Latent Variable Modelling of Categorical Data

Tools of Analysis for Cross-National Surveys

Training workshop May 2012

Class 4 instructions: Latent trait and latent class models for multiple groups, with differential item functioning

The data for these exercises are taken from the latest (2008-9) wave of the European Values Survey:

EVS (2011): European Values Study 2008: Integrated Dataset (EVS 2008). GESIS Data Archive, Cologne. ZA4800 Data file Version 3.0.0, doi:10.4232/1.11004
http://info1.gesis.org/dbksearch/sdesc2.asp?no=4800&db=e&doi=10.4232/1.11004

We will focus on a set of items asking respondents how much confidence they have in a range of institutions. We focus on a subset of the items for the analyses here, but include the full list of items in the data set, so that you may explore models using other items if you wish. The institutions are:

Variable name	**Description**
church	church
army	armed forces
educ	education system
press	the press
tu	trade unions
police	the police
parlt	parliament
civil	civil service
welfare	social security system
eu	European Union
nato	NATO
un	United Nations
health	health care system
justice	justice system
company	major companies
environ	environmental organisations
party	political parties
govt	government

In the data set they are provided in two forms. Variable names followed by a '4' (e.g. church4) are items with 4-point scales (with coding reversed from the original) so that:

1 = none at all
2 = not very much
3 = quite a lot
4 = a great deal

Variable names followed by a '2' (e.g. church2) are binary items recoded from the originals so that:

0 = none or not very much
1 = quite a lot or a great deal.

Project funded by the Economic and Social Research Council grant RES-239-25-0022 under the Comparative Cross-National Research Methods initiative

The models fitted in Class 3 and Class 4 are multiple-group models, where the groups are countries. The data file that you are using contains responses for the complete set of countries in the European Values Survey, but we will focus on those for Great Britain (country code 45), Northern Ireland (country code 46) and Greece (country code 19).

If you would like to try the analyses with different sets of countries you may wish to know the country codes:

1 Albania
2 Azerbaijan
3 Austria
4 Armenia
5 Belgium
6 Bosnia Herzegovina
7 Bulgaria
8 Belarus
9 Croatia
10 Cyprus
11 Northern Cyprus
12 Czech Republic
13 Denmark
14 Estonia
15 Finland
16 France
17 Georgia
18 Germany
19 Greece
20 Hungary
21 Iceland
22 Ireland
23 Italy
24 Latvia
25 Lithuania
26 Luxembourg
27 Malta
28 Moldova
29 Montenegro
30 Netherlands
31 Norway
32 Poland
33 Portugal
34 Romania
35 Russian Federation
36 Serbia
37 Slovak Republic
38 Slovenia
39 Spain
40 Sweden
41 Switzerland
42 Turkey
43 Ukraine
44 Macedonia
45 Great Britain
46 Northern Ireland
47 Kosovo

1. Review models fitted under item invariance

- Print the list of models fitted in Class 3 (we will only work with binary items in this session, so just review that set of models):

```
bin3gbnigr
```

2. 1 latent trait; 3 groups; 3 binary items: trait means vary across groups; intercept for one item varies across groups

- Mplus input file: **1trait_1c_binary_gbnigr.inp**.

- Model: 1 trait, 3 binary items (police2 parlt2 justice2), 3 groups = respondents from Great Britain ('gb'), Northern Ireland ('ni') and Greece (gr); trait variances fixed to 1 for all three countries (groups); means allowed to differ. In addition, intercept or constant for one item (police2) allowed to vary across countries

- R command to read in the data and apply the LCAT function, and read the results in two different formats:

```
bin3gbnigr<-lcat("1trait_1c_binary_gbnigr.out",sessionpath,addto=
bin3gbnigr)
print(bin3gbnigr,6)
print(bin3gbnigr,6,alt=T)
```

- Plotting ICCs for the higher level of response (trust quite a lot or a great deal) for all 3 countries, for the item police2 whose intercept is allowed to vary by country:

```
plot(bin3gbnigr,6,groups=1:3,items=1,levels=2)
```

- Compare with the same for the model where intercepts are constrained to be equal across groups:

```
plot(bin3gbnigr,2,groups=1:3,items=1,levels=2)
```

3. 1 latent trait; 3 groups; 3 binary items: trait means vary across groups; intercept and slope for one item varies across groups

- Mplus input file: **1trait_1cs_binary_gbnigr.inp**.

