The Rasch Model as a Loglinear Model

Gideon J. Mellenbergh and Pieter Vijn
University of Amsterdam

The Rasch model is formulated as a loglinear model. The goodness of fit and parameter estimates of the Rasch model can be obtained using the iterative proportional fitting algorithm for loglinear models. It is shown in an example that the relation between the estimates of the iterative proportional fitting algorithm and the unconditional maximum likelihood Rasch algorithm are almost perfectly linear. The Rasch model can be extended with a design for the items, which can be formulated as a loglinear model.

In the Rasch model for binary scored items the probability that person j gives a response scored one to item i is written as (Rasch, 1960)

$$P_{ij} = \exp(a_j - d_i)/\{1 + \exp(a_j - d_i)\} \quad , \qquad [1]$$

where a_j represents the person's ability and d_i the item's difficulty. From the model in Equation 1, it follows that the natural logarithm of the response ratio is

$$\ln\{P_{ij}/(1 - P_{ij})\} = a_j - d_i \quad . \qquad [2]$$

Brogden (1977) has shown the relation between the Rasch model, the law of comparative judgment, and additive conjoint measurement; and Perline, Wright, and Wainer (1979) have presented the Rasch model as a special case of additive conjoint measurement. Additive conjoint measurement applies to mental test data when a monotonic transformation of the P_{ij}'s yields an additive representation:

$$f(P_{ij}) = a_j - d_i \quad , \qquad [3]$$

where f is a monotonic function. When f is the inverse logistic transformation of Equation 1, the Rasch model follows. This shows that the Rasch model is a special case of additive conjoint measurement.

In the Bradley-Terry model for paired comparisons, the probability that stimulus j is preferred to stimulus i is written as (Fienberg, 1977, p. 119):

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 5, No. 3, Summer 1981, pp. 369–376

0146-6216/81/030369-08$1.40

$$P_{ij} = \Pi_j/(\Pi_i + \Pi_j) \quad , \tag{4}$$

where $\pi_i \geqslant 0$, and $\pi_j \geqslant 0$. Defining $d_i^* = \ln\pi_i$, and $d_j^* = \ln\pi_j$, it follows that the natural logarithm of the response ratio is

$$\ln\{P_{ij}/(1 - P_{ij})\} = d_j^* - d_i^* \quad . \tag{5}$$

This shows the analogy between the Rasch model for mental tests and the Bradley-Terry model for paired comparisons. In both models two objects are compared. In the Rasch model the person's ability is compared with the item difficulty, whereas in the Bradley-Terry model the comparison is between the "difficulties" of two items. In both models the response ratio of the comparison has an additive structure.

Fienberg and Larntz (1976) have shown that the Bradley-Terry model is a special case of the loglinear model; this has also been discussed by Fienberg (1977, sec. 8.4). In this paper the Rasch model is described as a special case of the loglinear model.

The Rasch Model and the Loglinear Model

Suppose that a test is administered to a group of people. After removing all items that are scored 0 or 1 by everyone, the test consists of n binary items. Under the Rasch model a person's score, i.e., the number of 1's, is the sufficient statistic for his or her ability (Fischer, 1974, p. 203). This means that all information on the person's ability is contained in the score. Consequently, all people with the same score will obtain the same ability estimate. It does *not* imply that all people with the same score do have the same ability. It only implies that if the Rasch model is valid for a set of data, people with the same score will obtain the same ability estimate and no further differentiation in ability estimates can be made between them. Therefore, as Perline, Wright, and Wainer (1979, p. 239) have stated, "statistical estimates of abilities and item parameters can be obtained proceeding *as if* everyone with the same raw score has exactly the same ability." The number of score groups is $(n + 1)$, i.e., the scores run from 0 to n. In general, p of these score groups do not yield any information. This is always the case for the score groups 0 and n because all people in these groups have obtained a zero or one, respectively, on all items. Moreover, it is possible that some score groups do not contain people at all. Therefore, only $(n + 1 - p)$ score groups yield relevant information. From the sufficiency property of the score, it follows that for parameter estimation, the continuous latent ability space can be partitioned into $(n + 1 - p)$ latent classes. The response variable has only two categories, 0 or 1. The information can be summarized in an $n \times (n + 1 - p) \times 2$ Item × Score × Response table.

