Item Response Theory in R using Package `ltm`

Dimitris Rizopoulos
Department of Biostatistics, Erasmus University Medical Center, the Netherlands

`d.rizopoulos@erasmusmc.nl`

Department of Statistics and Mathematics
WU Wirtschaftsuniversität Wien

January 12th, 2010

1.1 Introduction

- Item Response Theory (IRT) plays nowadays a central role in the analysis and study of tests and item scores

- Application of IRT models can be found in many fields
  - ▷ psychometrics
  - ▷ educational sciences
  - ▷ sociometrics
  - ▷ medicine
  - ▷ ...

1.1 Introduction

- A number of item response models exist in the statistics and psychometric literature for the analysis of multiple discrete responses

- Goals of this talk:
  - ▷ brief review of standard IRT models
  - ▷ estimation using marginal maximum likelihood
  - ▷ implementation in the freely available R package `ltm`

2.1 IRT Models for Dichotomous Data

- Notation
  - ▷ x_{im}: response of the mth subject in the ith item
  - ▷ z_m: latent variable (e.g., latent ability) – typically, $z_m \in (-\infty, \infty)$

- We are interested in the relation between the probability of a positive (e.g., correct) response in item i from subject m, and the value of her latent ability z_m

$$\text{Pr}(x_{im} = 1 \mid z_m; \theta)$$

θ: parameters describing this relation

2.1 IRT Models for Dichotomous Data (cont'd)

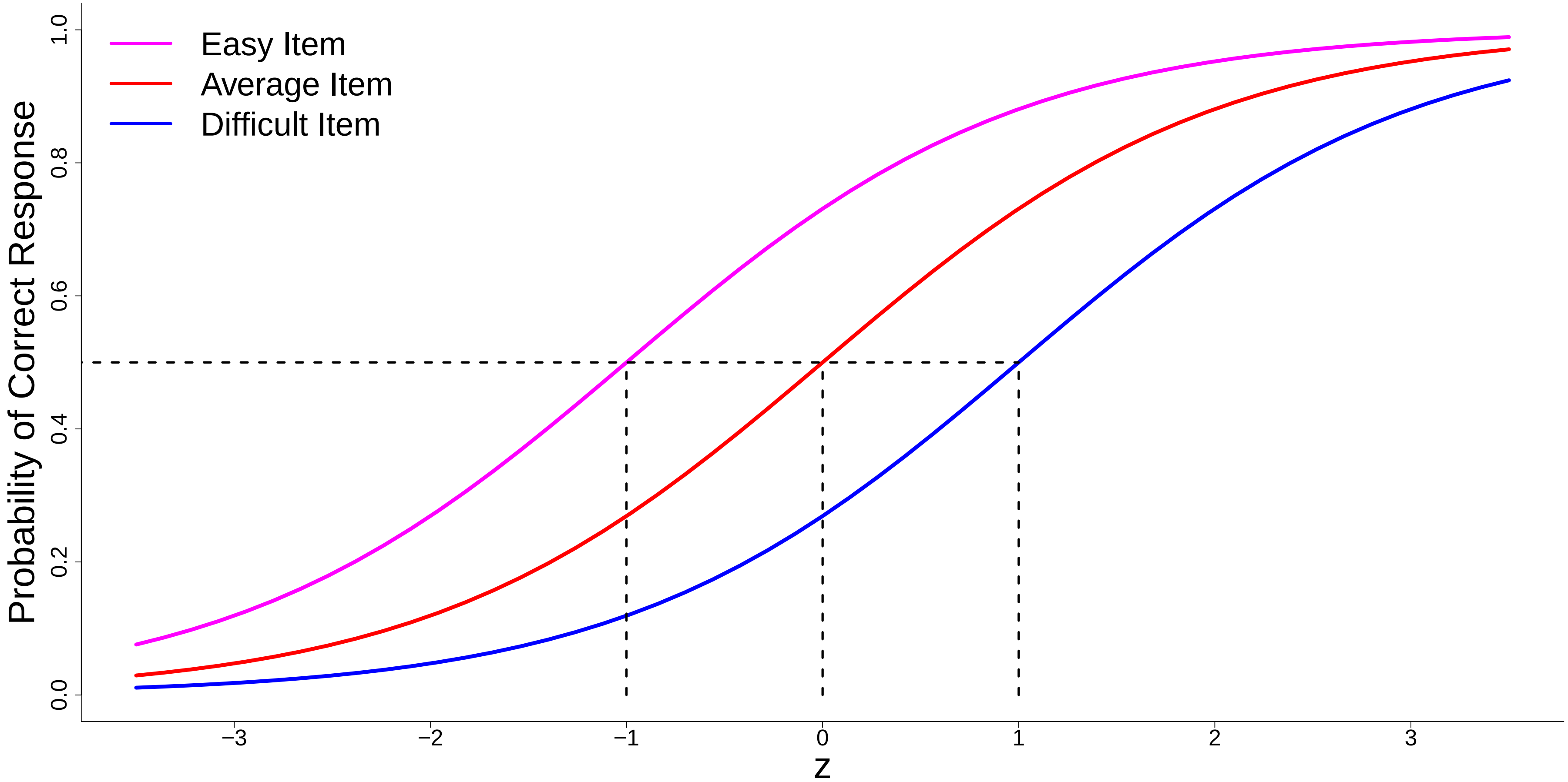
- The Rasch model

$$\mathsf{Pr}(x_{im} = 1 \mid z_m; \theta) = \frac{\exp(z_m - \beta_i)}{1 + \exp(z_m - \beta_i)}$$

where β_i is the difficulty parameter

- Properties and Features
  - ▷ closed-form sufficient statistics
  - ▷ all items have the same discrimination power
  - ▷ widely used

2.1 IRT Models for Dichotomous Data (cont'd)



2.1 IRT Models for Dichotomous Data (cont'd)

- The one-parameter logistic model

$$\Pr(x_{im} = 1 \mid z_m; \theta) = \frac{\exp\{\alpha(z_m - \beta_i)\}}{1 + \exp\{\alpha(z_m - \beta_i)\}}$$

where α is a common discrimination parameter

- Properties and Features
  - ▷ common discrimination not fixed at one ⇒
  - ▷ a bit more flexible than the Rasch model

