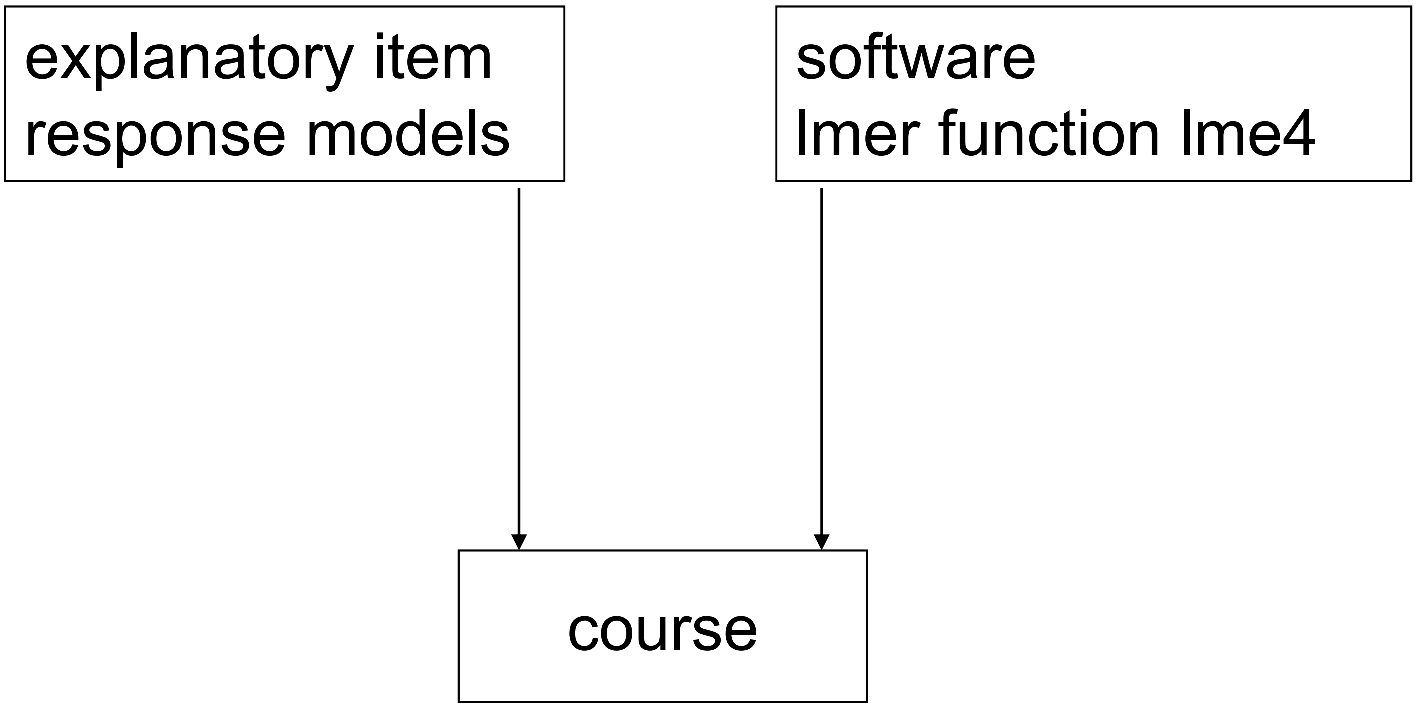
A Generalized Linear Mixed Model Approach to Item Response Modeling

Paul De Boeck
U. Amsterdam
& K.U.Leuven

Sun-Joo Cho
Peabody College
Vanderbilt U.

Vienna, June 6-10, 2011


explanatory item response models
software
lmer function lme4
course

1. Rijmen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods*, *8*, 185-205.

2. De Boeck, P., & Wilson, M. (Eds.) (2004). *Explanatory item response models: A generalized linear and nonlinear approach.* New York: Springer.

3. De Boeck, P. et al. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software.*

Website : http://bearcenter.berkeley.edu/EIRM/

Statistics for Social Science and Public Policy

Paul De Boeck, Mark Wilson, Editors

Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach

This edited volume gives a new and integrated introduction to item response models (predominantly used in measurement applications in psychology, education, and other social science areas) from the viewpoint of the statistical theory of generalized linear and nonlinear mixed models. Moreover, this new framework allows the domain of item response models to be co-ordinated and broadened to emphasize their *explanatory* uses beyond their standard *descriptive* uses.

The basic explanatory principle is that item responses can be modelled as a function of predictors of various kinds. The predictors can be (a) characteristics of items, of persons, and of combinations of persons and items; they can be (b) observed or latent (of either items or persons); and they can be (c) latent continuous or latent categorical. Thus, a broad range of models is generated, including a wide range of extant item response models as well as some new ones. Within this range, models with explanatory predictors are given special attention in this book, but we also discuss descriptive models. Note that the "item responses" that we are referring to are not just the traditional "test data," but are broadly conceived as categorical data from a repeated observations design. Hence, data from studies with repeated observations experimental designs, or with longitudinal designs, may also be modelled.

The book starts with a four-chapter section containing an introduction to the framework. The remaining chapters describe models for ordered-category data, multilevel models, models for differential item functioning, multidimensional models, models for local item dependency, and mixture models. It also includes a chapter on the statistical background and one on useful software. In order to make the task easier for the reader, a unified approach to notation and model description is followed throughout the chapters, and a single data set is used in most examples to make it easier to see how the many models are related. For all major examples, computer commands from the SAS package are provided which can be used to estimate the results for each model. In addition, sample commands are provided for other major computer packages.

Paul De Boeck is Professor of Psychology at K.U. Leuven (Belgium), and Mark Wilson is Professor of Education at UC Berkeley (USA). They are also co-editors (along with Pamela Moss) of a new journal entitled *Measurement: Interdisciplinary Research and Perspectives*. The chapter authors are members of a collaborative group of psychometricians and statisticians centered on K.U. Leuven and UC Berkeley.

ISBN 0-387-40275-6

EAN

9 780387 402758 >

springeronline.com

De Boeck
Wilson, Editors

Explanatory Item Response Models

Statistics for Social Science and Public Policy

Paul De Boeck
Mark Wilson
Editors

Explanatory Item Response Models

A Generalized Linear and Nonlinear Approach

Springer

- In 1 and 2 mainly SAS NLMIXED
- In 3 lmer function from lme4

Modeling data

- A basic principle
  Data are seen as resulting from a true part and an error part.

  binary
  $Y_{pi}=1$ if $V_{pi} \geq 0$
  $Y_{pi}=0$ if $V_{pi}<0$
  V_{pi} is a real defined on the interval $-\infty$ to $+\infty$
  $V_{pi} = \eta_{pi} + \varepsilon_{pi}$ $\varepsilon_{pi} \sim N(0,1)$ probit, normal-ogive
  $\varepsilon_{pi} \sim \text{logistic}(0,3.29)$ logit, logistic

Logistic models

- Standard logistic instead of standard normal
  Logistic model – logit model
  vs
  Normal-ogive model – probit model

  density general logistic distribution:
  $f(x)=k \exp(-kx)/(1+\exp(-kx))^2$

  $var = \pi^2/3k^2$

  standard logistic: $k=1$,
  $\sigma = \pi/\sqrt{3} = 1.814$
  setting $\sigma=1$, implies that $k=1.814$

  best approximation from standard normal: $k=1.7$
  this is the famous $D=1.7$ in early IRT formulas

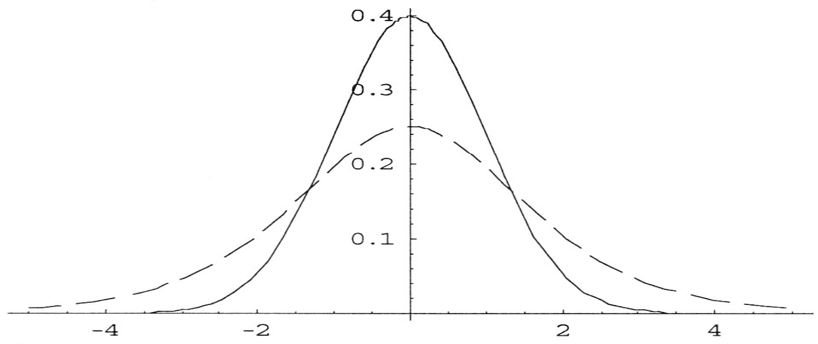


FIGURE 1.
The logistic distribution with $k = 1$ and the standard normal (solid line).

standard (k=1) logistic
vs
standard normal

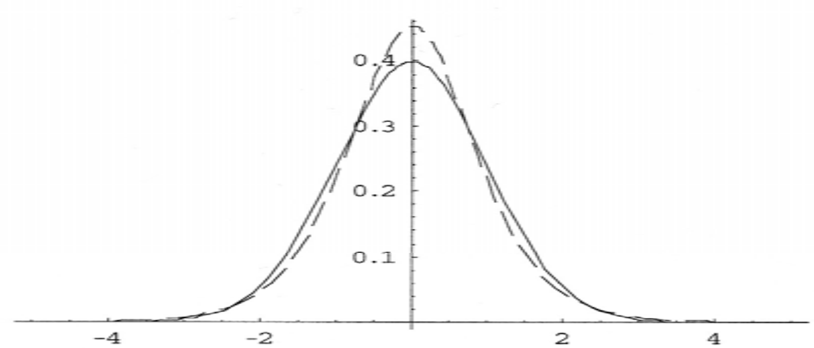
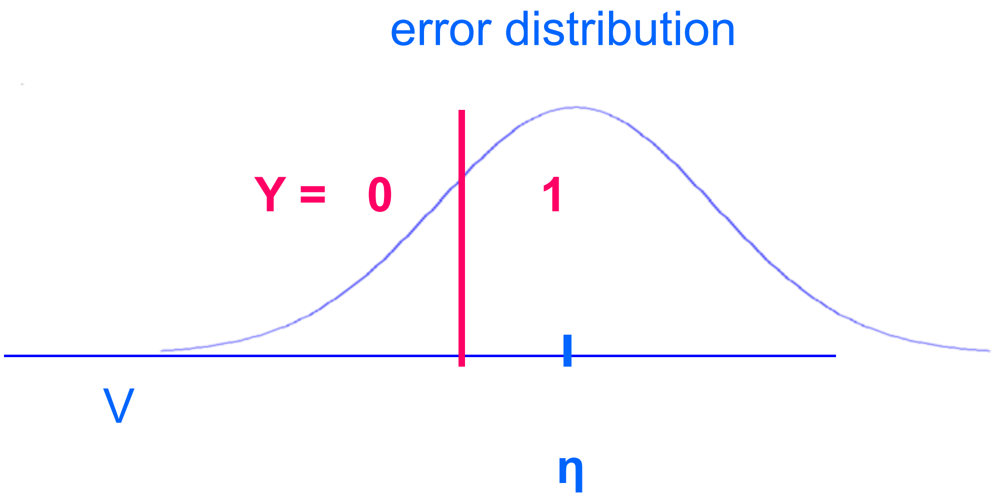


FIGURE 2.
The logistic distribution with $k = 1.8$ and the standard normal (solid line).

