Modelling multiple choice items with the two-parameter logistic model

Alejandra Osses 25 July 2015

The Rasch's simple logistic model specifies the probability of a correct response in a given item as a function of on the individual's ability and the difficulty of the item. The model assumes that all items have equal discrimination power in measuring the latent trait by fixing the slope parameter to ´1´ (Rasch, 1980). The two-parameter logistic model (2PL) is a more general model that estimates a discrimination parameter for each item. In ConQuest we refer to these additional parameters as scoring parameters, or scores. In the 2PL, items have different levels of difficulty and also different capabilities to discriminate among individuals of different proficiency (Birnbaum, 1968). Thus, the 2PL model 'frees' the slope of each parameter, allowing different discrimination power for each item. This tutorial exemplifies how to fit a 2PL model for dichotomously scored data in ConQuest. The actual form the model that is fit for dichotomous data is provided as equation (3) in Note 6 (Adams & Macaskill, 2012).

The files used in this sample analysis are:[1]

`ex10.cqc`	The command statements.
`ex1.dat`	The data.
`ex1.lab`	The variable labels for the items on the multiple choice test.
`ex10_show`	The results of the two-parameter analysis.
`ex10_itanal`	The results of the traditional item analyses.

The data used in this tutorial comes from a 12-item multiple-choice test that was administered to 1000 students. The data have been entered into the file `ex1.dat`, using one line per student. A unique student identification code has been entered in columns 1 through 5, and the students' responses to each of the items have been recorded in columns 12 through 23. The response to each item has been allocated one column; and the codes a, b, c and d have been used to indicate which alternative the student chose for each item. If a student failed to respond to an item, an M has been entered into the data file. An extract from the data file is shown in Figure 1.

[1] The last two files are created when the command file is executed.

```
          1         2
12345678901234567890123(column numbers)[2]
40016      acdabaeadacd
655        acdcccecbaca
31140      eccdbcebbacb
.          .
.          .
50321      dabcMcebdaca
30782      acddbcebbacc
.          .
.          .
```

Figure 1. Extract from the Data File `ex1.dat`

In this sample analysis, the generalised model for dichotomously-scored items will be fitted to the data. Traditional item analysis statistics are generated. The contents of a command file that can be used to analyse these data are shown in Figure 2.

```
1.[3]  Datafile ex1.dat;
2.     Format id 1-5 responses 12-23;
3.     Labels << ex1.lab;
4.     set constraints=cases;
5.     Key acddbcebbacc! 1;
6.     Model item! scoresfree;
7.     Estimate;
8.     Show! filetype=xlsx  >> ex10_show.xlsx;
9.     Itanal! filetype=xlsx  >> ex10_itanal.xlsx;
10.    Plot icc! filesave=yes >> ex10_;
11.    Plot mcc! legend=yes, filesave=yes >> ex10_;
12.    plot icc! gins=all, raw=no, overlay=yes, filesave=yes >> ex10_;
```

Figure 2. Sample Command File for a Dichotomous Test

1. The `datafile` statement indicates the name and location of the data file. Any file name that is valid for the operating system you are using can be used here.

2. The `format` statement describes the layout of the data in the file `ex1.dat`. This `format` statement indicates that a field that will be called `id` is located in columns 1 through 5 and that the responses to the items are in columns 12 through 23 of the data

[2] In each of the listings of the data file, column so you can easily identify the data column. The actual ConQuest data file does not have any column labels.

[3] In each of the listings of the command file, each statement is labelled so that it can be easily referred to in the text. The actual ConQuest command file does not have any statement labels.

file. Every `format` statement must give the location of the responses. In fact, the explicit variable `responses` must appear in the `format` statement or ConQuest will not run. In this particular sample analysis, the responses are those made by the students to the multiple choice items; and, by default, `item` will be the implicit variable name that is used to indicate these responses. The levels of the `item` variable (that is, item 1, item 2 and so on) are implicitly identified through their location within the set of responses (called the response block) in the `format` statement; thus, in this sample analysis, the data for item 1 is located in column 12, the data for item 2 is in column 13, and so on.