2.1 IRT Models for Dichotomous Data (cont'd)

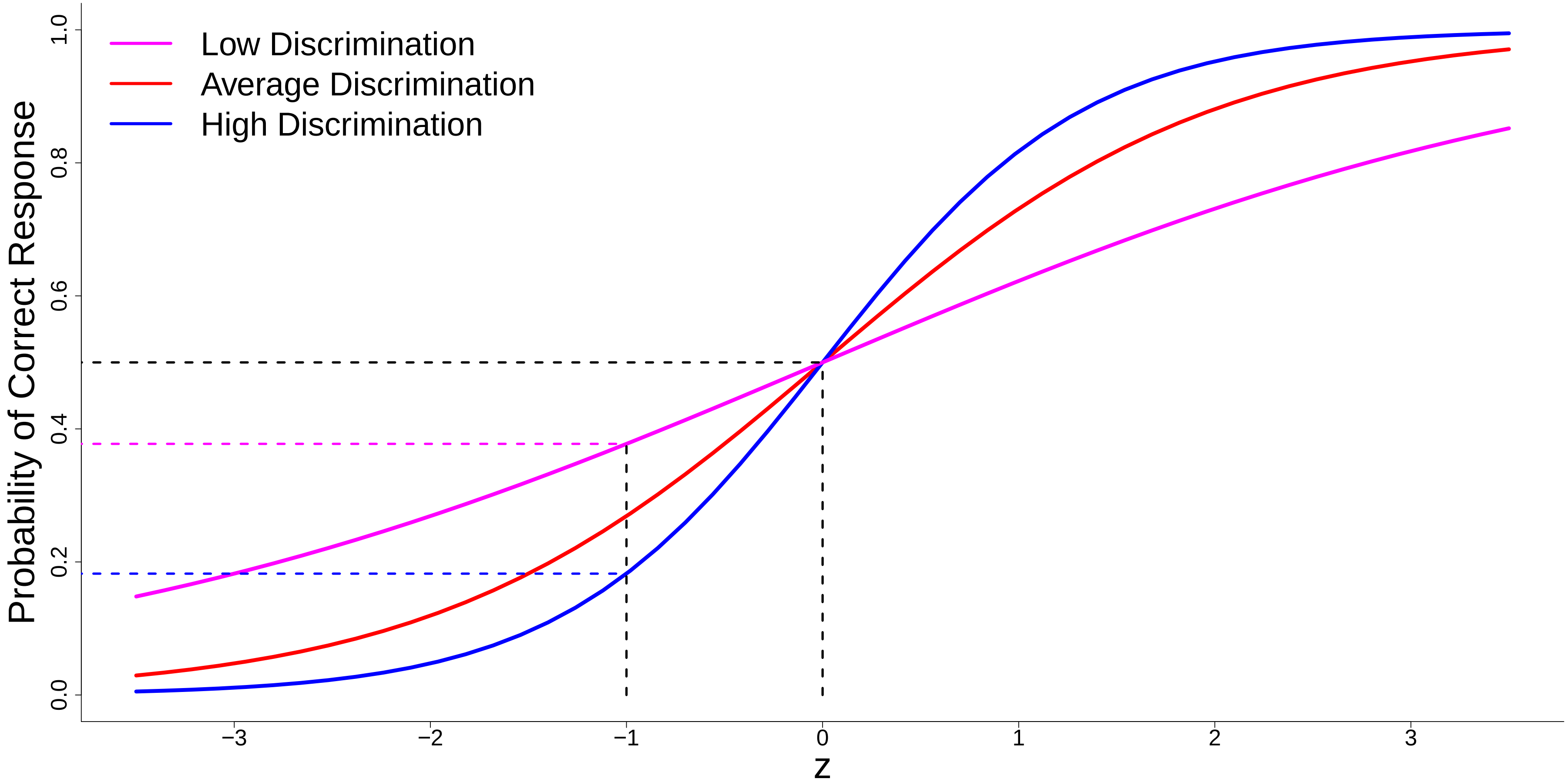
- The two-parameter logistic model

$$\text{Pr}(x_{im} = 1 \mid z_m; \theta) = \frac{\exp\{\alpha_i(z_m - \beta_i)\}}{1 + \exp\{\alpha_i(z_m - \beta_i)\}}$$