logistic k=1.8
vs
standard normal

copied from Savalei, *Psychometrika* 2006

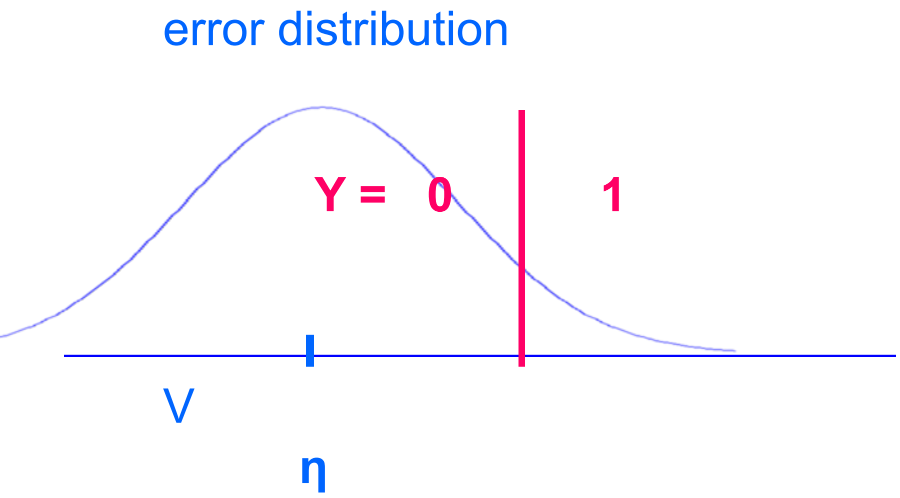
binary data

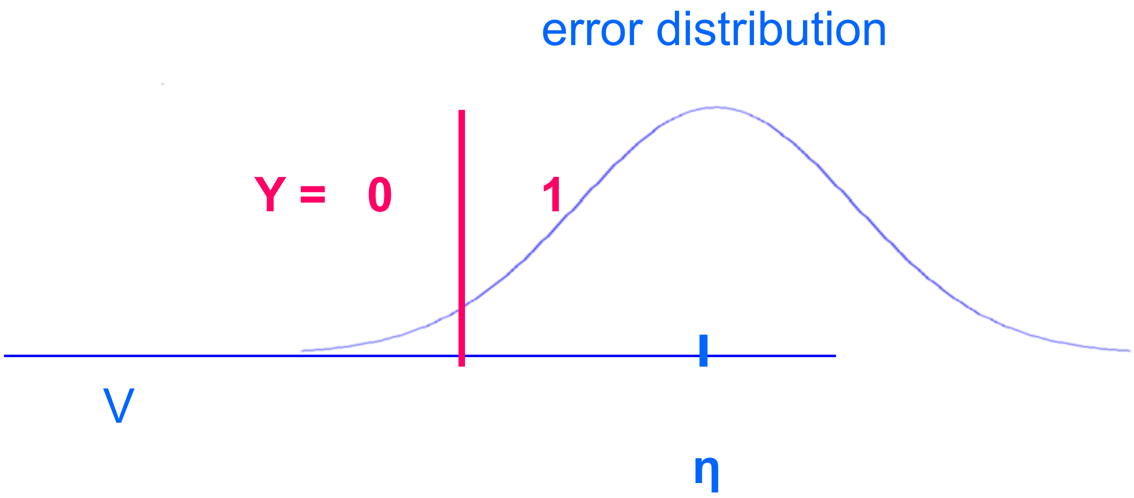


moving hat model

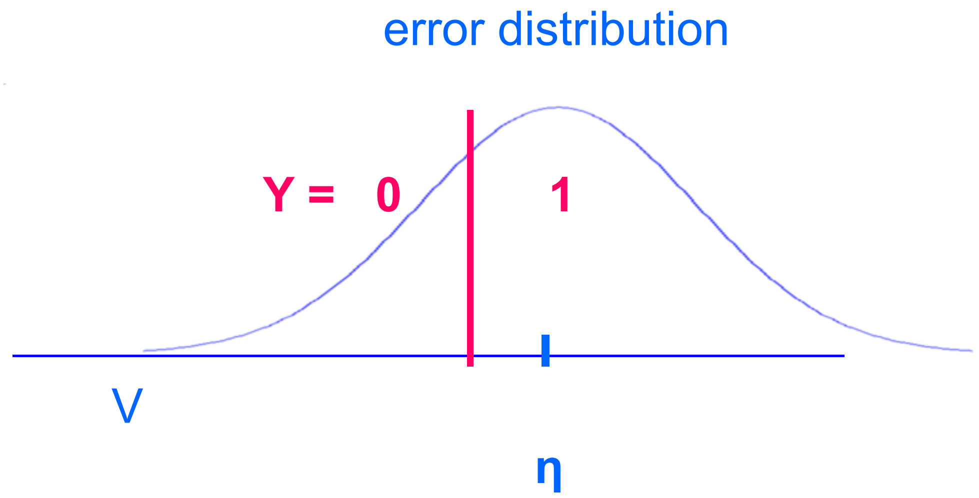
$$V_{pi} = \eta_{pi} + \varepsilon_{pi}$$

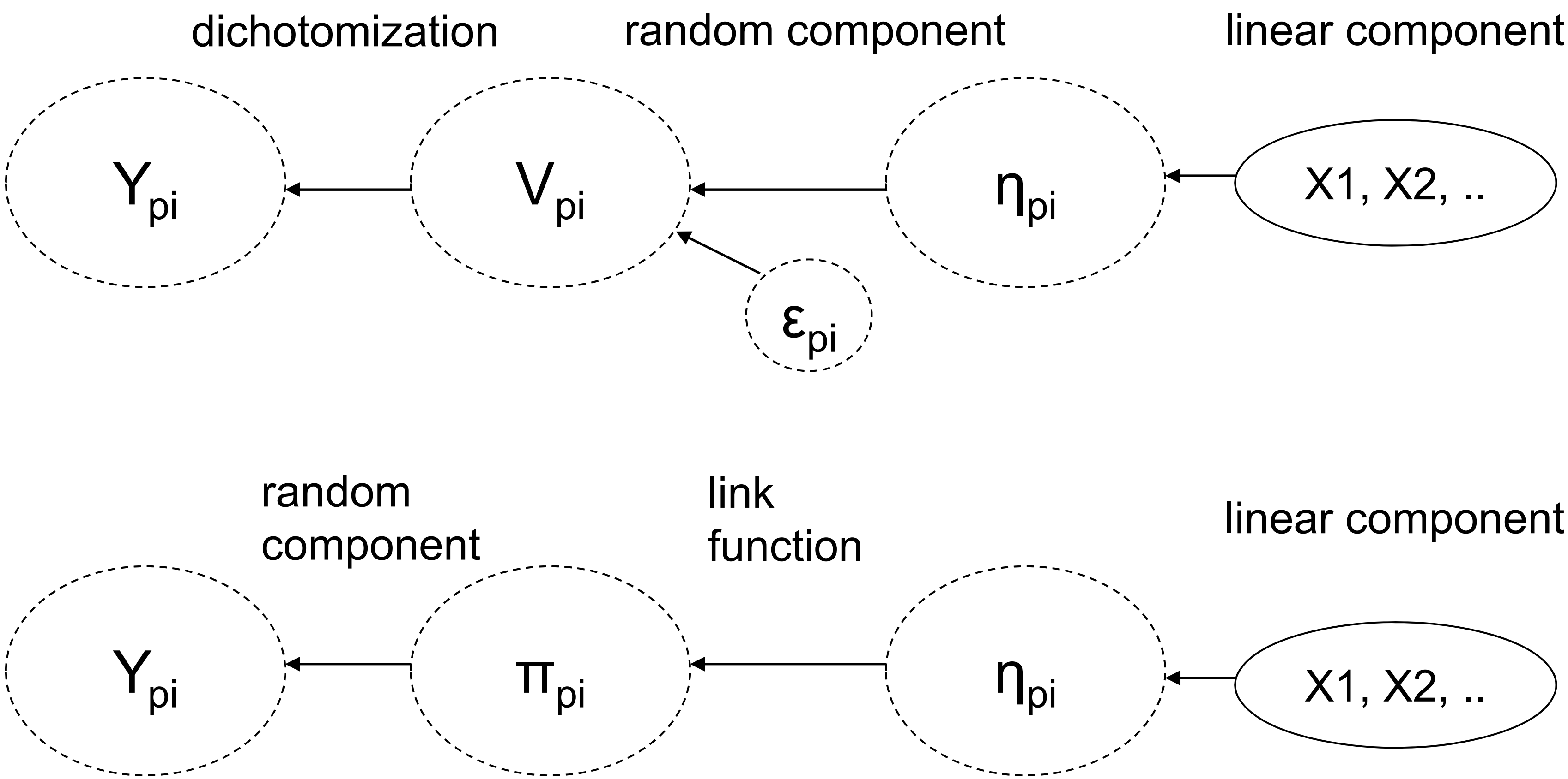
$$\eta_{pi} = \Sigma_k \beta_{k(r)} X_{pik}$$


error distribution
Y = 0
1
V
η

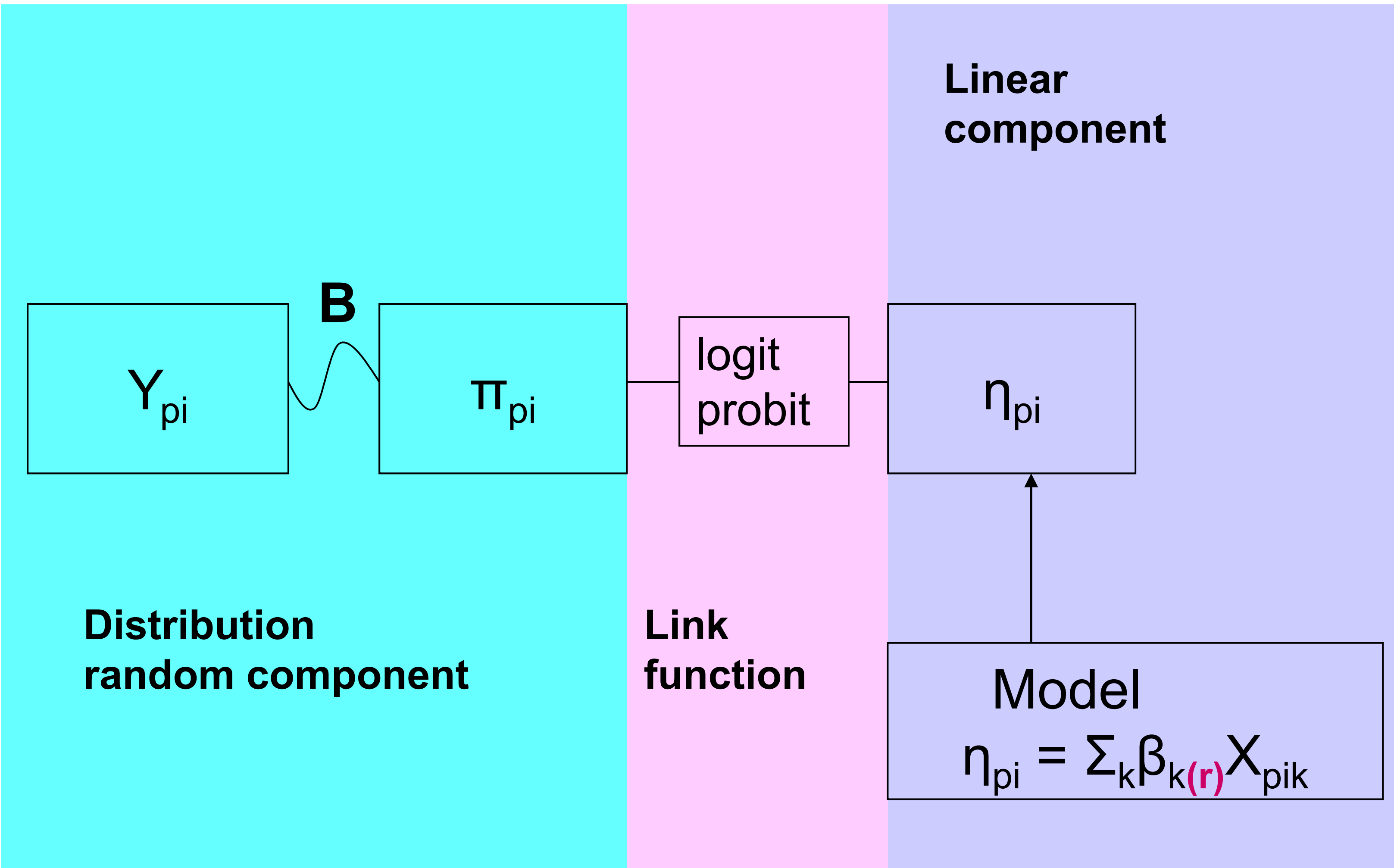


error distribution
Y = 0
1
V
η


error distribution
Y = 0
1
V
η


dichotomization
random component
linear component
Y pi
V pi
η pi
X1, X2, ..
ε pi
random
component
link
function
linear component
Y pi
π pi
η pi
X1, X2, ..

