IRT Models in SAS NLMIXED

Example data: 635 older adults (age 80-100) self-reporting on 7 items assessing the Instrumental Activities of Daily Living (IADL) as follows:

1. Housework (cleaning and laundry)
2. Bedmaking
3. Cooking
4. Everyday shopping
5. Getting to places outside of walking distance
6. Handling banking and other business
7. Using the telephone

Two versions of a response format were available:

Binary → 0 = "needs help", 1 = "does not need help"
Categorical → 0 = "can't do it", 1="big problems", 2="some problems", 3="no problems"

Higher scores indicate greater function. We will look at each response format in turn.

These models were estimated using Mplus (see Class 10a and 10c examples), and the results are similar.

The first step is to convert the data from the format in which each person's responses are on a single row (as separate variables) to a "stacked" format in which each item for each person is a separate row. We then create dummy code indicators for each item that take the value of 1 if the item is on that line, or 0 otherwise.

SAS Code to read in data and run NLMIXED:

```
* Generic session options - set page and line size, no page breaks;
OPTIONS nonumber nodate nocenter pagesize=MAX linesize=120 formdlim='-';
* Kill default titles on output;
TITLE;

* Location for files to be used or saved to - CHANGE THIS;
%GLOBAL filesave; %LET filesave=F:\09_Psyc948\Examples\OCTO;
LIBNAME octo "F:\09_Psyc948\Examples\OCTO";

* Importing IADL binary item data and stacking into a single variable;
DATA diadl; SET octo.iadl;
      ARRAY adia(7) dia1-dia7;
            DO i=1 TO 7;
                  item=i; dia=adia(i); OUTPUT;
            END;
RUN;
* Importing categorical IADL item data and stacking into a single variable;
DATA ciadl; SET octo.iadl;
      ARRAY acia(7) cia1-cia7;
            DO i=1 TO 7;
                  item=i; cia=acia(i); OUTPUT;
            END;
RUN;
* Create dummy indicator variables for each binary item;
DATA diadl; SET diadl;
      ARRAY dummy (7) i1-i7;
            DO d=1 TO 7;
                  IF item=d THEN dummy(d)=1; ELSE dummy(d)=0;
            END;
      DROP i d dia1-dia7 cia1-cia7;
RUN;
* Create dummy indicator variables for each categorical item;
DATA ciadl; SET ciadl;
      ARRAY dummy (7) i1-i7;
            DO d=1 TO 7;
                  IF item=d THEN dummy(d)=1; ELSE dummy(d)=0;
            END;
      DROP i d dia1-dia7 cia1-cia7;
RUN;
```

```
*********************************************************************************;
/* PROC NLMIXED syntax commands used below:

DATA gives file used for analysis -- udatafile is named above
ITDETAILS gives iteration history for examining convergence
METHOD=GAUSS specifies Gauss-Hermite quadrature (default)
TECHNIQUE=QUANEW specifies Quasi-Newton optimization (default)
TECHNIQUE=NEWRAP specifies Newton-Raphson optimization (also used)
NOAD specifies non-adaptive quadrature for theta (use if have problems estimating)
QPOINTS=15 specifies 15 quadrature points for theta
      The more points, the longer but more precise the estimation
NOITPRINT suppresses iteration history print

PARMS must list ALL parameters to be estimated with start values
MODEL defines equation, where [outcome] is distributed [parameter]
RANDOM lists random effects with distribution, here normal [mean, variance]
              SUBJECT defines level-2 nesting (here, persons)
              OUT lists dataset random effects (thetas) saved to
ODS OUTPUT ParameterEstimates saves all estimates to named dataset
*/
*********************************************************************************;

TITLE "1PL Binary Rasch Model with Theta Variance=1";
PROC NLMIXED DATA=diadl METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* Add better start values if you want to reduce estimation time or difficulty;
      PARMS b01-b07=0 a=1;
* A placeholder for item difficulty is created by multiplying each item difficulty
  by a dummy code for that item, such that we get 24 separate estimates;
      b = b01*i1 + b02*i2 + b03*i3 + b04*i4 + b05*i5 + b06*i6 + b07*i7;
* The model equation is defined here -- same a for all items;
      p = exp((1.7*a)*(theta-b)) / (1+exp((1.7*a)*(theta-b)));
* The response variable is defined here;
      MODEL dia ~ binary(p);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_1PLbinary;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_1PLBinary;
RUN;
```

1 minute later…

Binary 1-PL (Rasch) Model Syntax and Truncated Output:

```
1PL Binary Rasch Model with Theta Variance=1
The NLMIXED Procedure

                         Specifications
Data Set                                    WORK.DIADL
Dependent Variable                          dia
Distribution for Dependent Variable         Binary
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature

               Dimensions
Observations Used                    4367
Observations Not Used                 106
Total Observations                   4473
Subjects                              635
Max Obs Per Subject                     7
Parameters                              8
Quadrature Points                      15


          NOTE: GCONV convergence criterion satisfied.


             Fit Statistics
-2 Log Likelihood                  2927.9
AIC (smaller is better)            2943.9
AICC (smaller is better)           2944.0
BIC (smaller is better)            2979.6
```

This -2LL value is very close to what Mplus provides in LL form (SAS multiples LL by -2 for you).