The saturated loglinear model for the natural logarithm of the expected value of the frequency of the k^{th} response ($k = 1, 2$) for item i ($i = 1, 2, ..., n$) in score group j ($j = 1, 2, ..., n + 1 - p$) is (Bishop, Fienberg, & Holland, 1975, chap. 2):

$$\ln F_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)} \quad , \tag{6}$$

with constraints:

$$\sum_{i=1}^{n} u_{1(i)} = \sum_{j=1}^{n+1-p} u_{2(j)} = \sum_{k=1}^{2} u_{3(k)} = 0 \quad , \tag{7}$$

$$\sum_{i=1}^{n} u_{12(ij)} = \sum_{j=1}^{n+1-p} u_{12(ij)} = \sum_{i=1}^{n} u_{13(ik)} = \sum_{k=1}^{2} u_{13(ik)}$$

$$= \sum_{j=1}^{n+1-p} u_{23(jk)} = \sum_{k=1}^{2} U_{23(jk)} = 0 \quad , \qquad [8]$$

$$\sum_{i=1}^{n} u_{123(ijk)} = \sum_{j=1}^{n+1-p} u_{123(ijk)} = \sum_{k=1}^{2} u_{123(ijk)} = 0 \quad . \qquad [9]$$

Note that the index j refers to the score group, whereas j in model Equation 1 indicates the person.

Using P_{ij} for the probability that a person from score group j gives a correct response to item i, it follows from these Equations that the natural logarithm of the response ratio is

$$\ln\{P_{ij}/(1 - P_{ij})\} = \ln(F_{ij1}/F_{ij2}) = (u_{3(1)} - u_{3(2)})$$

$$+ (u_{13(i1)} - u_{13(i2)}) + (u_{23(j1)} - u_{23(j2)})$$

$$+ (u_{123(ij1)} - u_{123(ij2)})$$

$$= 2u_{3(1)} + 2u_{13(i1)} + 2u_{23(j1)} + 2u_{123(ij1)}$$

$$= C + D_i + A_j + DA_{ij} \quad , \qquad [10]$$

with constraints

$$\sum_{i=1}^{n} D_i = \sum_{j=1}^{n+1-p} A_j = 0 \quad , \qquad [11]$$

$$\sum_{i=1}^{n} DA_{ij} = \sum_{j=1}^{n+1-p} DA_{ij} = 0 \quad . \qquad [12]$$

The model is the general logit model (Fienberg, 1977, chap. 6). The parameter C is a constant, D_i is the effect of the i^{th} item, A_j the effect of the j^{th} score group, and DA_{ij} the combined effect of i^{th} item and j^{th} score group on the response ratio.

The model is saturated, which implies that the fit of the model to the data is perfect. A special nonsaturated case is obtained by setting the parameters $u_{123(ijk)}$ in Equation 6 equal to zero:

$$\ln F_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} \quad , \qquad [13]$$

with the constraints of Equations 7 and 8. The corresponding model for the response ratio is the linear logit model:

$$\ln\{P_{ij}/(1 - P_{ij})\} = C + D_i + A_j \quad , \qquad [14]$$

with the constraints of Equation 11. The relations between the parameters of Equations 13 and 14 are

$$C = 2u_{3(1)}; \; D_i = 2u_{13(i1)}; \; A_j = 2u_{23(j1)} \quad . \qquad [15]$$

