Computing Conditional Maximum Likelihood Estimates for Generalized Rasch Models Using Simple Loglinear Models with Diagonals Parameters

Author(s): Alan Agresti

Source: *Scandinavian Journal of Statistics*, 1993, Vol. 20, No. 1 (1993), pp. 63-71

Published by: Wiley on behalf of Board of the Foundation of the Scandinavian Journal of Statistics

Stable URL: https://www.jstor.org/stable/4616260

Scand J Statist 20: 63–71, 1993

Computing Conditional Maximum Likelihood Estimates for Generalized Rasch Models Using Simple Loglinear Models with Diagonals Parameters

ALAN AGRESTI
University of Florida

ABSTRACT. Generalized Rasch models for multiple-response items proposed by Andersen (1973) are related to quasi-symmetric loglinear models. The loglinear models are obtained by treating subject parameters in the Rasch models as random effects. Fitting the loglinear models yields estimates of item parameters in the generalized Rasch models that are also conditional maximum likelihood estimates when the subject effects are treated as fixed. For models that apply naturally when there are ordinal response categories, the related loglinear models are simple quasi-symmetric models having diagonals parameters. Our results generalize Tjur's (1982) observation about the connection between binary-response Rasch models and loglinear models.

Key words: latent class models, logit model, matched pairs, ordinal responses, quasi symmetry, square contingency tables

1. Introduction

Suppose n subjects respond to k items, such as k questions on an exam. Let ϕ_{ij} denote the probability that subject i answers item j correctly. The Rasch model (Rasch, 1961) is a simple logistic model

$$\log \{\phi_{ij}/(1-\phi_{ij})\} = \alpha_i - \beta_j$$

that describes how that probability depends on subject abilities $\{\alpha_i\}$ and item difficulties $\{\beta_j\}$. Normally, primary interest focuses on estimating $\{\beta_j\}$. Under the usual sampling assumption that the observations are independent Bernoulli random variables, the maximum likelihood (ML) estimates of $\{\beta_j\}$ are inconsistent as $n \to \infty$ because of the concomitant increase in the number of subject parameters (Andersen, 1980, p. 246).

One obtains consistency by estimating $\{\beta_j\}$ conditional on sufficient statistics for $\{\alpha_i\}$. For matched-pairs ($k=2$), this provides justification for McNemar's test (Cox, 1958). Tjur (1982) noted that, when k is not large, one can easily obtain the conditional estimates by fitting a certain loglinear model to the 2^k contingency table that cross-classifies the n subjects' responses on the k items. Inspection of the sufficient statistics for Tjur's model reveals that it is simply the quasi-symmetry model. For related discussion, see Darroch (1981), Fienberg (1981) and McCullagh (1982).

For multiple-response items, Andersen (1973, 1990, sect. 12.3) discussed multinomial logit generalizations of the Rasch model. We consider special cases of one of Andersen's models that are relevant for ordered categorical responses. We show that one can obtain conditional ML estimates for the generalized Rasch models by fitting parsimonious quasi-symmetry models. For $k=2$ items, the ordinal-response Rasch models relate to diagonals-parameter symmetry models introduced by Goodman (1979, 1985) and Agresti (1983).

2. An ordinal Rasch model

Let r denote the number of possible responses for each item, and assume the response scale is the same for each item. For subject i and item j, let ϕ_{hij} denote the probability of response h, $h = 1, \ldots, r$. Rasch (1961) and Andersen (1973, 1990) discussed models of form

$$\phi_{hij} = \frac{\exp(\alpha_{hi} - \beta_{hj})}{\sum_a \exp(\alpha_{ai} - \beta_{aj})}. \tag{1}$$

Suppose that the response scale is ordinal. For instance, each item may have a correct response, a response that receives partial credit, and an incorrect response. For each item we order the responses so that response r is the "best" response, and we let scores $v_1 \leqslant \ldots \leqslant v_r$ denote the credit assigned to the responses.