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
b01	-0.3761	0.05463	634	-6.89	<.0001	0.05	-0.4834	-0.2689	-0.00006
b02	-1.0562	0.06712	634	-15.74	<.0001	0.05	-1.1880	-0.9244	-0.00023
b03	-0.7979	0.06045	634	-13.20	<.0001	0.05	-0.9166	-0.6792	-0.00021
b04	-0.4369	0.05502	634	-7.94	<.0001	0.05	-0.5450	-0.3289	-0.00002
b05	-0.4308	0.05485	634	-7.85	<.0001	0.05	-0.5385	-0.3231	0.000111
b06	-0.6802	0.05775	634	-11.78	<.0001	0.05	-0.7937	-0.5668	0.000107
b07	-1.6939	0.09239	634	-18.33	<.0001	0.05	-1.8753	-1.5124	0.000117
a	2.6143	0.1637	634	15.97	<.0001	0.05	2.2929	2.9358	0.000152

The common discrimination (slope) parameter is “a”, which is 2.6143. This matches closely with the common discrimination parameter of 2.578 estimated by Mplus for the same model.

The item difficulties (location) parameters are the b’s, which gives the latent trait location at which a response of 1 (instead of 0) has a probability of 50%. These also match very closely with the b’s that were obtained from Mplus (b’s converted from the thresholds, b=threshold/loading).

Mplus version:

```
TITLE:  1PL Binary Model under ML
DATA:   FILE IS ADL.dat;
VARIABLE:   NAMES ARE case dia1-dia7 cia1-cia7;
            USEVARIABLES ARE dia1-dia7;
            CATEGORICAL ARE dia1-dia7;
            MISSING ARE .;

ANALYSIS:   ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all constrained equal in 1PL
    IADL dia1-dia7* (1);
! Item thresholds all free
    [dia1$1-dia7$1];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

TESTS OF MODEL FIT

Loglikelihood
          H0 Value                      -1464.457

Information Criteria
          Number of Free Parameters             8
          Akaike (AIC)                   2944.915
          Bayesian (BIC)                 2980.544
          Sample-Size Adjusted BIC       2955.144
            (n* = (n + 2) / 24)
```

```
IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC
WHERE THE LOGIT IS 1.7*DISCRIMINATION*(THETA - DIFFICULTY)

 Item Discriminations

 IADL     BY
    DIA1               2.578      0.160     16.094      0.000
    DIA2               2.578      0.160     16.094      0.000
    DIA3               2.578      0.160     16.094      0.000
    DIA4               2.578      0.160     16.094      0.000
    DIA5               2.578      0.160     16.094      0.000
    DIA6               2.578      0.160     16.094      0.000
    DIA7               2.578      0.160     16.094      0.000

 Item Difficulties
    DIA1$1            -0.383      0.053     -7.182      0.000
    DIA2$1            -1.069      0.065    -16.331      0.000
    DIA3$1            -0.809      0.059    -13.761      0.000
    DIA4$1            -0.444      0.054     -8.302      0.000
    DIA5$1            -0.438      0.053     -8.209      0.000
    DIA6$1            -0.690      0.056    -12.303      0.000
    DIA7$1            -1.711      0.091    -18.758      0.000
```

Binary 2-PL (Unequal Slopes) Model Syntax and Truncated Output:

```
TITLE "2PL Binary Model with Theta Variance=1";
PROC NLMIXED DATA=diadl METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* This time, this includes the discrimination parameters per item;
* Add better start values if you want to reduce estimation time or difficulty;
      PARMS b01-b07=0 a01-a07=1;
* A placeholder for item difficulty is created by multiplying each item difficulty
  by a dummy code for that item, such that we get 24 separate estimates;
* We do the same to get an item discrimination separately per item;
      b = b01*i1 + b02*i2 + b03*i3 + b04*i4 + b05*i5 + b06*i6 + b07*i7;
      a = a01*i1 + a02*i2 + a03*i3 + a04*i4 + a05*i5 + a06*i6 + a07*i7;
* The model equation is defined here -- now different a across items;
      p = exp((1.7*a)*(theta-b)) / (1+exp((1.7*a)*(theta-b)));
* The response variable is defined here;
      MODEL dia ~ binary(p);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_2PLbinary;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_2PLBinary;
RUN;
```

8 minutes later…

```
2PL Binary Model with Theta Variance=1
The NLMIXED Procedure

                         Specifications
Data Set                                    WORK.DIADL
Dependent Variable                          dia
Distribution for Dependent Variable         Binary
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature
               Dimensions
Observations Used                    4367
Observations Not Used                 106
Total Observations                   4473
Subjects                              635
Max Obs Per Subject                     7
Parameters                             14
Quadrature Points                      15

          NOTE: GCONV convergence criterion satisfied.

             Fit Statistics
-2 Log Likelihood                  2907.4
AIC (smaller is better)            2935.4
AICC (smaller is better)           2935.5
BIC (smaller is better)            2997.7
```

Fit Statistics from 1 PL Binary Model

-2 Log Likelihood	2927.9
AIC (smaller is better)	2943.9
AICC (smaller is better)	2944.0
BIC (smaller is better)	2979.6

2927.9 – 2907.4 = 20.5 on df=6 is p =.0023, so the 2PL is an improvement over the 1PL (we need different slopes across items)

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
b01	-0.3733	0.05513	634	-6.77	<.0001	0.05	-0.4816	-0.2651	0.000179
b02	-1.0307	0.06815	634	-15.12	<.0001	0.05	-1.1646	-0.8969	-0.00174
b03	-0.7899	0.06231	634	-12.68	<.0001	0.05	-0.9122	-0.6675	0.001153
b04	-0.4067	0.05225	634	-7.78	<.0001	0.05	-0.5093	-0.3041	-0.00339
b05	-0.4261	0.05557	634	-7.67	<.0001	0.05	-0.5353	-0.3170	0.003258
b06	-0.6978	0.06267	634	-11.13	<.0001	0.05	-0.8209	-0.5747	0.001762
b07	-1.7959	0.1259	634	-14.26	<.0001	0.05	-2.0432	-1.5486	-0.00078
a01	2.6016	0.3388	634	7.68	<.0001	0.05	1.9363	3.2670	-0.00385
a02	2.9554	0.4797	634	6.16	<.0001	0.05	2.0134	3.8974	-0.00044
a03	2.5793	0.3401	634	7.58	<.0001	0.05	1.9114	3.2471	0.006694
a04	4.7349	1.1221	634	4.22	<.0001	0.05	2.5313	6.9384	-0.00013
a05	2.5853	0.3263	634	7.92	<.0001	0.05	1.9445	3.2261	0.001994
a06	2.0700	0.2423	634	8.54	<.0001	0.05	1.5941	2.5459	0.001393
a07	1.9530	0.3574	634	5.46	<.0001	0.05	1.2511	2.6548	-0.00283

The item discriminations (slope) parameters are the a's, which gives one slope per item. These match closely numerically to the discriminations estimated by Mplus.

