Fixed-effects estimation of 2PL models

Ruggero Bellio

Department of Economics & Statistics, University of Udine (Italy)

Joint work with: I. Kosmidis (UCL), N. Sartori (Padova)

Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Item response data

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
3	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0
4	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1
5	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
6	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0
7	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
9	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
10	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
11	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1
12	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1
13	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
14	0	0	0	1	0	0	1	0	0	1	0	0	1	1	0	1	1	0	0	0
15	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0
16	0	1	0	1	1	1	1	0	1	1	1	0	1	0	0	1	1	1	0	0
17	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0
20	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	0	1	0
21	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0
22	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	1	1	1	0	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0
24	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1
25	0	1	0	1	0	0	1	1	1	1	0	0	1	1	1	0	1	0	0	1

The 2PL model

- I items, S subjects, $S \times I$ independent binary responses Y_{si}
- Let $\pi_{si} = P(Y_{si} = 1)$, then

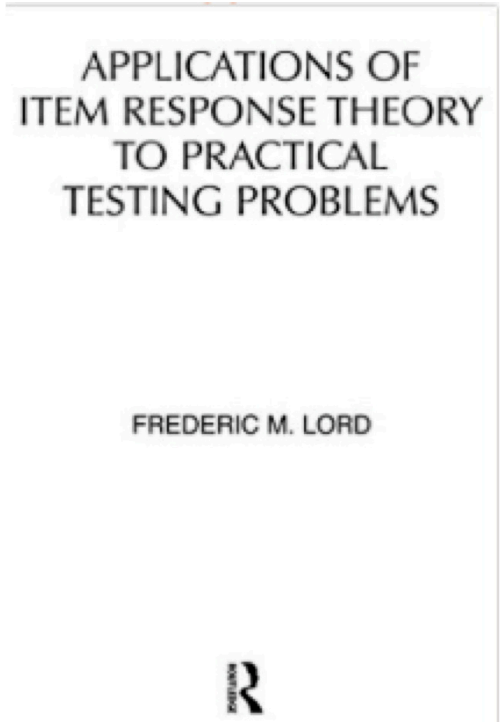
$$\boxed{\text{logit}(\pi_{si}) = \beta_{0i} + \beta_{1i}\,\theta_s}$$

  - β_{0i} and β_{1i} **item parameters** for item i
  - θ_s **ability parameter** for subject s

(Slightly different from usual formulation with difficulty parameter $\beta^*_{0i} = (-\beta_{0i}/\beta_{1i})$).

Historical development: JML

- Early usage of 2PL model took the abilities $\theta_s, s = 1, \ldots, S$ as **fixed parameters** (e.g. Lord, 1980), estimated by **(Joint) Maximum Likelihood (JML)** that jointly estimates the item parameters and the person parameters.



- The fixed-effects approach is logically very simple (San Martín et al., 2015), yet JML is hampered by numerical difficulties, and it was soon abandoned.

MML took the spotlight

Marginal Maximum Likelihood (MML) assuming a normal distribution for θ_s (Bock and Aitkin, 1981)

$$\boxed{\theta_s \sim N(0,1)}, \qquad s = 1, \ldots, S$$

became the standard in applications.

Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm
RD **Bock**, M **Aitkin** - Psychometrika, 1981 - Springer
Abstract Maximum likelihood estimation of item parameters in the marginal distribution, integrating over the distribution of ability, becomes practical when computing procedures based on an EM algorithm are used. By characterizing the ability distribution empirically, ...
Cited by 1915 Related articles All 11 versions Cite Save More

Pros and cons of JML and MML

JML

Pros Simplicity, robustness

Cons Considerable probability to result in infinite estimates, incidental parameter problem (Neyman and Scott, 1948)

MML

Pros Shrinkage from normality of abilities, which is a reasonable assumption in many cases (and it washes away with large I)

Cons Requires integration of latent abilities, at times normality unjustified

Aim of our proposal: getting the pros of both methods, with only a small fraction of the cons!

Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Our proposal

Orthodox fixed-effects approach, θ_s treated as fixed parameters.

Made of two parts:

1. Joint estimation of item and person parameters by **bias-reduced estimation (BR)**.
2. Elimination of item parameters by means of the **Modified Profile Likelihood (MPL)**, built upon the BR estimates.

Note: BR already provides a valid set of estimates. *Why do we need also MPL*? Two reasons

- Better mathematical properties (in principle), easier to study.
- MPL-based estimates are obtained by minimizing an objective function, whereas the BR ones by solving estimating equations, and this has some advantages (e.g. lasso-based model selection).

Some notation

- Model $\boxed{\text{logit}(\pi_{si}) = \beta_{0i} + \beta_{1i}\theta_s}$
- Model parameters

Easyness parameters	$\boldsymbol{\beta}_0$	$=$	$(\beta_{01}, \ldots, \beta_{0I})$
Discrimination parameters	$\boldsymbol{\beta}_1$	$=$	$(\beta_{11}, \ldots, \beta_{1I})$
Item parameters	$\boldsymbol{\beta}$	$=$	$(\boldsymbol{\beta}_0, \boldsymbol{\beta}_1)$
Abilities	$\boldsymbol{\theta}$	$=$	$(\theta_1, \ldots, \theta_S)$
All together	$\boldsymbol{\omega}$	$=$	$(\boldsymbol{\beta}, \boldsymbol{\theta})$

- For identification purposes $\beta_{01} = 0$ and $\beta_{11} = 1$, total number of parameters is $2(I-1) + S$.
- Log-likelihood function

$$\ell(\boldsymbol{\omega}) = \sum_{s=1}^{S} \sum_{i=1}^{I} \left[y_{si} \left(\beta_{0i} + \theta_s \, \beta_{1i} \right) - \log \left\{ 1 + \exp(\beta_{0i} + \theta_s \, \beta_{1i}) \right\} \right] .$$

Bias-Reduced estimation

- Smaller asymptotic bias than maximum likelihood estimation (Firth, 1993; Kosmidis and Firth, 2009).
- For categorical responses, **finiteness and shrinkage properties** (Heinze, G. and Schemper, 2002; Kosmidis, 2014).
- Solves the **adjusted score equations**

$$\frac{\partial \ell(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} + \boldsymbol{A}(\boldsymbol{\omega}) = 0$$

with $\boldsymbol{A}(\boldsymbol{\omega})$ depending on the expected Fisher information and on higher moments of the log-likelihood derivatives.