A special case of (1) for ordinal responses,

$$\phi_{hij} = \frac{\exp\{\lambda_h + v_h(\alpha_i - \beta_j)\}}{\sum_a \exp\{\lambda_a + v_a(\alpha_i - \beta_j)\}}, \tag{2}$$

has simple interpretations for the item and subject effects. For instance, when $\{v_{h+1} - v_h = 1$ for all $h\}$, the model has the adjacent-categories logit representation

$$\log(\phi_{h+1,ij}/\phi_{hij}) = \gamma_h + \alpha_i - \beta_j. \tag{3}$$

For each subject, the odds of making response $h+1$ instead of response h for item a are $\exp(\beta_b - \beta_a)$ times the odds for item b, and this holds uniformly in h. It follows that, for each subject, the response distributions for the various items are stochastically ordered according to $\{\beta_j\}$. Similarly, for each item, the odds of making response $h+1$ instead of h for subject a are $\exp(\alpha_a - \alpha_b)$ times the odds for subject b. Conditional on response in category h or $h+1$, the probability of the "better" response increases to 1.0 as $\alpha_i \to \infty$ or as $\beta_j \to -\infty$. See Agresti (1989, 1990) for motivation regarding the use of adjacent-categories logit models with effects independent of the cutpoints, for analyzing ordinal data.

Andersen (1973, 1980, pp. 272–274) discussed estimation of parameters for model (2). We now present a simple way to obtain conditional ML estimates of item parameters by ordinary ML fitting of a loglinear model. For a given subject with ability parameter α, let $y_{hj} = 1$ if the subject makes response h to item j, and let $y_{hj} = 0$ if the subject's response is other than h, so $\Sigma_h y_{hj} = 1$. Given α, we assume that responses on separate items by the same subject are independent. For model (2), the probability of a particular sequence of responses on the k items for that subject is then

$$\prod_j \frac{\prod_h \exp[y_{hj}\{\lambda_h + v_h(\alpha - \beta_j)\}]}{\sum_a \exp\{\lambda_a + v_a(\alpha - \beta_j)\}}. \tag{4}$$

The sufficient statistic for α is the subject's "total score" $s = \Sigma_j \Sigma_h y_{hj} v_h$. The conditional distribution of $\{y_{hj}\}$ for that subject, given s, equals

$$\frac{\exp\left\{\sum_h y_{h+}\lambda_h - \sum_j \left(\sum_h y_{hj} v_h\right)\beta_j\right\}}{\sum_{A_s} \exp\left\{\sum_h y^*_{h+}\lambda_h - \sum_j \left(\sum_h y^*_{hj} v_h\right)\beta_j\right\}},$$

where A_s denotes the set of possible data $\{y^*_{hj}\}$ having $\Sigma_j \Sigma_h y^*_{hj} v_h = s$, and where a + subscript denotes summation over that index. The conditional likelihood is the product of such terms over all n subjects. The sufficient statistic for β_j in that likelihood is $\Sigma_h v_h n_j(h)$, where $n_j(h)$ denotes the number of times that response h occurs for item j.

A likelihood related to this conditional likelihood results also from a random effects representation for the subjects. Let $F(\alpha)$ denote a (unknown) distribution for the subject effects. From (4), the marginal probability of a particular set of responses is

$$\exp\left\{\sum_h \lambda_h\left(\sum_j y_{hj}\right) - \sum_j \beta_j\left(\sum_h y_{hj}v_h\right)\right\}\int \frac{\exp\left\{\alpha\left(\sum_h\sum_j y_{hj}v_h\right)\right\}}{g(\alpha)}\,dF(\alpha),$$

where $g(\alpha)$ depends on other parameters but not on $\{y_{hj}\}$. Letting $\pi(h_1, \ldots, h_k)$ denote the marginal probability of response h_j on item j, $j = 1, \ldots, k$, we have

$$\pi(h_1, \ldots, h_k) = \exp\left(\sum_h \lambda_h t_h - \sum_j \beta_j v_{h_j}\right)\int \frac{\exp\left\{\alpha\left(\sum_j v_{h_j}\right)\right\}}{g(\alpha)}\,dF(\alpha), \tag{5}$$

where t_h denotes the number of $\{h_j\}$ that equal h. From (5), we see that apart from randomness for the sample size n, this distribution satisfies a Poisson loglinear model for expected frequencies $\{m(h_1, \ldots, h_k)\}$ in a r^k contingency table,

$$\log m(h_1, \ldots, h_k) = \mu + \sum_{h=1}^{r} \lambda_h t_h - \sum_{j=1}^{k} \beta_j v_{h_j} + \rho_s, \tag{6}$$

where $s = \Sigma_j v_{h_j}$.

