Bayesian Estimation of Item Response Models
Edps 590BAY

Carolyn J. Anderson

Department of Educational Psychology

©Board of Trustees, University of Illinois

Spring 2018
Overview

- A multilevel logistic regression model — end of last lecture
- 1 PL model and fit to GSS 5 items
- 2PL model and fit to GSS 5 items
1PL model

The 1 parameter logistic regression model (or Rasch model):

\[
P(Y_j = 1) = \frac{1}{1 + \exp(-(\theta_i - b_j))}
\]

or

\[
P(Y_j = 1) = \frac{1}{1 + \exp((\theta_i - b_j))}
\]

or

\[
P(Y_j = 1) = \frac{1}{1 + \exp(-(\theta_i + b_j))}
\]

where

- \( P(Y_j = 1) \) = the probability that item \( j \) is answered correctly
- \( \theta_i \) is the value of latent variable for individual \( i \)
- Exact interpretation of \( b_j \) depends a bit on the parametrization: “easiness”, “difficulty”....or I think of it as an intercept.
Item Characteristic Curve—show from file
Item Characteristic Curve

Example of 1PL Item Characteristic Curve

Value of Latent Variable

Probability \( Y=1 \)

Item
- \( b_1 = -0.80 \)
- \( b_2 = 2.30 \)
- \( b_3 = -5.27 \)
- \( b_4 = 1.24 \)
- \( b_5 = -1.15 \)
# 1PL model: 5 GSS items

JAGS model summary statistics from 80000 samples (thin = 5; chains = 4; adapt+burnin = 5000):

<table>
<thead>
<tr>
<th></th>
<th>Lower95</th>
<th>Median</th>
<th>Upper95</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>item.difficulty[1]</td>
<td>1.9819</td>
<td>2.2105</td>
<td>2.436</td>
<td>2.2117</td>
<td>0.11576</td>
</tr>
<tr>
<td>item.difficulty[2]</td>
<td>3.3345</td>
<td>3.6551</td>
<td>3.9916</td>
<td>3.6579</td>
<td>0.16803</td>
</tr>
<tr>
<td>item.difficulty[3]</td>
<td>-1.7268</td>
<td>-1.5289</td>
<td>-1.3387</td>
<td>-1.5302</td>
<td>0.099269</td>
</tr>
<tr>
<td>item.difficulty[4]</td>
<td>3.5237</td>
<td>3.874</td>
<td>4.2222</td>
<td>3.8763</td>
<td>0.17843</td>
</tr>
<tr>
<td>item.difficulty[5]</td>
<td>2.1385</td>
<td>2.3709</td>
<td>2.6087</td>
<td>2.373</td>
<td>0.11996</td>
</tr>
<tr>
<td>sd.theta</td>
<td>1.351</td>
<td>1.5298</td>
<td>1.71</td>
<td>1.5317</td>
<td>0.091315</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MCerr</th>
<th>MC%ofSD</th>
<th>SSeff</th>
<th>AC.50</th>
<th>psrf</th>
</tr>
</thead>
<tbody>
<tr>
<td>item.difficulty[1]</td>
<td>0.00067605</td>
<td>0.6</td>
<td>29319</td>
<td>0.020975</td>
<td>1.0001</td>
</tr>
<tr>
<td>item.difficulty[2]</td>
<td>0.0010107</td>
<td>0.6</td>
<td>27639</td>
<td>0.021191</td>
<td>1</td>
</tr>
<tr>
<td>item.difficulty[3]</td>
<td>0.00049271</td>
<td>0.5</td>
<td>40593</td>
<td>0.0069649</td>
<td>1</td>
</tr>
<tr>
<td>item.difficulty[4]</td>
<td>0.0010295</td>
<td>0.6</td>
<td>30039</td>
<td>0.017578</td>
<td>1.0001</td>
</tr>
<tr>
<td>item.difficulty[5]</td>
<td>0.00071393</td>
<td>0.6</td>
<td>28233</td>
<td>0.021779</td>
<td>1.0001</td>
</tr>
<tr>
<td>sd.theta</td>
<td>0.00083251</td>
<td>0.9</td>
<td>12031</td>
<td>0.058506</td>
<td>1.0001</td>
</tr>
</tbody>
</table>

Total time taken: 2.2 hours ← there must be a faster way.
5 GSS items: Item Difficulty 1

![Graphs showing item difficulty over iterations.](image-url)
5 GSS items: Item Difficulty 2
5 GSS items: Item Difficulty 3

- Iteration
- Item Difficulty
- ECDF
- Lag
- Autocorrelation of Item Difficulty

C.J. Anderson (Illinois)
5 GSS items: Item Difficulty 4
5 GSS items: Item Difficulty 5
5 GSS items: sd of theta

- Iteration
- sd.theta
- ECDF
- % of total
- Autocorrelation of sd.theta
5 GSS items: ICCs
Sub-set of Data to $N=250$

Since $N=1155$ take a long time, I took a random sub-set of 250 examinees. The code for this was:

```r
# Takes the random sample
examinees ← unique(vo5$id)[sample(1:1155, 250)]
v5.n250 ← vo5[vo5 $ id %in% examinees,]
```

Need id that goes from 1 to length(v5.n250)

```r
new.id ← rep(1:250, each=5)
v5.n250 ← cbind(new.id, v5.n250)
```

The results are similar, but much faster (depends on how many sample and whether I thinned to not). and thin=5.
The two parameter logistic model

In the 1PL model, the item characteristic curves all increase at the same rate; that is, the slope of $\theta$ was the same. This is actually a very strict assumption.

$$P(Y_j = 1) = \frac{1}{1 + \exp(-(a_j \theta_i - b_j))}$$

or $$= \frac{1}{1 + \exp(-a_j(\theta_i - b_j))}$$

or $$= \frac{1}{1 + \exp(-a_j \theta_i + b_j)}$$

where

- $P(Y_j = 1)$ = the probability that item $j$ is answered correctly
- $a_j$ is the discrimination parameter or slope of $\theta_i$
- $\theta_i$ is the value of latent variable for individual $i$
- Exact interpretation of $b_j$ depends a bit on the parametrization: “easiness”, “difficulty”....or I think of it as an intercept.
Bayesian Estimation of 2pl

The 2pl:

\[ y_j \sim \text{Bernoulli}(p_j) \]
\[ p_j = \frac{1}{1 + \exp(-(a_j \theta_i - b_j))} \]

where (in jags)

\[ \theta \sim N(0, 1) \]
\[ b_j \sim N(\mu_b, 1/\sigma_b^2) \]
\[ \log(a_j) \sim N(\mu_a, 1/\sigma_a^2) \]

And for the hyper-parameters:

\[ \mu_a \sim N(0, 0.1) \quad \text{and} \quad \mu_b \sim N(0, 0.1) \]
\[ 1/\sigma_a^2 \sim \text{dgamma}(0.1, 0.1) \quad \text{and} \quad \sigma_b^2 \sim \text{dgamma}(0.1, 0.1) \]

so \( \sigma_a \text{ and } \sigma_b \approx \sqrt{10} = 3.16 \).
GSS and the 2PL, \( N = 700 \)
5 GSS items: Item Difficulty 1
5 GSS items: Item Difficulty 2
5 GSS items: Item Difficulty 3
5 GSS items: Item Difficulty 4
5 GSS items: Item Difficulty 5
5 GSS items: Item Discrimination 1
5 GSS items: Item Discrimination 2
5 GSS items: Item Discrimination 3
5 GSS items: Item Discrimination 4
5 GSS items: Item Discrimination 5
5 GSS items: ICCs from 2pl