Some more details

- The vector $\boldsymbol{A}(\boldsymbol{\omega})$ has t-th component

$$A_t(\boldsymbol{\omega}) = \frac{1}{2} \operatorname{tr}\left[\{i(\boldsymbol{\omega})\}^{-1} \{p_t(\boldsymbol{\omega}) + q_t(\boldsymbol{\omega})\}\right]$$

  where $i(\boldsymbol{\omega}) = -E_{\boldsymbol{\omega}}\{\ell_{\boldsymbol{\omega}\boldsymbol{\omega}}(\boldsymbol{\omega})\}$ is the Fisher information matrix for $\boldsymbol{\omega}$, and

$$p_t(\boldsymbol{\omega}) = E_{\boldsymbol{\omega}}\left\{\ell_{\boldsymbol{\omega}}(\boldsymbol{\omega})\ell_{\boldsymbol{\omega}}(\boldsymbol{\omega})^\top \ell_{\boldsymbol{\omega},t}(\boldsymbol{\omega})\right\}$$
$$q_t(\theta) = E_{\boldsymbol{\omega}}\left\{\ell_{\boldsymbol{\omega}\boldsymbol{\omega}}(\boldsymbol{\omega})\ell_{\boldsymbol{\omega},t}(\boldsymbol{\omega})\right\},$$

  are higher-order joint null moments of log-likelihood derivatives.

- Can be expressed in compact form for 2PL models.

Inference on item parameters based on MPL

- In many settings $S >> I$, thus $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are estimated with different precision.
- We treat $\boldsymbol{\beta}$ as the **parameter of interest** and $\boldsymbol{\theta}$ as **nuisance parameter**, and we resort to suitable methodology.
- The Modified Profile Likelihood (MPL) is a general method for removing the effect of nuisance parameters. Here we use the version defined in Severini (1998)

$$\ell_M(\boldsymbol{\beta}) = \ell(\boldsymbol{\beta}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{\beta}}) + \sum_{s=1}^{S} M_s(\widehat{\boldsymbol{\beta}}, \widehat{\theta}_s; \boldsymbol{\beta}, \widehat{\theta}_{s,\boldsymbol{\beta}})$$

  where $M_s(\cdot)$ is an additive adjustment.
- The estimates employed for the parameters are unconstrained and constrained BR estimates.

Some more details

$M_s(\cdot)$ is a simple function of moments involving the score function for θ_s

$$M(\boldsymbol{\beta}) = \sum_{s=1}^{S} \left[\frac{1}{2} \log\{i_{\theta_s\theta_s}(\boldsymbol{\beta}, \widehat{\theta}_{s,\boldsymbol{\beta}})\} - \log\{\widehat{I}_{\theta_s\theta_s}(\boldsymbol{\beta}, \widehat{\theta}_{s,\boldsymbol{\beta}})\} \right]$$

with

$$\begin{aligned} i_{\theta_s\theta_s}(\boldsymbol{\beta}, \widehat{\theta}_{s,\boldsymbol{\beta}}) &= \sum_{i=1}^{I} \beta_{1i}^2 \, \pi(\beta_{0i}, \beta_{1i}, \widehat{\theta}_{s,\boldsymbol{\beta}}) \, \{1 - \pi(\beta_{0i}, \beta_{1i}, \widehat{\theta}_{s,\boldsymbol{\beta}})\} \, , \\ \widehat{I}_{\theta_s\theta_s}(\boldsymbol{\beta}, \widehat{\theta}_{s,\boldsymbol{\beta}}) &= \sum_{i=1}^{I} \widehat{\beta}_{i1} \, \beta_{i1} \, \pi(\widehat{\beta}_{i0}, \widehat{\beta}_{i1}, \widehat{\theta}_s) \, \{1 - \pi(\widehat{\beta}_{i0}, \widehat{\beta}_{i1}, \widehat{\theta}_s)\} \, . \end{aligned}$$

The case of 1PL (Rasch) model

- In the 1PL model $\beta_{1s} = 1$, a classical fixed-effects approach is given by the **Conditional Likelihood (CL)**.
- The modified profile likelihood in general approximates marginal or conditional likelihoods, when available, so in 1PL models it (essentially) recovers the CL estimation.
- The BR method always gives finite estimates, being equal to Weighted Likelihood Estimation (WLE) (Warm, 1989).
- In 1PL models, we endorse the classical strategy

  CL (MPL) for item parameters + BR for person parameters

  Our proposal for 2PL models is very similar!

Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Modified profile likelihood: properties

- In the Rasch model, the CL method is $\sqrt{S}$-consistent, with estimation accuracy improving with S regardless of I.
- In 2PL models, the MPL approach is not $\sqrt{S}$-consistent, and for formal consistency it is required that
  - i) the number of subjects S grows to infinity
  - ii) the number of items I grows to infinity, but possibly with a slower rate than S.

Modified profile likelihood: properties

- Following Sartori (2003), it is possible to prove that for any element ψ of $\boldsymbol{\beta}$
  - (a) Let $\widehat{\psi}$ be the JML estimator. The score test for ψ is asymptotically $N(0,1)$ with error of order $O_p(\sqrt{S}/I)$, and the usual asymptotic inferential results are obtained when

$$\boxed{S = o(I^2)}$$

  - (b) For the estimator obtained from the MPL method $\widehat{\psi}_M$, the error is $O_p(\sqrt{S}/I^2)$, the usual asymptotic inferential results are valid when

$$\boxed{S = o(I^4)}$$

  - (c) The asymptotic bias is of order $O(I^{-1})$ for $\widehat{\psi}$, whereas it is of order $O(I^{-2})$ for $\widehat{\psi}_M$.

Practical implications

- **For a given number of subjects** S, a much larger number of items is required for the JML method to obtain the same accuracy of the MPL method.
- **For a given number of items** I, if $S \to \infty$ then inference based on the JML method will eventually break down.

  This will happen also for the MPL method, but the number of subjects handled by it for fixed number of items will be much larger.

Remarks

- The accuracy of MPL is better than for fixed-effects panel data models reported in Bartolucci et al. (2015).

  This is not surprising, as the IRT setting is more favourable (Haberman, 1977).
- The fact that the accuracy of MPL depends on I may appear to be in favor of MML, which achieves $\sqrt{S}$-consistency *when the model is correctly specified*.

  For non-normal θ_s, the accuracy of MML may be as good as that of JML (Arellano and Bonhomme, 2009).