Let $n(h_1, \ldots, h_k)$ denote the number of subjects making response h_j to item j, $j = 1, \ldots, k$, and let n_s denote the number of subjects having $\Sigma_j v_{h_j} = s$. One can express the Poisson likelihood for (6) as the product of the Poisson probability for n, a multinomial distribution (given n) for $\{n_s\}$ that is parametrized by probabilities $\{q_s\}$ for the possible total scores for a subject, and the conditional distribution of $\{n(h_1, \ldots, h_k)\}$ given those totals. The term involving $\{\beta_j\}$ and $\{\lambda_h\}$, which appears in the third part of that expression for the likelihood, is identical to the term involving those parameters in the conditional likelihood for the generalized Rasch model. It follows that ML estimates and second derivatives of the log likelihood with respect to those parameters for the loglinear model are identical to those for the conditional approach to fitting the generalized Rasch model. The derivation of the Poisson likelihood factorization is a direct extension of Tjur's (1982) presentation for the case $r = 2$.

The sufficient statistics for model (6) are

$$\sum_j n_j(h), \quad h = 1, \ldots, r$$

$$\sum_h v_h n_j(h), \quad j = 1, \ldots, k$$

$$n_s, s = kv_1, (k-1)v_1 + v_2, \ldots, kv_r.$$

The likelihood equations equate these to their fitted values. The sufficient statistics for the item parameters may be interpreted (when divided by n) as "mean" responses in the first-order margins of $\{n(h_1, \ldots, h_k)\}$. The third set of sufficient statistics are equivalent to those for the subject parameters. Identifiability of parameters requires constraints such as $\Sigma\,\lambda_h = \Sigma\,\beta_j = \Sigma\,\rho_s = 0$.

3. Other versions of the generalized Rasch model

Model (1), the most general version of the Rasch model for multiple response categories, is more appropriate than (2) when the response categories are nominal. For that model, integrating with respect to an unknown distribution for the subject effects yields a marginal distribution satisfying quasi symmetry; that is, $m(h_1, \ldots, h_k)$ has form

$$m(h_1, \ldots, h_k) = b_1(h_1) \ldots b_k(h_k)c(h_1, \ldots, h_k), \tag{7}$$

where $c(\cdot)$ is permutationally invariant. In addition, the conditional ML estimates of item effects in (1) can be obtained from ML estimates of main effects in the loglinear version of quasi-symmetry model (7). This is the result for r response categories that corresponds to Tjur's result for $r = 2$. See also Darroch & McCloud (1986) and Conaway (1989).

The ordinal model (2) we have discussed is a special case of the general Rasch model (1) that uses the ordinality of the response both in terms of describing item effects and subject effects. Other special cases of (1) exist that are less parsimonious than (2), retaining a general form for one of these types of effects. For instance, the model

$$\phi_{hij} = \frac{\exp(\alpha_{hi} - v_h \beta_j)}{\sum_a \exp(\alpha_{ai} - v_a \beta_j)} \tag{8}$$

retains simple interpretations for item effects. When $\{v_{h+1} - v_h = 1 \text{ for all } h\}$, for each subject the odds of making response $h + 1$ instead of response h for item a are $\exp(\beta_b - \beta_a)$ times the odds for item b. When primary interest focuses on the item effects, it is unnecessary to assume that subject effects are identical for the logit for each pair of adjacent response categories, and model (8) sometimes fits well when model (2) does not.

Using the same arguments given in the previous section, one can show that conditional ML estimates of item effects $\{\beta_j\}$ for model (8) are identical to the ML estimates for the corresponding parameters in the loglinear model

$$\log m(h_1, \ldots, h_k) = \mu + \sum_{h=1}^{r} \lambda_h t_h - \sum_{j=1}^{k} \beta_j v_{h_j} + \rho(h_1, \ldots, h_k), \tag{9}$$

where $\rho(\cdot)$ is permutationally invariant. This model is the special case of quasi symmetry (7) satisfying $b_j(h) = \exp(\lambda_h - \beta_j v_h)$. Model (6) is a further special case having a coarser partition of cells on which the ρ parameter relating to "equivalent" subjects is constant. Both models are equivalent to quasi symmetry when $r = 2$. The model of complete symmetry, for which $m(h_1, \ldots, h_k)$ is permutationally invariant, is the special case of (9) in which $\beta_1 = \ldots = \beta_k$. For reference purposes, we shall refer to the general model (1) as the *multinomial* Rasch model, the fully-ordinal model (2) as the *ordinal* Rasch model, and the partly-ordinal model (8) as the *ordinal item-effect* Rasch model.

