Chapter 18
Other Models of Modern Test Theory for Dichotomous Responses

There are a number of models in modern test theory for analysing dichotomous responses. Dichotomous responses are scored into two categories, for example correct (1) and incorrect (0) or agree (1) and disagree (0). Three common unidimensional models are the Rasch model of Rasch Measurement Theory (RMT), and the two-parameter logistic (2PL) model and three-parameter logistic (3PL) model of item response theory (IRT). The distinction between Rasch measurement and item response theories is explained in Andrich (2004, 2011).

The Rasch Model

The Rasch model for dichotomous responses takes the form

$$\Pr\{X_{ni} = 1|\beta_n, \delta_i\} = e^{\beta_n - \delta_i}/\left(1 + e^{\beta_n - \delta_i}\right), \tag{18.1}$$

where β_n is the proficiency of person n and δ_i is the difficulty of item i (Rasch, 1960).

Figure 18.1 shows the item characteristic curves from the Rasch model for three items of different difficulty.

The Rasch model has a single person parameter and a single item parameter. The following features of the Item Characteristic Curve (ICC) in the Rasch model have been studied in Chap. 7 and are relevant to recall here. First, the probability of answering an item correctly gradually increases with proficiency level. Second, the slopes of the curves are equal producing parallel curves that do not cross. Third, the point of inflection of the curve occurs where the probability of answering the item correctly is 0.5. The relevance of non-crossing ICCs for dichotomous items is described in Wright (1997). The central relevance is that for all values of a person location, the items have the same order of difficulty. From the perspective of IRT, the Rasch model for dichotomous responses is known simply as the one-parameter logistic (1PL) model.

D. Andrich and I. Marais, *A Course in Rasch Measurement Theory*,
Springer Texts in Education, https://doi.org/10.1007/978-981-13-7496-8_18

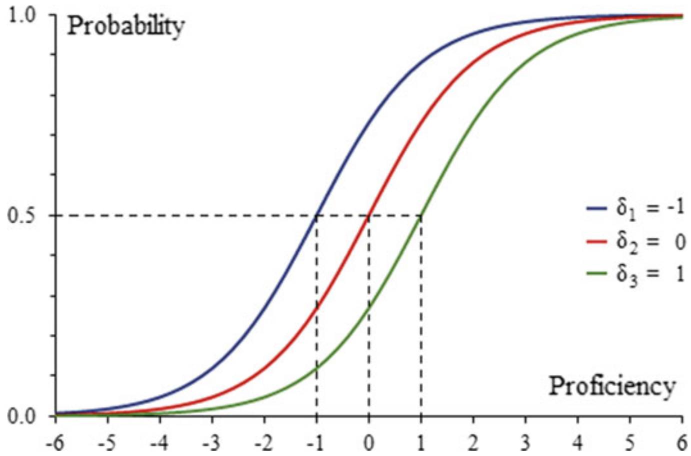

Fig. 18.1 Item characteristic curves from the Rasch model

2PL Model

The two-parameter logistic (2PL) model for dichotomous responses takes the form

$$\Pr\{X_{ni} = 1|\beta_n, \delta_i, \alpha_i\} = e^{\alpha_i(\beta_n - \delta_i)}/\left[1 + e^{\alpha_i(\beta_n - \delta_i)}\right], \tag{18.2}$$

where β_n is the proficiency of person n, δ_i is the difficulty and α_i is the discrimination parameter for item i (Birnbaum, 1968). The discrimination parameter characterizes the slope of the ICC.

The addition of the discrimination parameter means the 2PL can model items, which are not equally related to the latent trait (Embretson & Reise, 2000). Because they may not have equal slopes, as shown in Fig. 18.2, it is possible for the ICCs of the 2PL model to cross. The point of inflection of the curve still occurs where the probability of answering the item correctly is 0.5.

A consequence of including the discrimination parameter in the 2PL model is that the interpretation of item difficulties becomes ambiguous (Ryan, 1983). In particular, because the relative ordering of the items depends on the proficiency of the person, it is not possible to order the items according to difficulty when items vary significantly in discrimination (Ryan, 1983). Figure 18.2 illustrates how the probability of answering an item correctly depends on the proficiency of the person, despite all three items having the same difficulty. For example, for a person with a proficiency of −2 logits on the scale in Fig. 18.2 the ordering of the difficulties is items 1, 2, 3, while for a person with a proficiency of 2 logits the ordering is items 3, 2, 1. In addition, unlike the Rasch model, the total score is not a sufficient statistic for the person parameter.

