The estimation of item response models …

Paul De Boeck
University of Amsterdam & K.U.Leuven
Ivailo Partchev
Cito, Arnhem

Psychoco 2011, Tübingen Feb 24-25

1. Intro

see

De Boeck, P., et al. *Journal of Statistical Software,* to appear

2. Latent variables
3. Categorical data

1. Intro

data

flat binary data
Y_{pi} = 0, 1
p = 1, .., P persons
i = 1, .., I items

multilevel binary data
$Y_{p(g)i}$ = 0, 1
p(g) persons nested within groups
g = 1, .., G groups

long format →

format

p=1 i=1	Y_{11}=1
p=1 i=2	Y_{12}=0
..	..
p=1 i=I	Y_{1I}= 1
p=2 i=1	Y_{21}=1
p=2 i=2	Y_{22}=0
..	..
p=2 i=I	Y_{2I}= 1
..	..
..	..
p=P i=1	Y_{P1}=0
p=P i=2	Y_{P2}=0
..	..
p=P i=I	Y_{PI}= 0

data

flat binary data
Y_{pi} = 0, 1
p = 1, .., P persons
i = 1, .., I items

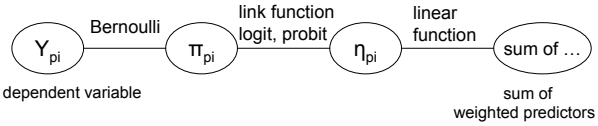
multilevel binary data
$Y_{p(g)i}$ = 0, 1
p(g) persons nested within groups
g = 1, .., G groups

dependent variable

format

p=1 i=1	Y_{11}=1
p=1 i=2	Y_{12}=0
..	..
p=1 i=I	Y_{1I}= 1
p=2 i=1	Y_{21}=1
p=2 i=2	Y_{22}=0
..	..
p=2 i=I	Y_{2I}= 1
..	..
..	..
p=P i=1	Y_{P1}=0
p=P i=2	Y_{P2}=0
..	..
p=P i=I	Y_{PI}= 0

Generalized Linear Mixed Models three components



three types of *covariates (predictors)*

item covariates	X_{ik}	(k=1, .., K)
person covariates	Z_{pj}	(j=1, .., J)
person-by-item covariates	W_{pih}	(h=1, .., H)

examples are partition covariates

origin

external: externally given, exogenous

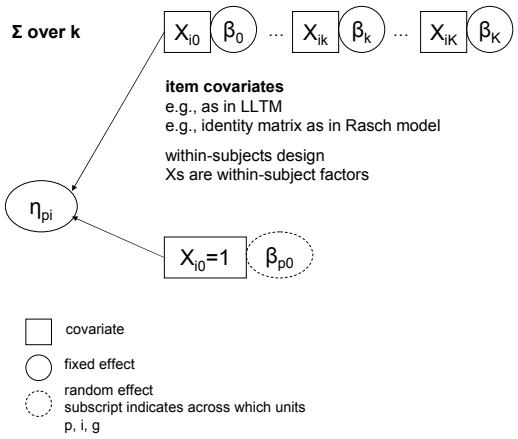
internal: from the data, endogenous

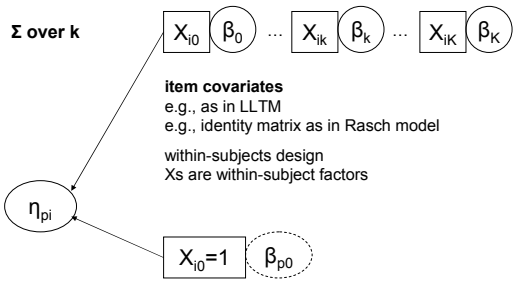
***effects (weights)* are fixed of random**

if random

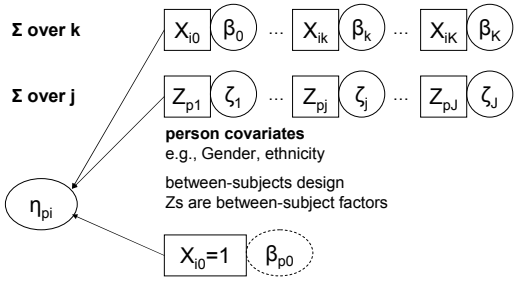
- across persons
- across items
- across groups

