Approximate expressions for parameter estimates in the Rasch model

Leslie Cohen

The Wright–Panchapakesan (1969) maximum-likelihood equations for estimating the ability and difficulty parameters in the Rasch model have hitherto required an iterative procedure for their solution. In this paper the equations are solved approximately and simple closed forms are established which give values close enough for most practical purposes to those obtained by iteration. The results are illustrated with reference to a 50-item physics test.

1. Introduction

The Rasch model of item analysis is a one-parameter latent trait model applicable to tests comprised of dichotomous items. Associated with each item i is a difficulty parameter δ_i and with each person v an ability parameter β_v. The probability of person v succeeding on item i is assumed to be given by

$$P_{vi} = \Psi(\beta_v - \delta_i), \tag{1}$$

where the function Ψ is the cumulative distribution function for the logistic distribution and is defined by

$$\Psi(x) = e^x/(1+e^x).$$

The maximum-likelihood estimation procedure for the model has been essentially resolved (Andersen, 1972, 1973), leading to sufficient and consistent estimators for the parameters β_v and δ_i. Wright & Douglas (1977) refer to this procedure as 'conditional'. They show that the simpler approach of Wright & Panchapakesan (1969), hereafter referred to as W–P, in which item parameters and ability parameters are estimated simultaneously, gives results which are close enough for all practical purposes to those obtained using the conditional procedure.

W–P solved their maximum-likelihood equations by an iteration procedure, UCON. Convergence of the iteration is fairly slow, and Wright & Douglas (1975*b*) report an approximate technique, PROX, based on an integral approximation developed by the present author, which gives results 'equivalent to the estimates obtained from UCON for all practical purposes', but with a very substantial saving in computer time. The present paper takes the technique a stage further and establishes closed forms for the PROX values; the formulae have been quoted by Wright (1977, p. 100). Simple closed forms are also obtained for the standard errors of estimate of the ability and difficulty parameters. A simple approximate expression is also established for the internal consistency reliability coefficient.

2. Outline of the approximation procedure

In the Rasch model the test score is a sufficient statistic for estimating the ability parameters. Persons may therefore be grouped by their test score in the equations for

0007-1102/79/0501-0113$02.00/0

maximum-likelihood estimation. Let

b_r = the ability estimate for any person with score r,

d_i = the difficulty estimate for item i,

N_r = the number of persons with score r.

We denote the estimated probability that a person with score r will succeed on item i by P_{ri}, so that

$$P_{ri} = \Psi(b_r - d_i). \tag{2}$$

In order to avoid infinite parameter estimates it is first necessary to edit the data to remove items on which all persons succeed or all persons fail, and to exclude persons who either succeed on all items or fail on all items. If L is the total number of items after editing, and S_i is the total score on item i, the maximum-likelihood equations of W–P are

$$\sum_{r=1}^{L-1} N_r P_{ri} = S_i \quad (i = 1, \ldots, L), \tag{3}$$

$$\sum_{i=1}^{L} P_{ri} = r \quad (r = 1, \ldots, L-1), \tag{4}$$

with the subsidiary condition

$$\sum_{i=1}^{L} d_i = 0. \tag{5}$$

W–P proceeded to solve the equations by iteration, taking as initial values $b_r^{(0)} = y_r$ and $d_i^{(0)} = x_i - \xi$, where

$$y_r = \ln\{r/(L-r)\}, \tag{6}$$

$$x_i = \ln\{(N-S_i)/S_i\}, \tag{7}$$

$$\xi = \sum_{i=1}^{L} x_i/L, \tag{8}$$

and N is the total number of persons taking the test.

The aim of the present paper is to obtain simple approximate solutions of the W–P equations. The starting point is the empirical observation that the iterated values of the ability parameters b_r are, to quite a high degree of accuracy, proportional to y_r. The following steps are then involved in the approximation:

(i) b_r is replaced by λy_r in equation (3) where λ is a constant to be determined,

(ii) the summation over discrete values of r in equation (3) is replaced by an integral over a continuum,

(iii) the integral is evaluated using the normal approximation to the logistic distribution,

(iv) the resulting equation is inverted, leading to an approximate expression for d_i in terms of x_i,

(v) the approximations for b_r and d_i are substituted into equation (4) and the summation over i is replaced by an integral over a continuum,

(vi) this leads to explicit expressions for λ and for the corresponding constant occurring in the approximation for d_i, and to a further condition which must be satisfied if the approximations are to be consistent.