Logit and probit models



Long form

- Wide form is P x I array

items

persons

```
111001000
000101010
001100101
101011000
110101100
```

- Long form is vector with length PxI

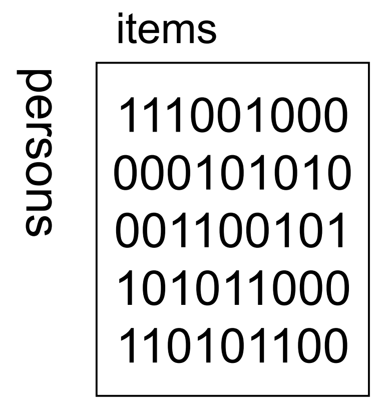
Y_{pi} covariates

pairs (person, item)

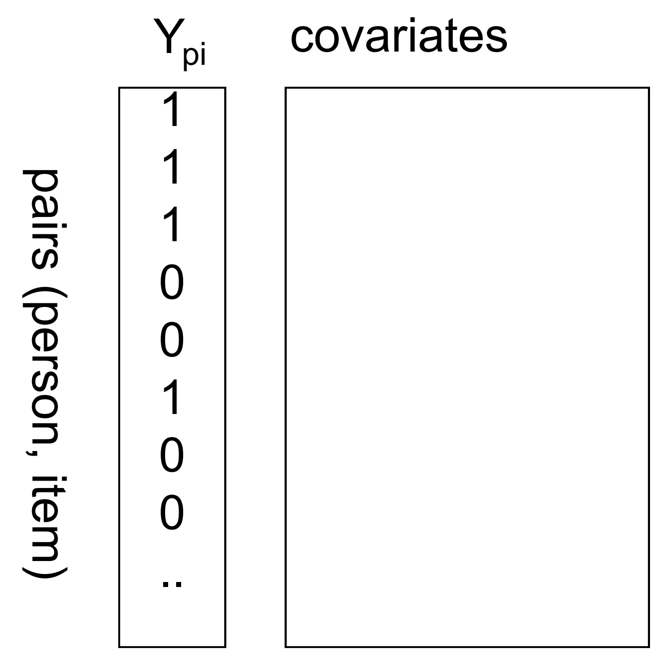
```
1
1
1
0
0
1
0
0
..
```

Long form

- Wide form is P x I array



- Long form is vector with length PxI



Content

1. Item covariate models
2. Person covariate models
3. Person x item covariate models
4. Random item models
   Models for ordered-category data
5. Estimation and testing

1. Item covariate models

- open R console
- setwd(“ ”)
- library(lme4)

Data

?VerbAgg

24 items with a 2 x 2 x 3 design

- situ: other vs self
  two frustrating situations where an*other* person is to be blamed
  two frustrating situations where one is *self* to be blamed
- mode: want vs do
  wanting to be verbally agressive vs doing
- btype: cursing, scolding, shouting
  three kinds of being verbally agressive

e.g., “A bus fails to stop. I would want to curse” yes perhaps no

316 respondents

- Gender: F (men) vs M (women)
- Anger: the subject's Trait Anger score as measured on the State-Trait Anger Expression Inventory (STAXI)

str(VerbAgg)   head(VerbAgg)

1. Rasch model
1PL model

1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1	β_1 β_2 β_3 β_4		fixed
1 1 1 1		θ_p	random
Y			

$$\eta_{pi} = \theta_p X_{i0} - \Sigma_k \beta_i X_{ik}$$

$$\eta_{pi} = \theta_p - \beta_i$$

$$\pi_{pi} = \exp(\eta_{pi})/(1+\exp(\eta_{pi}))$$

$$\theta_p \sim N(0, \sigma^2_\theta)$$

Note on 2PL: Explain that in 2PL the constant X_{i0} is replaced with discrimination parameters

lmer(r2 ~ …… , family=binomial(“logit”), data=VerbAgg)
lmer(r2 ~ …… , family=binomial, VerbAgg) logistic model

lmer(r2 ~ …… , family=binomial(“probit”), data=VerbAgg) normal-ogive
lmer(r2 ~ …… , family=binomial(“probit”), VerbAgg) probit model

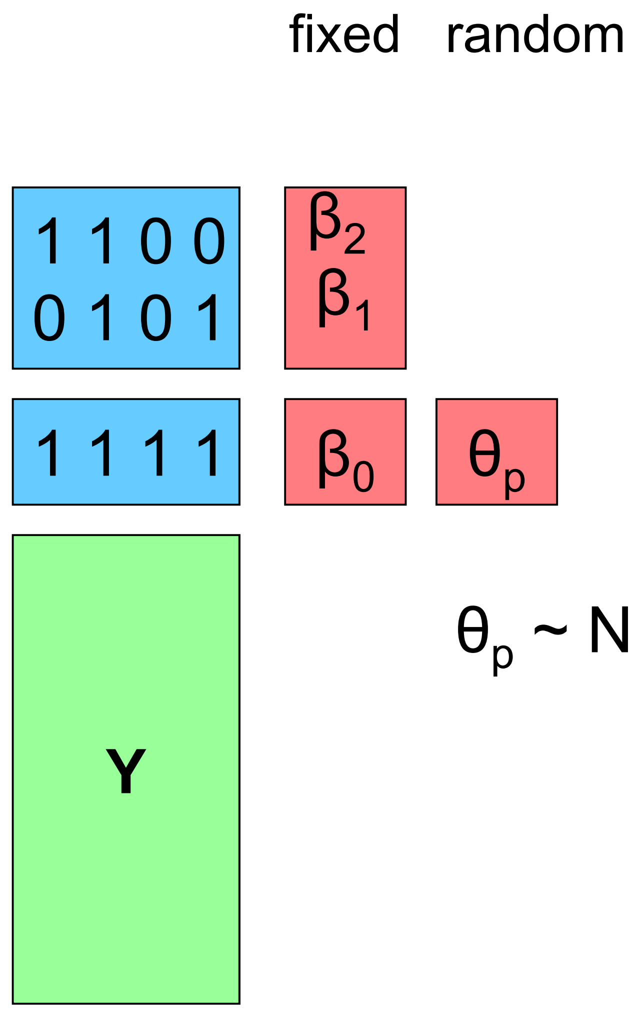
……
item + (1 |id), first item is intercept, other item parameters are differences with first
$\beta_0 = \beta_1, \beta_2 - \beta_1, \beta_3 - \beta_1, ..$
or
-1 + item + (1 |id) no intercept, only the common item parameters
0 + item + (1 |id)

item is item factor
id is person factor
1 is 1-covariate
(a|b) effect of a is random across levels of b

- to avoid correlated error output:

  ```
  print(modelname, cor=F)
  ```

2. LLTM model



$$\eta_{pi} = \theta_p X_{i0} - \Sigma_k \beta_k X_{ik}$$

$$\eta_{pi} = \theta_p - \Sigma_k \beta_k X_{ik}$$

$$\theta_p \sim N(0, \sigma^2_\theta)$$

-1+mode+situ+btype+(1|id), family=binomial, VerbAgg

contrasts

treatment	sum	helmert	poly
dummy	effect		
00	1 0	-1 -1	linear
10	0 1	1 -1	quadratic
01	-1-1	0 2	

without intercept always

100

010

001

S-J

- <u>lmer treatment coding with intercept</u>

want	other	curse	0 0 0 0
want	other	scold	0 0 1 0
want	other	shout	0 0 0 1
want	self	curse	0 1 0 0
want	self	scold	0 1 1 0
want	self	shout	0 1 0 1
do	other	curse	1 0 0 0
do	other	scold	1 0 1 0
do	other	shout	1 0 0 1
do	self	curse	1 1 0 0
do	self	scold	1 1 1 0
do	self	shout	1 1 0 1

- <u>lmer treatment coding without intercept</u>

want	other	curse	0 1 0 0 0
want	other	scold	0 1 0 1 0
want	other	shout	0 1 0 0 1
want	self	curse	0 1 1 0 0
want	self	scold	0 1 1 1 0
want	self	shout	0 1 1 0 1
do	other	curse	1 0 0 0 0
do	other	scold	1 0 0 1 0
do	other	shout	1 0 0 0 1
do	self	curse	1 0 1 0 0
do	self	scold	1 0 1 1 0
do	self	shout	1 0 1 0 1

btype	treatment	sum	helmert	mode	treatment	sum	helmert
curse	0 0	1 0	-1-1	want	0	1	-1
scold	1 0	0 1	1-1	do	1	-1	1
shout	0 1	-1-1	0 2				

main effects and interactions

✓ mode:btype is for cell means independent of coding

✓ dummy coding

main effects: mode+btype or
C(mode,treatment) + C(btype,treatment)

main effects & interaction: mode*btype or
C(mode,treatment) *C(btype,treatment)

✓ effect coding

main effects: 1+C(mode,sum)+ C(btype,sum)

main effects & interaction: C(mode,sum)*C(btype,sum)

3. LLTM + error model

remember there are two items per cell

$$\eta_{pi} = \theta_p - \Sigma_k \beta_k X_{ik} + \varepsilon_i$$

fixed random

$$\begin{matrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{matrix} \quad \begin{matrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{matrix}$$

$$\begin{matrix} 1 & 1 & 1 & 1 \end{matrix} \quad \theta_p \;\; \varepsilon_i$$

Y

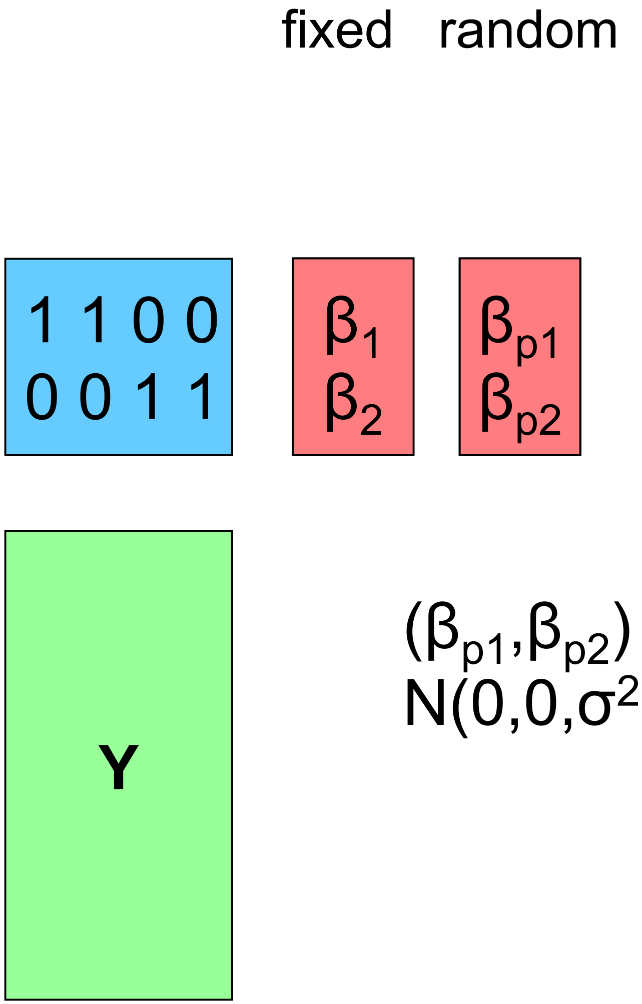
$$\theta_p \sim N(0, \sigma^2_\theta)$$

$$\varepsilon_i \sim N(0, \sigma^2_\varepsilon)$$

```
lmer(r2 ~ mode + situ + btype + (1 |id) + (1|item),
or
lmer(r2 ~ - 1 + mode + situ + btype + (1 |id) + (1|item),
family=binomial, data=VerbAgg)
```

- two types of multidimensional models
  - random-weight LLTM
  - multidimensional 1PL

4. Random-weight LLTM



$$\eta_{pi} = \Sigma_k \beta_{pk} X_{ik} - \Sigma_k \beta_k X_{ik}$$

$$(\beta_{p1}, \beta_{p2}) \sim N(0, 0, \sigma^2_{\theta 1}, \sigma^2_{\theta 2}, \sigma_{\theta 1 \theta 2})$$

```
lmer(r2 ~ mode + situ + btype + (-1 + mode|id),
family=binomial, data=VerbAgg)
```

5. multidimensional 1PL model

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} \beta_{p1} \\ \beta_{p2} \end{pmatrix}$$