where we now have a different discrimination parameter per item

- Properties and Features
  - ▷ no more closed-form sufficient statistics

2.1 IRT Models for Dichotomous Data (cont'd)



2.1 IRT Models for Dichotomous Data (cont'd)

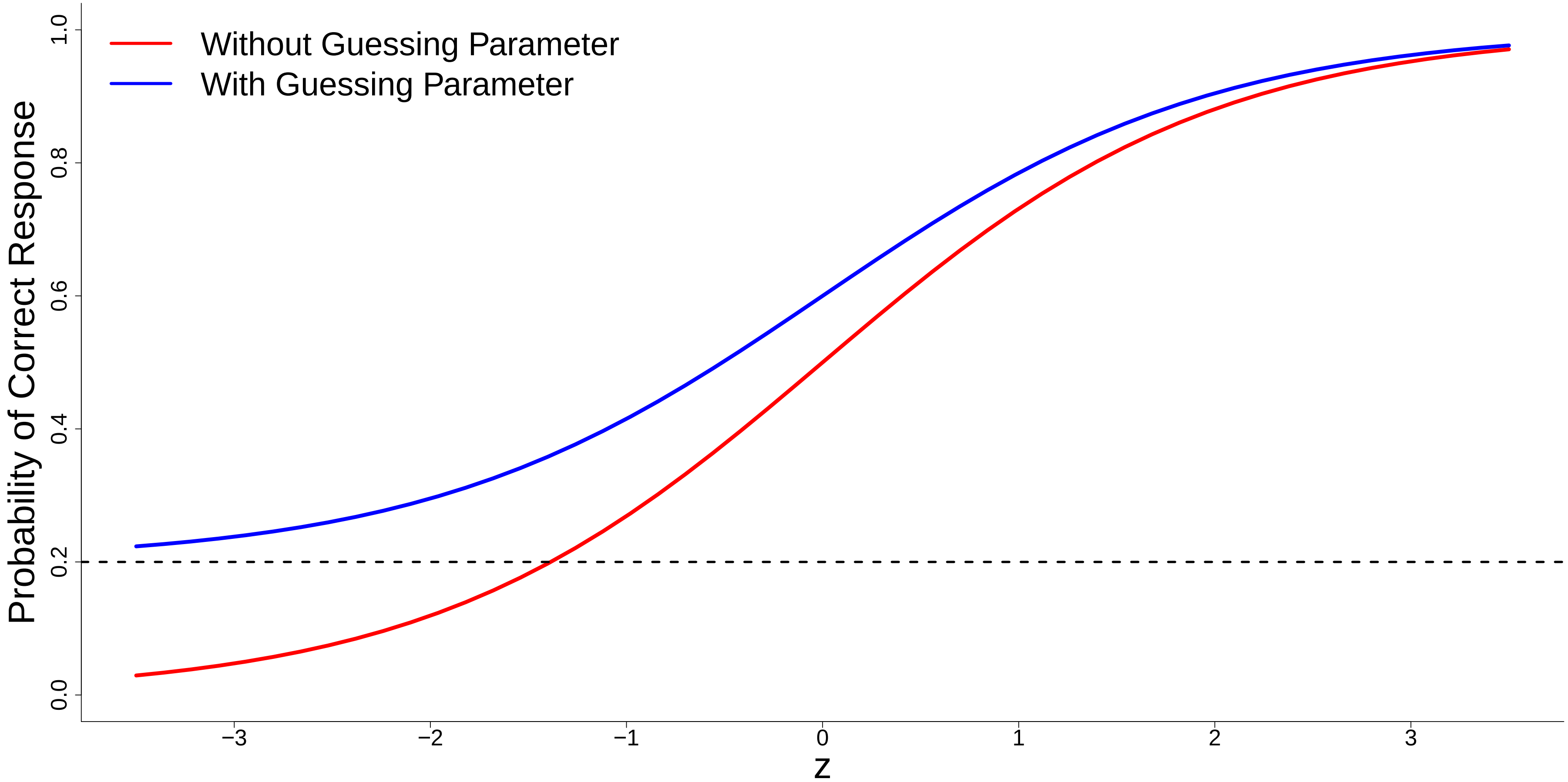
- The three-parameter model

$$\text{Pr}(x_{im} = 1 \mid z_m; \theta) = c_i + (1 - c_i)\frac{\exp\{\alpha_i(z_m - \beta_i)\}}{1 + \exp\{\alpha_i(z_m - \beta_i)\}}$$

where c_i is a guessing parameter

- Properties and Features
  - ▷ numerically less stable

2.1 IRT Models for Dichotomous Data (cont'd)



2.2 IRT Models for Polytomous Data

- Notation: the same as for dichotomous data but now x_{im} can take K_i possible values

- Examples:
  - ▷ "bad", "good", "very good", and "excellent"
  - ▷ "very concerned", "slightly concerned" and "not very concerned"
  - ▷ . . .

2.2 IRT Models for Polytomous Data (cont'd)

- The generalized partial credit model (Masters, 1982; Muraki, 1992)

$$\Pr(x_{im} = k \mid z_m; \theta) = \frac{\exp \sum_{c=0}^{k} \alpha_i (z_m - \beta_{ic})}{\sum_{r=0}^{K_i} \exp \sum_{c=0}^{r} \alpha_i (z_m - \beta_{ic})}$$

- Properties and Features
  - ▷ Rasch version: $\alpha_i = 1$ for all items
  - ▷ 1PL version: $\alpha_i = \alpha$ for all items
  - ▷ GPCM version: different α_i per item

2.2 IRT Models for Polytomous Data (cont'd)

- The graded response model (Samejima, 1962)

