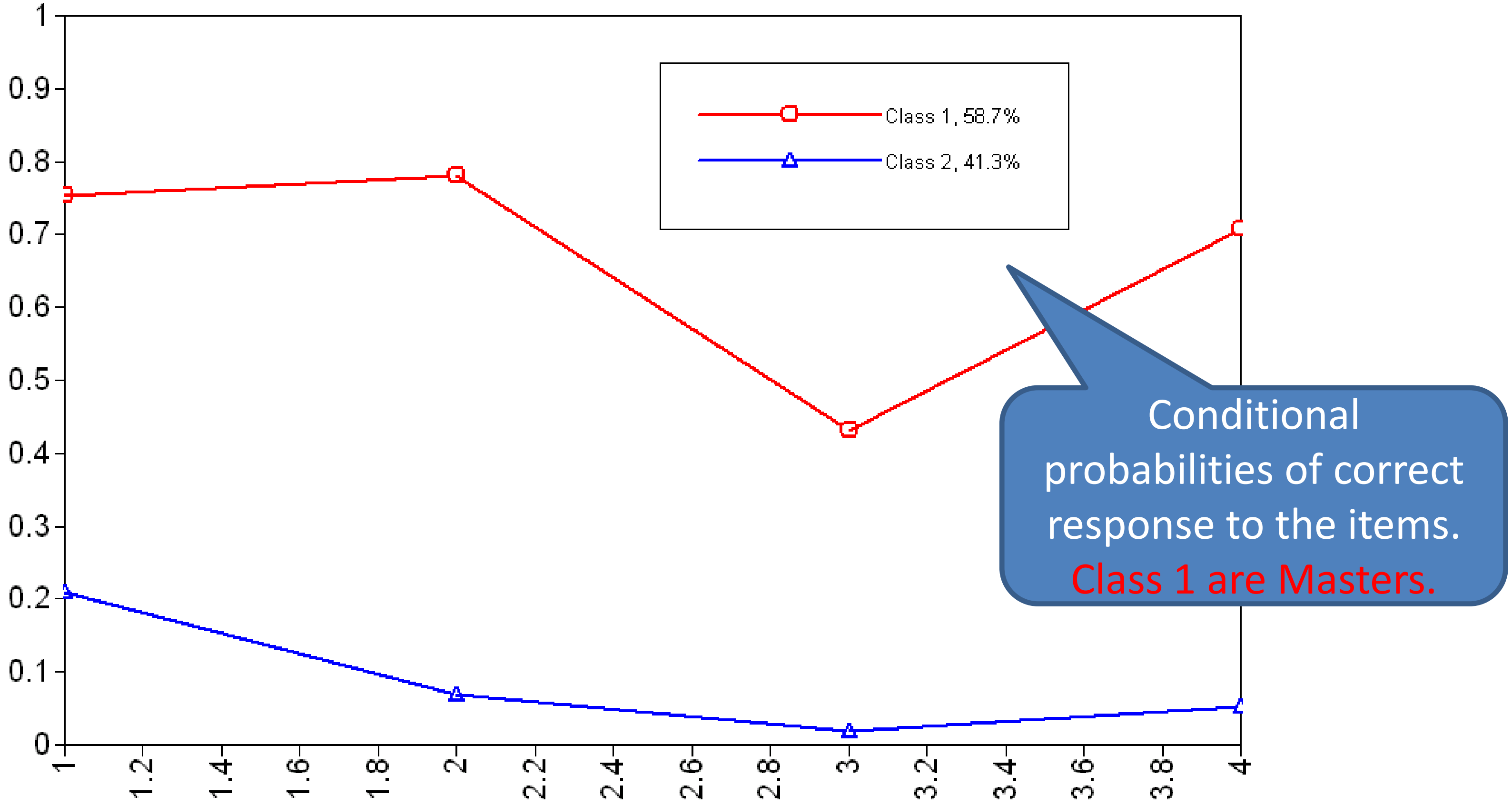
Practical: Mastery model

- Macready and Dayton's Mastery model
- Four test items selected at random from a domain of items testing mastery in the multiplication of a two-digit number by a three- or four-digit number.
- Items are coded 0=fail, 1=pass
- N=142 respondents are expected to belong to one of the two groups: Masters and Non-Masters.

  - Bartholomew, D.J., Steele, F., Moustaki, I. and Galbraith, J. (2008) Analysis of Multivariate Data for Social Scientists. Chapman and Hall/CRC.