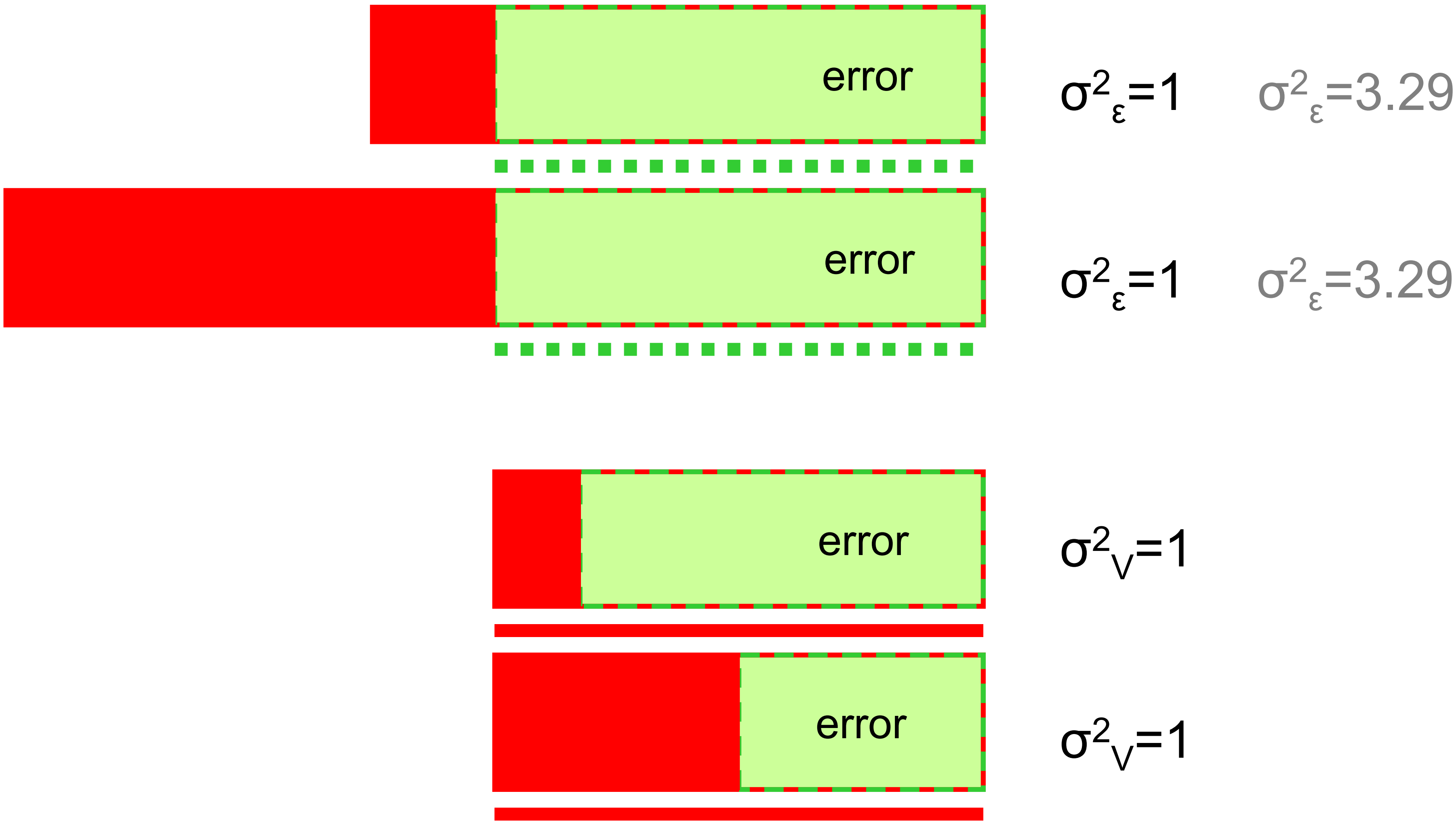
$$\eta_{pi} = \Sigma_k \beta_{pk} X_{ik} - \beta_i$$

Y

$$(\beta_{p1},\beta_{p2}) \sim N(0,0,\sigma^2_{\theta 1},\sigma^2_{\theta 2},\sigma_{\theta 1 \theta 2})$$

Note on factor models, how they differ from IRT models
Note on rotational positions

variance partitioning



- item covariate based multidimensional models
  a non-identified model
  and four possible identified models

(crossed out)	1	2	3	3	4
1 0 1 0	1 0 0	1 0 0	1 0	1 0	1 0 0 0
1 0 1 0	1 0 0	1 0 0	1 0	1 0	1 0 0 0
1 0 0 1	1 0 1	1 0 1	1 0	0 1	0 1 0 0
1 0 0 1	1 0 1	1 0 1	1 0	0 1	0 1 0 0
0 1 1 0	1 1 0	0 1 0	0 1	1 0	0 0 1 0
0 1 1 0	1 1 0	0 1 0	0 1	1 0	0 0 1 0
0 1 0 1	1 1 1	0 1 1	0 1	0 1	0 0 0 1
0 1 0 1	1 1 1	0 1 1	0 1	0 1	0 0 0 1

1 -1 + item + (mode + situ|id)

2 -1 + item + (-1 + mode + situ|id)

3 -1 + item + (-1 + mode |id) + (-1 + situ |id)

4 -1 + item + (mode:situ|id)

Illustration of non-identified model

```
VerbAgg$do=(VerbAgg$mode==“do”)+0
VerbAgg$want=(VerbAgg$mode==“want”)+0
VerbAgg$self=(VerbAgg$mode==“self”)+0
VerbAgg$other=(VerbAgg$mode==“other”)+0
mMIR1=lmer(r2~-1+item+
    (-1+do+want+self+other|id),family=binomial,VerbAgg)
mMIR2=lmer(r2~-1+item+
    (-1+want+do+self+other|id),family=binomial,VerbAgg)
```

compare with identified model

```
mMIR3=lmer(r2~-1+item+(-1+mode+situ|id), family=binomial, VerbAgg)
```

-1 + item + (mode + situ + btype |id)
-1 + item + (-1 + mode + situ + btype |id)
-1 + item + (-1 + mode |id) + (-1 + situ |id) + (-1 + btype |id)
-1 + item + (mode:situ:btype |id)

how many dimensions?

rotations

VerbAgg$do=(VerbAgg$mode==“do”)+0.
VerbAgg$want=(VerbAgg$mode==“want”)+0.
VerbAgg$dowant=(VerbAgg$mode==“do”)-1/2.

1. simple structure orthogonal
(-1+do|id)+(-1+want|id)
2. simple structure correlated
(-1+mode|id)
3. general plus bipolar
(1+dowant|id)
4. general plus bipolar uncorrelated
(1|id)+(-1+dowant|id)

2 and 3 are equivalent
1 and 4 are constrained solutions
all four are confirmatory

S-J

estimation of person parameters and random effects in general

three methods

- ML maximum likelihood – flat prior
- MAP mode a posteriori – normal prior, mode of posterior
- EAP expected a posteriori – normal prior, mean of posterior, and is therefore a prediction

irtoys does all three

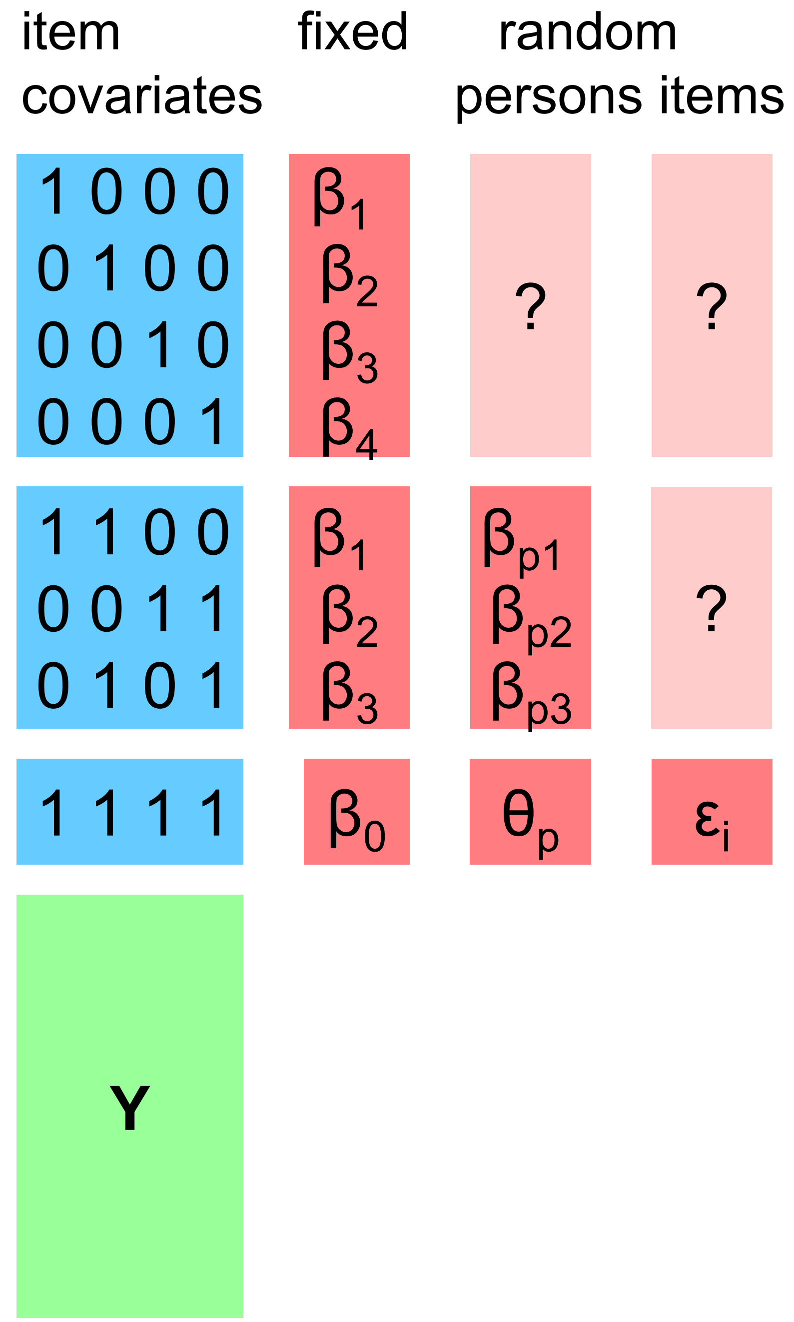
lmer does MAP

ranef(model)

se.ranef(model) for standard errors

Conflicts?

Extensions?



3. Person-by-item covariate models

i1	i2	i3
1	1	1
1	0	1
0	0	0
0	0	0
1	0	1
1.00	0.50	1.00

i1	i2	i3
1	1	1
0	1	0
0	0	0
0	0	0
0	1	0
0.50	1.00	0.50

DIF as discrepancy between within-group structure of differences between-group structure of differences

- covariates of person-item pairs

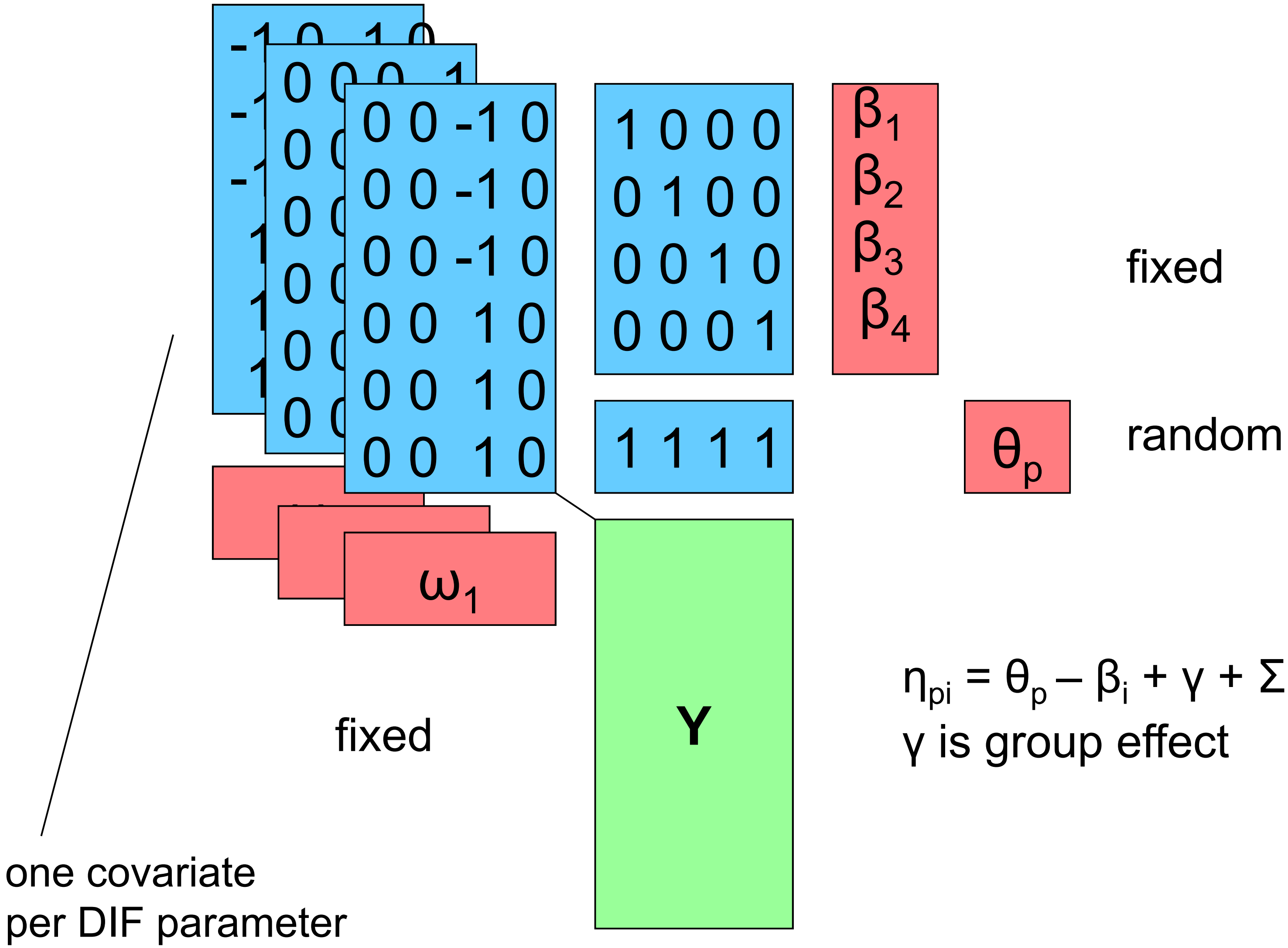
  external covariates

  e.g., differential item functioning
  an item functioning differently depending on the group
  person group x item
  e.g., strategy information per pair person-item