The item difficulties (location) parameters are the b's, which gives the latent trait location at which a response of 1 (instead of 0) has a probability of 50%. These match fairly closely numerically to the b's estimated by Mplus.

Mplus version:

```
TITLE:  2PL Binary Model under ML
DATA:   FILE IS ADL.dat;
VARIABLE:   NAMES ARE case dia1-dia7 cia1-cia7;
            USEVARIABLES ARE dia1-dia7;
            CATEGORICAL ARE dia1-dia7;
            MISSING ARE .;
ANALYSIS:   ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all free in 2PL
    IADL dia1-dia7*;
! Item thresholds all free
    [dia1$1-dia7$1];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

TESTS OF MODEL FIT
Loglikelihood
          H0 Value                       -1454.634
Information Criteria
          Number of Free Parameters             14
          Akaike (AIC)                    2937.268
          Bayesian (BIC)                  2999.619
          Sample-Size Adjusted BIC        2955.170
            (n* = (n + 2) / 24)
```

```
IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC
WHERE THE LOGIT IS 1.7*DISCRIMINATION*(THETA - DIFFICULTY)

 Item Discriminations

 IADL     BY
    DIA1                 2.546      0.330      7.725      0.000
    DIA2                 2.928      0.475      6.159      0.000
    DIA3                 2.543      0.335      7.579      0.000
    DIA4                 4.418      0.998      4.429      0.000
    DIA5                 2.499      0.310      8.062      0.000
    DIA6                 2.030      0.236      8.600      0.000
    DIA7                 1.931      0.353      5.467      0.000

 Item Difficulties
    DIA1$1              -0.376      0.052     -7.298      0.000
    DIA2$1              -1.045      0.065    -15.978      0.000
    DIA3$1              -0.801      0.059    -13.562      0.000
    DIA4$1              -0.415      0.047     -8.849      0.000
    DIA5$1              -0.432      0.052     -8.296      0.000
    DIA6$1              -0.708      0.060    -11.889      0.000
    DIA7$1              -1.816      0.126    -14.454      0.000
```

Polytomous (4-category) Constrained Graded Response Model (Equal Slope across Items)
Model Syntax and Truncated Output:

```
TITLE "1PL (constrained) Graded Response Model with Theta variance=1";
PROC NLMIXED DATA=ciadl METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* 3 difficulties per item, same discrimination across items;
* Add better start values if you want to reduce estimation time or difficulty;
      PARMS b101-b107=-3 b201-b207=-2 b301-b307=-1 a=2;
      b1 = b101*i1 + b102*i2 + b103*i3 + b104*i4 + b105*i5 + b106*i6 + b107*i7;
      b2 = b201*i1 + b202*i2 + b203*i3 + b204*i4 + b205*i5 + b206*i6 + b207*i7;
      b3 = b301*i1 + b302*i2 + b303*i3 + b304*i4 + b305*i5 + b306*i6 + b307*i7;
* The cumulative model is now a series of equations conditional on response;
* eta1 refers to 0 vs. 123;
* eta2 refers to 01 vs. 23;
* eta3 refers to 012 vs, 3;
      eta1 = exp((1.7*a)*(theta-b1)) / (1+exp((1.7*a)*(theta-b1)));
      eta2 = exp((1.7*a)*(theta-b2)) / (1+exp((1.7*a)*(theta-b2)));
      eta3 = exp((1.7*a)*(theta-b3)) / (1+exp((1.7*a)*(theta-b3)));
* Calculate probabilities per category via subtraction of cumulative model;
           IF cia=0 THEN p =    1-eta1;
      ELSE IF cia=1 THEN p = eta1-eta2;
      ELSE IF cia=2 THEN p = eta2-eta3;
      ELSE IF cia=3 THEN p = eta3;
* The response variable is defined here;
* Because NLMIXED doesn't have a standard distribution for polytomous data,
  we specify a general one as log(p);
      ll = log(p);
      MODEL cia ~ general(ll);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_1PLcategorical;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_1PLcategorical;
RUN;
```