The quasi-symmetry model is often interpreted as an adjustment of the complete symmetry model in which the marginal distributions are not required to be identical. The ordinal Rasch models imply a particular type of departure from symmetry in which the marginal heterogeneity has a simple structure reflecting the ordering of response categories. In terms of the sufficient statistics, this refers to heterogeneity in the k marginal means for the chosen scores $\{v_j\}$. Whatever the choice of monotone scores for the ordinal Rasch models, the k first-order marginal distributions of the cross classification of $\{m(h_1, \ldots, h_k)\}$ are stochastically ordered according to the $\{\beta_j\}$. This follows from the results that (1) for each subject, the item response distributions $\{\phi_{hij}, h = 1, \ldots, r\}$ for $j = 1, \ldots, k$ are stochastically ordered according to $\{\beta_j\}$, (2) the k first-order marginal distributions of $\{m(h_1, \ldots, h_k)\}$ are equivalent to the $r \times k$ two-way marginal table of $\{\phi_{hij}\}$ collapsed over subjects, and (3) the two-way subject-by-item marginal table of $\{\phi_{hij}\}$ has each cell entry equal to 1.

For each choice of $h_1 \leqslant h_2 \leqslant \ldots \leqslant h_k$, let $n_p(h_1, \ldots, h_k)$ denote the sum of all cell counts for cells having index given by a permutation of $(h_1, \ldots, h_k)$. These sums are the sufficient statistics for the final term in model (9). This is also the finest possible partition one can have for $\{n_s\}$ in model (6), corresponding to a choice of scores for which different sums of scores occur whenever different cells have indices that are not permutationally related. For such scores, the two models are equivalent.

The residual degrees of freedom for asymptotic chi-squared goodness-of-fit tests equal $r^k-(r-1)(k-1)-(r+k-1)!/\{(r-1)!k!\}$ for the quasi-symmetry model. They equal $r^k-(k-1)-(r+k-1)!/\{(r-1)!k!\}$ for model (9), just $k-1$ less than *df* for the complete symmetry model. When the scores are equal-interval, the simpler model (6) has residual degrees of freedom equal to $r^k-(k+1)r+2$. In practice, $\{n(h_1, \ldots, h_k)\}$ are often too sparse to conduct direct goodness-of-fit tests of models (6) and (9). However, one can compare their fits to that of the quasi-symmetry model, when the data support the latter model. An advantage of the simpler model is that when $n_s > 0$ for each s, but there are some $(h_1, \ldots, h_k)$ such that $n(h_1, \ldots, h_k)=0$ for all permutations of the indices, then ML estimates exist for model (6) but not for model (9) and the quasi-symmetry model.

4. Relation to models for square tables

Suppose there are only $k=2$ items. Letting $\beta=\beta_1=-\beta_2$, model (6) has form

$$\log m(a, b)=\mu+\lambda_a+\lambda_b+\beta(v_b-v_a)+\rho_{v_a+v_b}. \tag{10}$$

When the scores are equal-interval, this model is a special case of the double-diagonals-parameter model

$$m(a, b)=\alpha_a\alpha_b\,\delta_{a-b}\,\delta^*_{a+b}, \tag{11}$$

which relates to models discussed by Goodman (1985). Similarly, the more general model (9) has form

$$\log m(a, b)=\mu+\lambda_a+\lambda_b+\beta(v_b-v_a)+\rho(a, b), \tag{12}$$

where $\rho(a, b)=\rho(b, a)$ for all a and b. This special case is a model introduced by Agresti (1983), which for equal-interval scores has logit characterization

$$\log\{m(a, b)/m(b, a)\}=\delta(b-a). \tag{13}$$

Models (11) and (13) are special cases of a family of models satisfying

$$\log\{m(a, b)/m(b, a)\}=\delta_{a-b}, \tag{14}$$

discussed by Goodman (1979). Model (14) is called a *diagonals-parameter symmetry* model, and model (13) is called a *linear diagonals-parameter symmetry* model. Jørgensen (1985) used models with diagonals parameters indexed by total scores to describe inter-rater agreement.

