Logit Models with Random Effects and Quasi-Symmetric Loglinear Models

Alan Agresti

ABSTRACT:
Logit models allowing subject heterogeneity, such as Rasch-type models with random effects, are useful for comparing distributions of repeated categorical responses. Some simple models of this type imply a quasi-symmetric loglinear model for the repeated responses, regardless of the distribution of the random effect. In many cases, the estimates for this loglinear model are also conditional maximum likelihood estimates for a fixed effects treatment of the logit model. This paper surveys connections among these various approaches and presents an extension for repeated measurement of a multivariate vector of binary responses.

KEYWORDS:
Binary data; conditional maximum likelihood; Item response model; Nonparametric; Ordinal response; Rasch model; Repeated measurement.

1 Introduction

This article discusses models for repeated responses of subjects to a set of similar categorical items. We illustrate for Table 1, presented by Coleman (1964), from a study that interviewed a sample of schoolboys twice, several months apart. The boys were asked about their self-perceived membership in the "leading crowd" (yes, no) and about whether one must sometimes go against his principles in order to be part of that leading crowd (agree, disagree). The table summarizes responses on the two variables (membership in the leading crowd, attitude toward the leading crowd) at two interview times.

We discuss logit models with random effects for subjects that focus on comparing the repeated response distributions, simultaneously for each variable. In Table 1, for instance, the logit model describes subject-specific changes in membership and changes in attitude between the two interview times. We survey a variety of such models, for binary responses, nominal or ordinal responses, or multivariate binary responses. For each model, effect parameters relate to main effects in certain loglinear models, called *quasi-symmetry* models.

Section 2 introduces a multivariate logit model for repeated responses and derives a loglinear model implied by a nonparametric treatment of random effects. Section 3 discusses simpler random effects structures and analyzes Table 1. Section 4 presents extensions for multiple-category responses. That section

(M, A) for First Interview	(M, A) for Second Interview (Yes, Agree)	(Yes, Disagree)	(No, Agree)	(No, Disagree)
Yes Agree	458 (458)	140 (141.8)	110 (119.5)	49 (49.1)
Yes Disagree	171 (169.2)	182 (182)	56 (58.6)	87 (74.8)
No Agree	184 (174.5)	75 (71.7)	531 (531)	281 (282.3)
No Disagree	85 (85.6)	97 (109.2)	338 (336.7)	554 (554)

TABLE 1. Membership (M) and Attitude (A) toward the "Leading Crowd" for Boys, with Fitted Values for Multivariate Quasi-Symmetry Model.

also surveys literature about connections among the parametric and nonparametric random effects logit model, conditional maximum likelihood estimates for a fixed effects version of the logit model, and ordinary estimates for quasi-symmetric loglinear models.

2 A Multivariate Logit Model with Repeated Measurement

Suppose n subjects respond to T items having the same binary scale. For subject s and item t, let Y_{st} denote the response category. The *Rasch model* is

$$\text{logit}[P(Y_{st} = 1)] = \alpha_s + \beta_t, \; s = 1, ..., n, \; t = 1, ..., T.$$

This is the most basic of the *item response models* commonly used in educational testing (Rasch 1961) for the probability that student s makes a correct response on question t. The $\{\alpha_s\}$ permit heterogeneity among students. The two common approaches to estimating $\{\beta_t\}$ eliminate $\{\alpha_s\}$ (i) using conditional maximum likelihood, conditioning on sufficient statistics for $\{\alpha_s\}$ (Rasch 1961), (ii) using a random effects approach, assuming a particular form of distribution for $\{\alpha_s\}$ (Bock and Aitkin 1981).