3. The `labels` statement indicates that a set of labels for the variables (in this case, the items) is to be read from the file `ex1.lab`. An extract of `ex1.lab` is shown in Figure 3. (This file must be text only; if you create or edit the file with a word processor, make sure that you save it using the text only option.)

   The first line of the file contains the special symbol `===>` (a string of three equals signs and a greater than sign) followed by one or more spaces and then the name of the variable to which the labels are to apply (in this case, `item`). The subsequent lines contain two pieces of information separated by one or more spaces. The first value on each line is the level of the variable (in this case, `item`) to which a label is to be attached, and the second value is the label. If a label includes spaces, then it must be enclosed in double quotation marks `(" ")`. In this sample analysis, the label for item 1 is `BSMMA01`, the label for item 2 is `BSMMA02`, and so on.

```
===> item
1      BSMMA01
2      BSMMA02
3      BSMMA03
4      BSMMA04
5      BSMMA05
6      BSMMA06
```

Figure 3. Contents of the Label File `ex1.lab`

4. The `set` statement specifies new values for a range of ConQuest system variables. In this case, the use of the `constraints` argument is setting the identification constraints to `cases`. Therefore, the constraints will be set through the population model by forcing the means of the latent variables to be set to zero and allowing all item parameters (difficulty and discrimination) to be free. The use of cases as the identification constraint is required when estimating a 2PL.

5. The `key` statement identifies the correct response for each of the multiple choice test items. In this case, the correct answer for item 1 is `a`, the correct answer for item 2 is `c`, the correct answer for item 3 is `d`, and so on. The length of the argument in the `key` statement is 12 characters, which is the length of the response block given in the `format` statement.

If a `key` statement is provided, ConQuest will recode the data so that any response `a` to item 1 will be recoded to the value given in the key statement option (in this case, `1`). All other responses to item 1 will be recoded to the value of the `key_default` (in this case, `0`). Similarly, any response `c` to item 2 will be recoded to `1`, while all other responses to item 2 will be recoded to `0`; and so on.

6. The `model` statement must be provided before any traditional or item response analyses can be undertaken. In this example, the argument for the `model` statement is the name of the variable that identifies the response data that are to be analysed (in this case, item). The option `scoresfree` indicates that a score is to be estimated for each scoring category. In this case the data are dichotomously coded, so the resulting model is the 2PL model.

7. The `estimate` statement initiates the estimation of the item response model.

8. The `show` statement produces a sequence of tables that summarise the results of fitting the item response model. The option `filetype` sets the format of the results file, in this case an Excel file. The redirection symbol (>>) is used so that the results will be written to the file `ex10_show.xlsx` in your current directory.

9. The `itanal` statement produces a display of the results of a traditional item analysis. As with the `show` statement, the results are redirected to a file (in this case, `ex10_itanal.xlsx`).

10. The `plot icc` statement will produce 12 item characteristic curve plots, one for each item. The plots will compare the modelled item characteristic curves with the empirical item characteristic curves. The option `filesave` indicates that the resulting plot will be saved into a file in your working directory. The redirection symbol (>>) is used so that the plots will be written to png files named `ex10_`. The name of the file will be completed with 'ítem X' where the X represents the number of the item (eg '`ex10_item 7`'). Note that the `plot` command is not available in the console version of ConQuest.

11. The `plot mcc` statement will produce 12 category characteristic curve plots, one for each item. The plots will compare the modelled item characteristic curves with the empirical item characteristic curves (for correct answers) and will also show the behaviour of the distractors. As with the `plot icc` statement, the results are redirected to a file (in this case, `ex10_`). Note that this command is not available in the console version of ConQuest.

