IRT workshop

Christopher David Desjardins

19 February 2014

Overview

Review

- 1-PL and Rasch model
- IRF, IIF, TRF, and TIF
- Person and item estimation
- Ran R code
- Today: 2-PL and 3-PL models

Proposed topics

2-PL and 3-PL models (19/2)
Polytomous models (5/3)
Differential Item Function (19/3)

Multidimensional models? LCA? (2/4)
Scaling, linking & Presentations (16/4)

First half will be lecture and introduction to **mirt** package in R.

Shift in thinking

- For the 1-PL and the Rasch, the probability of getting an item correct was a function of the distance an item was located from a person.
- For the 2-PL, the probability of a correct response to an item is also a function of how well the item differentiates among people at different locations.
- Items will now be allowed to have different slopes.
- From a Rasch to a 2-PL represents a philosophical shift.
  - Rasch, construct an instrument that is that is consistent with the Rasch model.
  - 2-PL, develop a model that is consistent with our data.

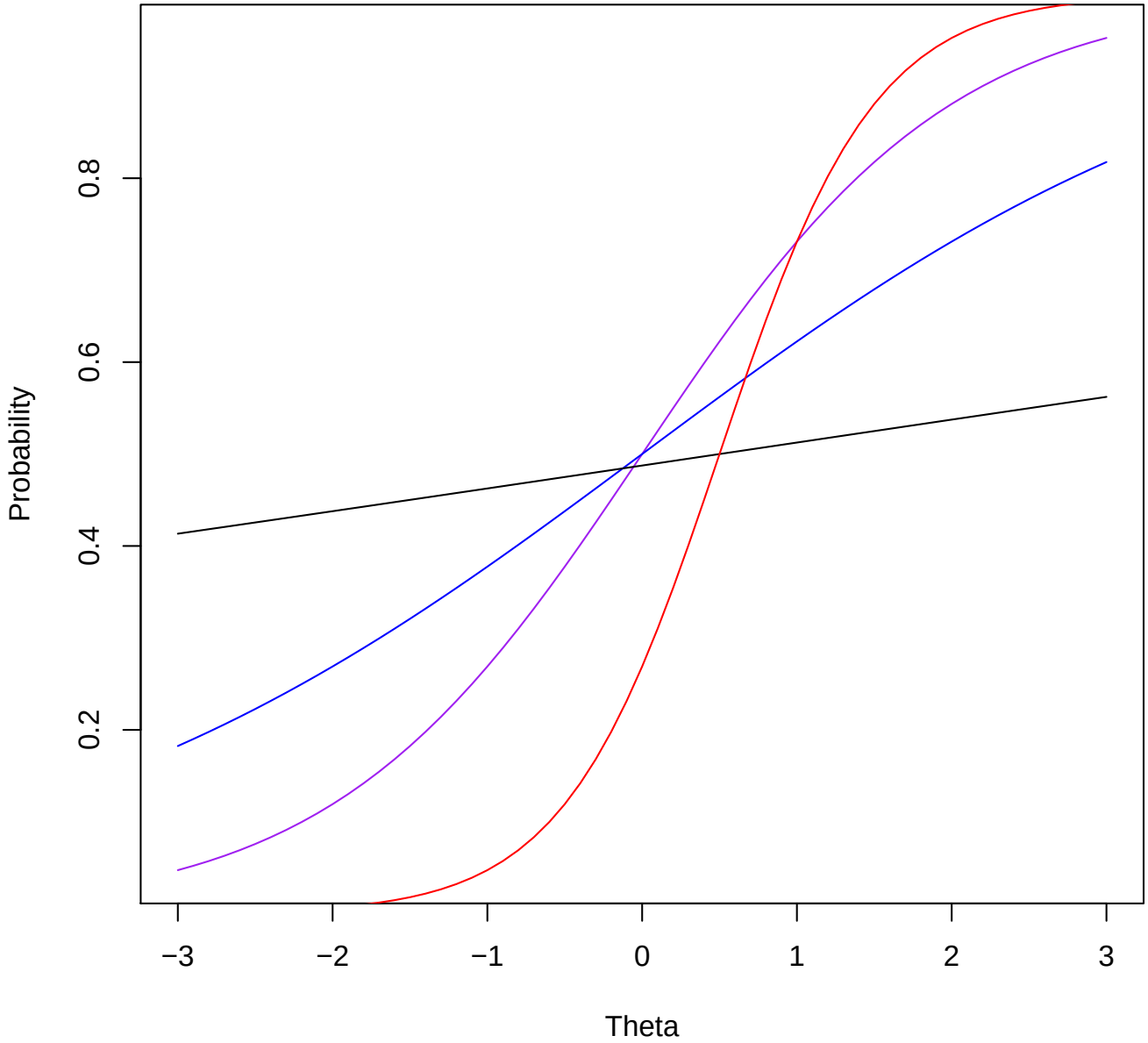
The 2-PL model

$$p(x_j = 1|\theta, a_j, d_j) = \frac{e^{a_j(\theta - d_j)}}{1 + e^{a_j(\theta - d_j)}}$$