7 minutes later…

```
1PL (constrained) Graded Response Model with Theta variance=1
The NLMIXED Procedure

                         Specifications
Data Set                                    WORK.CIADL
Dependent Variable                          cia
Distribution for Dependent Variable         General
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature

               Dimensions

Observations Used                   4230
Observations Not Used                243
Total Observations                  4473
Subjects                             634
Max Obs Per Subject                    7
Parameters                            22
Quadrature Points                     15

NOTE: GCONV convergence criterion satisfied.

             Fit Statistics

-2 Log Likelihood                 5182.4
AIC (smaller is better)           5226.4
AICC (smaller is better)          5226.6
BIC (smaller is better)           5324.3
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
b101	-1.5218	0.08581	633	-17.73	<.0001	0.05	-1.6903	-1.3533	0.010086
b102	-1.6077	0.08986	633	-17.89	<.0001	0.05	-1.7841	-1.4312	-0.0075
b103	-1.5041	0.08610	633	-17.47	<.0001	0.05	-1.6731	-1.3350	0.004928
b104	-1.3889	0.08101	633	-17.14	<.0001	0.05	-1.5480	-1.2298	-0.00468
b105	-1.7737	0.09650	633	-18.38	<.0001	0.05	-1.9632	-1.5842	-0.00547
b106	-1.7379	0.09717	633	-17.88	<.0001	0.05	-1.9287	-1.5470	-0.00678
b107	-2.3355	0.1508	633	-15.48	<.0001	0.05	-2.6317	-2.0393	0.00014
b201	-1.0270	0.06694	633	-15.34	<.0001	0.05	-1.1585	-0.8956	-0.01087
b202	-1.2847	0.07571	633	-16.97	<.0001	0.05	-1.4334	-1.1360	0.007809
b203	-1.1656	0.07207	633	-16.17	<.0001	0.05	-1.3071	-1.0241	0.008566
b204	-0.8759	0.06325	633	-13.85	<.0001	0.05	-1.0001	-0.7517	0.004199
b205	-0.8440	0.06146	633	-13.73	<.0001	0.05	-0.9647	-0.7233	-0.00238
b206	-1.1431	0.07088	633	-16.13	<.0001	0.05	-1.2823	-1.0039	0.009929
b207	-1.8182	0.1012	633	-17.96	<.0001	0.05	-2.0170	-1.6195	-0.00638
b301	-0.1734	0.05471	633	-3.17	0.0016	0.05	-0.2809	-0.06601	0.014153
b302	-0.7613	0.05997	633	-12.69	<.0001	0.05	-0.8791	-0.6435	-0.00275
b303	-0.5765	0.05745	633	-10.03	<.0001	0.05	-0.6893	-0.4637	0.003288
b304	-0.2442	0.05444	633	-4.49	<.0001	0.05	-0.3511	-0.1373	-0.0008
b305	-0.1806	0.05416	633	-3.33	0.0009	0.05	-0.2869	-0.07421	-0.00007
b306	-0.6691	0.05843	633	-11.45	<.0001	0.05	-0.7838	-0.5543	-0.00521
b307	-1.3076	0.07517	633	-17.40	<.0001	0.05	-1.4552	-1.1600	0.001121
a	2.3274	0.1138	633	20.45	<.0001	0.05	2.1038	2.5509	0.004489

The common item discrimination parameters is a= 2.3274. The item difficulties (location) parameters are the b's, interpreted as the difficulty for each submodel. B1 is the item difficulty (where probability is 50%) of moving from a response of 0 to 123, B2 is 01 vs 23, and B3 is 012 vs 3.

Mplus does not provide IRT parameters from categorical response models, but they can be calculated from the item factor model loadings and thresholds it does provide. The b's match pretty well, and a would have been 2.29 from Mplus.

Polytomous (4-category) Typical Graded Response Model (Unequal Slopes across Items) Model Syntax and Truncated Output:

```
TITLE "2PL (Typical) Graded Response Model with Theta variance=1";
PROC NLMIXED DATA=ciadl ITDETAILS METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* 3 difficulties per item, 1 discrimination per item now;
* Add better start values if you want to reduce estimation time or difficulty;
      PARMS b101-b107=-3 b201-b207=-2 b301-b307=-1 a01-a07=2;
      b1 = b101*i1 + b102*i2 + b103*i3 + b104*i4 + b105*i5 + b106*i6 + b107*i7;
      b2 = b201*i1 + b202*i2 + b203*i3 + b204*i4 + b205*i5 + b206*i6 + b207*i7;
      b3 = b301*i1 + b302*i2 + b303*i3 + b304*i4 + b305*i5 + b306*i6 + b307*i7;
      a =  a01*i1 + a02*i2 + a03*i3 + a04*i4 + a05*i5 + a06*i6 + a07*i7;
* The cumulative model is now a series of equations conditional on response;
* eta1 refers to 0 vs. 123;
* eta2 refers to 01 vs. 23;
* eta3 refers to 012 vs, 3;
      eta1 = exp((1.7*a)*(theta-b1)) / (1+exp((1.7*a)*(theta-b1)));
      eta2 = exp((1.7*a)*(theta-b2)) / (1+exp((1.7*a)*(theta-b2)));
      eta3 = exp((1.7*a)*(theta-b3)) / (1+exp((1.7*a)*(theta-b3)));
* Calculate probabilities per category via subtraction of cumulative model;
           IF cia=0 THEN p =    1-eta1;
      ELSE IF cia=1 THEN p = eta1-eta2;
      ELSE IF cia=2 THEN p = eta2-eta3;
      ELSE IF cia=3 THEN p = eta3;
* The response variable is defined here;
* Because NLMIXED doesn't have a standard distribution for polytomous data,
  we specify a general one as log(p);
      ll = log(p);
      MODEL cia ~ general(ll);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_2PLcategorical;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_2PLcategorical;
RUN;
```

32 minutes later…

```
2PL (Typical) Graded Response Model with Theta variance=1
The NLMIXED Procedure

                         Specifications
Data Set                                    WORK.CIADL
Dependent Variable                          cia
Distribution for Dependent Variable         General
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature

               Dimensions
Observations Used                    4230
Observations Not Used                 243
Total Observations                   4473
Subjects                              634
Max Obs Per Subject                     7
Parameters                             28
Quadrature Points                      15

NOTE: GCONV convergence criterion satisfied.