This section presents a multivariate extension of the Rasch model and shows its connection to loglinear models. It refers to I separate binary variables, each measured for T items ($I = T = 2$ in Table 1, with items being the interview times). For subject s, denote the response under item t for variable i by Y_{sit}, with observed value 1 or 0. We consider the model

$$\text{logit}[P(Y_{sit} = 1)] = \alpha_{is} + \beta_{it} \qquad (1)$$

For each variable i, this model has the form of the Rasch model, assuming a lack of subject-by-item interaction. Given the model parameters, we treat the observations as independent Bernoulli variates. Identifiability requires a constraint such as $\beta_{i1} = 0$ for each variable. The $\{\beta_{i1}, ..., \beta_{iT}\}$ for each i describe the item effects for each variable. The $\{\alpha_{is}\}$ parameters reflect the heterogeneity among subjects that induces the correlations among repeated responses on a variable.

For subject s, the probability of a particular sequence of responses $\mathbf{y} = (y_{11}, ..., y_{IT})$ for the IT variable–item combinations equals

$$\Pi_i \Pi_t \left(\frac{e^{\alpha_{is}+\beta_{it}}}{1+e^{\alpha_{is}+\beta_{it}}} \right)^{y_{it}} \left(\frac{1}{1+e^{\alpha_{is}+\beta_{it}}} \right)^{1-y_{it}} = \frac{\exp[\sum_i \alpha_{is}(\sum_t y_{it}) + \sum_i \sum_t \beta_{it} y_{it}]}{\Pi_i \Pi_t [1+\exp(\alpha_{is}+\beta_{it})]}.$$

Let $\alpha_s = (\alpha_{1s}, ..., \alpha_{Is})$. We treat this vector as a random effect, permitting correlated components. In Table 1, for instance, subjects having a relatively high random effect for the membership variable, thus having a propensity to be members regardless of the interview time, probably tend to have a relatively high random effect for the attitude variable.

Suppose that $\alpha_1, ..., \alpha_n$ are independent with cumulative distribution function F. Denote the marginal probability of responses $\mathbf{y}$, averaged over the subjects, by $\pi(\mathbf{y})$. For model (1), the marginal probability equals

$$\pi(\mathbf{y}) = \exp(\sum_i \sum_t \beta_{it} y_{it}) \int_{\alpha_s} \frac{\exp[\sum_i \alpha_{is}(\sum_t y_{it})]}{\Pi_i \Pi_t [1+\exp(\alpha_{is}+\beta_{it})]} dF(\alpha_{1s}, ..., \alpha_{Is})$$

Regardless of F, the integral determining this marginal probability yields a complex function of $\{\beta_{it}\}$. Note, however, that that function depends on the data only through the values of $(\sum_t y_{1t}, ..., \sum_t y_{It})$. Thus, the model for the marginal probability is a special case of one that provides a separate parameter for each possible value of that vector of sums. This more general marginal model has form

$$\pi(\mathbf{y}) = \exp(\sum_i \sum_t \beta_{it} y_{it}) \gamma(\sum_t y_{1t}, ..., \sum_t y_{It}),$$

where γ is an unspecified positive parameter that can assume a different value for each combination of the arguments.

The sample of n observations on the binary responses $\mathbf{y}$ for the IT variable–item combinations form a multinomial sample with probabilities $\{\pi(\mathbf{y})\}$. The form just derived that these probabilities satisfy is a loglinear model for expected frequencies $\{\mu(\mathbf{y})\}$ in a 2^{IT} contingency table that cross classifies the responses for the IT variable–item combinations. That model has form

$$\log[\mu(\mathbf{y})] = \sum_i \sum_t \beta_{it} y_{it} + \lambda(\sum_t y_{1t}, ..., \sum_t y_{It}). \qquad (2)$$

For this model, the interaction involving any set of items for a particular variable has term that is invariant for any permutation of the response outcomes for those items.

No matter what form the random effects distribution F takes, the implied marginal model has the same main effects structure, and it has an interaction term that is a special case of the one in (2). Thus, one can consistently estimate $\{\beta_{it}\}$ in a nonparametric manner using the ordinary ML estimates for the loglinear model. One can fit that model with standard software for generalized linear models. The usual goodness-of-fit statistics have large-sample chi-squared distributions with $df = 2^{IT} - [I(T-1) + (T+1)^I]$. For the univariate case, model (2) is the *quasi-symmetry model* (Caussinus 1966). It adds an interaction term to the mutual independence model that is invariant to permutations of the indices. We refer to (2) as a *multivariate quasi-symmetry model*.