a_j, item discrimination

d_j, item difficulty

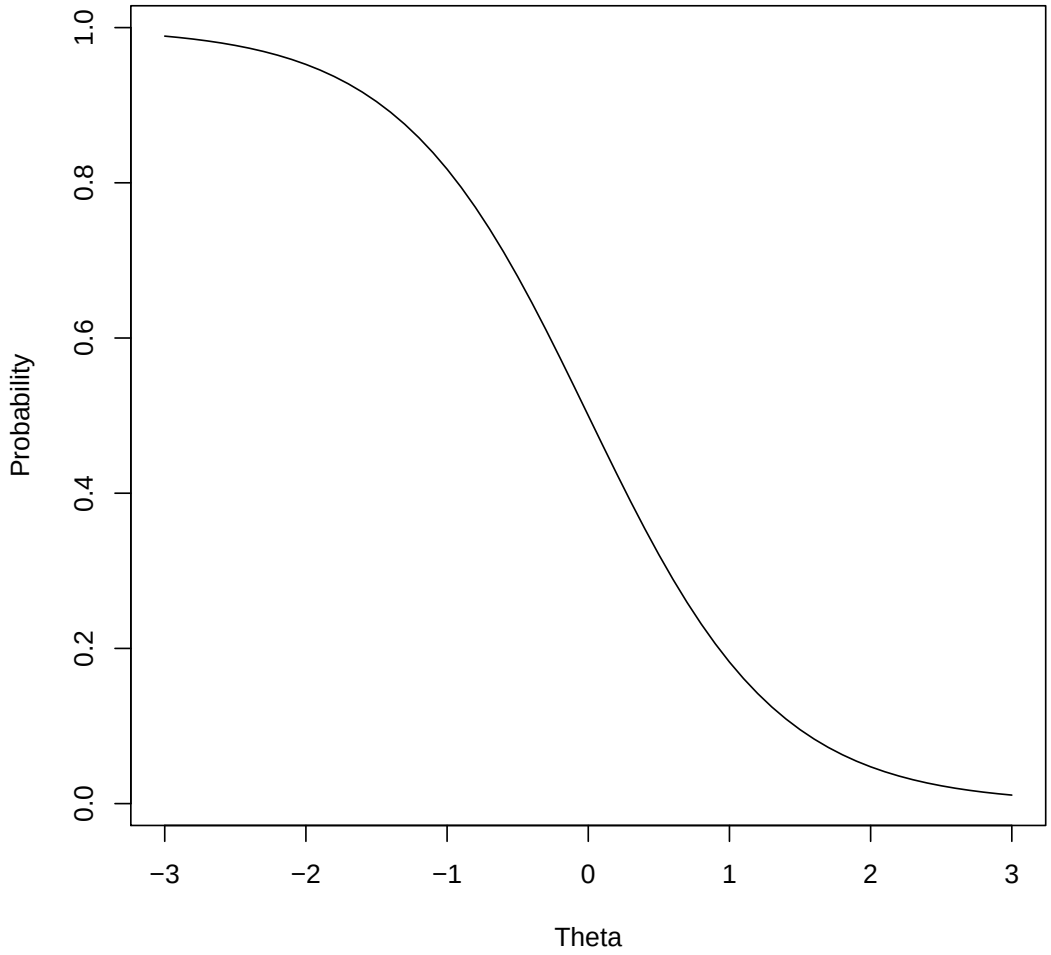
θ, person ability


Probability
0.2
0.4
0.6
0.8
−3
−2
−1
0
1
2
3
Theta

Item discrimination

- How well can an item differentiate among individuals located at different abilities
- Can range between $-\infty$ to ∞
- Generally, between 0.8 - 2.5 are good values
- Slope at d_j is $a_j/4$
- Negative discrimination item that typically need to be discarded or are inconsistent.

Negative discrimination



Information in the 2PL

- The item information function for the 2PL can be written as:
  $$I_j(\theta) = a_j^2 p_j(1 - p_j)$$
- This function is at a maximum when $p_j = .5$
  $$I_j(\theta) = a_j^2 0.25$$
- Values of a_j greater than 1 provide more information at d_j than a Rasch model