3. Approximate solutions of the W–P equations

We assume that the ability estimates are given approximately by the expression

$$b_r \simeq \lambda y_r. \tag{9}$$

Substituting this into equation (3) gives, on division by N,

$$S_i/N \simeq \sum_{r=1}^{L-1} \Psi(\lambda y_r - d_i)\, N_r/N. \tag{10}$$

The summation over discrete values of r may be evaluated approximately by means of the following device: it is assumed that the ys (and hence also the abilities) are approximately normally distributed, the first and second moments being denoted by η and σ^2 where

$$\eta = \sum_{r=1}^{L-1} y_r N_r/N,$$

$$\sigma^2 = \sum_{r=1}^{L-1} (y_r - \eta)^2 N_r/N.$$

The proportion of values of y to be found in the interval $(y, y+\delta y)$ is then given by $[\mathrm{d}\Phi\{(y-\eta)/\sigma\}/\mathrm{d}y]\,\delta y$, where Φ is the cumulative distribution function for the normal distribution. The sum in equation (10) can then be replaced by an integral giving

$$S_i/N \simeq \int_{-\infty}^{\infty} \Psi(\lambda y - d_i) \frac{\mathrm{d}}{\mathrm{d}y}\left\{\Phi\left(\frac{y-\eta}{\sigma}\right)\right\} \mathrm{d}y. \tag{11}$$

The standard deviation σ is determined by the actual population of candidates taking the test, and no statistical estimation is involved in obtaining its value. In practice this integral approximation for the sum in equation (10) seems to be a good one even when the actual distribution of abilities diverges substantially from the normal.

The next step is to use the normal approximation to the logistic distribution, viz.

$$\Psi(x) \simeq \Phi(x/1{\cdot}7).$$

Equation (11) then becomes

$$S_i/N \simeq \int_{-\infty}^{\infty} \Phi\left(\frac{\lambda y - d_i}{1{\cdot}7}\right) \frac{\mathrm{d}}{\mathrm{d}y}\left\{\Phi\left(\frac{y-\eta}{\sigma}\right)\right\} \mathrm{d}y.$$

This integral can be evaluated using the general result

$$\int_{-\infty}^{\infty} \Phi(a+bt)\,\Phi'(t)\,\mathrm{d}t = \Phi\left(\frac{a}{\sqrt{(1+b^2)}}\right).$$

Hence

$$S_i/N \simeq \Phi\left(\frac{\lambda\eta - d_i}{(\lambda^2\sigma^2 + 2{\cdot}89)^{\frac{1}{2}}}\right)$$

$$\simeq \Psi\left(\frac{\lambda\eta - d_i}{(1+\lambda^2\sigma^2/2{\cdot}89)^{\frac{1}{2}}}\right) \tag{12}$$

on using again the approximate relationship between Φ and Ψ. Inversion of this equation gives

$$\lambda\eta - d_i \simeq (1+\lambda^2\sigma^2/2{\cdot}89)^{\frac{1}{2}} \ln\{(N-S_i)/S_i\};$$

that is,

$$d_i \simeq \gamma x_i + \lambda\eta, \tag{13}$$

where

$$\gamma^2 = 1 + \lambda^2\,\sigma^2/2{\cdot}89. \tag{14}$$

If the approximations (9) and (13) for b_r and d_i are now substituted into equation (4) we obtain

$$r/L \simeq \sum_{i=1}^{L} \Psi(\lambda y_r - \gamma x_i - \lambda\eta)/L. \tag{15}$$