12. The `plot icc` statement will produce 12 item characteristic curve plots, one for each item. The option `gins=all` indicates that one plot is provided for each listed generalised item. The use of the `raw=no` option prevents the display of the raw data in the plot. The `overlay=yes` option allows the requested plots to be shown in a single window. As with the previous `plot` statements, the resulting plots are saved to png files in the working directory.

RUNNING THE TWO-PARAMETER MODEL

To run this sample analysis, start the GUI version. Open the file `ex10.cqc` and choose Run➔Run All. ConQuest will begin executing the statements that are in the cqc file; and as they are executed they will be echoed in the Output Window. When it reaches the estimation command ConQuest will begin fitting the two-parameter model to the data. This analysis will converge in 23 iterations.

After the estimation is completed, the two statements that produce Excel files output (`show` and `itanal`) will be processed. The `show` statement will produce an Excel file (ex10_show.xlsx) with nine tabs summarising the results of fitting the item response model. The `itanal` statement will produce an Excel file (ex10_itanal.xlsx) with one tab showing items statistics. In the case of the GUI version, the `plot` statements will produce 25 plots altogether. 12 plots will contain the item characteristic curve by score category for each of the items in the data. 12 plots will contain the item characteristic curve by response category for each of the items in the data. The last `plot` statement will produce one plot with the ICC by score category for all items.

RESULTS OF FITTING THE TWO PARAMETER MODEL

As mentioned above, the show file will contain nine tabs. The first tab in the `ex10_show.xlsx` file shows a summary of the estimation. An extract is shown in Figure 4. The table indicates the data set that set that was analysed and provides summary information about the model fitted (eg the number of parameters estimated, the number of iterations that the estimation took, the reason for the estimation termination).

```
============================================================================
ConQuest: Generalised Item Response Modelling Software    Wed Jul 08 09:21 2015
SUMMARY OF THE ESTIMATION
=============================================================Build: Jul  8 2015===

Estimation method was: Gauss-Hermite Quadrature with 15 nodes
No node filtering
Assumed population distribution was: Gaussian
Constraint was: CASES
The Data File: ex1.dat
The format:  id 1-5 responses 12-23
No case weights
The regression model:
Grouping Variables:
The item model: item
Two parameter logistic model
Cases in file: 1000  Cases in estimation: 1000
Final Deviance:     13154.79108
Akaike Information Criterion (AIC): 13202.79108
Total number of estimated parameters: 24
The number of iterations: 23
Termination criteria:  Max iterations=1000, Parameter Change= 0.00010
                       Deviance Change= 0.00010
Iterations terminated because the deviance convergence criteria was reached
Random number generation seed:     1.00000
Number of nodes used when drawing PVs: 2000
Number of nodes used when computing fit: 200
Number of plausible values to draw: 5
Maximum number of iterations without a deviance improvement: 100
Maximum number of Newton steps in M-step: 10
Value for obtaining finite MLEs for zero/perfects:     0.30000

key 1 scored as 1: acddbcebbacc
============================================================================
```

Figure 4. Summary of estimation Table

The second tab in the `ex10_show` Excel file gives the parameter difficulty estimates for each of the items along with their standard errors and some diagnostics tests of fit (Figure 5[4]). The difficulty parameter estimates the "delta" values in equation (3) of Note 6. The last column in the table (*2PL scaled estimate*) shows the two-parameter scaled estimate of the item. Each value in this column is the delta value divided by the estimate of the score and is a common *alternative expression* of item difficulty for 2PL models. At the bottom of the table an item separation reliability and chi-squared test of parameter equality are reported.