In the matched-pairs case ($T = 2$), model (2) has fitted values in the 2×2 marginal table for each variable that are identical to the observed counts. The estimate of $\exp(\beta_{i2} - \beta_{i1})$ then equals the number of cases with $(y_{i1}, y_{i2}) = (0, 1)$ divided by the number of cases with $(y_{i1}, y_{i2}) = (1, 0)$. In the univariate case ($I = 1$), this is also the conditional ML estimate for the logit model, and Neuhaus et al. (1994) showed that it is also normally the estimate for a parametric random effects approach.

For the fixed effects approach with model (1), the conditional likelihood factors into a product of I terms, one for each variable. It follows that the conditional ML estimates of $\{\beta_{it}\}$ are identical to those obtained using conditional ML separately with the data for each variable. Using the approach of Tjur (1982), one can show that the conditional ML estimates are identical to the ordinary ML estimates of $\{\beta_{it}\}$ obtained by fitting the loglinear model (2).

3 Rasch-Type Models for "Leading-Crowd" Example

For Table 1, loglinear model (2) fits fairly well. The goodness-of-fit statistics are $G^2 = 4.92$ for the likelihood-ratio statistic (deviance) and $X^2 = 4.95$ for the Pearson statistic, with $df = 5$. Table 1 also displays the fitted values. The ML estimates of the item effects are $\hat{\beta}_{A1} - \hat{\beta}_{A2} = .176$ (ase = .058) for attitude and $\hat{\beta}_{M1} - \hat{\beta}_{M2} = .379$ (ase = .075) for membership. For instance, for each subject, the estimated odds of membership in the leading crowd at the first interview equal exp(.379) = 1.46 times the estimated odds of membership at the second interview.

Goodman (1974) and Haber (1985) presented alternative models for these data. Goodman used a latent class model with four latent classes that cross classify two associated binary latent variables, one of which affects the membership responses and one of which affects the attitude responses. Haber (1985) fitted a model that assumes solely that the marginal odds ratio between attitude and membership is identical for each interview. The sample odds ratios are 1.53 and 1.71, and Haber's model yielded fitted odds ratios of 1.62.

The fit of model (2) also suggests that these marginal odds ratios are similar, as the fitted odds ratios equal 1.63 and 1.61. Using the methodology described by Lang and Agresti (1994) for simultaneous fitting of generalized loglinear models to joint and marginal distributions of contingency tables, we fitted the

Model	G^2	X^2	*df*
a. Mutual Independence	1421.7	1572.6	11
b. 4-item Quasi Symmetry	616.6	680.3	8
c. Indep. Random Effects	97.5	96.8	9
d. Multiv. Symmetry	40.3	40.0	7
e. Multiv. Quasi Symmetry	4.9	5.0	5
f. (e) + Common Odds Ratio	5.3	5.4	6

NOTE: Models result from logit model (1) with (a) degenerate random effects, (b) perfectly correlated random effects, (c) independent random effects, (d) identical item effects for each variable, (e) unspecified distribution of random effects, and (f) case (e) with identical odds ratio between variables at each time.

TABLE 2. Summary of Loglinear Model Fits to Table 1.

simpler version of model (2) that constrains these marginal odds ratios to be identical. The fit, also shown in Table 1, has $G^2 = 5.31$ and $X^2 = 5.41$ with $df = 6$. The fitted common odds ratio equals 1.62, with $\hat{\beta}_{A1} - \hat{\beta}_{A2} = .176$ (ase = .058) and $\hat{\beta}_{M1} - \hat{\beta}_{M2} = .378$ (ase = .075). In summary, this analysis describes Table 1 using three parameters: One parameter compares the attitude responses at the two interviews, estimated by an odds ratio of exp(.176) = 1.19; a second parameter compares the membership responses at the two interviews, estimated by an odds ratio of exp(.378) = 1.46; and a third parameter describes the association between the attitude and membership responses at each interview, estimated by an odds ratio of 1.62.