$$\mathsf{Pr}(x_{im} \leq k \mid z_m; \theta) = \frac{\exp\{\alpha_i(z_m - \beta_{ik})\}}{1 + \exp\{\alpha_i(z_m - \beta_{ik})\}}$$

$$\mathsf{Pr}(x_{im} = k \mid z_m; \theta) = g(\eta_{ik}) - g(\eta_{i,k+1})$$

where $\eta_{ik} = \alpha_i(z_m - \beta_{ik})$, $g(\eta) = \exp(\eta)/\{1 + \exp(\eta)\}$, and $\beta_1 < \beta_2 < \ldots < \beta_{K_i}$

- Properties and Features
  - ▷ constrained version: $\alpha_i = \alpha$ for all items
  - ▷ unconstrained version: different α_i per item

3.1 Marginal Maximum Likelihood Estimation

- We assume that the subjects represent a random sample from a population and their ability is distributed according to a distribution function $F(z)$ (e.g., Normal)

- We integrate out the latent abilities to obtain the marginal likelihood for the responses of the mth subject

$$\ell_m(\theta) = \log p(x_m; \theta) = \log \int p(x_m | z_m; \theta) \, p(z_m) \, dz_m$$

- Conditional independence assumption:
  - ▷ given the latent variable value z_m, the responses $x_m = (x_{1m}, \ldots, x_{pm})$ in the p items are assumed independent

3.1 Marginal Maximum Likelihood Estimation

- To estimate the parameters we require a combination of
  - ▷ numerical integration because the above integral does not have a closed-form solution (Gaussian quadrature)
  - ▷ numerical optimization (EM, Newton-Raphson, quasi-Newton)