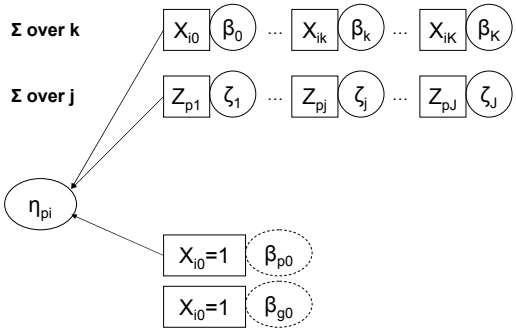
intro examples



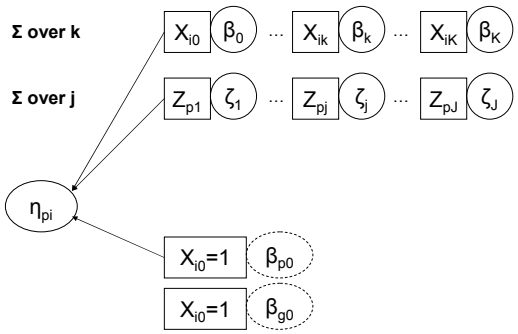



```
RaschM=lmer(r2~-1+item+(1|id),family=binomial, VerbAgg)
```



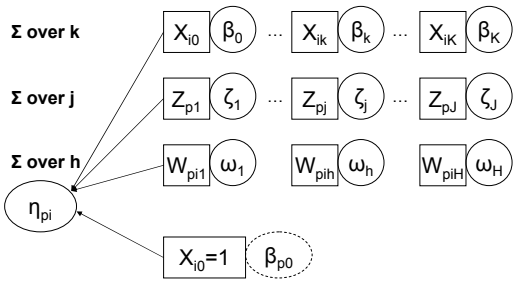


simple multilevel



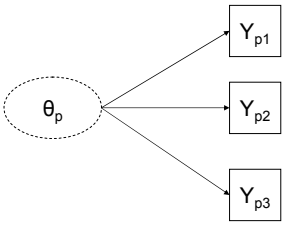
simple multilevel

```
multilevelM=lmer(r2~
-1+item+(1|id)+(1|Gender),family=binomial, VerbAgg)
```

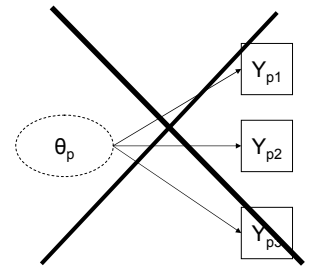


interaction covariates
e.g. DIF covariates: combination of focal group and a set of items, possibly just one
e.g., LID covariates: response to another item, the response depends on the pair (p,i)