Outline

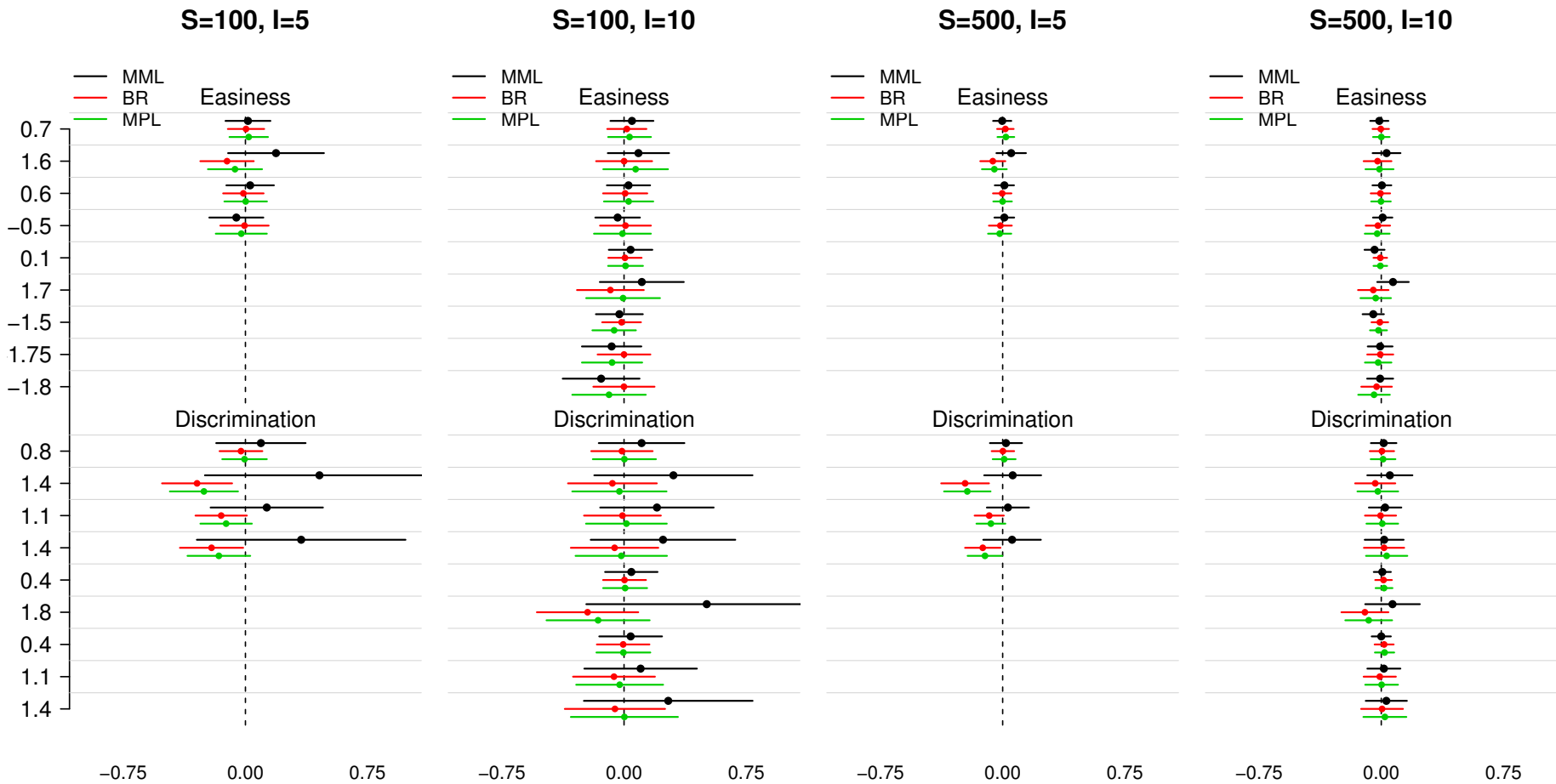
- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Simulation study

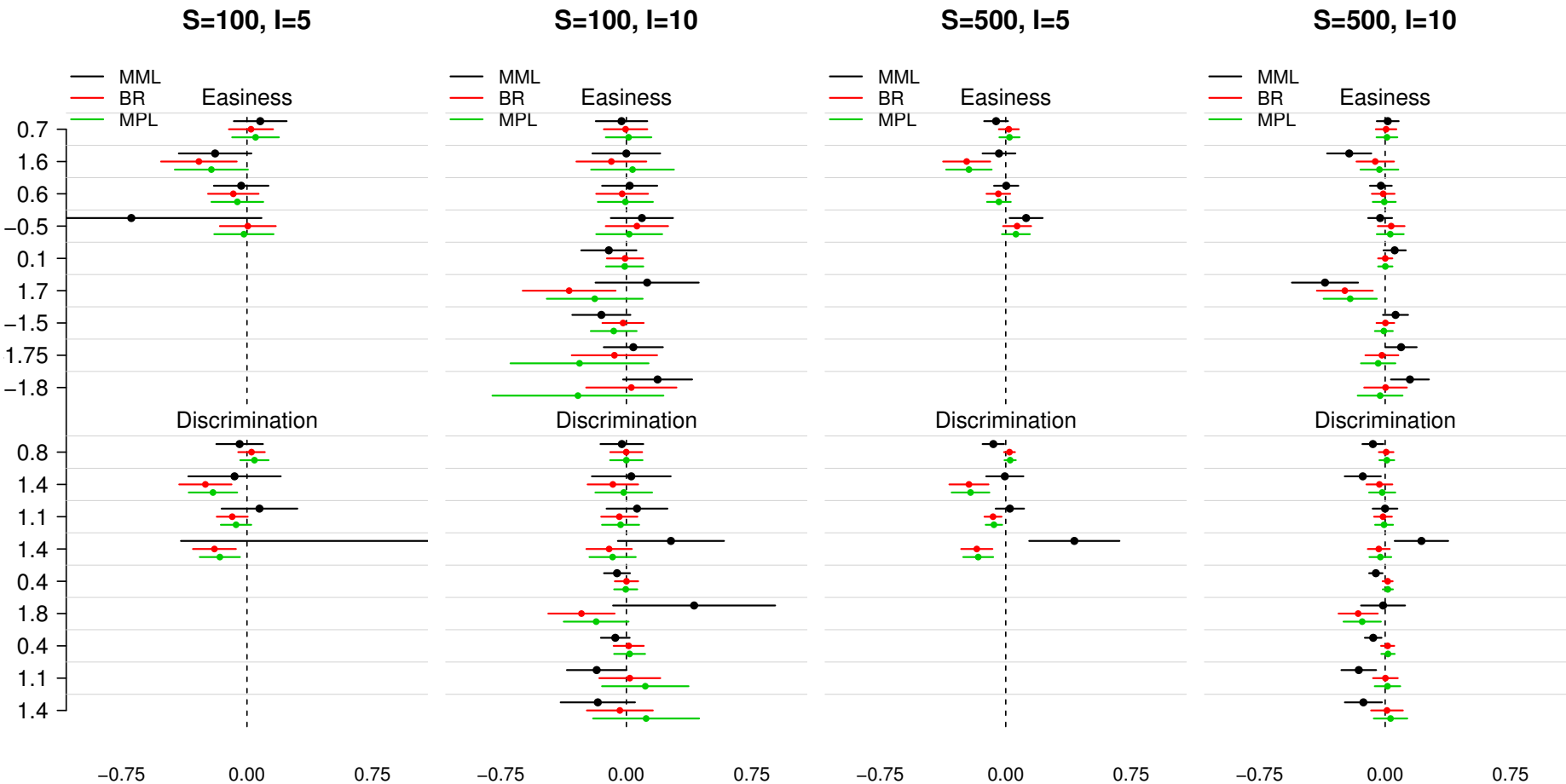
- Three different scenarios for abilities
  1. $\theta_s \sim N(0, 1)$, the most favourable setting for MML;
  2. θ_s from a mixture of two normal distributions;
  3. θ_s from a zero-inflated mixture, following Wall et al. (2015). This is a setting with 50-60% subjects with $y_{si} = 0$, where MML essentially breaks down.
- Item parameters chosen following early literature on 2PL models, some large discrimination parameters.
- Simulations (1,000 draws) for $S = 100, 500$ and $I = 5, 10$.
- MML estimates from the `mirt` package (Chalmers, 2012).
- There are some occasional large estimates, especially for $S = 100$ and the MML method, so the plots that follow report