Observed	Response pattern
15	1111
23	1101
7	1110
4	0111
1	1011
7	1100
6	1001
5	0101
3	1010
2	0110
4	0011
13	1000
6	0100
4	0001
1	0010
41	0000

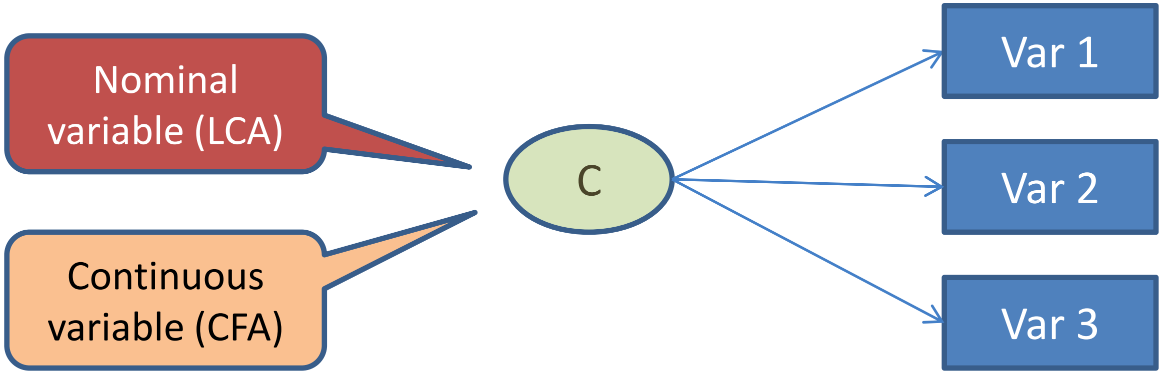
Mastery model: Estimated conditional probabilities



Model fit: Pearson Chi-square 9.459 (df=6)

LCA versus CFA

- An alternative model to explain the variation in the item responses is the latent trait model
- Variation in the latent factor (continuous variable) explains the variation in item responses
- In this example, the responses are binary and the logistic regression is used to link the responses to the latent trait – this is actually an IRT model!



Thank you

- Please give us your feedback
- Our contact details are on the slide 1 of each day
- The Psychometric Centre website

http://www.psychometrics.ppsis.cam.ac.uk/

Appendix

EFA WITH TARGET ROTATION

Target rotations

- Target rotation (Browne, 2001) is used to specify target factor loading values to guide the rotation of the factor loading matrix
- More control than in EFA but more freedom than CFA
- Used for cross-validation with more flexibility than CFA
  - Checking similarity of factor structure

Target rotation – technical detail

- For TARGET rotation, a minimum number of target values must be given for identification
  - For oblique rotation, the minimum is $m(m-1)$ where m is the number of factors.
  - For orthogonal rotation, the minimum is $m(m-1)/2$.
- The ROTATION = TARGET option has been available from version 5.1

TARGET rotation syntax

- The target values are specified in a BY statement using the tilde (~) symbol, for example:

  f1 BY y1-y6 y1~0 (*1);

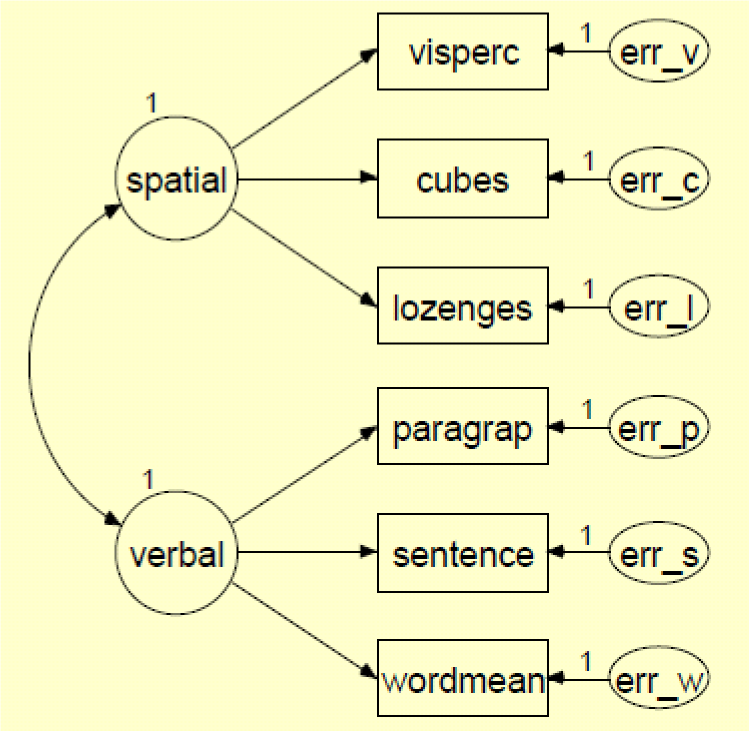
  f2 BY y1-y6 y6~0 (*1);

  – here the target factor loading values for indicator y1 for factor f1 and y5 for factor f2 are zero;

  – (*1) tells Mplus that f1 and f2 belong to the same loading matrix – i.e. one rotation is sought here.

Intelligence test data

- Holzinger-Swineford data
- Six intelligence tests
- Two groups – boys and girls
- Let's use this simple teaching example for practicing target rotation



Loadings to be used as target

- First we run EFA for boys only

PROMAX ROTATED LOADINGS

	1	2
VISPERC	0.529	0.073
CUBES	0.459	-0.044
LOZENGES	0.736	-0.043
PARAGRAP	0.231	0.698
SENTENCE	-0.095	0.925
WORDMEAN	0.216	0.663

Specifying the target loadings

```
ANALYSIS:  ESTIMATOR IS ML;
    ROTATION=TARGET; !oblique is default

MODEL:
spatial BY visperc* cubes lozenges paragrap
   sentence~0 wordmean (*1);
verbal BY visperc~0 cubes lozenges paragrap
   sentence wordmean (*1);

OUTPUT: STAND;
```