A similar device to that used earlier can be adopted for replacing the sum in this equation by an integral. We assume that the xs (and hence the difficulty parameters) are approximately normally distributed, the first and second moments being ξ (equation 8) and τ^2 where

$$\tau^2 = \sum_{i=1}^{L} (x_i - \xi)^2/L.$$

The corresponding integral approximation for the sum once again appears to be very satisfactory in practice, except in those cases where the number of items is small (< 20) and the difficulty distribution is very skewed. Equation (15) now becomes

$$\begin{aligned}
r/L &\simeq \int_{-\infty}^{\infty} \Psi(\lambda y_r - \gamma x - \lambda\eta)\,\frac{\mathrm{d}}{\mathrm{d}x}\left\{\Phi\left(\frac{x-\xi}{\tau}\right)\right\}\mathrm{d}x \\
&\simeq \int_{-\infty}^{\infty} \Phi\left(\frac{\lambda y_r - \gamma x - \lambda\eta}{1{\cdot}7}\right)\frac{\mathrm{d}}{\mathrm{d}x}\left\{\Phi\left(\frac{x-\xi}{\tau}\right)\right\}\mathrm{d}x \\
&= \Phi\left(\frac{\lambda y_r - \gamma\xi - \lambda\eta}{(\gamma^2\tau^2 + 2{\cdot}89)^{\frac{1}{2}}}\right) \\
&\simeq \Psi\left(\frac{\lambda y_r - \gamma\xi - \lambda\eta}{(1+\gamma^2\tau^2/2{\cdot}89)^{\frac{1}{2}}}\right).
\end{aligned}$$

Finally inversion of this equation gives

$$\lambda y_r - \gamma\xi - \lambda\eta \simeq (1+\gamma^2\tau^2/2{\cdot}89)^{\frac{1}{2}} \ln\{r/(L-r)\}.$$

It therefore follows that for the equations to be consistent we must have

$$\lambda^2 = 1 + \gamma^2\tau^2/2{\cdot}89, \tag{16}$$

and

$$\lambda\eta + \gamma\xi \simeq 0. \tag{17}$$

Equations (14) and (16) may be solved to give expressions for λ and γ entirely in terms of the observed standard deviations σ, τ, viz.

$$\lambda = \left(\frac{1+\tau^2/2{\cdot}89}{1-\sigma^2\tau^2/8{\cdot}352}\right)^{\frac{1}{2}}, \tag{18}$$

$$\gamma = \left(\frac{1+\sigma^2/2{\cdot}89}{1-\sigma^2\tau^2/8{\cdot}352}\right)^{\frac{1}{2}}. \tag{19}$$

Using equation (17) the approximation (13) for the estimates of the difficulty parameters becomes

$$d_i \simeq \gamma(x_i - \xi). \tag{20}$$

The subsidiary condition (5) is therefore automatically satisfied by these values.

4. Standard errors of estimate

The standard errors of estimate of the ability parameters are given in the W–P procedure by the formula

$$\sigma_E{}^2(b_r) = \left(\sum_{i=1}^{L} P_{ri}(1-P_{ri})\right)^{-1}$$

(Wright & Douglas, 1975*a*).

Now the probability of person v of ability β_v succeeding on item i is P_{vi} (equation 1). The score obtained by this person will therefore have an expected value $\sum_{i=1}^{L} P_{vi}$, while its standard error will be given by $\sum_{i=1}^{L} P_{vi}(1-P_{vi})$. Hence the estimated value of the expected score for persons whose total score is r is given by $\sum_{i=1}^{L} P_{ri}$, which is equal to r by equation (4); the estimated value of the standard error of the expected score for such persons is $\sum_{i=1}^{L} P_{ri}(1-P_{ri})$. The standard error of score r, $\sigma_E{}^2(r)$, is therefore given by the approximate equation

$$\sigma_E{}^2(r) \simeq \sum_{i=1}^{L} P_{ri}(1-P_{ri}).$$

It therefore follows that

$$\sigma_E(b_r) \simeq 1/\sigma_E(r). \tag{21}$$

Since $b_r \simeq \lambda y_r$ there is another approximate relation between the standard errors which can be obtained by differentiation, namely

$$\sigma_E(b_r) \simeq \frac{\lambda L}{r(L-r)}\,\sigma_E(r). \tag{22}$$

From (21) and (22) we obtain

$$\sigma_E(b_r) \simeq \left(\frac{\lambda L}{r(L-r)}\right)^{\frac{1}{2}}. \tag{23}$$