Mean bias and RMSE computed with 5% trimming

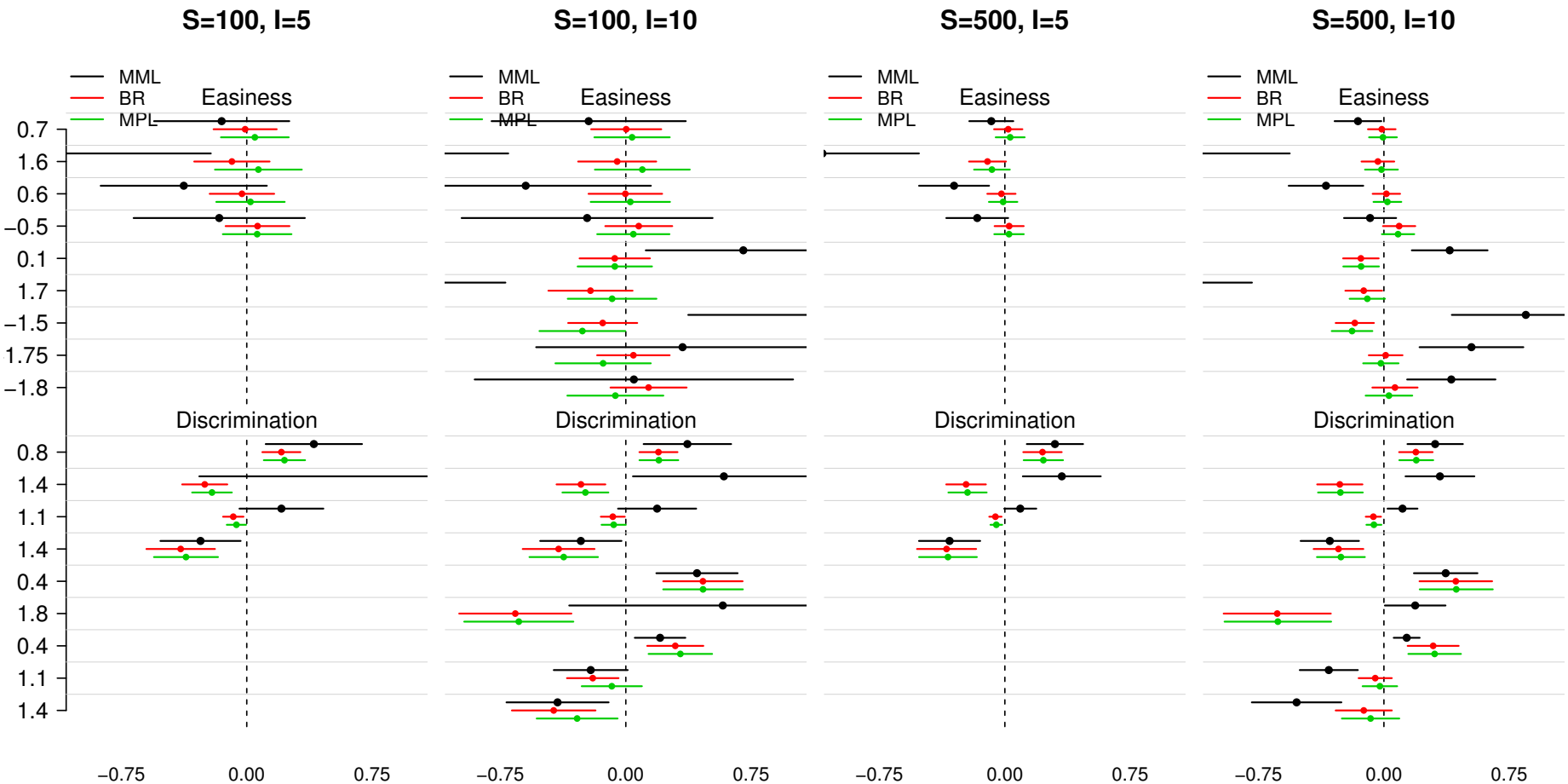
Setting 1., normal abilities



Setting 2., two-component mixture



Setting 3., zero-inflated mixture

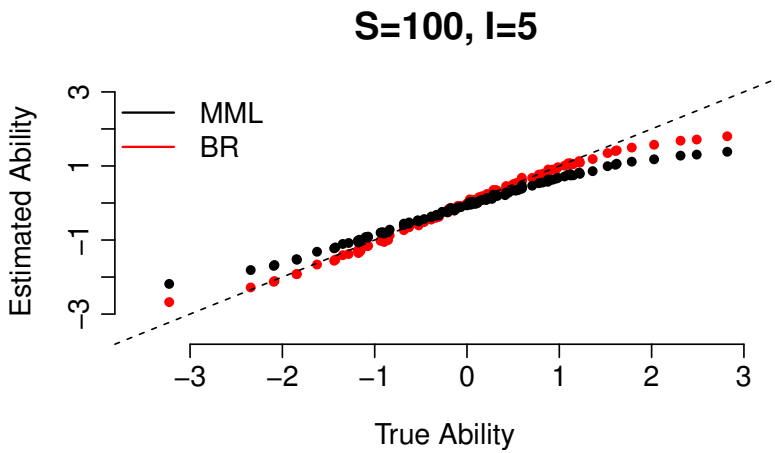


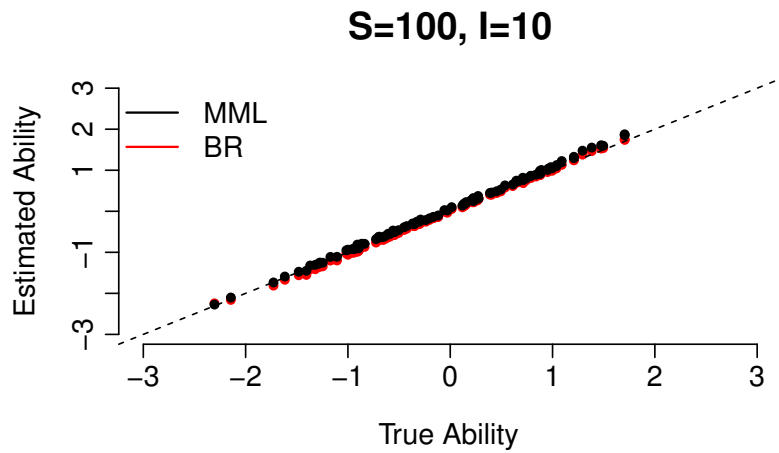
What about estimated abilities?