Comparing Equations 2 and 14 shows the equivalence of the Rasch model and the linear logit model. In Equation 14 P_{ij} is the probability that a person from score group j gives a correct response to item i, whereas in Equation 2 P_{ij} is the probability that a person with ability a_j gives a correct response to item i. As stated before, however, under the Rasch model people with the same score have the same ability estimate and can be treated *as if* they had the same ability. Therefore, although the interpretation of P_{ij} in both equations is different, for parameter estimation they can be considered to be equivalent. The differences between both equations are the constant term C and the constraints of Equation 11 in the linear logit model. In the Rasch model the constant term is absorbed in the parameters. The constraints are necessary to fix the scale of the parameters. In the Rasch model this is done setting one item difficulty parameter or one score ability parameter equal to a fixed value. Consequently, the parameters of the Rasch model and the linear logit model are linearly related to each other.

The parameters in the model of Equation 13, and their asymptotic standard errors, can be estimated from a sample using the maximum likelihood method (Bishop, Fienberg, & Holland, 1975, chap. 3, sec. 4.4.2). For the model of Equation 13 no direct estimates of the parameters are possible (Bishop et al., 1975, sec. 3.4.3), but the estimates can be obtained using the iterative proportional fitting algorithm (Bishop et al., 1975, sec. 3.5.1). The computations can be done, for example, with the programs ECTA (Goodman & Fay, 1974) and BMD (Dixon & Brown, 1977). The parameter estimates and their standard errors of the model in Equation 14 are twice the value of the corresponding estimates of the model of Equation 13.

The model of Equation 13 can be tested using Pearson's χ^2 or the likelihood ratio G^2 statistic:

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{n+1-p} \sum_{k=1}^{2} \{(f_{ijk} - \hat{f}_{ijk})^2/\hat{f}_{ijk}\} \quad , \qquad [16]$$

$$G^2 = 2 \sum_{i=1}^{n} \sum_{j=1}^{n+1-p} \sum_{k=1}^{2} \{f_{ijk} \ln(f_{ijk}/\hat{f}_{ijk})\} \quad , \qquad [17]$$

where f_{ijk} and $\hat{f}_{ijk}$ are the observed and model expected frequencies in the cell corresponding to the i^{th} item, the j^{th} score group, and the k^{th} response category. Both statistics are asymptotically chi-square distributed with $(n - 1) \times (n - p)$ degrees of freedom (Bishop et al., 1975, sec. 4.2). The contributions of item i to the χ^2 and G^2 statistics are, respectively,

$$\chi_i^2 = \sum_{j=1}^{n+1-p} \sum_{k=1}^{2} \{(f_{ijk} - \hat{f}_{ijk})^2/\hat{f}_{ijk}\} \quad , \qquad [18]$$

$$G_i^2 = 2 \sum_{j=1}^{n+1-p} \sum_{k=1}^{2} \{f_{ijk} \ln(f_{ijk}/\hat{f}_{ijk})\} \quad . \qquad [19]$$

These quantities can be used to show which items have the largest contributions to the χ^2 and G^2 statistics. They cannot be used as statistics to test the fit of the items to the Rasch model. The model expected frequencies are estimated using the Score × Response table. This table is used for all items simultaneously. The item contributions to the total χ^2 or G^2 statistics are therefore dependent.

Facet designs can be used to construct items. For example, Dekking and Raadsheer (reported by Mellenbergh, Kelderman, Stijlen, & Zondag, 1979) used a situational and a behavioral facet for the construction of items measuring social anxiety in children. An example is an item for the combination of the elements Social from the situational facet and Avoidance from the behavioral facet: "During recreation I don't play with other children" (yes/no). The logit model for the Rasch model with v elements of the first facet and w elements of the second facet is

$$\ln\{P_{ghj}/(1 - P_{ghj})\} = C + G_g + H_h + GH_{gh} + A_j \quad , \qquad [20]$$

with constraints

$$\sum_{g=1}^{v} G_g = \sum_{h=1}^{w} H_h = \sum_{j=1}^{n+1-p} A_j = 0 \quad , \qquad [21]$$

$$\sum_{g=1}^{v} GH_{gh} = \sum_{h=1}^{w} GH_{gh} = 0 \quad . \qquad [22]$$

The parameters G_g and H_h are the effects of the g^{th} element of the first facet, and the h^{th} element of the second facet, while GH_{gh} is the combined effect of these elements. The item difficulty parameters D_i of the model in Equation 14 are replaced by the parameters G_g, H_h, and GH_{gh}.