             Fit Statistics

-2 Log Likelihood                  5045.4
AIC (smaller is better)            5101.4
AICC (smaller is better)           5101.8
BIC (smaller is better)            5226.1
```

```
Fit Statistics for 1PL version of GRM
-2 Log Likelihood                  5182.4
AIC (smaller is better)            5226.4
AICC (smaller is better)           5226.6
BIC (smaller is better)            5324.3
```

5182.4 – 5045.4 = 137.0 on df=6 is p <.0001, so the full GRM is an improvement over the constrained GRM (we need different slopes across items).

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
b101	-1.4127	0.08086	633	-17.47	<.0001	0.05	-1.5715	-1.2539	0.002948
b102	-1.5483	0.08953	633	-17.29	<.0001	0.05	-1.7241	-1.3725	0.004991
b103	-1.4645	0.08777	633	-16.69	<.0001	0.05	-1.6369	-1.2921	-0.00216
b104	-1.2883	0.07796	633	-16.53	<.0001	0.05	-1.4414	-1.1352	-0.01104
b105	-1.8322	0.1057	633	-17.33	<.0001	0.05	-2.0398	-1.6246	-0.0032
b106	-1.8937	0.1209	633	-15.66	<.0001	0.05	-2.1311	-1.6563	-0.00475
b107	-3.2529	0.3204	633	-10.15	<.0001	0.05	-3.8821	-2.6236	0.001156
b201	-0.9244	0.06116	633	-15.11	<.0001	0.05	-1.0445	-0.8043	0.005896
b202	-1.1959	0.07378	633	-16.21	<.0001	0.05	-1.3408	-1.0511	-0.00276
b203	-1.1074	0.07192	633	-15.40	<.0001	0.05	-1.2486	-0.9662	-0.00349
b204	-0.7943	0.05969	633	-13.31	<.0001	0.05	-0.9115	-0.6771	0.003012
b205	-0.8375	0.06492	633	-12.90	<.0001	0.05	-0.9650	-0.7101	0.006577
b206	-1.2190	0.08445	633	-14.43	<.0001	0.05	-1.3848	-1.0531	0.013277
b207	-2.4576	0.2159	633	-11.39	<.0001	0.05	-2.8815	-2.0337	0.003548
b301	-0.1611	0.05071	633	-3.18	0.0016	0.05	-0.2607	-0.06154	0.007689
b302	-0.7007	0.05693	633	-12.31	<.0001	0.05	-0.8125	-0.5889	-0.00427
b303	-0.5394	0.05563	633	-9.70	<.0001	0.05	-0.6487	-0.4302	-0.01173
b304	-0.2314	0.05153	633	-4.49	<.0001	0.05	-0.3326	-0.1303	0.009098
b305	-0.1746	0.05478	633	-3.19	0.0015	0.05	-0.2822	-0.06705	-0.01052
b306	-0.6878	0.06484	633	-10.61	<.0001	0.05	-0.8152	-0.5605	-0.02726
b307	-1.6428	0.1367	633	-12.02	<.0001	0.05	-1.9112	-1.3744	-0.0004
a01	4.0914	0.5038	633	8.12	<.0001	0.05	3.1022	5.0807	0.006923
a02	3.0578	0.3250	633	9.41	<.0001	0.05	2.4196	3.6961	-0.01184
a03	2.7159	0.2668	633	10.18	<.0001	0.05	2.1919	3.2399	-0.00774
a04	3.3736	0.3604	633	9.36	<.0001	0.05	2.6659	4.0814	0.011136
a05	2.0930	0.1754	633	11.93	<.0001	0.05	1.7485	2.4375	0.002814
a06	1.7043	0.1532	633	11.12	<.0001	0.05	1.4034	2.0052	-0.00164
a07	1.0448	0.1226	633	8.52	<.0001	0.05	0.8040	1.2855	-0.00986

Each item now has a separate “a” discrimination parameter. The item difficulties (location) parameters are the b’s, interpreted as the difficulty for each submodel. B1 is the item difficulty (where probability is 50%) of moving from a response of 0 to 123, B2 is 01 vs 23, and B3 is 012 vs 3.

Mplus does not provide IRT parameters from categorical response models, but they can be calculated from the item factor model loadings and thresholds it does provide. The a’s and b’s match pretty well the converted estimates from Mplus.

Polytomous (4-category) Partial Credit Model (Equal Slopes) Model Syntax and Truncated Output:

```
TITLE "1PL (typical) Partial Credit Model Model with Theta variance=1";
PROC NLMIXED DATA=ciadl METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* 3 step thresholds per item, same discrimination across items;
      PARMS d101-d107=-2 d201-d207=-1 d301-d307=0 a=2;
      d1 = d101*i1 + d102*i2 + d103*i3 + d104*i4 + d105*i5 + d106*i6 + d107*i7;
      d2 = d201*i1 + d202*i2 + d203*i3 + d204*i4 + d205*i5 + d206*i6 + d207*i7;
      d3 = d301*i1 + d302*i2 + d303*i3 + d304*i4 + d305*i5 + d306*i6 + d307*i7;
* The adjacent-category model is also a series of equations conditional on response;
* eta1 refers to 0 vs. 1;
* eta2 refers to 1 vs. 2;
* eta3 refers to 2 vs. 3;
      eta1 = exp((1.7*a)*((theta-d1)));
      eta2 = exp((1.7*a)*((theta-d1)+(theta-d2)));
      eta3 = exp((1.7*a)*((theta-d1)+(theta-d2)+(theta-d3)));
* Probabilities per category estimated directly;
           IF cia=0 THEN p = 1    / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=1 THEN p = eta1 / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=2 THEN p = eta2 / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=3 THEN p = eta3 / (1 + eta1 + eta2 + eta3);
* The response variable is defined here;
* Because NLMIXED doesn't have a standard distribution for polytomous data,
  we specify a general one as log(p);
      ll = log(p);
      MODEL cia ~ general(ll);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_1PCM;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_1PCM;
RUN;
```

4 minutes later…