ConQuest: Generalised Item Response Modelling Software Fri Jul 24 14:34 2015
TABLES OF RESPONSE MODEL PARAMETER ESTIMATES
Build: Jul 22 2015

TERM 1: item

VARIABLES				UNWEIGHTED FIT				WEIGHTED FIT				2PL SCALED
item		ESTIMATE	ERROR^	MNSQ	Confidence Interval		T	MNSQ	Confidence Interval		T	ESTIMATE
1	BSMMA01	-0,985	0,125	1,00	0,91	1,09	0,1	1,00	0,91	1,09	0,0	-0,491
2	BSMMA02	-1,276	0,096	1,04	0,91	1,09	0,8	1,00	0,92	1,08	0,0	-1,260
3	BSMMA03	-1,284	0,113	1,02	0,91	1,09	0,5	1,01	0,92	1,08	0,2	-0,882
4	BSMMA04	-0,252	0,082	0,99	0,91	1,09	-0,1	1,00	0,94	1,06	-0,1	-0,214
5	BSMMA05	0,097	0,067	1,00	0,91	1,09	0,0	1,00	0,97	1,03	0,1	0,203
6	BSMMA06	-1,412	0,100	1,01	0,91	1,09	0,1	1,00	0,92	1,08	0,1	-1,406
7	BSMSA07	-1,369	0,089	1,01	0,91	1,09	0,2	1,00	0,92	1,08	0,0	-1,998
8	BSMSA08	-1,237	0,080	1,00	0,91	1,09	-0,1	1,00	0,92	1,08	0,0	-2,956
9	BSMSA09	-2,076	0,128	0,97	0,91	1,09	-0,6	1,00	0,88	1,12	0,0	-2,068
10	BSMSA10	-1,338	0,086	0,99	0,91	1,09	-0,2	1,00	0,92	1,08	-0,1	-2,283
11	BSMSA11	-1,694	0,118	0,97	0,91	1,09	-0,6	1,01	0,91	1,09	0,2	-1,417
12	BSMSA12	-0,357	0,075	1,00	0,91	1,09	-0,1	1,00	0,95	1,05	0,0	-0,409

A parameter estimate in italics indicates that it is constrained
Separation Reliability = 0.975
Chi-square test of parameter equality = 1788.60, df = 12, Sig Level = 0.000
^ Empirical standard errors have been used

Figure 5. Item Parameter Estimates

The sixth and seventh tabs provide the item map of the item difficulty parameters (not shown here). The first of these maps provides an item difficulty plot according to the estimate displayed in the *2PL scaled estimate* column in Figure 5. The second map is based on the unscaled estimate (*estimate* column in Figure 5).

For the purpose of this Tutorial, the tab of interest in the `ex10_show` Excel file is the *scores* tab. Here, the item discrimination parameters are presented (Figure 6). The 'score' column displays the different score assigned to the correct response in each item (discrimination parameter). The error associated to the estimate is also presented.

[4] Note: the tables in Figs. 5–7 show decimal commas in the parameter estimates. Different versions of Excel might render decimal marks differently (e.g., as 'dot').

ConQuest: Generalised Item Response Modelling Software Fri Jul 24 14:34 2015
TABLE(S) OF GIN SCORES
Build: Jul 22 2015

GIN	Score	Error	GIN Labels		
1.1	2,003814	0,22667	item	1	BSMMA01
2.1	1,012438	0,122578	item	2	BSMMA02
3.1	1,455735	0,158226	item	3	BSMMA03
4.1	1,173713	0,127218	item	4	BSMMA04
5.1	0,4782	0,087131	item	5	BSMMA05
6.1	1,004396	0,124657	item	6	BSMMA06
7.1	0,685228	0,108473	item	7	BSMSA07
8.1	0,418354	0,097187	item	8	BSMSA08
9.1	1,003912	0,141661	item	9	BSMSA09
10.1	0,586227	0,104229	item	10	BSMSA10
11.1	1,195995	0,142907	item	11	BSMSA11
12.1	0,873232	0,106369	item	12	BSMSA12

Average Score 0.99094

Figure 6. Score estimates for each item

The item analysis is shown on the `ex1.itanal` output file. The `itanal` output includes a table showing classical difficulty, discrimination, and point-biserial statistics for each item. Figure 7 shows the results for items 2 and 3. The 2PL discrimination estimate for each is shown in the *score* column. Summary results, including coefficient alpha for the test as a whole, are printed at the end of the spreadsheet.