A good approximation for the standard error of estimate of d_i can be obtained in a similar way. W–P use the formula

$$\sigma_E{}^2(d_i) = \left(\sum_{r=1}^{L-1} N_r P_{ri}(1-P_{ri})\right)^{-1}.$$

The standard error of estimate of S_i can be shown to be given by

$$\sigma_E{}^2(S_i) \simeq \sum_{r=1}^{L-1} N_r P_{ri}(1-P_{ri}).$$

Hence

$$\sigma_E(S_i) \simeq 1/\sigma_E(d_i).$$

Using $d_i \simeq \gamma(x_i - \xi)$ it follows on differentiation that

$$\sigma_E(d_i) \simeq \frac{\gamma N}{S_i(N-S_i)}\,\sigma_E(S_i).$$

Hence

$$\sigma_E(d_i) \simeq \left(\frac{\gamma N}{S_i(N-S_i)}\right)^{\frac{1}{2}}. \tag{24}$$

We can also obtain an estimate of internal consistency reliability (cf. Lawley, 1943). The mean value of $\sigma_E{}^2(r)$ averaged over the whole population is

$$\sigma_E{}^2 = \frac{1}{N}\sum_{r=1}^{L-1} N_r\,\sigma_E{}^2(r).$$

Using equations (21) and (23) this gives

$$\sigma_E^2 \simeq (LM - M^2 - \sigma_r^2)/\lambda L,$$

where M is the mean total score and σ_r^2 is the total score variance. Hence the internal consistency reliability coefficient is given by

$$\rho \simeq 1 - (LM - M^2 - \sigma_r^2)/\lambda L \sigma_r^2. \tag{25}$$

5. Results and discussion

By way of illustration we consider a 50-item test in physics taken by 703 candidates.† The values of the constants λ, η, γ and ξ obtained in this case are

$$\lambda = 1{\cdot}074, \quad \eta = 0{\cdot}419, \quad \gamma = 1{\cdot}146, \quad \xi = -0{\cdot}358.$$

Equation (17) is reasonably well satisfied in this example. The approximations for the ability and difficulty parameter estimates are given in Tables 1 and 2 alongside the iterated solutions of the W–P equations. Also given are the approximations for the standard errors from equations (23) and (24).

Table 1. Item difficulties

Item no.	Estimates d_i obtained by iteration	Approximate values of d_i	Standard errors of estimate	Item no.	Estimates d_i obtained by iteration	Approximate values of d_i	Standard errors of estimate
1	−1·58	−1·63	0·11	26	−0·65	−0·65	0·09
2	−0·15	−0·15	0·08	27	0·16	0·16	0·08
3	0·38	0·38	0·08	28	−0·28	−0·27	0·09
4	0·07	0·07	0·08	29	−0·35	−0·35	0·09
5	−0·36	−0·36	0·09	30	0·93	0·93	0·09
6	−0·49	−0·49	0·09	31	−0·20	−0·19	0·08
7	−0·73	−0·74	0·09	32	1·18	1·17	0·09
8	−0·84	−0·85	0·09	33	−0·80	−0·80	0·09
9	1·61	1·59	0·09	34	0·11	0·12	0·08
10	−0·16	−0·16	0·08	35	−0·50	−0·50	0·09
11	0·27	0·28	0·08	36	−0·62	−0·62	0·09
12	−0·05	−0·04	0·08	37	0·91	0·92	0·08
13	−0·23	−0·23	0·08	38	0·88	0·88	0·08
14	0·13	0·13	0·08	39	−0·38	−0·38	0·09
15	0·40	0·41	0·08	40	0·21	0·22	0·08
16	−0·59	−0·59	0·09	41	−0·47	−0·47	0·09
17	−0·03	−0·02	0·08	42	−0·68	−0·69	0·09
18	1·31	1·30	0·09	43	−0·29	−0·29	0·09
19	−0·12	−0·11	0·08	44	−0·46	−0·45	0·09
20	−0·67	−0·67	0·09	45	−0·49	−0·49	0·09
21	1·01	1·01	0·09	46	0·74	0·74	0·08
22	1·47	1·46	0·09	47	0·10	0·11	0·08
23	0·53	0·54	0·08	48	0·80	0·80	0·08
24	−0·21	−0·20	0·08	49	0·11	0·11	0·08
25	−0·56	−0·56	0·09	50	−0·38	−0·37	0·09

† I am grateful to the Joint Matriculation Board for making the data available to me and for allowing me to quote the results of the W–P procedure for this test.