  internal covariates

  responses being depending on other responses

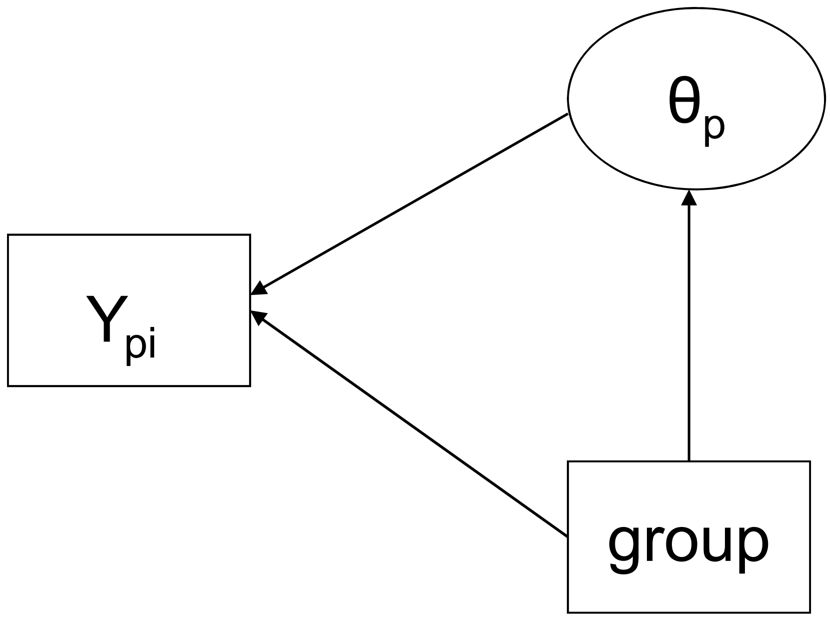
  e.g., do responses depending on want responses
  local item dependence – LID;
  e.g., learning during the test, during the experiment
  dynamic Rasch model



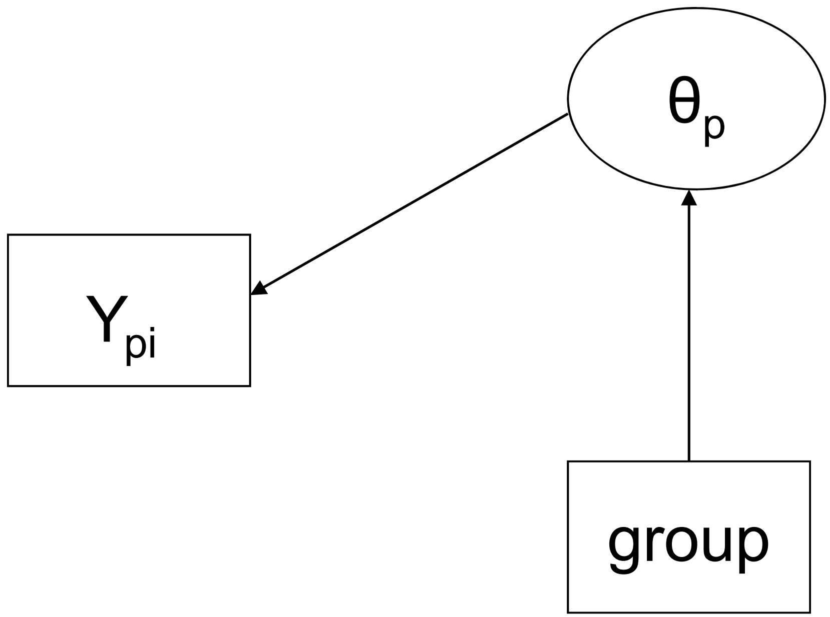
$\eta_{pi} = \theta_p - \beta_i + \gamma + \Sigma_h \omega_h W_{(p,i)h}$

γ is group effect

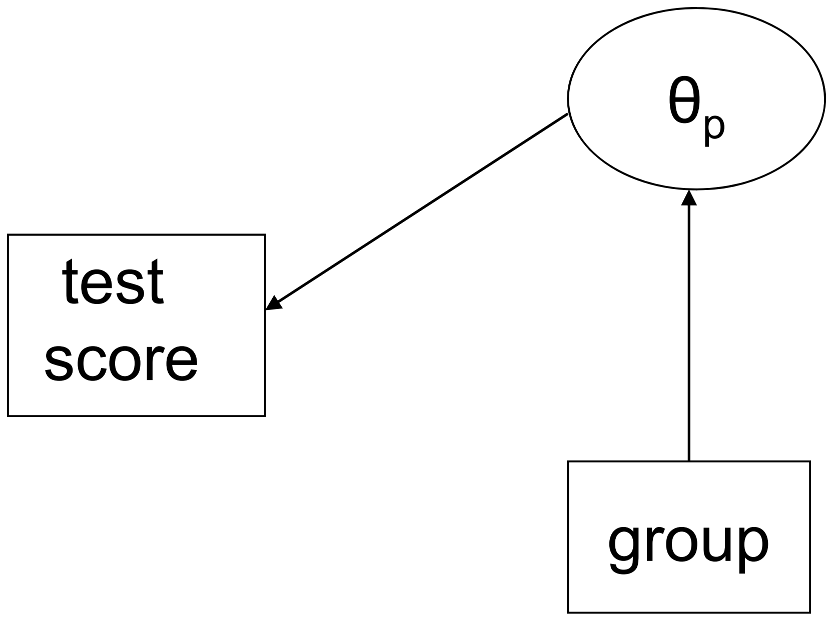
1. DIF model Differential item functioning unfair (because of DIF)



fair (no DIF)



fair (lack of differential test funtioning)



- DIF approaches with lmer

  1. simultaneous test:
  - interaction items x group
  to identify DIF items, use effect coding
  to see the item difficulties in the two groups, use Gender:item
  2. itemwise test
  - interaction of each item in turn with group
  item1 x group,
  next item 2 x group,
  etc.
  3. random item approach

DIF approaches

difficulties in the two groups – *equal mean abilities*

VerbAgg$M=(VerbAgg$Gender==“M”)+0.
VerbAgg$F=(VerbAgg$Gender==“F”)+0.

-1+Gender:item+(-1+M|id)+(-1+F|id)

simultaneous test all items – *equal mean difficulties*

-1+C(Gender,sum)*C(item,sum)+(-1+M|id)+(-1+F|id)

-- *difference with reference group*

-1+Gender*item+(-1+M|id)+(-1+F|id)

itemwise test

VerbAgg$i1=(VerbAgg$item==“S1wantcurse”)+0.
VerbAgg$i2=(VerbAgg$item==“S1WantScold”)+0. (pay attention to item labels)
…

e.g., item 3

-1+Gender+i1+i2+i4+i5…+i24+Gender*i3+(-1+M|id)+(-1+F|id)

S-J

result depends on equating

therefore a LR test is recommended

gender DIF for all do items of the curse and scold type

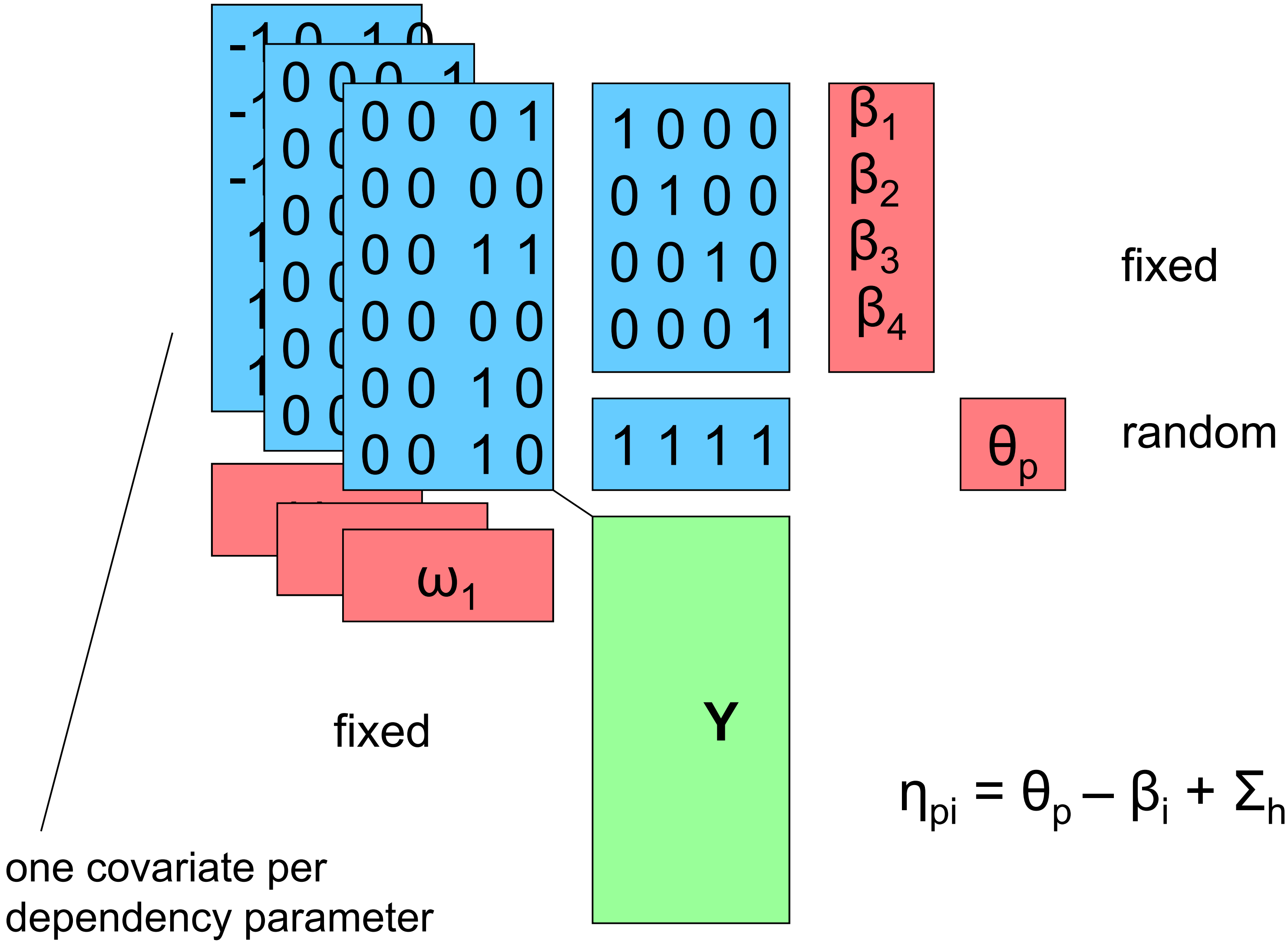
dif=with(VerbAgg, factor(0 + (Gender==“F” & mode==“do” & btype!=“shout”)))

-1 +item + Gender + dif + (1|id)

random across persons
-1 +item + Gender + dif + (1 + dif|id)