Item 2

item:2 (BSMMA02)

Cases for this item 1000 Item-Rest Cor. 0.31 Item-Total Cor. 0.47
Item Threshold(s): NOT AVAILABLE Weighted MNSQ 1.00
Item Delta: -1.28 2-PL scaled delta: -1.26 Slope: 1.01

Label	Score	Count	% of tot	Pt Bis	t	sig	PV1Avg:1	PV1 SD:1
M	0	5	0,5	-0,08	-2,51	0,012	-0,928	1,339
a	0	59	5,9	-0,17	-5,3	0,000	-0,762	0,984
b	0	152	15,2	-0,19	-6,22	0,000	-0,626	0,972
c	1,01	743	74,3	0,31	10,13	0,000	0,185	0,882
d	0	41	4,1	-0,1	-3,12	0,002	-0,574	1,041

Scores assigned to each of the codes

Item slope value

Item 3

item:3 (BSMMA03)

Cases for this item 1000 Item-Rest Cor. 0.41 Item-Total Cor. 0.56
Item Threshold(s): NOT AVAILABLE Weighted MNSQ 1.01
Item Delta: -1.28 2-PL scaled delta: -0.88 Slope: 1.46

Label	Score	Count	% of tot	Pt Bis	t	sig	PV1Avg:1	PV1 SD:1
M	0	9	0,9	-0,11	-3,47	0,001	-0,919	0,815
a	0	117	11,7	-0,26	-8,61	0,000	-0,879	0,863
b	0	119	11,9	-0,21	-6,66	0,000	-0,736	0,875
c	0	38	3,8	-0,11	-3,54	0,000	-0,848	0,854
d	1,46	717	71,7	0,41	14,03	0,000	0,279	0,843

Figure 7. Item Analysis Results

Figure 8 shows plots that were produced by the `plot icc` and the `plot mcc` command for items 1 and item 5. In the left panel, the ICC plot shows a comparison of the empirical item characteristic curve (the broken line, which is based directly upon the observed data) with the modelled item characteristic curve (the smooth line).

The right panel shows a matching plot produced by the `plot mcc` command. In addition to showing the modelled curve and the matching empirical curve, this plot shows the characteristics of the incorrect responses — the distractors. In particular it shows the proportion of students in each of a sequence of ten ability groupings[5] that responded with each of the possible responses.