- These information functions are steeper and provide more information near d_j

How big of a sample size?

- There have been various simulation studies to attempt to answer this question.
- Factors include sample size, number of items, prior distribution, assumed distribution of θ.
- de Ayala (2009) recommends at least 500 people with 20 items, using MMLE (an estimation technique) where your prior matches the posterior distribution.
  - However, this is not a rule!
- I would suggest running a simulation study
- Can simulate data from `simdata` in the **mirt** package

Indeterminancy

- Recall that for the Rasch model

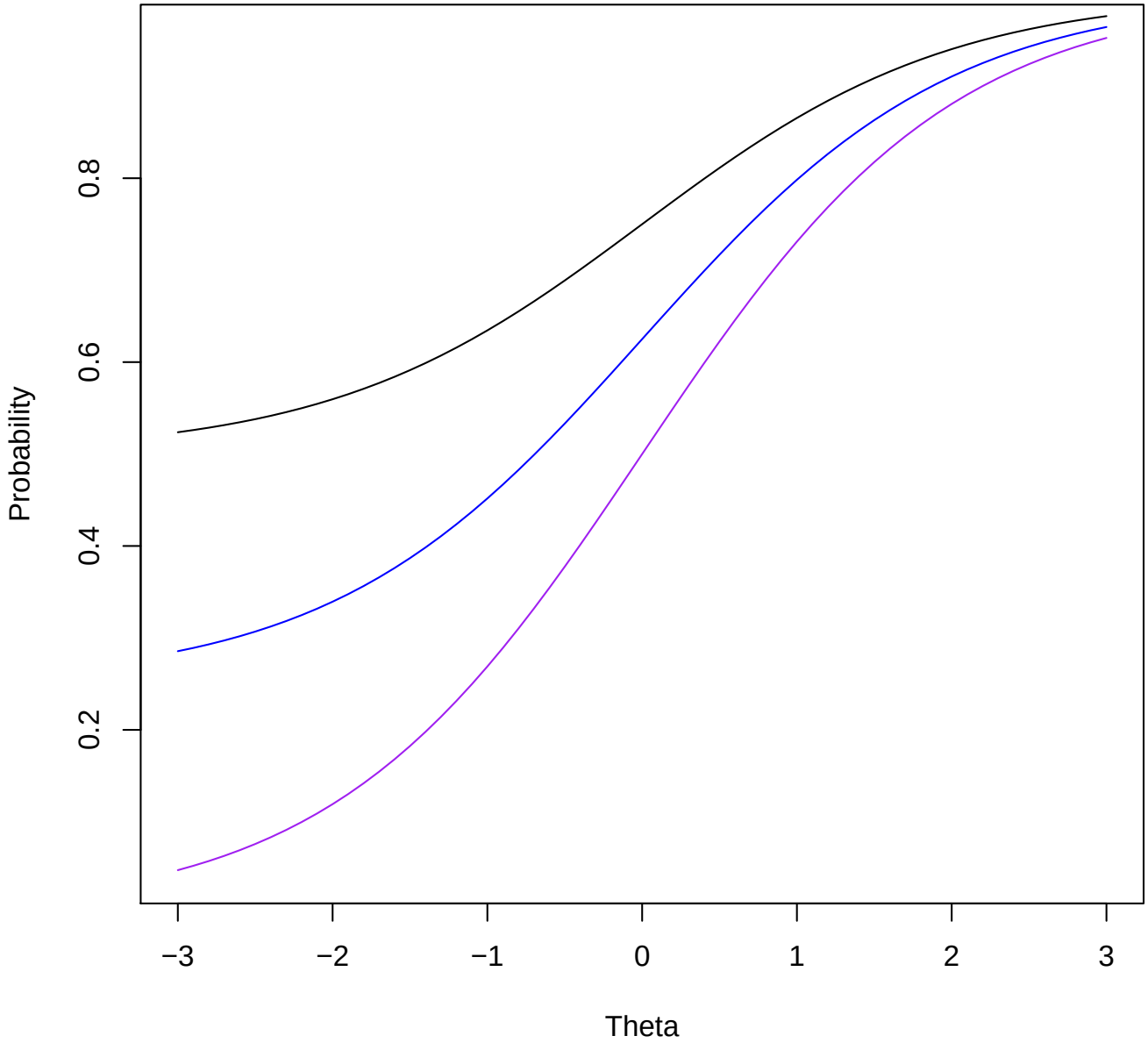
$$p(x_j = 1|\theta, d_j) = \frac{e^{(\theta - d_j)}}{1 + e^{(\theta - d_j)}}$$

