SAS code for “Application of Item Response Theory Models for Intensive Longitudinal Data”

Donald Hedeker, Robin J. Mermelstein, and Brian R. Flay
In T.A. Walls & J.L. Schafer (Eds.), *Models for Intensive Longitudinal Data* (pp. 84-108).

SAS PROC NLMIXED can be used to perform IRT analysis, and code is given below for analysis of the LSAT-6 data. This code assumes that the data are at the individual level. An individual identifier must be present in the data and here it is named *id*. The dependent variable is named *lsat6* and coded 1 for a correct response and 0 for an incorrect response. The item indicators are named *item1* to *item5*. For example, the data are as follows for an individual with *id* 1001 who did not get any of the five items correct (*id*, *lsat6*, *item1*, *item2*, *item3*, *item4*, *item5*):

<table>
<thead>
<tr>
<th></th>
<th>1001</th>
<th>1001</th>
<th>1001</th>
<th>1001</th>
<th>1001</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>id</em></td>
<td>1001</td>
<td>1001</td>
<td>1001</td>
<td>1001</td>
<td>1001</td>
</tr>
<tr>
<td><em>lsat6</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>item1</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><em>item2</em></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>item3</em></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>item4</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><em>item5</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Because the mixed model does not need to assume an equal number of observations per individual, individuals missing a particular item would have less than five lines of data (or have a missing value code for the missed item response). In the LSAT-6 dataset, all of the 1000 subjects had responses on the five items, so the data are complete. Below is the PROC NLMIXED code to estimate, respectively, a Rasch logistic model and a two-parameter probit model. NLMIXED is somewhat slow in running these analyses. As an alternative, one can use the freeware MIXOR program, which is available from www.uic.edu/~hedeker/mix.html.
/* Rasch logistic model in mixed regression formulation */
proc nlmixed;
parms c1=0 c2=0 c3=0 c4=0 c5=0 a=1;
    z = c1*item1 + c2*item2 + c3*item3 + c4*item4 + c5*item5 + a*theta;
if (lsat6=0) then p = 1 - (1 / (1 + exp(-z)));
else p = 1 / (1 + exp(-z));
if (p > 1e-8) then ll = log(p);
else ll = -1e20;
model lsat6 ~ general(ll);
random theta ~ normal(0,1) subject=id;
run;

/* 2 parameter probit model in mixed regression formulation */
proc nlmixed;
parms a1=1 a2=1 a3=1 a4=1 a5=1 c1=0 c2=0 c3=0 c4=0 c5=0;
bounds a1>0, a2>0, a3>0, a4>0, a5>0;
    z = (c1*item1 + c2*item2 + c3*item3 + c4*item4 + c5*item5) +
      (a1*item1 + a2*item2 + a3*item3 + a4*item4 + a5*item5) * theta;
if (lsat6=0) then p = probnorm(0-z);
else p = probnorm(z);
if (p > 1e-8) then ll = log(p);
else ll = -1e20;
model lsat6 ~ general(ll);
random theta ~ normal(0,1) subject=id;
run;