```
1PL (typical) Partial Credit Model Model with Theta variance=1
The NLMIXED Procedure

                         Specifications

Data Set                                    WORK.CIADL
Dependent Variable                          cia
Distribution for Dependent Variable         General
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature

               Dimensions

Observations Used                   4230
Observations Not Used                243
Total Observations                  4473
Subjects                             634
Max Obs Per Subject                    7
Parameters                            22
Quadrature Points                     15

             Fit Statistics

-2 Log Likelihood                 5161.6
AIC (smaller is better)           5205.6
AICC (smaller is better)          5205.8
BIC (smaller is better)           5303.5


NOTE: GCONV convergence criterion satisfied.
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
d101	-1.3104	0.09796	633	-13.38	<.0001	0.05	-1.5028	-1.1181	0.000585
d102	-1.2928	0.1107	633	-11.68	<.0001	0.05	-1.5102	-1.0754	0.003055
d103	-1.2077	0.1045	633	-11.56	<.0001	0.05	-1.4129	-1.0025	-0.00385
d104	-1.2004	0.09198	633	-13.05	<.0001	0.05	-1.3810	-1.0198	-0.00362
d105	-1.7265	0.1051	633	-16.43	<.0001	0.05	-1.9329	-1.5202	-0.00053
d106	-1.6211	0.1098	633	-14.77	<.0001	0.05	-1.8366	-1.4055	0.002483
d107	-2.2538	0.1879	633	-11.99	<.0001	0.05	-2.6229	-1.8847	0.000143
d201	-1.1278	0.08263	633	-13.65	<.0001	0.05	-1.2901	-0.9655	-0.00127
d202	-1.3612	0.1036	633	-13.14	<.0001	0.05	-1.5646	-1.1577	0.003924
d203	-1.2661	0.09706	633	-13.04	<.0001	0.05	-1.4567	-1.0755	-0.00418
d204	-0.9251	0.07890	633	-11.73	<.0001	0.05	-1.0801	-0.7702	-0.00271
d205	-0.8135	0.07010	633	-11.60	<.0001	0.05	-0.9511	-0.6758	0.002007
d206	-1.0849	0.08619	633	-12.59	<.0001	0.05	-1.2542	-0.9157	-0.00041
d207	-1.8568	0.1285	633	-14.45	<.0001	0.05	-2.1091	-1.6045	0.00132
d301	-0.2453	0.06003	633	-4.09	<.0001	0.05	-0.3632	-0.1274	-0.00024
d302	-0.9249	0.07138	633	-12.96	<.0001	0.05	-1.0651	-0.7848	-0.00063
d303	-0.7038	0.06659	633	-10.57	<.0001	0.05	-0.8346	-0.5731	0.004363
d304	-0.3653	0.06186	633	-5.91	<.0001	0.05	-0.4868	-0.2438	-0.00243
d305	-0.2778	0.06045	633	-4.60	<.0001	0.05	-0.3965	-0.1591	-0.00162
d306	-0.8312	0.06971	633	-11.92	<.0001	0.05	-0.9681	-0.6943	0.001843
d307	-1.4324	0.08674	633	-16.51	<.0001	0.05	-1.6028	-1.2621	0.000294
a	1.8293	0.09707	633	18.85	<.0001	0.05	1.6387	2.0200	-0.0022

The common discrimination (slope) parameter is “a”, which is 1.83. The item step parameters (locations) are the d’s. The d1’s give the latent trait location at which, given the answer was 0 or 1, 1 becomes more likely than 0 (not 50%). The d2’s give the latent trait location at which, given that the answer was 1 or 2, 2 becomes more likely than 1 (not 50%). The d3’s give the latent trait location at which, given that the answer was 2 or 3, 3 becomes more likely than 1 (not 50%). This model is not currently estimable in Mplus.

Polytomous (4-category) Generalized Partial Credit Model (Unequal Slopes) Model Syntax and Truncated Output:

```
TITLE "2PL (generalized) Partial Credit Model Model with Theta variance=1";
PROC NLMIXED DATA=ciadl METHOD=GAUSS TECHNIQUE=QUANEW QPOINTS=15; *NOAD;
* All model parameters must be listed here WITH start values;
* 2 step thresholds per item, 1 discrimination per item;
      PARMS d101-d107=-2 d201-d207=-1 d301-d307=0 a01-a07=2;
      d1 = d101*i1 + d102*i2 + d103*i3 + d104*i4 + d105*i5 + d106*i6 + d107*i7;
      d2 = d201*i1 + d202*i2 + d203*i3 + d204*i4 + d205*i5 + d206*i6 + d207*i7;
      d3 = d301*i1 + d302*i2 + d303*i3 + d304*i4 + d305*i5 + d306*i6 + d307*i7;
      a =  a01*i1 + a02*i2 + a03*i3 + a04*i4 + a05*i5 + a06*i6 + a07*i7;
* The adjacent-category model is also a series of equations conditional on response;
* eta1 refers to 0 vs. 1;
* eta2 refers to 1 vs. 2;
* eta3 refers to 2 vs. 3;
      eta1 = exp((1.7*a)*((theta-d1)));
      eta2 = exp((1.7*a)*((theta-d1)+(theta-d2)));
      eta3 = exp((1.7*a)*((theta-d1)+(theta-d2)+(theta-d3)));
* Probabilities per category estimated directly;
           IF cia=0 THEN p = 1    / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=1 THEN p = eta1 / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=2 THEN p = eta2 / (1 + eta1 + eta2 + eta3);
      ELSE IF cia=3 THEN p = eta3 / (1 + eta1 + eta2 + eta3);
* The response variable is defined here;
* Because NLMIXED doesn't have a standard distribution for polytomous data,
  we specify a general one as log(p);
      ll = log(p);
      MODEL cia ~ general(ll);
* The random intercept (theta) is defined here, saved to named dataset;
      RANDOM theta ~ normal(0,1) SUBJECT = case OUT=octo.Theta_2PCM;
* All parameter estimates saved to named dataset;
      ODS OUTPUT ParameterEstimates=octo.Item_2PCM;
RUN;
```