For arbitrary k, models (6) and (9) may be regarded as diagonals-parameter models in k dimensions. For instance, for model (9) consider the log odds $\log\{m(a, b, h_3, \ldots, h_k)/m(b, a, h_3, \ldots, h_k)\}$. For unit-spaced scores, this equals $(\beta_1-\beta_2)(b-a)$ for all pairs of cells that are $b-a$ diagonals away from the main diagonal in the first two dimensions.

An alternative to the models of this article uses the same linear predictor as in models such as (3) but utilizes logits of *cumulative* probabilities. Such models have the advantage of invariance under groupings of response categories, but conditional ML does not apply because of the lack of sufficient statistics. McCullagh (1977) described weighted least squares estimates for the cumulative logit model, for the case $k=2$. Alternative forms of estimation include maximizing the marginal likelihood after assuming a particular parametric form $F(\alpha;\theta)$ for the subject effects or assuming a latent class structure, such as discussed for $r=2$ by Cressie & Holland (1983) and by Lindsay *et al.* (1991).

Table 1. *Opinions* about teenage sex, premarital sex, and extramarital sex, with fitted values for ordinal item-effect Rasch model in parentheses*

Teen sex	Premarital sex	Extramarital sex 1	2	3	4
1	1	140 (140)	1 (1.5)	0 (0.2)	0 (0.0)
	2	30 (30.3)	3 (2.8)	1 (0.4)	0 (0.0)
	3	66 (66.5)	4 (7.7)	2 (1.6)	0 (0.4)
	4	83 (83.0)	15 (15.5)	10 (7.0)	1 (2.0)
2	1	3 (2.2)	1 (0.2)	0 (0.0)	0 (0.0)
	2	3 (4.0)	1 (1)	1 (0.4)	0 (0.0)
	3	15 (11.1)	8 (8.0)	0 (0.8)	0 (0.2)
	4	23 (22.3)	8 (7.9)	7 (4.2)	0 (0.6)
3	1	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)
	2	0 (0.8)	0 (0.6)	0 (0.1)	0 (0.0)
	3	3 (3.4)	2 (1.1)	3 (3)	1 (0.4)
	4	13 (14.6)	4 (6.1)	6 (7.1)	0 (0.8)
4	1	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	2	0 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)
	3	0 (1.1)	0 (0.4)	1 (0.5)	0 (0.1)
	4	7 (6.0)	2 (1.3)	2 (1.1)	4 (4)

*Data from 1989 General Social Survey, with categories: 1, always wrong; 2, almost always wrong; 3, wrong only sometimes; 4, not wrong.

5. Example

Table 1 is taken from the 1989 General Social Survey, conducted by the National Opinion Research Center at the University of Chicago. Subjects in the sample were asked their opinion on (1) early teens (age 14–16) having sex relations before marriage, (2) a man and a woman having sex relations before marriage, (3) a married person having sexual relations with someone other than the marriage partner. The response scale was (always wrong, almost always wrong, wrong only sometimes, not wrong at all). We denote the classifications by T for teenage sex, P for premarital adult sex, and X for extramarital sex. Table 2 shows results of goodness-of-fit tests for fitting several loglinear models to the responses for the 475 subjects who responded to these items. Table 1 contains many empty cells and small counts, so the likelihood-ratio statistic (denoted by G^2) is useful mainly for comparing models. The quasi-symmetry model, which corresponds to the multinomial Rasch model, fits well ($G^2 = 20.8$, df $= 38$). Simpler models that utilize the ordering also fit reasonably well, and the reduction in quality of fit compared to the general multinomial model is compensated by ease of interpretation. Table 2 reports results for ordinal models using $\{v_{h+1} - v_h = 1\}$.

Inspection of residuals indicates that the loglinear model (6) corresponding to the ordinal Rasch model fits decently except in the cell for responses (2, 3, 2) for (T, P, X), which had an observed count of 8 and fitted value of 2.9. Loglinear model (9), which corresponds to the ordinal item-effects Rasch model, fits reasonably well in all cells ($G^2 = 27.8$, df $= 42$). Table 1 also displays fitted values for this model. These models necessarily have perfect fits in cells $(1, \ldots, 1)$ and $(r, \ldots, r)$, and model (9) and the quasi-symmetry model have perfect fits in all diagonal cells $(h, \ldots, h)$, by virtue of the likelihood equations for the final term in the models. Both ordinal Rasch models have very strong evidence of heterogeneous item effects. For model (9), for instance, the difference in likelihood-ratio statistics between the complete