- Model: 1 trait, 3 binary items (police2 parlt2 justice2), 3 groups = respondents from Great Britain ('gb'), Northern Ireland ('ni') and Greece (gr); trait variances fixed to 1 for all three countries (groups); means allowed to differ. In addition, intercept and slope for one item (police2) allowed to vary across countries

- R command to read in the data and apply the LCAT function, and read the results in two different formats:

```
bin3gbnigr <-lcat("1trait_1cs_binary_gbnigr.out",sessionpath,addto=
bin3gbnigr)
print(bin3gbnigr,7)
print(bin3gbnigr,7,alt=T)
```

- Plotting ICCs for the higher level of response (trust quite a lot or a great deal) for all 3 countries, for the item police2 whose intercept and slope is allowed to vary by country:

```
plot(bin3gbnigr,7,groups=1:3,items=1,levels=2)
```

- Compare with the same for the previous model in which the intercepts can vary but the slopes are the same across groups:

```
plot(bin3gbnigr,6,groups=1:3,items=1,levels=2)
```

4. Fit statistics

- Review fit statistic summaries for the models fitted so far:
```
bin3gbnigr
```

- Now look more closely at any that interest you. You may find it helpful to order them from largest to smallest (using `sort=T`). For example:
```
resid(bin3gbnigr,2,item2way=T,over4=T,sort=T)
resid(bin3gbnigr,6,item2way=T,over4=T,sort=T)
resid(bin3gbnigr,7,item2way=T,over4=T,sort=T)
```

- Try a likelihood ratio comparison test of a pair of nested models, for example:
```
lcat.lrtest(bin3gbnigr,2,6)
lcat.lrtest(bin3gbnigr,6,7)
```

5. 2 latent classes; 3 groups; 3 binary items: intercept for one item varies across groups

- Mplus input file: **2class_1c_binary_gbnigr.inp**.

- Model: 2 latent classes, 3 binary items (police2 parlt2 justice2), 3 groups = respondents from Great Britain ('gb'), Northern Ireland ('ni') and Greece (gr); prior probabilities (i.e. proportions estimated to belong to each class) allowed to vary across countries; intercepts free to vary across countries for one item (police2)

- R command to read in the data and apply the LCAT function, and read the results:
```
bin3gbnigr <-lcat("2class_1c_binary_gbnigr.out",sessionpath,addto=
bin3gbnigr)
print(bin3gbnigr,8)
```

- Plotting response probabilities for groups 1 to 3, for the higher response option (quite a lot or a great deal) for item 1 which now may have different constants for different countries:
```
plot(bin3gbnigr,8,groups=1:3,items=1,levels=2)
```

- Compare with the model in which intercepts are fixed to be equal across groups:
```
plot(bin3gbnigr,5,groups=1:3,items=1,levels=2)
```

6. 2 latent classes; 3 groups; 3 binary items: intercept and slope for one item varies across groups

- Mplus input file: **2class_1cs_binary_gbnigr.inp**.

- Model: 2 latent classes, 3 binary items (police2 parlt2 justice2), 3 groups = respondents from Great Britain ('gb'), Northern Ireland ('ni') and Greece (gr); prior probabilities (i.e. proportions estimated to belong to each class) allowed to vary across countries; intercepts and slopes free to vary across countries for one item (police2)

- R command to read in the data and apply the LCAT function, and read the results:
```
bin3gbnigr <-lcat("2class_1cs_binary_gbnigr.out",sessionpath,addto=
bin3gbnigr)
print(bin3gbnigr,9)
```

- Plotting response probabilities for groups 1 to 3, for the higher response option (quite a lot or a great deal) for item 1 which now may have different constants for different countries:
```
plot(bin3gbnigr,9,groups=1:3,items=1,levels=2)
```

- Compare with the previous models:
```
plot(bin3gbnigr,8,groups=1:3,items=1,levels=2)
plot(bin3gbnigr,5,groups=1:3,items=1,levels=2)
```

7. Fit statistics

- Review fit statistic summaries for the models fitted so far:
```
binary3gbnigr
```

- Now look more closely at any that interest you. You may find it helpful to order them from largest to smallest (using `sort=T`). For example:
```
resid(bin3gbnigr,5,item2way=T,over4=T,sort=T)
resid(bin3gbnigr,8,item2way=T,over4=T,sort=T)
resid(bin3gbnigr,9,item2way=T,over4=T,sort=T)
```

- Try a likelihood ratio comparison test of a pair of nested models, for example:
```
lcat.lrtest(bin3gbnigr,5,8)
lcat.lrtest(bin3gbnigr,8,9)
```