Estimating Parameters

Two methods have mainly been used in estimating the parameters of the Rasch model: the unconditional (UML) and conditional maximum likelihood (CML) method. (For a review of estimation procedures see Wainer, Morgan, & Gustafsson, 1980.)

In Wright and Panchapakesan's (1969) UML method, the likelihood as a function of the item and person parameters is considered. The estimates are obtained by maximizing the likelihood function with respect to the item and person parameters. These estimates can, however, be inconsistent (Andersen, 1980, p.p. 244, 245). Maximum likelihood estimates may be inconsistent in models with incidental parameters, i.e., in models where the number of parameters increases as the sample size increases. In Wright and Panchapakesan's (1969) UML method the person parameters are incidental because each person adds one person parameter to the model. Wright and Douglas (1977) showed in simulation studies that the inconsistency can be removed by multiplying the item parameters with the correction factor $(n - 1)/n$. For reasonable test lengths the correction factor is about equal to 1 and can be neglected. At the theoretical level, Haberman (1977) showed that under certain conditions the UML estimates converge to their population values: Let n_i be a strictly increasing sequence of item samples, N_i an increasing sequence of person samples such that N_i goes to infinity as i goes to infinity, and $N_i \geqslant n_i$ for each $i (i = 1, 2, \ldots)$. The UML estimates for the person and item parameters converge in probability to their population values if ln (N_i/n_i) goes to zero as i goes to infinity. From these simulation studies and asymptotic arguments it is inferred that for reasonably long tests and a reasonably large number of persons the UML method will be sufficient for many practical applications.

The CML method is, however, theoretically a more sound method (Andersen, 1973, 1980; Fischer, 1974). In the CML method the likelihood function conditional on the sufficient statistics for the person parameters, i.e., the scores, is considered. Because the sufficient statistics contain all information on the person ability parameters, the conditional likelihood is a function of only the item parameters and is independent of the person parameters. Maximizing the conditional likelihood function with respect to the item parameters yields the CML estimates for the item parameters. These estimates are under some mild conditions consistent and asymptotically normally distributed with known covariance matrix (Andersen, 1980, p. 247). Finally, if the item parameters are estimated with small standard errors, person ability parameters are estimated, considering the estimated item parameters as given values (Andersen, 1980, p. 264). This means that the person ability parameters are estimated treating the item parameters as given population values.

Using the linear logit model, Equation 14, the likelihood is simultaneously maximized with respect to item and score group parameters. This means that the item and score group parameters are unconditionally estimated. In this respect the method resembles Wright and Panchapakesan's (1969) UML method. But in Wright and Panchapakesan's method the estimates are inconsistent because the person ability parameters are incidental. Using the linear logit model the score group parameters are not incidental because the number of score groups is $(n + 1 - p)$, which is independent of the number of persons. In this respect the method resembles the CML method. But in the CML method the person ability parameters are estimated treating the estimated item parameters as population values, whereas in the linear logit model both item and person ability parameters are simultaneously estimated. From standard theory on loglinear models for contingency tables (Fienberg, 1977, p. 130), it follows that the estimates in a linear logit model are asymptotically normal, consistent, and efficient. Thus, under the Rasch model, the estimation procedure in the linear logit model appears to be a sound method.