Table 2. *Goodness of fit of loglinear models for Table 1*

Model	Likelihood-ratio statistic	Pearson statistic	Degrees of freedom
Mutual independence	231.3	416.2	54
Complete symmetry	637.3	606.7	44
Ordinal Rasch	42.5	43.1	50
Ordinal item-effect Rasch	27.8	24.6	42
Quasi symmetry	20.8	24.6	38

Table 3. *Conditional maximum likelihood estimates of item parameters*

	Ordinal item-effect Rasch model		Ordinal Rasch model	
Item	Estimate	Std error	Estimate	Std error
Premarital	2.626	0.288	2.595	0.215
Extramarital	−0.364	0.125	−0.343	0.119
Teenage	0	0	0	0

symmetry model and this model tests the hypothesis of homogeneous item effects (i.e. $\beta_1 = \ldots = \beta_k$). The improvement in fit (609.5, df = 2) compared to the complete symmetry model is quite dramatic, and a similar reduction (640.9, df = 2) occurs in comparing model (6) to a reduced model with homogeneous effects.

Table 3 presents conditional ML estimates of item effects for the ordinal Rasch models, as well as their estimated standard errors, using constraints that equate the estimate for T to 0. There is considerably more tolerance for premarital adult sex than for other types. Using the ordinal Rasch model (3), for instance, for a given subject and each h, $h = 1, 2, 3$, the estimated odds of response $h+1$ instead of h for premarital sex are $\exp(2.595 + 0.343) = 18.9$ times as high as for extramarital sex. We obtained similar substantive results for other possible choices for the response scores $\{v_h\}$. For instance, if we regard the scores as reflecting approximate distances between the response categories, it might make sense to use scores such as $\{1, 1.5, 3.0, 4\}$, treating the distance between "almost always wrong" and "wrong only sometimes" as greater than those between the other adjacent pairs. The ordinal Rasch model then has conditional ML item estimates $\{2.91, -0.33, 0\}$.

Unconditional ML estimates are obtained by fitting generalized Rasch models directly to the $3 \times 4 \times 475$ table that cross classifies item-by-response-by-subject. For the ordinal Rasch model (3), the estimated item effects are the same as those obtained for the $3 \times 4 \times 10$ table stratified by values of the sufficient statistic (the total score s) for the subject parameter. These estimates are $(5.24, -0.62, 0)$, reflecting severe upward bias of unconditional estimates when the number of parameters has the same order as the sample size. When $k = 2$ and $r = 2$, the unconditional estimate is twice the conditional estimate (Andersen, 1973), which equals $\log\{n(1, 2)/n(2, 1)\}$.

Models discussed in this article have subject-specific item effects. An alternative model uses a simpler linear predictor deleting the subject effects. For instance, instead of (3) one uses the marginal model

$$\log(\phi_{h+1,+,j}/\phi_{h,+,j}) = \gamma_h - \beta_j. \tag{15}$$

Table 4. *GLIM code for fitting ordinal Rasch models to Table 1*

```
$units 64
$data count $read
140  1 0 0  3 1 0 0  1 0 0 0  0 0 0 0
30  3 1 0  3 1 1 0  0 0 0 0  0 0 0 0
66  4 2 0  15  8 0 0  3 2 3 1  0 0 1 0
83  15 10 1  23 8 7 0  13 4 6 0  7 2 2 4
$calc teen = %gl(4,4): xmar = %gl(4,1): pre = %gl(4,16) $
$calc t = teen: x = xmar: p = pre $
$calc la1 = %eq(t,1) + %eq(x,1) + %eq(p,1) $
$calc la2 = %eq(t,2) + %eq(x,2) + %eq(p,2) $
$calc la3 = %eq(t,3) + %eq(x,3) + %eq(p,3) $
$calc score = p + x + t - 2 $
$ass symm = 1,2,3,4,2,5,6,7,3,6,8,9,4,7,9,10,2,5,6,7,5,11,12,13,6,12,14,
15,7,13,15,16,3,6,8,9,6,12,14,15,8,14,17,18,9,15,18,19,4,7,9,10,7,13,15,
16,9,15,18,19,10,16,19,20
$fac teen 4 xmar 4 pre 4 symm 20 score 10 $
$yvar count $err pois
$fit symm: + p + x + t$! Fits symmetry and model (9)
$fit symm + pre + xmar + teen$! Fits quasi symmetry
$fit la1 + la2 + la3 + score + p + x + t$! Fits model (6)
```