It is clear that in this particular case the approximations for the parameter estimates are extremely good. The approximations for the standard errors also turn out to agree very closely with the W–P values (the latter are not given in the tables). The value of

Table 2. Ability parameters

Test score	Score count	Estimates b_r obtained by iteration	Approx. values of b_r	Standard errors of estimate	Test score	Score count	Estimates b_r obtained by iteration	Approx. values of b_r	Standard errors of estimate
1	0	−4·08	−4·18	1·01	26	32	0·08	0·09	0·30
2	0	−3·36	−3·41	0·73	27	32	0·16	0·17	0·30
3	0	−2·93	−2·95	0·60	28	30	0·25	0·26	0·30
4	0	−2·61	−2·62	0·53	29	30	0·34	0·35	0·30
5	0	−2·36	−2·36	0·48	30	21	0·44	0·44	0·30
6	0	−2·14	−2·14	0·44	31	25	0·53	0·53	0·30
7	0	−1·96	−1·95	0·42	32	27	0·62	0·62	0·31
8	1	−1·80	−1·78	0·40	33	35	0·72	0·71	0·31
9	1	−1·65	−1·63	0·38	34	22	0·82	0·81	0·32
10	2	−1·51	−1·49	0·36	35	23	0·92	0·91	0·32
11	1	−1·38	−1·36	0·35	36	15	1·03	1·01	0·33
12	5	−1·26	−1·24	0·34	37	18	1·14	1·12	0·33
13	7	−1·15	−1·12	0·33	38	28	1·26	1·24	0·34
14	11	−1·04	−1·01	0·33	39	26	1·38	1·36	0·35
15	16	−0·94	−0·91	0·32	40	14	1·51	1·49	0·36
16	10	−0·83	−0·81	0·32	41	16	1·65	1·63	0·38
17	12	−0·74	−0·71	0·31	42	10	1·81	1·78	0·40
18	15	−0·64	−0·62	0·31	43	8	1·97	1·95	0·42
19	25	−0·55	−0·53	0·30	44	13	2·16	2·14	0·44
20	17	−0·46	−0·44	0·30	45	5	2·38	2·36	0·48
21	23	−0·37	−0·35	0·30	46	7	2·64	2·62	0·53
22	28	−0·28	−0·26	0·30	47	6	2·96	2·96	0·60
23	27	−0·19	−0·17	0·30	48	6	3·40	3·41	0·73
24	23	−0·10	−0·09	0·30	49	5	4·12	4·18	1·01
25	25	−0·01	0·00	0·30					

the internal consistency reliability coefficient given by equation (25) is 0·871 which agrees well with the KR-20 value of 0·870.

In many other examples the approximations derived in this paper turn out to be as accurate as in the case of the physics test. Experience is now accumulating on the accuracy of the approximations in a wide range of situations and the results so far indicate that they are quite satisfactory provided the number of items in the test is not too small. It appears that the approximations for the ability parameter estimates are nearly always adequate; however, when the number of items is less than 20 and the distribution of difficulty parameters is skewed then equation (20) overestimates the magnitudes of the difficulty parameters. Even then the iterated values of d_i remain proportional to $(x_i - \xi)$ to a high degree of accuracy, and it might be possible to modify the expression for γ so that the approximations become adequate in all circumstances.

References

Andersen, E. B. (1972). The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistical Society, Series B*, **34** (1), 42–54.

Andersen, E. B. (1973). *Conditional Inference and Models for Measuring*. Copenhagen, Denmark: Mentalhygienjnisk Forlag.

Lawley, D. N. (1943). Problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh, A*, **61**, 273–287.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, **14**, 97–116.

Wright, B. D. & Douglas, G. A. (1975*a*). Best test design and self-tailored testing. Research Memorandum, no. 19. Statistical Laboratory, Department of Education, University of Chicago.

Wright B. D. & Douglas G. A. (1975*b*). Better procedures for sample-free item analysis. Research Memorandum, no. 20. Statistical Laboratory, Department of Education, University of Chicago.

Wright, B. D. & Douglas, G. A. (1977). Conditional versus unconditional procedures for sample-free item analysis. *Educational and Psychological Measurement*, **37**, 573–586.

Wright, B. D. & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement*, **29**, 28–37.

Received 25 September 1977; revised version received 30 May 1978

Requests for reprints should be addressed to Leslie Cohen, Faculty of Mathematical Studies, University of Southampton, Southampton SO9 5NH.