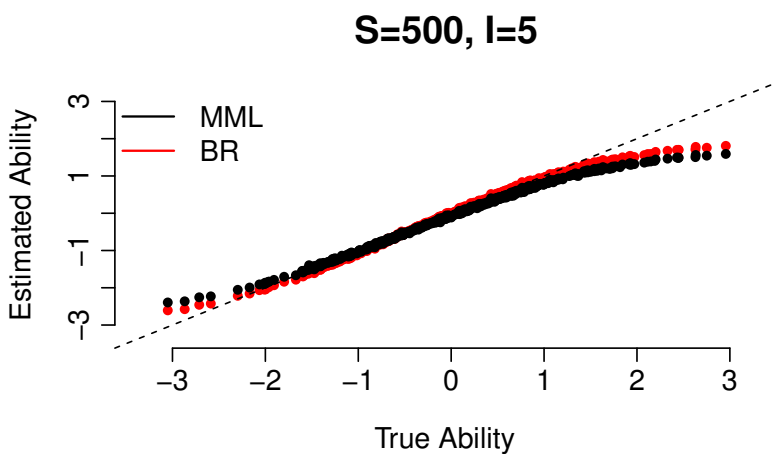
- After estimating item parameters, an estimate of abilities is usually required.
- Our proposal for this is the BR estimate, which extends the Weighted Likelihood Estimation by Warm (1989).
- Setting 1. and 2. of the simulation study provide useful results.

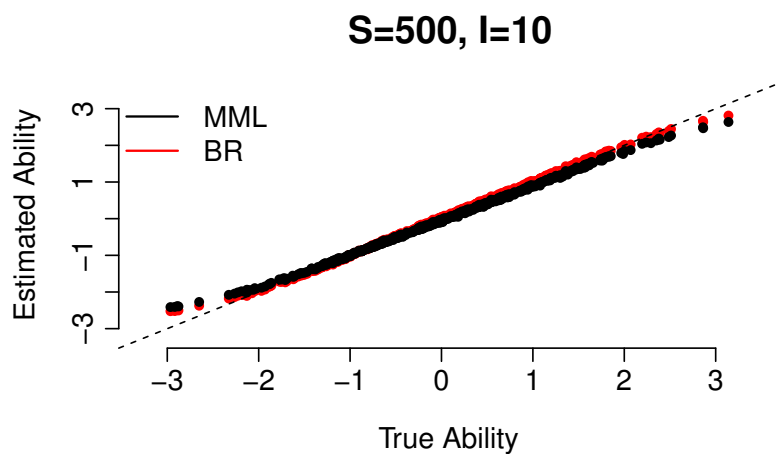
  For MML, ability estimates are computed by `mirt` with the option `"WLE"` (similar results obtained with other methods).