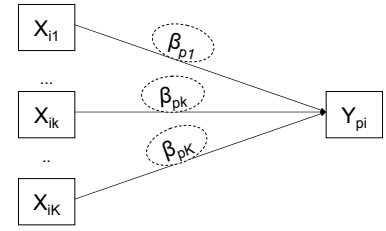
2. Latent variables



latent variable indicators



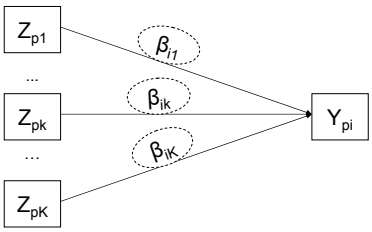
latent variable indicators

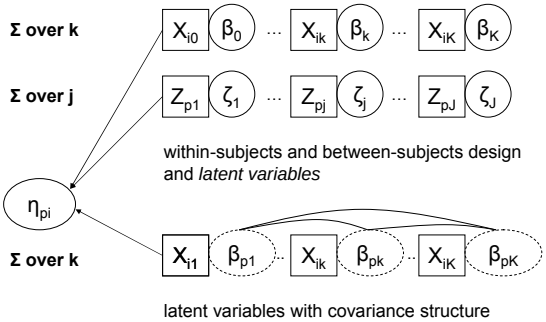


independent variables -- covariates

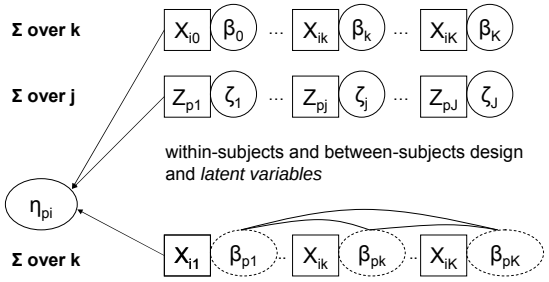
effect size

dependent variable



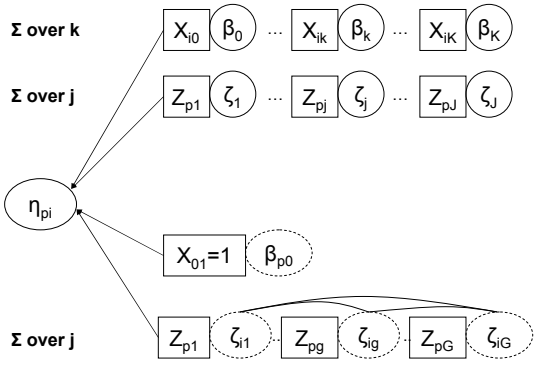


LANOVA



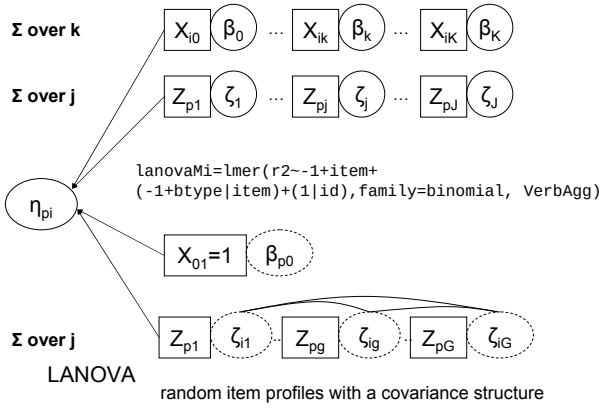
LANOVA

```
lanovaMp=lmer(r2~-1+item+(-1+btype|id),family=binomial,
VerbAgg)
```



LANOVA

random item profiles with a covariance structure

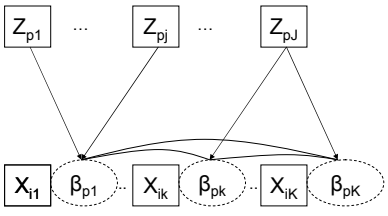



```
lanovaMi=lmer(r2~-1+item+
(-1+btype|item)+(1|id),family=binomial, VerbAgg)
```

LANOVA

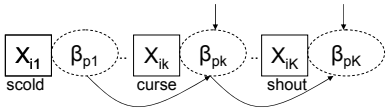
random item profiles with a covariance structure

explanation of latent person variables

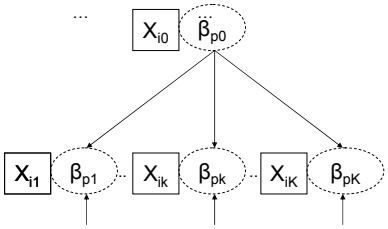


latent variables with covariance structure

LANOVA

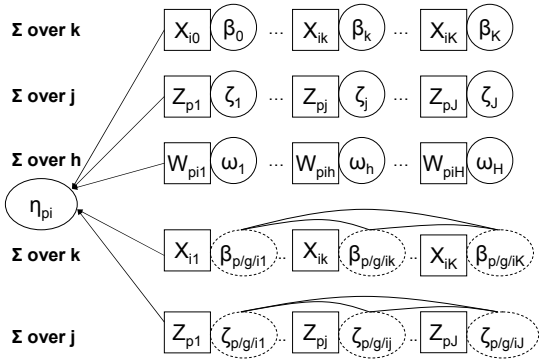


LANOVA



LANOVA

```
lanovaSEM2=lmer(r2~+item+(1|id)+(-1+curse|id)+(-1+scold|id)
+(-1+shout|id),family=binomial, VerbAgg)
```



3. Categorical data

data

flat ordered-category data
Y_{pi} = 1, 2, 3
p = 1, .., P persons
i = 1, .., I items

multilevel binary data
$Y_{p(g)i}$ = 1, 2, 3
g = 1, .., G groups

Continuation ratio model

Similar to discrete survival model

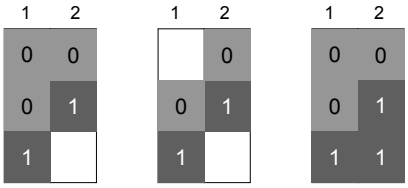
Ordered categories are like moments in time