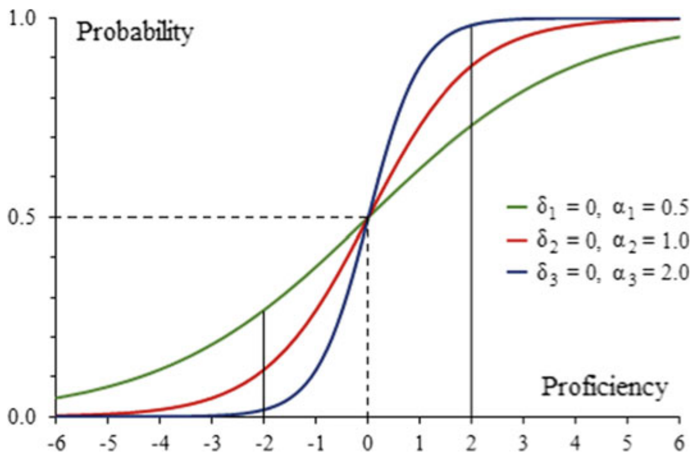

Fig. 18.2 Item characteristic curves from the 2PL model

3PL Model

The three-parameter logistic (3PL) model for dichotomous responses to multiple-choice items takes the form

$$\Pr\{X_{ni} = 1|\beta_n, \delta_i, \alpha_i, \gamma_i\} = \gamma_i + (1 - \gamma_i)P_{ni} = P_{ni} + \gamma_i(1 - P_{ni}), \qquad (18.3)$$

where $P_{ni} = e^{\alpha_i(\beta_n - \delta_i)} / \left[1 + e^{\alpha_i(\beta_n - \delta_i)}\right]$, β_n is the proficiency of person n, δ_i is the difficulty, α_i is the discrimination and γ_i is the guessing parameter for item i (Birnbaum, 1968).

How guessing can be considered from the perspective of the Rasch model was summarized in Chap. 17. The 3PL model involves a single person parameter and three item parameters; location, discrimination and guessing. The guessing parameter manifests as a lower asymptote on the ICC, as shown in Fig. 18.3. When an item can be guessed correctly, the probability of success is greater than zero (Embretson & Reise, 2000). Hence the ICC does not fall to zero, even for low proficiency persons, because guessing increases the probability of success on an item. The probability of success from random guessing is $1/C$ when there are C alternative responses to a multiple-choice (MC) item. Estimates of the lower asymptote in the 3PL model often differ from $1/C$ because persons can eliminate MC alternatives (Embretson & Reise, 2000) or be attracted to an incorrect MC alternative. Item difficulty occurs at the point of inflection in the ICC but is not necessarily associated with a probability of 0.5 (Embretson & Reise, 2000). This is illustrated in Fig. 18.3.

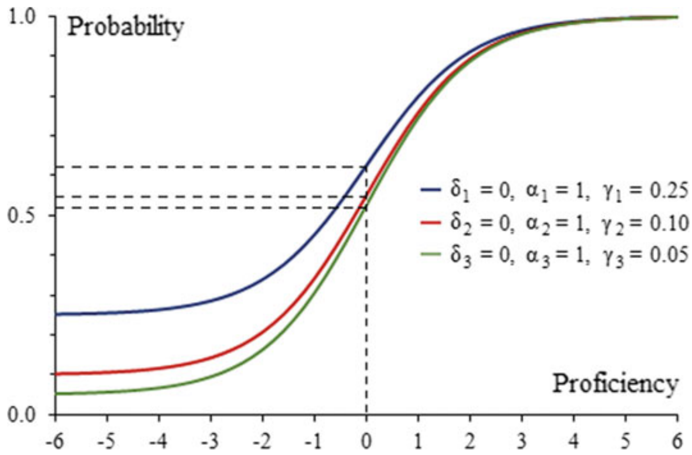

Fig. 18.3 Item characteristic curves from the 3PL model

The parameter estimates are dependent on the distribution of the persons who respond to the items (Maris & Bechger, 2009). Also, as in the 2PL, it is difficult to interpret an individual parameter because all the parameters are estimated simultaneously and influence each other (Han, 2012). Again, the total score is not a sufficient statistic for the person parameter.

References

Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care, 42*(1), i7–i16.

Andrich, D. (2011). Rating scales and Rasch measurement. *Expert Review of Pharmacoeconomics & Outcomes Research, 11*(5), 571–585.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Reading, Massachusetts: Addison-Wesley.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J: L. Erlbaum Associates.

Han, K. T. (2012). Fixing the c parameter in the three-parameter logistic model. *Practical Assessment, Research & Evaluation, 17*(1), 1–24.

Maris, G., & Bechger, T. (2009). On interpreting the model parameters for the three parameter logistic model. *Measurement, 7,* 75–88.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests.* Expanded edition (1980) with foreword and afterword by Wright, B. D. (Ed.). Chicago: The University of Chicago Press. Reprinted (1993) Chicago: MESA Press.

Ryan, J. P. (1983). Introduction to latent trait analysis and item response theory. In W. E. Hathaway (Ed.), *Testing in the schools: New directions for testing and measurement* (Vol. 19, pp. 49–64). San Fransisco: Jossey-Bass.

Wright, B. D. (1997). A history of social science measurement. *Educational Measurement: Issues and Practice, 16*(4), 33–45.