Setting 1., normal abilities

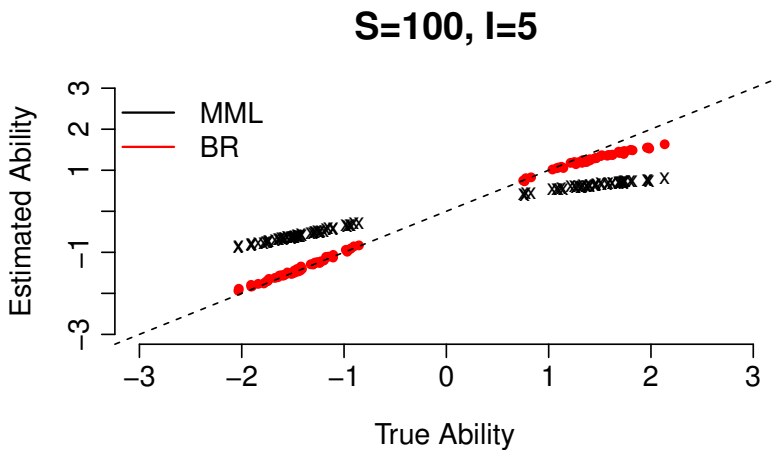


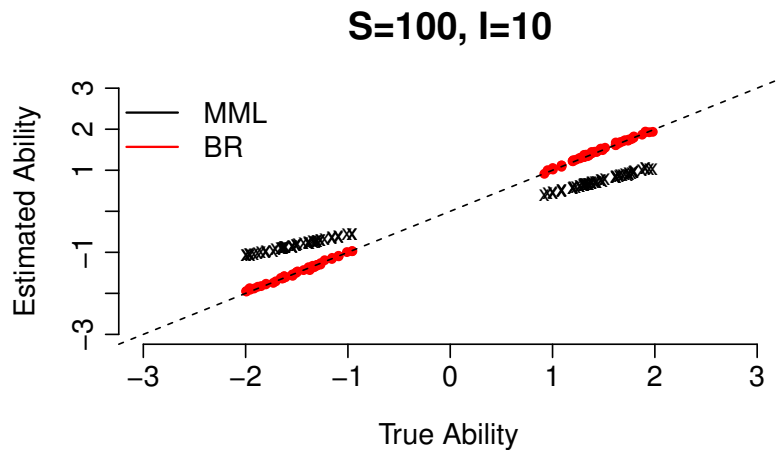


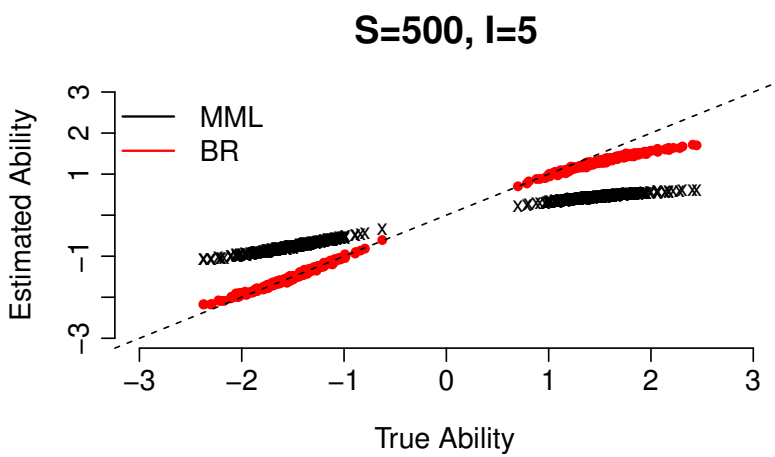


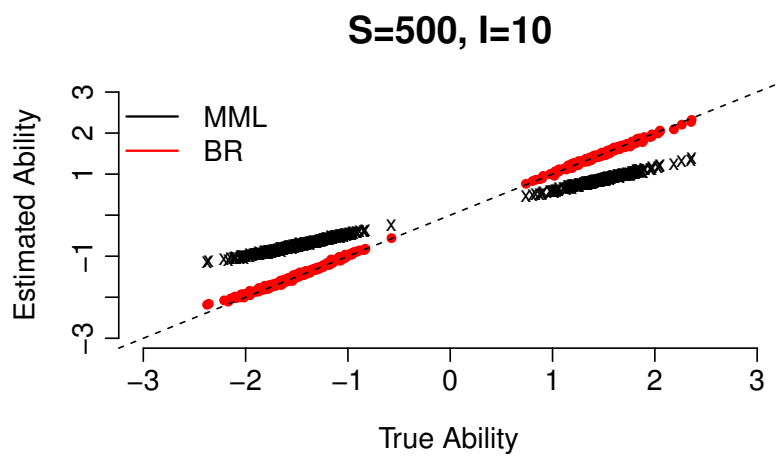


Setting 2., two-component mixture









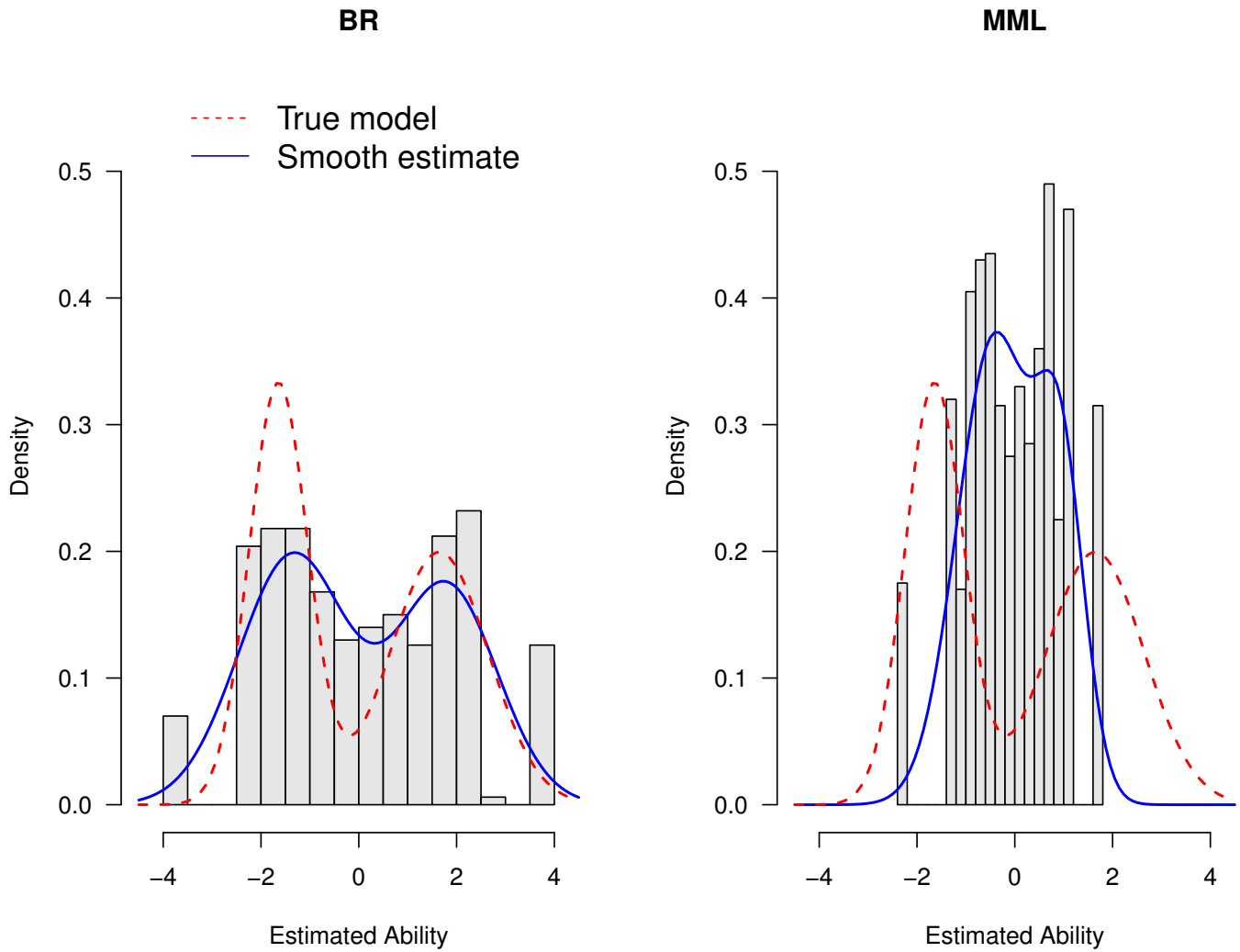
Example: Recovering abilities for simulated data

We simulated a data set with $S = 1000$ and $I = 10$, true model a normal mixture for θ_s.

We then fitted a two-component normal mixture based on the $1,000$ estimated abilities.