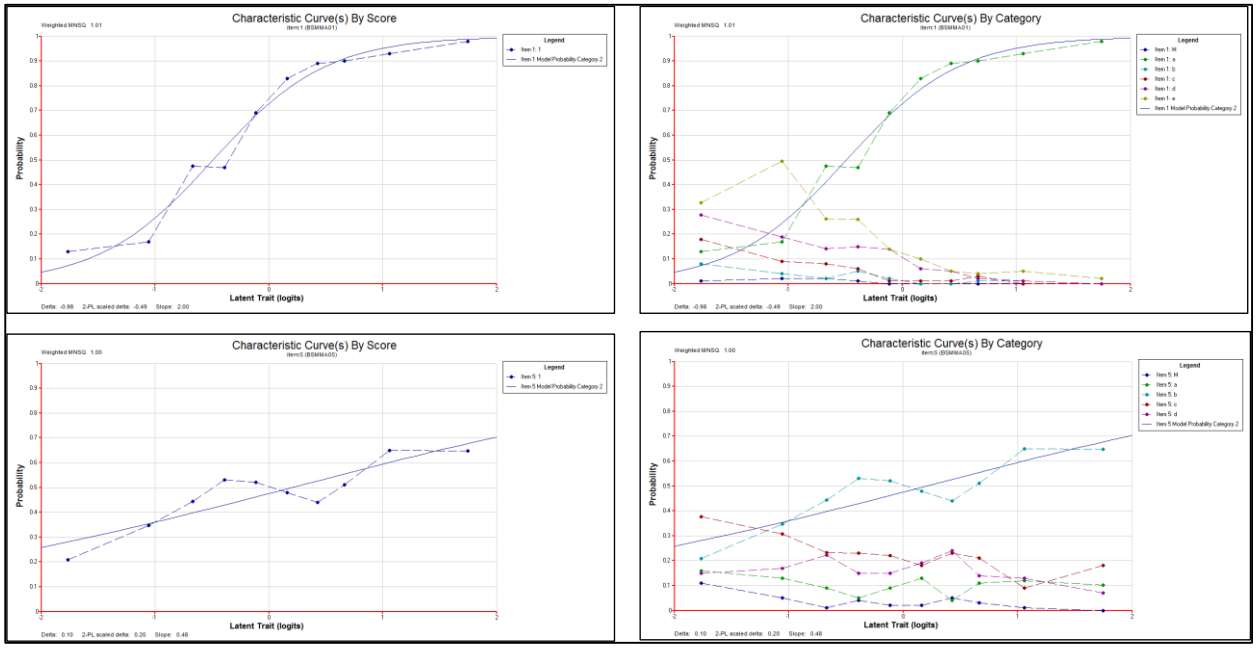

Figure 8. Plots for item 1 and item 5

The second `plot icc` command of the `ex10.cqc` file produces the plot shown in Figure 9. Here all ICCs are plotted in the same window, which allows the graphical comparison of the different discrimination capabilities of each item.

[5] Ten ability groupings is a default setting that can be altered.

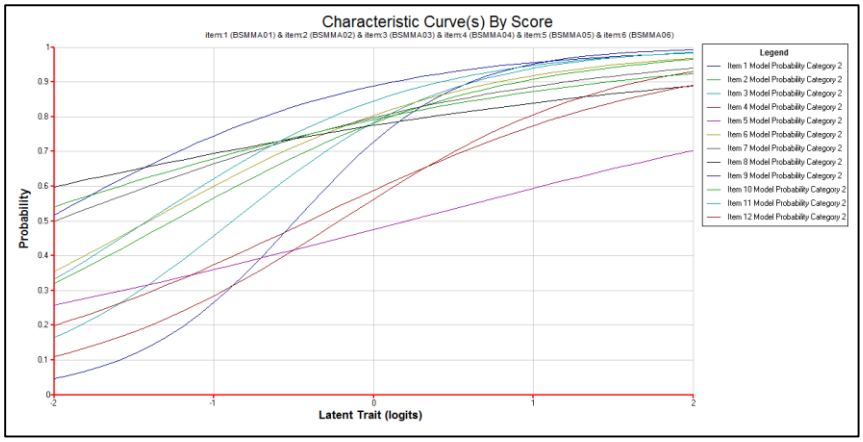

Figure 9. Item Characteristic Curve plot for all items in the data set

SUMMARY

This tutorial shows how ConQuest can be used to analyse a multiple-choice test with the 2PL model. Some key points covered in this tutorial are:

- the need to `set` constraints to cases when estimation of discrimination parameters is required.
- the `model` statement allows the estimation of different slopes (discrimination) for each item through the `scoresfree` option.
- the `itanal` statement provides information about the discrimination estimate for each item.
- the `plot` statement allows the graphical comparison of the discrimination power of each item.

REFERENCES

Adams, R.J., & Macaskill, G. (2012). Score Estimation and Generalised Partial Credit Models. Note 6. ACER ConQuest, Notes and Tutorials. Retrieved from http://www.acer.edu.au/files/Conquest-Notes-6-ScoreEstimationAndGeneralisedPartialCreditModels.pdf

Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*. In Lord, F.M. & Novick, M.R. (Eds.), Statistical theories of mental test scores. Reading, MA: AddisonWesley.

Rasch, G. 1980. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: University of Chicago Press (original work published 1960).

Wu, M. L., Adams, R. J., Wilson, M. R., Haldane, S.A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software* [computer program]. Camberwell: Australian Council for Educational Research.