11 minutes later…

```
2PL (generalized) Partial Credit Model Model with Theta variance=1
The NLMIXED Procedure

                         Specifications

Data Set                                    WORK.CIADL
Dependent Variable                          cia
Distribution for Dependent Variable         General
Random Effects                              theta
Distribution for Random Effects             Normal
Subject Variable                            case
Optimization Technique                      Dual Quasi-Newton
Integration Method                          Adaptive Gaussian
                                            Quadrature
               Dimensions

Observations Used                     4230
Observations Not Used                  243
Total Observations                    4473
Subjects                               634
Max Obs Per Subject                      7
Parameters                              28
Quadrature Points                       15

NOTE: GCONV convergence criterion satisfied.

             Fit Statistics

-2 Log Likelihood                   5040.2
AIC (smaller is better)             5096.2
AICC (smaller is better)            5096.6
BIC (smaller is better)             5220.9
```

Fit Statistics for 1PL version of PCM

```
-2 Log Likelihood                  5161.6
AIC (smaller is better)            5205.6
AICC (smaller is better)           5205.8
BIC (smaller is better)            5303.5
```

5161.6 – 5040.2 = 121.4 on df=6 is p <.0001, so the generalized PCM is an improvement over the original PCM (we need different slopes across items).

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
d101	-1.3539	0.08105	633	-16.70	<.0001	0.05	-1.5130	-1.1947	0.006363
d102	-1.3786	0.1024	633	-13.46	<.0001	0.05	-1.5796	-1.1775	-0.00682
d103	-1.2467	0.1040	633	-11.98	<.0001	0.05	-1.4510	-1.0424	-0.00171
d104	-1.2212	0.07967	633	-15.33	<.0001	0.05	-1.3776	-1.0648	0.006055
d105	-1.7787	0.1132	633	-15.72	<.0001	0.05	-2.0009	-1.5565	0.018844
d106	-1.6758	0.1449	633	-11.57	<.0001	0.05	-1.9603	-1.3913	0.000764
d107	-2.4545	0.3526	633	-6.96	<.0001	0.05	-3.1469	-1.7620	0.022102
d201	-0.9588	0.06568	633	-14.60	<.0001	0.05	-1.0877	-0.8298	0.019898
d202	-1.2677	0.09306	633	-13.62	<.0001	0.05	-1.4505	-1.0850	-0.04597
d203	-1.1994	0.09546	633	-12.56	<.0001	0.05	-1.3868	-1.0119	0.046948
d204	-0.8350	0.06601	633	-12.65	<.0001	0.05	-0.9646	-0.7054	-0.00287
d205	-0.7708	0.07235	633	-10.65	<.0001	0.05	-0.9129	-0.6287	0.028792
d206	-1.0384	0.1129	633	-9.20	<.0001	0.05	-1.2600	-0.8167	0.006261
d207	-2.0101	0.2360	633	-8.52	<.0001	0.05	-2.4735	-1.5466	0.010115
d301	-0.1806	0.05193	633	-3.48	0.0005	0.05	-0.2825	-0.07859	-0.01111
d302	-0.8080	0.06926	633	-11.67	<.0001	0.05	-0.9440	-0.6720	0.025557
d303	-0.6483	0.06906	633	-9.39	<.0001	0.05	-0.7840	-0.5127	-0.01451
d304	-0.2781	0.05580	633	-4.98	<.0001	0.05	-0.3877	-0.1686	-0.00459
d305	-0.2855	0.06542	633	-4.36	<.0001	0.05	-0.4140	-0.1570	-0.01613
d306	-1.0383	0.1149	633	-9.04	<.0001	0.05	-1.2639	-0.8127	-0.01129
d307	-2.1917	0.2476	633	-8.85	<.0001	0.05	-2.6780	-1.7054	-0.01206
a01	3.6663	0.5378	633	6.82	<.0001	0.05	2.6103	4.7224	-0.00447
a02	2.3915	0.3068	633	7.79	<.0001	0.05	1.7889	2.9940	-0.00353
a03	2.0288	0.2412	633	8.41	<.0001	0.05	1.5552	2.5024	-0.00384
a04	2.9261	0.3974	633	7.36	<.0001	0.05	2.1456	3.7065	0.00477
a05	1.6999	0.1721	633	9.88	<.0001	0.05	1.3620	2.0378	-0.01265
a06	1.1775	0.1302	633	9.05	<.0001	0.05	0.9219	1.4331	-0.00425
a07	0.7864	0.1059	633	7.42	<.0001	0.05	0.5784	0.9944	0.01225