- Notice that we are only concerned with the distance $(\theta - d_j)$
- But the choose of θ and d_j are arbitrary!
- Therefore, we have an infinite number of solutions, i.e. **interdeterminancy**
- To get around this:
  - Person centering, sets the mean of $\hat{\theta}s$ to 0 after each step of person location estimation
  - Item centering, sets the mean of $\hat{d}$s to 0 after each step of item estimation
- Note, we haven't discussed how we estimate these parameters in practice but it's iterative.

The 3-PL model

$$p(x_j = 1|\theta, a_j, d_j, g_j) = g_j + (1 - g_j)\frac{e^{a_j(\theta - d_j)}}{1 + e^{a_j(\theta - d_j)}}$$

- Same parameters as before, but g_j stands for the guessing parameter.
- It's really just a lower asympotate
- The point of inflection, i.e. the item location, now occurs half-way between the lower and upper asympotate


Probability
0.2
0.4
0.6
0.8
−3
−2
−1
0
1
2
3
Theta

Comments about the parameters

- The probability of getting an item correct at d_j is $1/2(1 - g_j)$
- The slope at d_j is now $0.25a_j(1 - g_j)$
- Note that g_j does not vary as a function of θ
- Non-zero g_j reduce the amount of item information

Final comments about the 3-PL

- The TIF is not guaranteed to be unimodel now
- The TCC is not necessarily going to be smooth but will still be monotonic, nondecreasing
- Can choose between the models empirically

Class Paper

- Introduction (1 - 3 pages)
  - Describe problem and what you are trying to accomplish
- Methods (3 - 5 pages)
  - Describe your data set, what model you're using, why, and describe the model
- Results (5 - 10 pages)
  - Report the results from your models considered
  - Show evidence that they are valid (i.e. do they fit)
  - Plot IRFs, IIFs, TCC, TIF
  - Report item and person estimates
- Discussion (3 - 5 pages)
  - What do your findings mean?
  - What limitations are there?
  - Where will you go next

Next time

- I will send out a paper or two
- Please continue to work with R