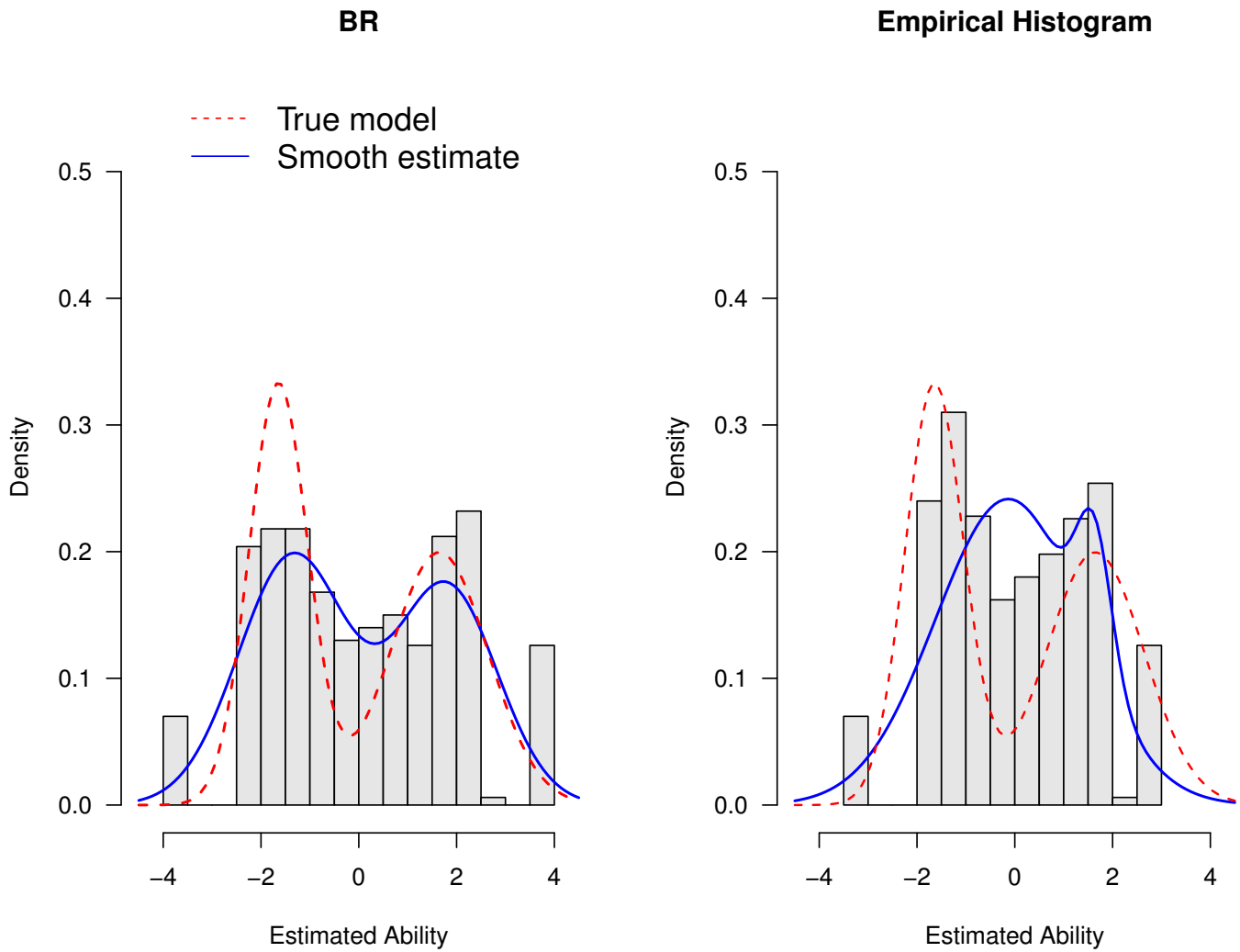
Recovering abilities for simulated data

The result is much better with BR, as the MML estimates display too much shrinkage



Recovering abilities for simulated data

BR does slightly better also when the MML estimates are based on the **Empirical Histogram** method (Knott and Tzamourani, 2007)



Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

MPL for lasso-type model selection

- The fixed-effects approach can be used also with moderate or large number of items.
- An important problem is to decide which discrimination parameters should be one, implying *neutral discrimination power* for that item.
- To this end, we may put a **penalty on the discrimination parameters** to select a model lying between the 1PL and 2PL ones. This is done by maximizing the objective function

$$l(\boldsymbol{\beta}) = \ell_M(\boldsymbol{\beta}) - \lambda \sum_{i=2}^{I} |\beta_{1i} - 1|$$

where λ is a tuning parameter, with larger values shrinking the discrimination parameters towards 1.

The lasso for 2PL models

- The lasso has been used for DIF detection using the JML in 1PL models (Tutz and Schauberger, 2015) and (approximated) 2PL models (Magis et al., 2015), the usage here appears novel.
- There are efficient ways to optimize $l(\boldsymbol{\beta})$ (Hastie et al., 2015), and the tuning parameter can be selected by BIC (Zhang et al., 2010).
- After selecting a model corresponding to the maximization of $l(\boldsymbol{\beta})$ for $\lambda = \lambda_{\mathrm{BIC}}$, it may be recommendable to refit the model to achieve some **debiasing**.

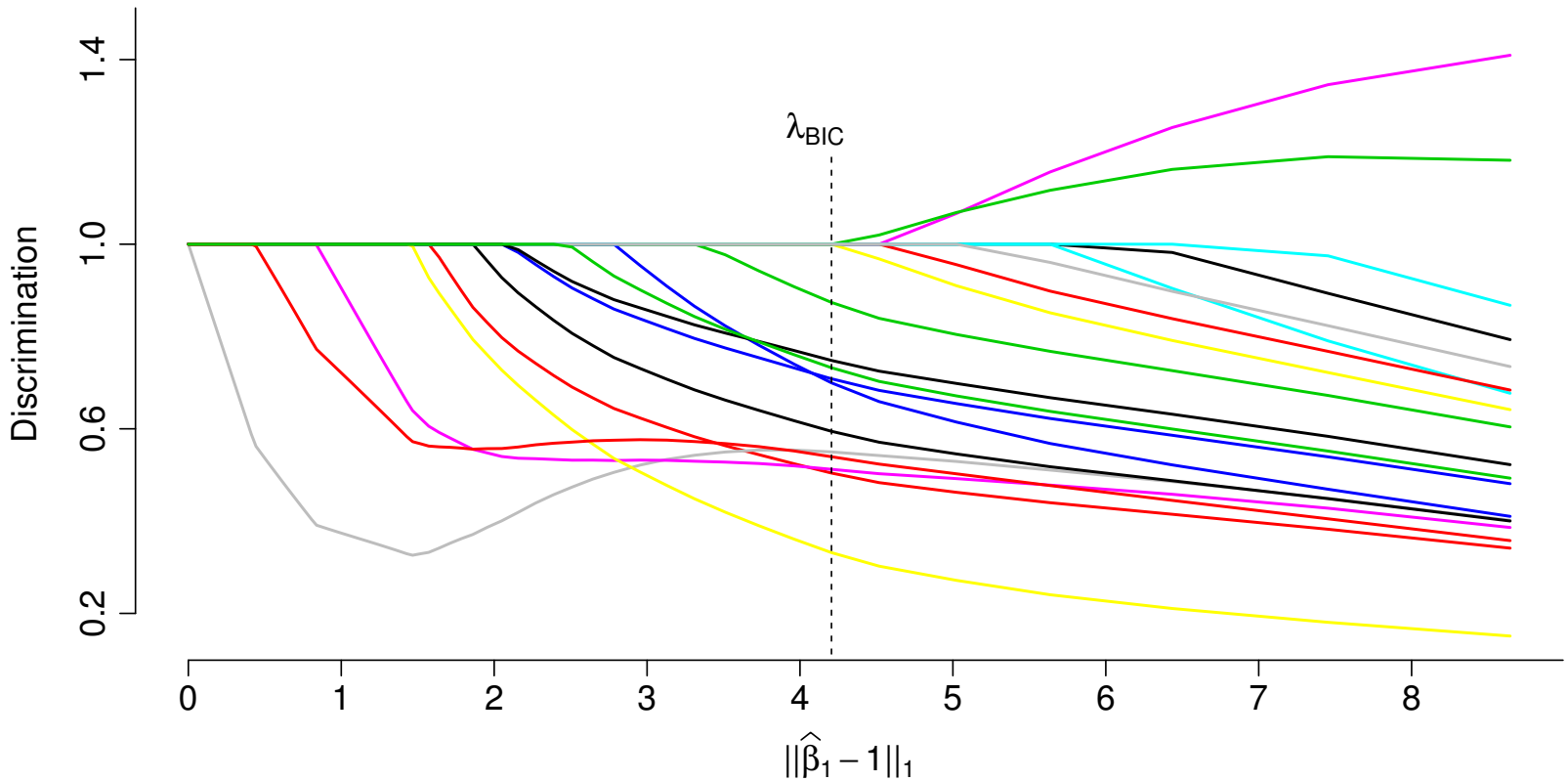
Example: `CTTdata`

Taken from `CTT` R package, data on $I = 20$ items for $S = 100$ subjects.