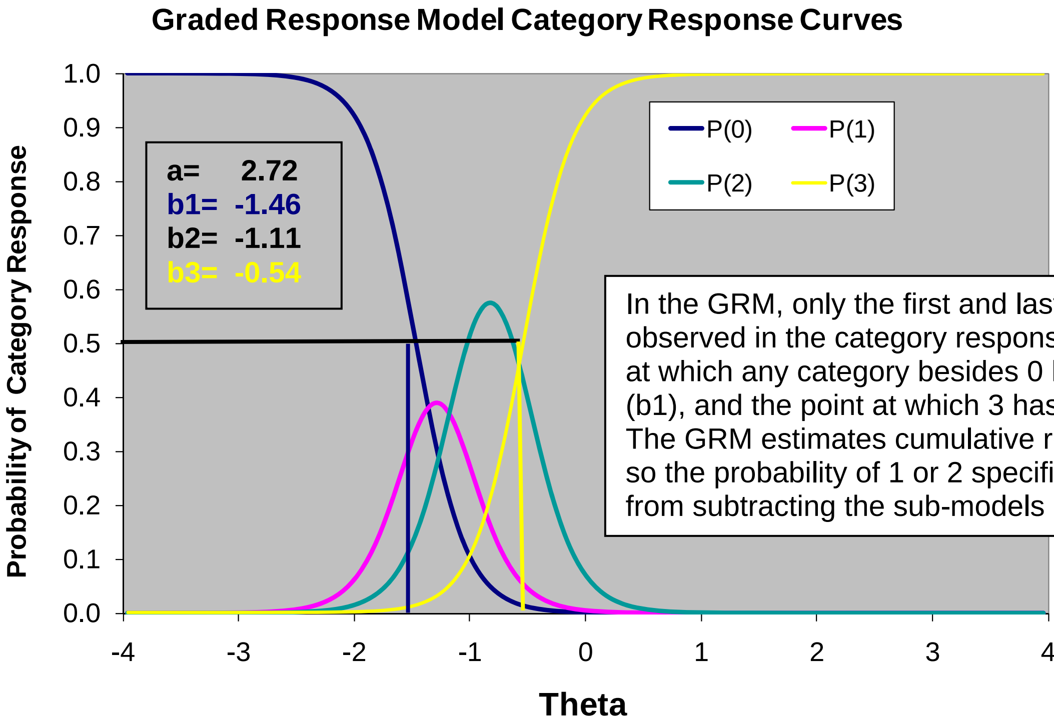
The item discriminations (slope) parameters are the a’s, which gives one slope per item.
The item step parameters (locations) are the d’s, interpreted as before. This model is also not currently estimable in Mplus.

Finally, we can compare information criteria to decide which fits better, the graded response model or the generalized partial credit model. They both use 28 parameters (unequal slopes):

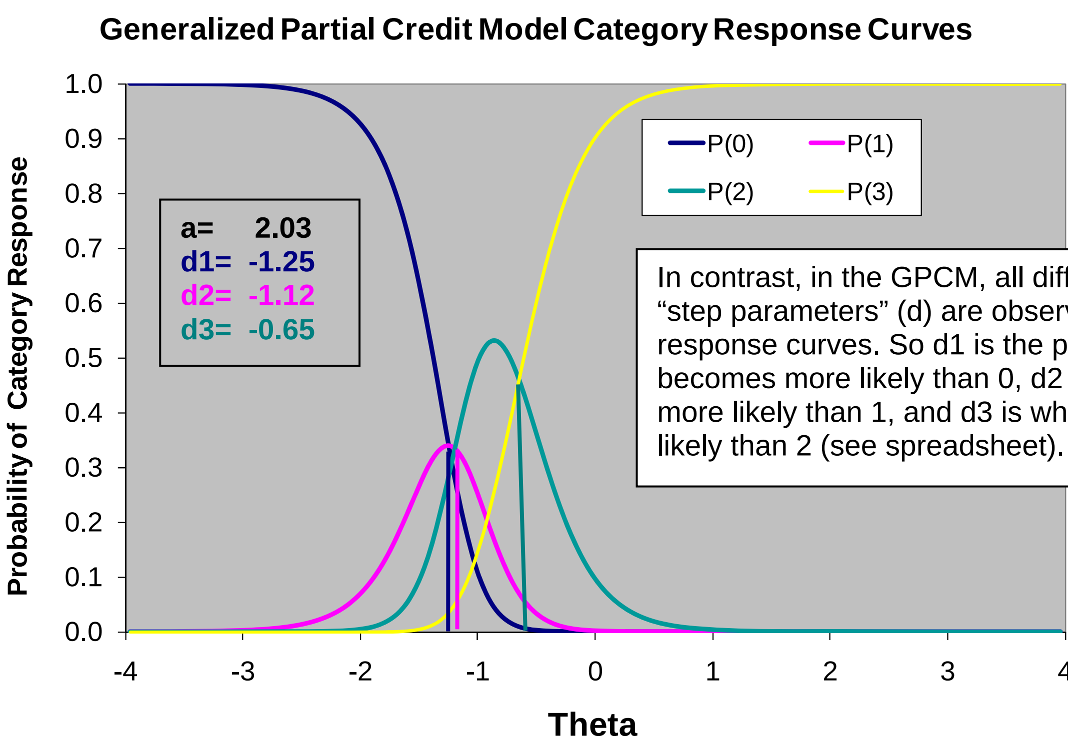
```
         Fit Statistics from GRM

-2 Log Likelihood                   5045.4
AIC (smaller is better)             5101.4
AICC (smaller is better)            5101.8
BIC (smaller is better)             5226.1
```

The AIC and BIC are slightly smaller for the generalized PCM, so apparently that one is better.

Comparison of Prediction from Graded Response vs. Generalized Partial Credit Model for Item 3:



In the GRM, only the first and last difficulty values are observed in the category response curves – the point at which any category besides 0 has 50% probability (b1), and the point at which 3 has 50% probability (b3). The GRM estimates cumulative response categories, so the probability of 1 or 2 specifically is only available from subtracting the sub-models (see spreadsheet).



In contrast, in the GPCM, all difficulty values, called “step parameters” (d) are observed in the category response curves. So d1 is the point at which 1 becomes more likely than 0, d2 is where 2 becomes more likely than 1, and d3 is where 3 becomes more likely than 2 (see spreadsheet).

The GRM and GPCM make very similar predictions here (and often in general). Further, both models suggest there is little distinction between 0 (“can’t do it”) and 1 (“big problems doing it”) for item 3.