Effects are then "population-averaged"; that is, the response odds refer to a randomly selected subject answering item a and another randomly selected subject answering item b (Agresti, 1989, 1990, sect. 11.4.3). For Table 1, the estimates of the population-averaged item effects for model (15) are (1.043, −0.195, 0), with estimated standard errors of (0.062, 0.063, 0). The estimates were obtained by using the methods of Aitchison & Silvey (1958) to maximize a multinomial likelihood for the 64 cells in Table 1, subject to the constraint that the marginal probabilities satisfy (15). These results illustrate the wide discrepancies that can occur among conditional estimates of subject-specific item effects, unconditional estimates of subject-specific item effects, and estimates of population-averaged item effects.

When k is not too large, loglinear models corresponding to ordinal Rasch models are easy to fit with most software that can handle loglinear models. To illustrate, Table 4 contains code for using GLIM to fit the models to Table 1. The factors denoted by "score" and "symm" generate loglinear terms for which the sufficient statistics correspond to those for the subject parameters in the Rasch models. The coefficients of the terms denoted by "t", "x", and "p" are the ordinal item parameters.

Acknowledgements

This research was partially supported by a grant from the National Institutes of Health. The author is appreciative of computational help from Joseph Lang in obtaining the data and calculating population-averaged estimates, and of comments from a referee that led to consideration of model (8).

References

Agresti, A. (1983). A simple diagonal-parameter symmetry and quasi-symmetry model. *Statist. Probab. Lett.* **1,** 313–316.

Agresti, A. (1989). A survey of models for repeated ordered categorical response data. *Statist. Medic.* **8,** 1209–1224.

Agresti, A. (1990). *Categorical data analysis*. Wiley, New York.

Aitchison, J. & Silvey, S. D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.* **29**, 813–828.

Andersen, E. B. (1973). Conditional inference for multiple-choice questionnaires. *Brit. J. Math. Statist. Psychol.* **26**, 31–44.

Andersen, E. B. (1980). *Discrete statistical models with social science application*. North-Holland, Amsterdam.

Andersen, E. B. (1990). *The statistical analysis of categorical data*. Springer, Berlin.

Conaway, M. (1989). Analysis of repeated categorical measurements with conditional likelihood methods. *J. Amer. Statist. Assoc.* **84**, 53–62.

Cox, D. R. (1958). Two further applications of a model for binary regression. *Biometrika* **45**, 562–565.

Cressie, N. & Holland, P. W. (1983). Characterizing the manifest probabilities of latent trait models. *Psychometrika* **48**, 129–141.

Darroch, J. N. (1981). The Mantel–Haenszel test and tests of marginal symmetry; fixed-effects and mixed models for a categorical response. *Intern. Statist. Rev.* **49**, 285–307.

Darroch, J. N. & McCloud, P. I. (1986). Category distinguishability and observer agreement. *Austral. J. Statist.* **28**, 371–388.

Fienberg, S. E. (1981). Recent advances in theory and methods for the analysis of categorical data: making the link to statistical practice. *Bull. Int. Statist. Inst.* **49** (book 2), 763–791.

Goodman, L. A. (1979). Multiplicative models for square contingency tables with ordered categories. *Biometrika* **66**, 413–418.

Goodman, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Ann. Statist.* **13**, 10–69.

Jørgensen, B. (1985). Estimation of inter-observer variation for ordinal rating scales. *Lecture Notes in Statistics: Generalized Linear Models* **32**, 93–104.

Lindsay, B., Clogg, C. C. & Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *J. Amer. Statist. Assoc.* **86**, 96–107.

McCullagh, P. (1977). A logistic model for paired comparisons with ordered categorical data. *Biometrika* **64**, 449–453.

McCullagh, P. (1982). Some applications of quasisymmetry. *Biometrika* **69**, 303–308.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the 4th Berkeley Symposium on Mathematics and Statistical Probability*, Vol. 4 (ed. J. Neyman), 321–333. University of California Press, Berkeley.

Tjur, T. (1982). A connection between Rasch's item analysis model and a multiplicative Poisson model. *Scand. J. Statist.* **9**, 23–30.

Received September 1991, in final form March 1992

Alan Agresti, Department of Statistics, University of Florida, Gainesville, FL 32611, USA.