A one indicates that the event occurs, so that later observations are missing

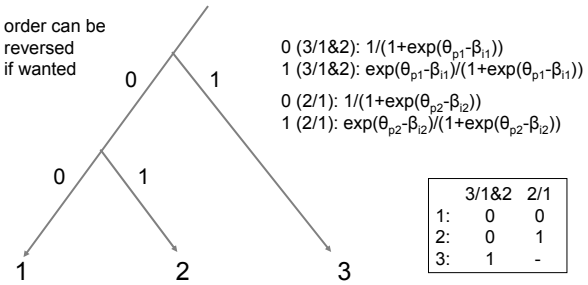
A zero indicates that the event has not yet occured, so that later observations are possible

Three models for ordered-category data

three types of odds ratios (dark vs light)
for example, three categories, two odds ratios

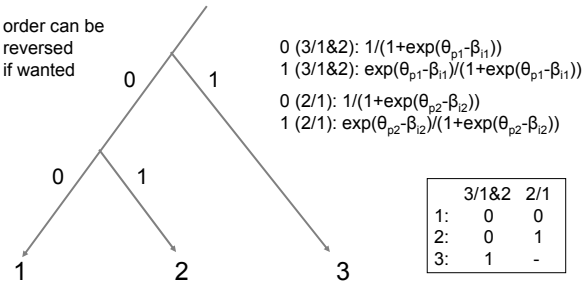


continuation ratio | partial credit | graded response



00: 1 / (1+exp(θ_{p1}-β_{i1})+exp(θ_{p2}-β_{i2})+exp(θ_{p1}+θ_{p2}-β_{i1}-β_{i2}))
01: exp(θ_{p2}-β_{i2}) / (1+exp(θ_{p1}-β_{i1})+exp(θ_{p2} β_{i2})+exp(θ_{p1}+θ_{p2}-β_{i1}-β_{i2}))
1- : exp(θ_{p1}-β_{i1}) / (1 +exp(θ_{p1} – β_{i1}))

```
TutzM=lmer(value~
-1+newItem+(-1+variable|id),family=binomial,NewVerbAgg)
```



0 (3/1&2): $1/(1+\exp(\theta_{p1}-\beta_{i1}))$
1 (3/1&2): $\exp(\theta_{p1}-\beta_{i1})/(1+\exp(\theta_{p1}-\beta_{i1}))$

0 (2/1): $1/(1+\exp(\theta_{p2}-\beta_{i2}))$
1 (2/1): $\exp(\theta_{p2}-\beta_{i2})/(1+\exp(\theta_{p2}-\beta_{i2}))$

	3/1&2	2/1
1:	0	0
2:	0	1
3:	1	-

00: $1 / (1+\exp(\theta_{p1}-\beta_{i1})+\exp(\theta_{p2}-\beta_{i2})+\exp(\theta_{p1}+\theta_{p2}-\beta_{i1}-\beta_{i2}))$
01: $\exp(\theta_{p2}-\beta_{i2}) / (1+\exp(\theta_{p1}-\beta_{i1})+\exp(\theta_{p2}\ \beta_{i2})+\exp(\theta_{p1}+\theta_{p2}-\beta_{i1}-\beta_{i2}))$
1- : $\exp(\theta_{p1}-\beta_{i1}) / (1 +\exp(\theta_{p1} - \beta_{i1}))$

extend dataset: replace each item response with two, except when missing:
1 00
2 01
3 1-

```
VerbAgg$a <- as.numeric(VerbAgg$resp == 'yes')
VerbAgg$b <- as.numeric(VerbAgg$resp == 'perhaps')
VerbAgg$b[which(VerbAgg$resp=='yes')] <- NA
NewVerbAgg <- melt(VerbAgg, measure.vars = c('a', 'b'))
NewVerbAgg$newItem<-interaction(NewVerbAgg$item,NewVerbAgg$variable)

mTutz=lmer(value~-1+newItem+(1|id),family=binomial,NewVerbAgg)
two-dimensional
mTutz=lmer(value~-1+newItem+(-1+variable|id),family=binomial,NewVerbAgg)
```

See R code in added Tutz script

Some important issues are not covered,
therefore,
or for other reasons,
you may have questions