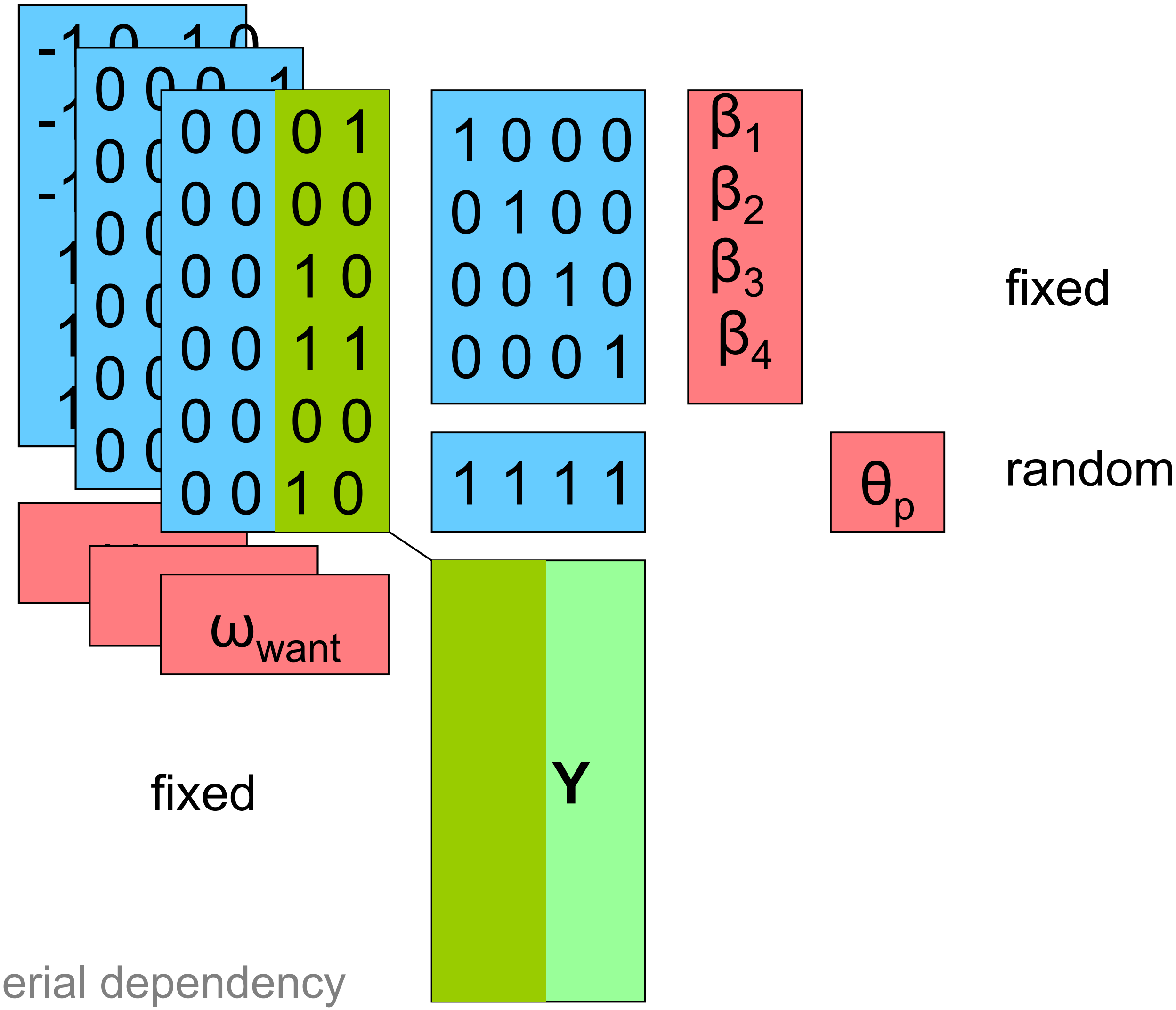
F = man
M = woman

dummy coding vs contrast coding
(treatment vs sum or helmert) makes
a difference for the item parameter estimates



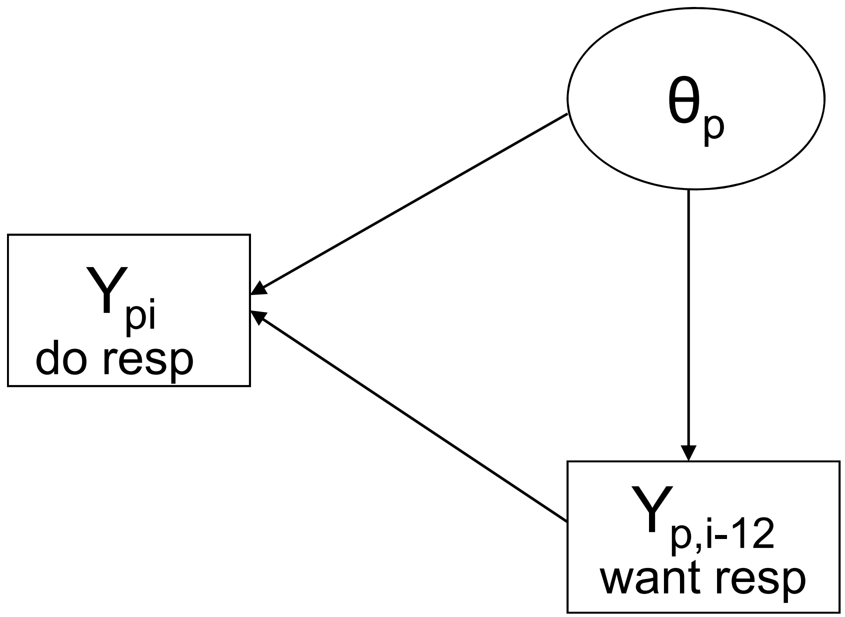
$$\eta_{pi} = \theta_p - \beta_i + \Sigma_h \omega_h W_{(p,i)h}$$

2. LID model Local item dependence



Note on serial dependency and stationary vs non-stantionary models (making use of random item models)

$$\eta_{pi} = \theta_p - \beta_i + \omega_{want} X_{i,do} Y_{p,i\ -12}$$


θp
Ypi
do resp
Yp,i-12
want resp

dep = with(VerbAgg, factor ((mode==“do”)*(r2 [mode==“want”]==“Y”)))

-1 + item + dep + (1|id)

random across persons

-1 + item + dep + (1+dep|id)

other forms of dependency

which other forms of dependency do you think are meaningful?
and how to implement them?

Remove for two examples

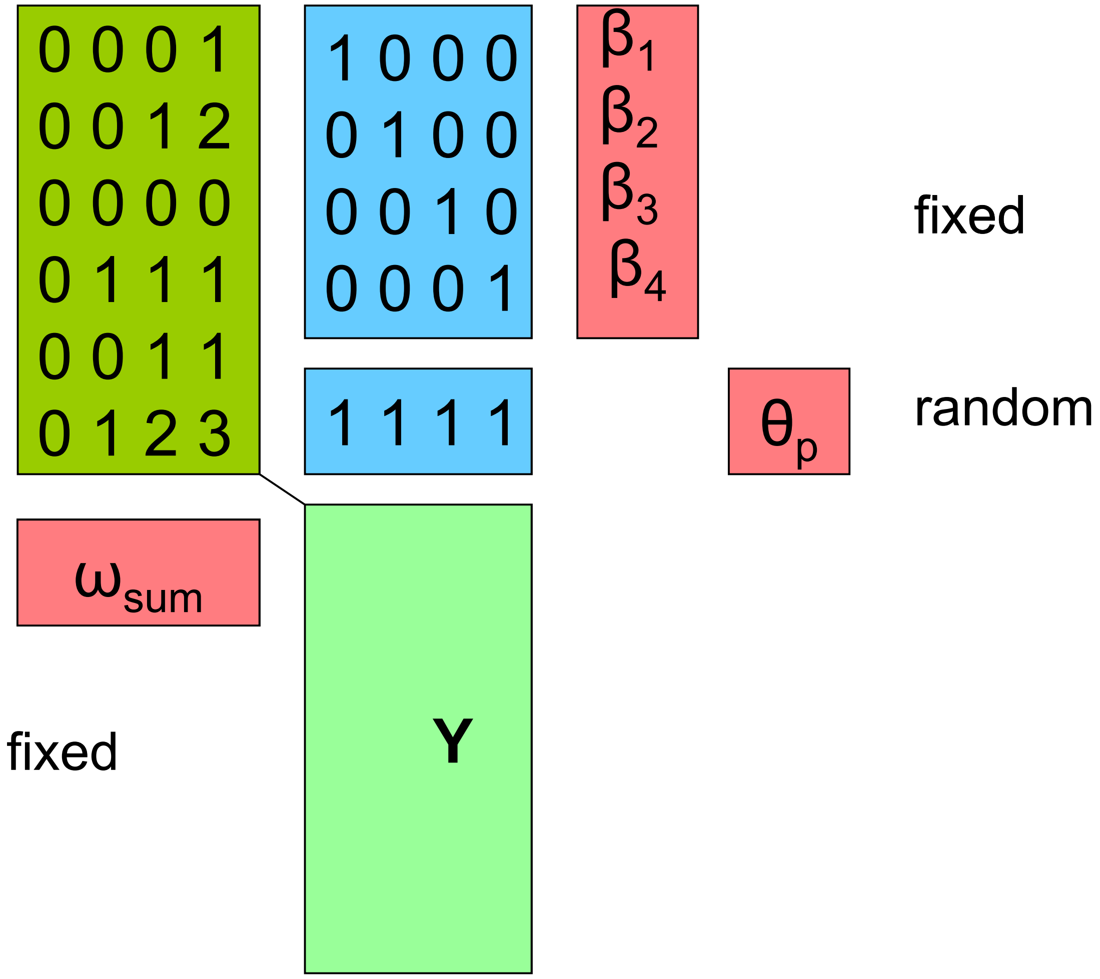
other forms of dependency

which other forms of dependency do you think are meaningful?
and how to implement them?

For example:

- serial dependency
  Y W
  1 -
- 1 1
  1 1
  0 1
  0 0
  0 0
  1 0
  1 1

- situational dependency
  random effect per situation
  after defining a new factor (situation)



3. Dynamic Rasch model

$$\eta_{pi} = \theta_p - \beta_i + \omega_{sum} W_{(p,i)sum}$$

```
long = data.frame(id=VerbAgg$id, item=VerbAgg$item, r2=VerbAgg$r2)
wide=reshape(long, timevar=c(“item”), idvar=c(“id”), dir=“wide”)[,-1]==“Y”
prosum=as.vector(t(apply(wide,1,cumsum)))
```

-1 + item + prosum + (1|id)
random across persons
-1 + item + prosum + (1+prosum|id)

Preparing a new dataset

- Most datasets have a wide format

Dataset

1 0 0 0 0 0 a

0 1 1 0 0 0 b

0 1 0 1 0 1 c

1 1 1 1 1 0 a

1 1 0 0 1 1 b

1 1 1 0 0 0 c

0 1 1 1 0 0 a

1 0 0 0 1 1 b

Type these data into a file “datawide.txt”

From wide to long

```
widedat=read.table(file="datawide.txt")
widedat$id=paste("id", 1:8, sep="")
   or
   widedat$id=paste("id",1:nrow(widedata),sep="")
library(reshape)
long=melt(widedat, id=7:8)
names(long)=c("con","id","item","resp")
```

S-J

Change type

from factor to numeric

```
long$connum=as.numeric(factor(long[,1]))
```

from numeric to factor

```
long$confac=factor(long[,5])
```

4a. Ordered-category data
4b. Structural Equation Models

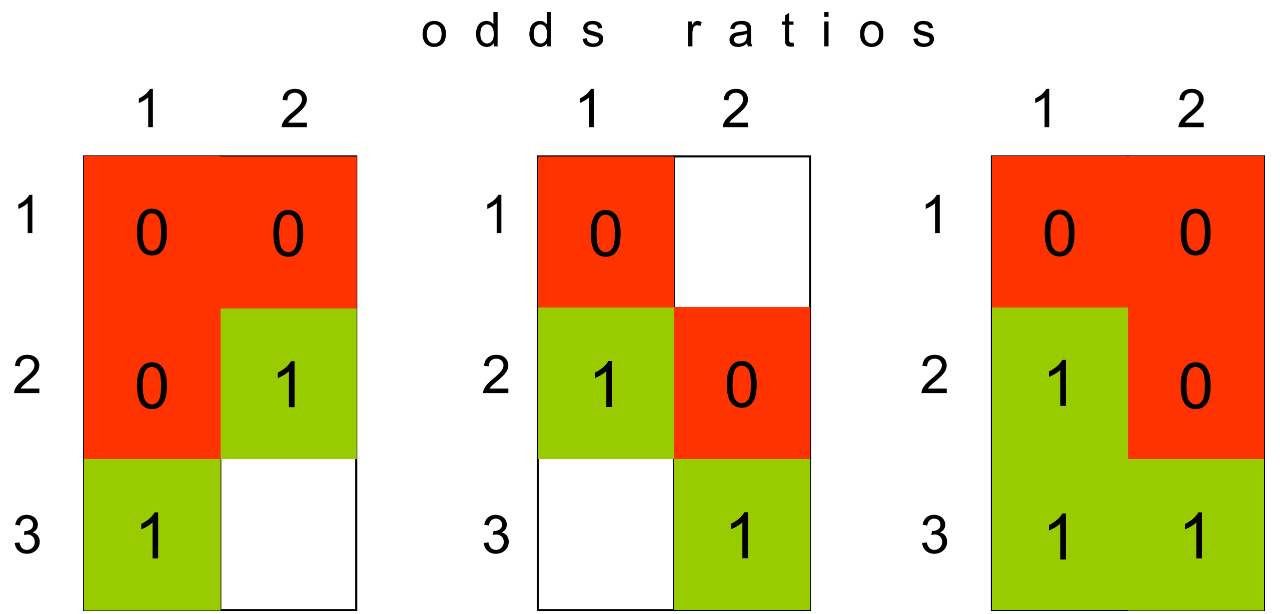
Vienna, June 6-10 2011

a. Ordered-category data

Models for ordered-category data

three types of odds ratios (green vs red)

for example, three categories, two odds ratios



continuation ratio	partial credit	graded response
P(Y=3)/P(Y=1,2)	P(Y=2)/P(Y=1)	P(Y=2,3)/P(Y=1)
P(Y=2)/P(Y=1)	P(Y=3)/P(Y=2)	P(Y=3)/P(Y=1,2)