Example

Perline, Wright, and Wainer (1979) reported data from a nine-item scale for parole decisions. They reported that the tests of the fit of the additive conjoint measurement model yielded poor results. The results of the Rasch model goodness-of-fit tests were somewhat better, but they, too, did not fit too well. Using the program BICAL, they computed unconditional maximum likelihood estimates of the parameters in the Rasch model, which were compared to estimates of the additive conjoint measurement model.

Perline et al.'s (1979) Table 2 shows that only the score groups 0 and 9 had the same response for all items and, therefore, that the value of p equals 2. From their Table 2 the $9 \times 8 \times 2$ Item × Score × Response table was calculated (Table 1). Using the program ECTA, the model of Equation 13 was fitted to the data. The values of χ^2 and G^2 were, respectively, 285.54 and 295.56 with 56 degrees of freedom. Consequently, the fit of the model was rather poor.

From Perline et al.'s (1979) Figures 2 and 3 the BICAL estimates of the item difficulty and subject ability estimates were read with a ruler as accurately as possible. As shown in Figure 1, the relation of the BICAL and ECTA estimates is almost perfectly linear. Finally, the item contributions to the χ^2 statistic were calculated (Table 2). Item 6 showed the largest contribution. Perline et al. (1979) also indicated that this item did not fit well.

Table 1
Frequencies in Item by Score by Response Table for Parole Data
(from Perline, Wright & Wainer, 1979)

	Response															
	1								0							
	Score								Score							
Item	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
6	0	3	4	15	11	11	10	8	15	44	57	69	71	75	50	39
1	0	2	9	20	27	24	28	40	15	45	52	64	55	62	32	7
8	0	2	5	10	25	55	51	47	15	45	56	74	57	31	9	0
7	0	9	24	34	42	50	49	46	15	38	37	50	40	36	11	1
4	0	3	11	44	60	82	60	47	15	44	50	40	22	4	0	0
9	0	11	20	43	56	78	56	47	15	36	41	41	26	8	4	0
2	4	24	37	54	56	66	54	47	11	23	24	30	26	20	6	0
3	0	10	32	57	69	83	58	47	15	37	29	27	13	3	2	0
5	11	30	41	59	64	67	54	47	4	17	20	25	18	19	6	0

Figure 1
Plot of Estimates Computed from BICAL (Perline, Wright, & Wainer, 1979) versus ECTA for Parole Data