Some highlights:

- For the standard MML, both AIC and BIC suggests the Rasch model over the 2PL one.
- The two information criteria give discordant results using the MML with the empirical histogram.
- In either case, the P-value based on the LRT is between 0.01 and 0.05.
- Using the MPL for fixed-effects approach, the 2PL model is **strongly preferred over the Rasch model**, and the lasso allows for a finer selection.

CTTdata: discrimination path



Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Implementation in R

Two main tasks:

1. *Solving the adjusted score equations in $\boldsymbol{\omega}$ to get the BR estimates.*
2. *Maximizing the expression of MPL.*

Worth noting: **no integrals involved, totally simulation-free** !

More on the BR part

- Estimation done by **quasi Fisher-scoring with careful steplength**, along the lines of Kosmidis and Firth (2010).
- Made complex by the need to handle simultaneously $2(I-1)+S$ parameters, with the additional complication that the adjusted score equations are **not equivariant** wrt reduction of data to *response patterns* + *their frequencies*.
- Efficient pure `R` implementation, can handle up to some thousands of subjects in reasonable time.
- Speed-up maybe possible by resorting to `Rcpp` and related linear algebra packages (Eddelbuettel and Francois, 2011).

More on the MPL part

- We get rid of the inner optimization required for each evaluation at $\boldsymbol{\beta}$ by a linear approximation to the constrained estimate (Cox and Wermuth, 1990).
- Very efficient `C++` implementation obtained via `Template Model Builder (TMB)` (Kristensen et al., 2016), fully embedded within R.
- TMB returns the coded gradient and Hessian of $\ell_M(\boldsymbol{\beta})$ $\Rightarrow$ quite fast optimization.

More on computation for the lasso

- $l(\boldsymbol{\beta})$ is quickly optimized by means of a cyclic coordinate descent algorithm, which is a standard approach for ℓ_1 penalties.
- An alternative approach employs the Orthant-Wise Limited-memory Quasi-Newton (OWL-QN) optimization algorithm implemented in the R package `lbfgs` (Coppola et al., 2014).
- For both alternatives, the fast coded returned by TMB is the key.

Outline

- Background on 2PL models
- Specifics of our proposal
- Asymptotia
- How good is it? Some simulation studies
- Model selection based on the lasso
- Implementation in R
- Winding up

Winding up

- The approach presented seems an improvement over both JML (indeed!) and MML (to some extent).

  Essentially, it replaces **shrinkage coming from the assumption of normality** with **likelihood-based shrinkage**.
- Good performances for small I, robustness, lasso-based model selection for larger settings: we recommend the fixed-effects approach as **default for 2PL models**.
- Extension to other models is straightforward, with obvious candidates given by graded response models and models with DIF.

What will be made available

- A research report will be released very soon, with software on `github` to apply the method.
- A more ambitious project aims to provide a user-friendly implementation, ideally with something like a Shiny-based web application (`http://shiny.rstudio.com`).

  Some changes would be required for handling very large data, with several thousands of subjects.

References:
Item response theory

- Bock, R.D, and Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46, 443-459.
- Haberman, S.J. (1977). Maximum likelihood estimates in exponential response models. *Annals of Statistics*, 5, 815-841.
- Lord, F.M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Erlbaum, Hillsdale NJ.
- Knott, M. and Tzamourani, P. (2007). Bootstrapping the estimated latent distribution of the two-parameter latent trait model. *British Journal of Mathematical and Statistical Psychology*, 60, 175-191.
- San Martín, E., González, J. and Tuerlinckx, F. (2015). On the unidentifiability of the fixed-effects 3PL model. *Psychometrika*, 80, 450-467.
- Wall, M.M., Park, J.Y. and Moustaki, I. (2015). IRT modeling in the presence of zero-inflation with application to psychiatric disorder severity. *Applied Psychological Measurement*, 39, 583-597.
- Warm, T.A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427-450.

References:
Likelihood methods

- Arellano, M. and Bonhomme, S. (2009). Robust priors in nonlinear panel data models. *Econometrica*, 77, 489-536.
- Bartolucci, F., Bellio, R., Salvan, A. and Sartori, N. (2015). Modified profile likelihood for fixed-effects panel data models. *Econ. Rev.*, in press.
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80, 27-38.
- Heinze, G. and Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, 21, 2409-2419.
- Kosmidis, I. (2014). Improved estimation in cumulative link models. *JRSS B*, 76, 169-196.
- Kosmidis, I. and Firth, D. (2009). Bias reduction in exponential family nonlinear models. *Biometrika*, 96, 793-804.
- Neyman, J. and Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16, 1-32.
- Sartori, N. (2003). Modified profile likelihoods in models with stratum nuisance parameters. *Biometrika*, 90, 533-549.
- Severini, T.A. (1998). An approximation to the modified profile likelihood function. *Biometrika*, 85, 403-411.

References:
Lasso

- Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC Press.
- Magis, D., Tuerlinckx, F., and De Boeck, P. (2015). Detection of differential item functioning using the lasso approach. *Journal of Educational and Behavioral Statistics*, 40, 111-135.
- Tutz, G. and Schauberger, G. (2015). A penalty approach to differential item functioning in rasch models. *Psychometrika*, 80, 21-43.
- Zhang, Y., Li, R., and Tsai, C.-L. (2010). Regularization parameter selections via generalized information criterion. *JASA*, 105, 312-323.

References:
Computing

- Chalmers, R.P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6).
- Coppola, A., Stewart, B. and Okazaki, N. (2014). lbfgs: Limited-memory BFGS Optimization. R package version 1.2.1.
- Cox, D.R., and Wermuth, N. (1990). An approximation to maximum likelihood estimates in reduced models. *Biometrika*, 77, 747-761.
- Eddelbuettel, D. and Francois, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8).
- Kosmidis, I., Firth, D. (2010). A generic algorithm for reducing bias in parametric estimation. *Electronic Journal of Statistics*, 4, 1097-1112.
- Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H.J. and Bell, B. (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, in press.

Thank you for your attention !

http://ruggerobellio.weebly.com