$P(Y=3)$ follows Rasch model	$P_1(\theta_1)$
$P(Y=2 \mid Y \neq 3)$ follows Rasch model and is independent of $P(Y=3)$	$P_2(\theta_2)$

$P(Y=3)$	$P_1(\theta_1)$
$P(Y=2)=P(Y \neq 3)P(Y=2 \mid Y \neq 3)$	$(1-P_1(\theta_1)) \times P_2(\theta_2)$
$P(Y=1)=P(Y \neq 3)P(Y \neq 2 \mid Y \neq 3)$	$(1-P_1(\theta_1)) \times (1-P_2(\theta_2))$

Continuation ratio model is similar to discrete survival model

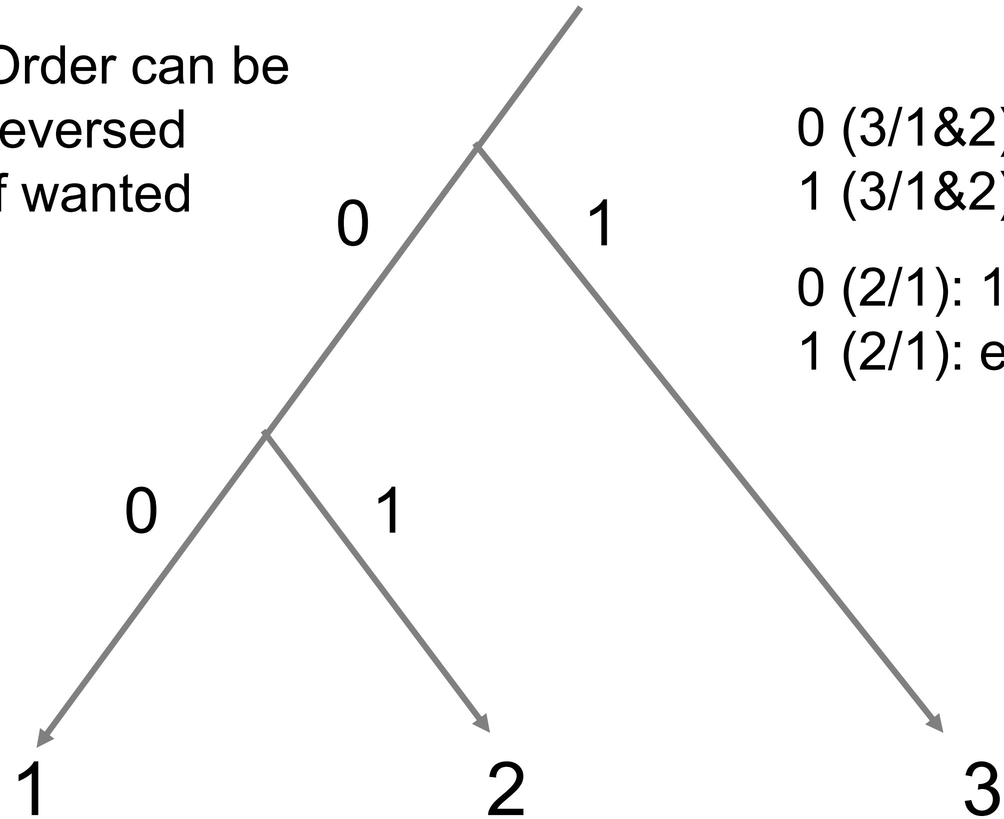
Choices are like decisive events in time

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

Continuation ratio model choice tree

Order can be reversed if wanted



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

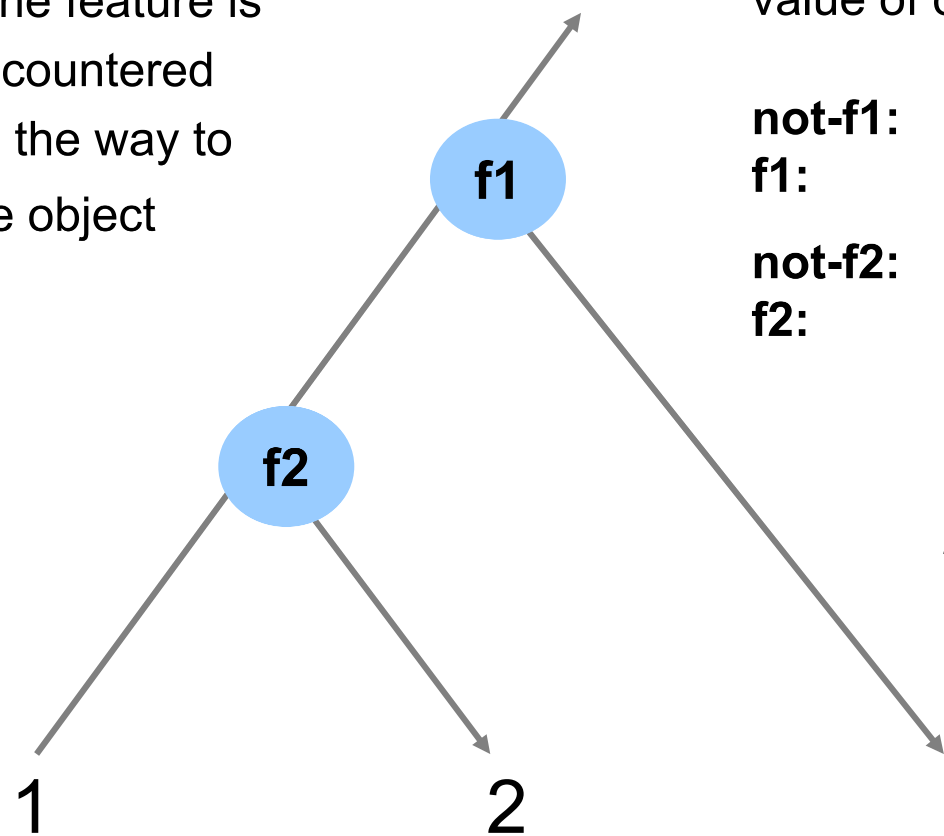
0 (2/1): $1/(1+\exp(\theta_{p2}-\beta_{i2}))$
1 (2/1): $\exp(\theta_{p2}-\beta_{i2})/(1+\exp(\theta_{p2}-\beta_{i2}))$

	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- : $\exp(\theta_{p1}-\beta_{i1}) \;/\; (1 \quad +\exp(\theta_{p1} - \beta_{i1}) \quad)$

partial credit model

An object has a feature
if the feature is
encountered
on the way to
the object



Choice probability
is value of object divided by sum of values of all objects
value of object = product of feature values

not-f1: $\exp(0)$
f1: $\exp(\theta_{p1})\exp(\beta_{i1})$

not-f2: $\exp(0)$
f2: $\exp(\theta_{p2})\exp(\beta_{i2})$

	f1	f2
1	0	0
2	0	1
3	1	1

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

extend dataset: replace each item response with two, except when missing:

1 00

2 01

3 1-

transformation can be done using `Tutzcoding` function in `R`.

VATutz=Tutzcoding(VerbAgg, “item”, “resp”) *label for*

recoded responses: tutz

subitems: newitems

subitem factor: category

estimation of common model

```
modelTutz=lmer(tutz~-1+newitem+(1|id),
   family=binomial,VATutz)
```

more Tutz models

rating scale version

    -1+item+category+(1|id)

gender specific rating scale model

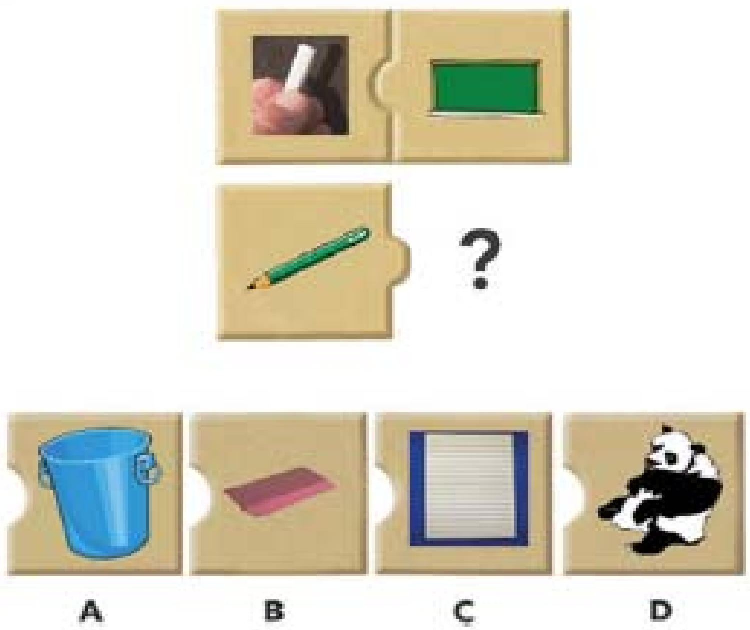
    -1+C(Gender,sum)*C(category,sum)+item+(1|id)

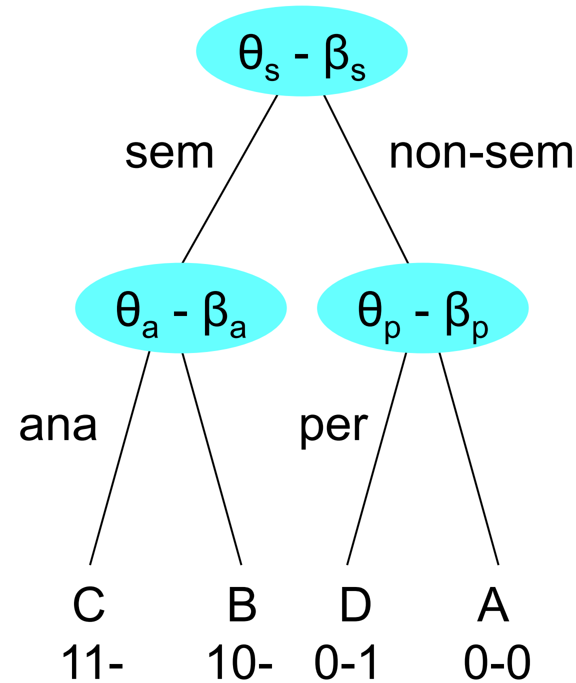
multidimensional: subitem specific dimensions

    -1+newitem+(-1+category|id)
    -1+item+category+(-1+category|id)   rating scale version

Branching variants

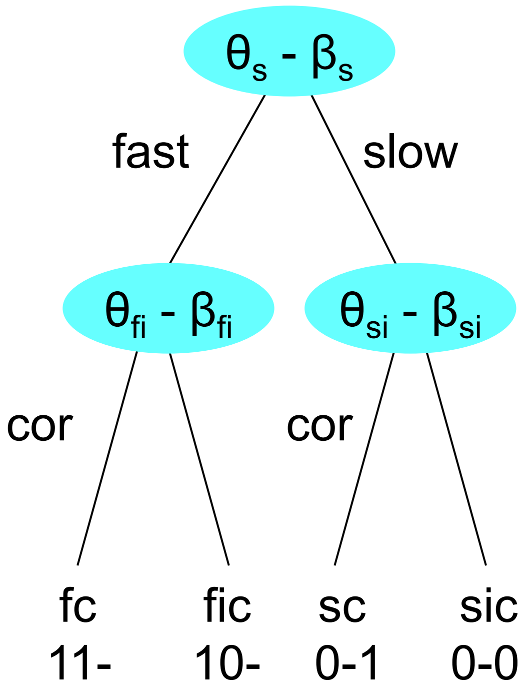
B.