3.2 Estimating Latent Abilities

- After fitting the desired IRT model we often require to obtain an estimate for z_m
  - ▷ that is, what is the most plausible value of the latent ability for subject m given his responses x_m and the assumed IRT model

- In particular, we are interested in

$$p(z_m \mid x_m; \theta) = \frac{p(x_m \mid z_m; \theta)\, p(z_m)}{p(x_m; \theta)}$$

3.2 Estimating Latent Abilities (cont'd)

- Note that $p(z_m \mid x_m; \theta)$ is a whole distribution $\Rightarrow$ we need a summary measure

- Maximum aposteriori

$$\hat{z}_m = \arg\max_z \{p(z \mid x_m; \theta)\} = \arg\max_z \{\log p(x_m \mid z; \theta) + \log p(z)\}$$

- Expected aposteriori

$$\tilde{z}_m = \int z \, p(z \mid x_m; \theta) \, dz$$

4 Design of Package `ltm`

- Package `ltm` can be freely downloaded from CRAN and perform IRT analysis under a marginal maximum likelihood approach – the design of the package is as follows

1. Descriptive Analysis
   - ▷ `descript()`: descriptive statistics relevant to IRT
   - ▷ `rcor.test()`: pairwise associations
   - ▷ `biserial.cor()`: biserial correlation
   - ▷ `cronbach.alpha()`: calculates Cronbach's alpha
   - ▷ `unidimTest()`: unidimensionality check

4 Design of Package `ltm` (cont'd)

2. Fitting IRT models

   ▷ `rasch()`: Rasch and 1PL models

   ▷ `ltm()`: 2PL and latent trait models with two latent variables (and nonlinear terms)

   ▷ `tpm()`: three parameter model

   ▷ `gpcm()`: generalized partial credit models (including the Rasch and 1PL versions)

   ▷ `grm()`: graded response model (including the constrained and unconstrained versions)

4 Design of Package `ltm` (cont'd)

3. Hypothesis testing & goodness-of-fit

   ▷ `anova()`: likelihood ratio test between nested IRT models

   ▷ `GoF.rasch()` & `GoF.gpcm()`: goodness-of-fit test based on Pearson's χ^2 (also Bootstrap approximation available)

   ▷ `margins()`: fit on the two- and three-way margins using Pearson's χ^2

   ▷ `item.fit()` & `person.fit()`: item & person fit statistics

4 Design of Package `ltm` (cont'd)

4. Functions to calculate or extract quantities from fitted IRT models

   ▷ `summary()`: summarize fitted IRT models

   ▷ `factor.scores()` ability estimates

   ▷ `plot()`: item characteristic & information curves

   ▷ `information()`: area under item or test information curves

   ▷ `residuals()`: χ^2 residuals

   ▷ `fitted()`: several types of fitted values

4 Design of Package `ltm` (cont'd)

5. Utility functions

   ▷ `rmvlogis()`: simulate random responses from IRT models for dichotomous data

   ▷ `rmvordlogis()`: simulate random responses from IRT models for polytomous data

   ▷ `mult.choice()`: converts multiple choice items to a matrix of binary responses

   ▷ `testEquatingData()`: combine items from different forms/tests

5.1 Application Dichotomous

- Data from the 1990 British social attitudes survey: questions on sexual attitudes

- The 10 questions asked were:

  (1) Should divorce be easier?

  (2) Do you support the law against sexual discrimination?

  (3) View on pre-marital sex: (wrong/not wrong)

  (4) View on extra-marital sex: (wrong/not wrong)

  (5) View on sexual relationship between individuals of the same sex: (wrong/not wrong)

  (6) Should gays teach in school?

5.1 Application Dichotomous (cont'd)

(7) Should gays teach in higher education?

(8) Should gays hold public positions?

(9) Should a female homosexual couple be allowed to adopt children?

(10) Should a male homosexual couple be allowed to adopt children?

- IRT analysis of this data set

5.2 Application Polytomous

- Data from the 1990 British social attitudes survey: questions on environmental issues

- The response options to the following questions were 'very concerned', 'slightly concerned' and 'not very concerned'

  (1) Lead from petrol

  (2) River and sea pollution

  (3) Transport and storage of radioactive waste

  (4) Air pollution

  (5) Transport and disposal of poisonous chemicals

  (6) Risks from nuclear power station

Thank you for your attention!

More Information for `ltm` is available at:

http://wiki.r-project.org/rwiki/doku.php?id=packages:cran:ltm