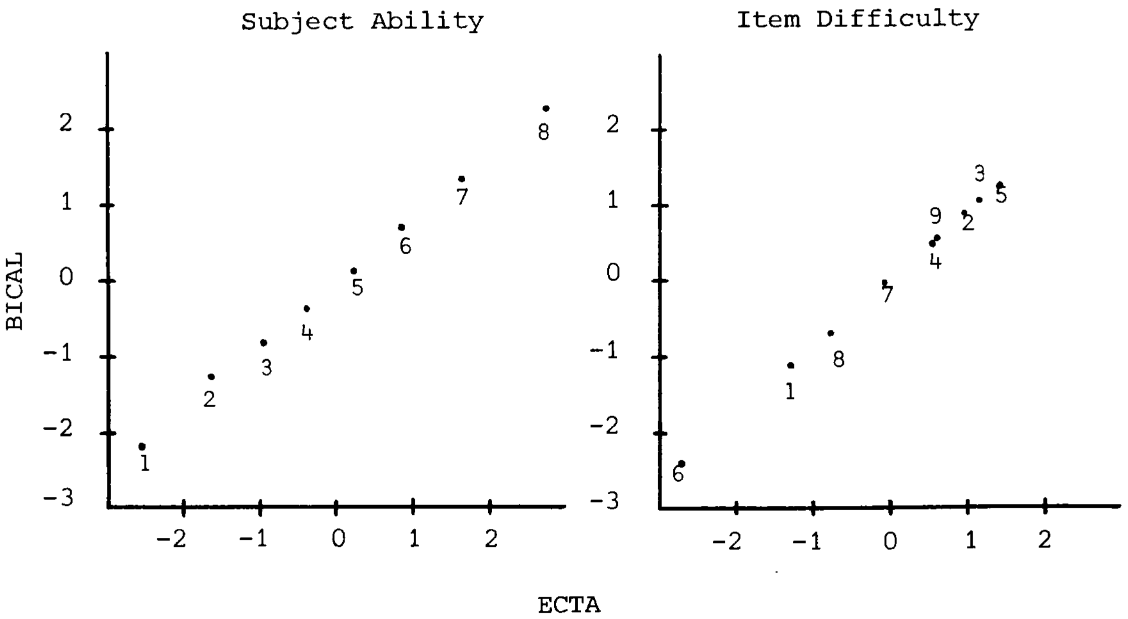

Table 2
Item Contributions to Pearson Chi-Square for Parole Data

Item	χ_i^2	Item	χ_i^2	Item	χ_i^2
1	17.32	4	45.19	7	13.08
2	20.78	5	50.16	8	29.46
3	18.34	6	78.98	9	12.28

References

Andersen, E. B. *Conditional inference and models for measuring.* Copenhagen, Denmark: Mental-hygienisk Forlag, 1973.

Andersen, E. B. *Discrete statistical models with social science applications.* Amsterdam: North-Holland Publishing Company, 1980.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. *Discrete multivariate analysis: Theory and practice.* Cambridge MA: MIT Press, 1975.

Brogden, H. E. The Rasch model, the law of comparative judgment, and additive conjoint measurement. *Psychometrika,* 1977, *42,* 631-634.

Dixon, W. J., & Brown, M. D. *Biomedical computer programs, P-series 1977.* Berkeley: University of California Press, 1977.

Fienberg, S. E. *The analysis of cross-classified data.* Cambridge MA: MIT Press, 1977.

Fienberg, S. E., & Larntz, K. Loglinear representation for paired and multiple comparisons models. *Biometrika,* 1976, *63,* 245-254.

Fischer, G. *Einführung in die Theorie Psychologischer Tests.* Bern: Huber, 1974.

Goodman, L. A., & Fay, R. *ECTA program, description for users.* Chicago: Department of Statistics, University of Chicago, 1974.

Haberman, S. J. Maximum likelihood estimates in exponential response models. *The Annals of Statistics,* 1977, *5,* 815-841.

Mellenbergh, G. J., Kelderman, H., Stijlen, J. G., & Zondag, E. Linear models for the analysis and construction of instruments in a facet design. *Psychological Bulletin,* 1979, *86,* 766-776.

Perline, R., Wright, B. D., & Wainer, H. The Rasch model as additive conjoint measurement. *Applied Psychological Measurement,* 1979, *3,* 237-255.

Rasch, G. *Probabilistic models for some intelligence and attainment tests.* Copenhagen, Denmark: The Danish Institute for Educational Research, 1960.

Wainer, H., Morgan, A., & Gustafsson, J.-E. A review of estimation procedures for the Rasch model with an eye towards longish tests. *Journal of Educational Statistics,* 1980, *5,* 35-64.

Wright, B. D., & Douglas, G. A. Best procedures for sample-free item analysis. *Applied Psychological Measurement,* 1977, *1,* 281-295.

Wright, B. D., & Panchapakesan, N. A procedure for sample-free item analysis. *Educational and Psychological Measurement,* 1969, *29,* 23-48.

Acknowledgments

The authors thank Fred N. Kerlinger, Wim J. van der Linden, and Henk Kelderman for their comments; Hans Verhey for doing the analysis with the ECTA-program; and Ineke Wesling for typing the manuscript.

Author's Address

Send requests for reprints or further information to Gideon J. Mellenbergh or Pieter Vijn, Vakgroep Methodenleer, Psychologisch Laboratorium, Universiteit van Amsterdam, Weesperplein 8, 1018 XA Amsterdam, The Netherlands.