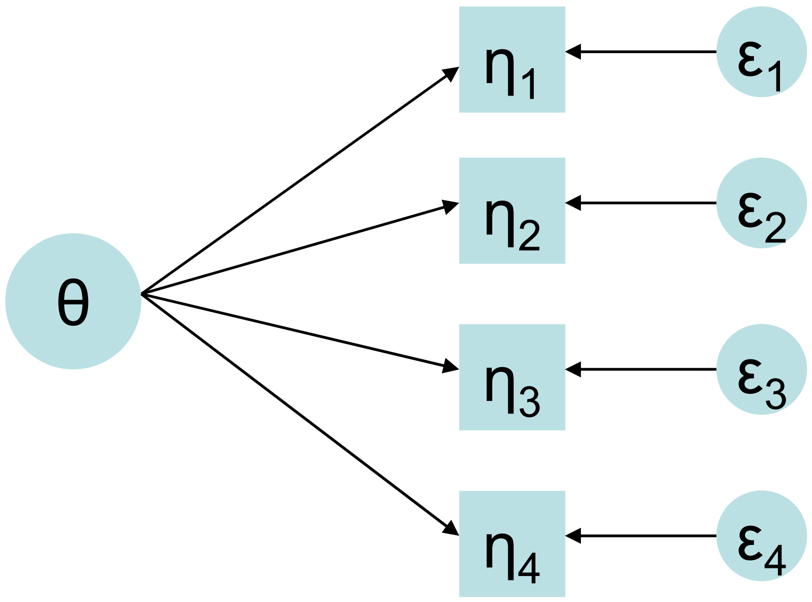
Branching variants

fast and slow intelligence



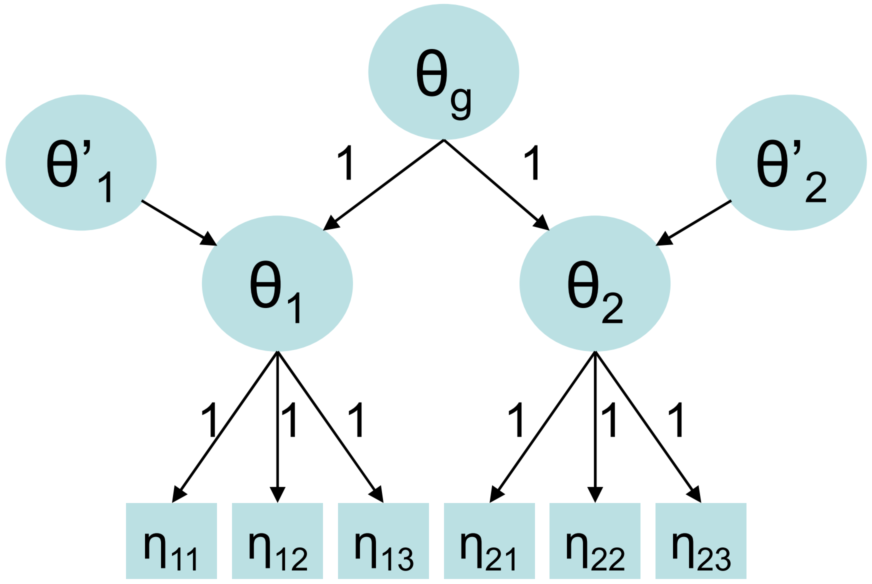
b. SEMs

- 2PL



-1 + item + (-1+item1|id) + (-1+item2|id) + .. + (1|id)

- Higher-order (Rasch?) models

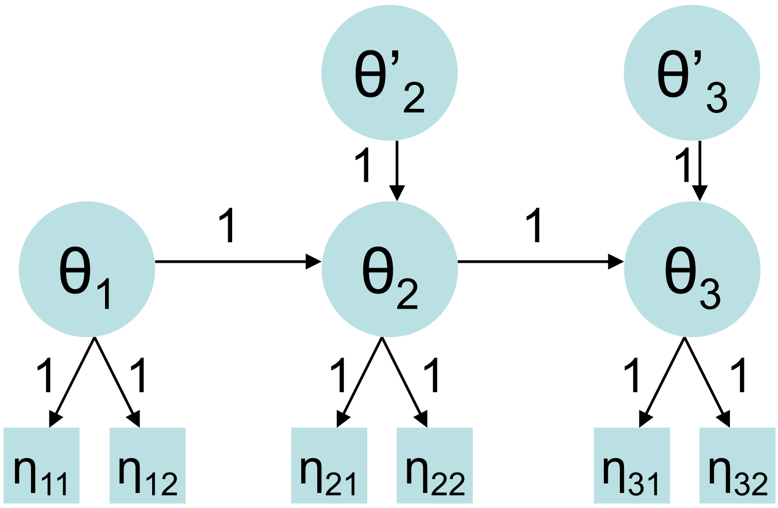


-1 + (1|item) + (-1+want|id) + (-1+do|id) + (1|id)

compare with

-1 + (1|item) + (-1+mode|id)

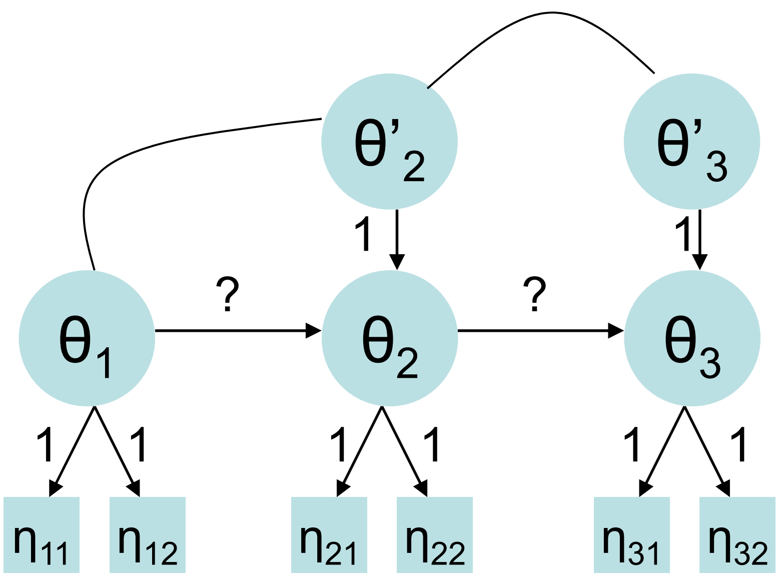
- Integrated process SEM



-1 + item + (1+scoldcurse+curse|id)

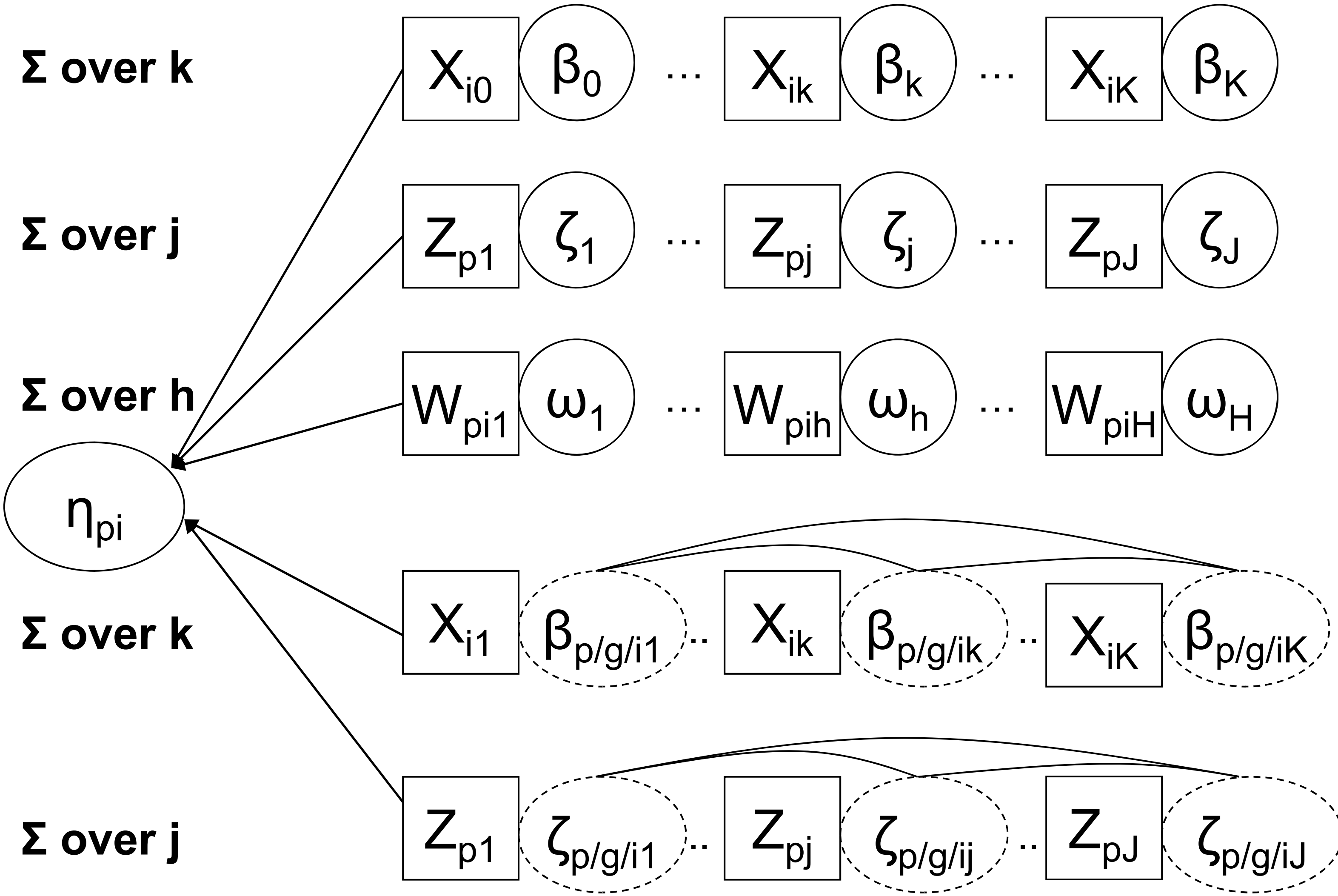
shout	1	0	0
scold	1	1	0
curse	1	1	1

- Mediation process SEM



-1 + item + (1+shoutscold|id)+(-1+shoutscold+curse|id)
    + (-1+curse|id)

shout	1	0	0
scold	1	1	0
curse	1	1	1


Σ over k
Σ over j
Σ over h
Σ over k
Σ over j
X_{i0} β_0 … X_{ik} β_k … X_{iK} β_K
Z_{p1} ζ_1 … Z_{pj} ζ_j … Z_{pJ} ζ_J
W_{pi1} ω_1 … W_{pih} ω_h … W_{piH} ω_H
η_{pi}
X_{i1} $\beta_{p/g/i1}$ .. X_{ik} $\beta_{p/g/ik}$ .. X_{iK} $\beta_{p/g/iK}$
Z_{p1} $\zeta_{p/g/i1}$ .. Z_{pj} $\zeta_{p/g/ij}$ .. Z_{pJ} $\zeta_{p/g/iJ}$

5. Estimation and testing

Vienna, June 6-10 2011

Estimation

- Laplace approximation of integrand

issue: integral is not tractable

solutions

1. approximation of integrand, so that it is tractable
2. approximation of integral
   Gaussian quadrate: non-adaptive or adaptive
3. Markov chain Monte Carlo

differences

- underestimation of variances using 1
- much faster using 1
- 1 is not ML, but most recent approaches are close

- approximation of integrand:
  PQL, PQL2, Laplace6
  MLwiN: PQL2
  HLM: Laplace6
  GLIMMIX: PQL
  lmer: Laplace
  Laplace6>Laplace>PQL2>PQL
- approximation of the integral
  SAS NLMIXED, gllamm, ltm, and many other
- MCMC
  WinBUGS, mlirt

Other R-programs

- **ltm** (Rizopoulos, 2006)
  1PL, 2PL, 3PL, graded response model
  included in **irtoys**
  Gaussian quadrature
- **eRm** (Mair & Hatzinger, 2007)
  Rasch, LLTM, partial credit model, rating scale model
  conditional maximum likelihood -- CML
- **mlirt** (Fox, 2007)
  2PNO binary & polytomous, multilevel

irtoys

calls among other things ltm

Illustration of ltm with irtoys

testing

problems

- strictly speaking no ML
- testing null hypothesis of zero variance
  LR Test does not apply
  - conservative test
  - mixture of $\chi^2(r)$ and $\chi^2(r+1)$ with mixing prob ½

```
m0=lmer( ..
m1=lmer( ..
anova(m0,m1)
```

z-tests

AIC, BIC $AIC = dev + 2N_{par}$

$BIC = dev + \log(P)N_{par}$

S-J