Four special cases of logit model (1) relate to loglinear models that are special cases of model (2). First, suppose the random effects distribution is degenerate, with variance equal to zero for each component. Then, the marginal model is the special case of (2) without the interaction term, which is the model of mutual independence among the responses for all the variable–item combinations. Second, suppose that the components of $\boldsymbol{\alpha}_s = (\alpha_{1s}, ..., \alpha_{Is})$ are mutually independent. Then, the marginal probability satisfies the loglinear model

$$\log[\mu(\mathbf{y})] = \sum_i \sum_t \beta_{it} y_{it} + \sum_i \lambda_i (\sum_t y_{it}).$$

For this model, responses on variable a for any item t_a and on a different variable b for any item t_b are independent, both marginally and also conditionally on other responses.

Third, suppose that the components of $\boldsymbol{\alpha}_s = (\alpha_{1s}, ..., \alpha_{Is})$ are perfectly positively correlated. Then, the marginal probability satisfies

$$\log[\mu(\mathbf{y})] = \sum_i \sum_t \beta_{it} y_{it} + \lambda(\sum_i \sum_t y_{it}).$$

In this case, the logit model (1) treats all the variable–item combinations symmetrically and is identical to the Rasch model applied to the IT separate responses. The loglinear model is the quasi-symmetry model for the 2^{IT} contingency table. Finally, suppose $\{\beta_j\}$ in the logit model (1) are identical. Then, so are they identical in the loglinear model. Model (2) then exhibits within-variable symmetry. Each cell having the same value of $(\sum_t y_{1t}, ..., \sum_t y_{It})$ has the same probability.

Each of these four simpler models is typically too simplistic to fit well. Table 2 summarizes their fits to Table 1. All of them fit poorly.

4 Survey of Connections Between Univariate Rasch-Type Models and Loglinear Models

In the univariate case of repeated binary responses on a single variable, Tjur (1982), Kelderman (1984), Hatzinger (1989) and others have discussed connections between the Rasch model and loglinear models. Tjur (1982) showed the equivalence between conditional ML estimates for the Rasch model and ordinary ML estimates for the quasi-symmetry model. He also showed that those estimates result from a slightly extended version of the likelihood obtained with the nonparametric random effects approach. Later work showed strong connections between the actual nonparametric ML estimates and conditional ML estimates for the Rasch model (de Leeuw and Verhelst 1986, Lindsay *et al.* 1991).

Logit model (1) generalizes to incorporate a group factor or to handle multiple-category responses. For simplicity, we concentrate on the multiple-category case with a single variable for a single group. For r unordered response categories, let Y_{st} denote the category outcome for subject s on item t. Rasch (1961) proposed the model

$$\log[P(Y_{st} = k)/P(Y_{st} = r)] = \alpha_{ks} + \beta_{kt}, \ k = 1, ..., r-1,$$

which maintains additivity of item and subject effects for each category. For this model, the conditional ML estimates of the item effects are identical to estimates of main effects in the quasi-symmetry model for expected frequencies $\{\mu_{ab...t}\}$ in a r^T contingency table (Conaway 1989),

$$\log[\mu_{ab...t}] = \beta_{a1} + \beta_{b2} + ... + \beta_{tT} + \lambda_{ab...t},$$

where the interaction term is symmetric in its indices.

The *complete symmetry* model is the special case in which the main effect terms are identical; that is, the response probability is identical for any permutation of $(a, b, ..., t)$. When the quasi-symmetry model holds, complete symmetry is equivalent to marginal homogeneity. The standard test of marginal homogeneity is based on comparing the fits of the quasi-symmetry and complete symmetry models (Caussinus 1966).

For the ordinal-response case, one generalization of the Rasch model has the adjacent-categories logit representation

$$\log[P(Y_{st} = k+1)/P(Y_{st} = k)] = \alpha_{ks} + \beta_t.$$

This is a special case of the nominal-scale model in which the item effects have the ordinal structure $\beta_{k+1,t} - \beta_{kt} = \beta_t$ for all k; that is, $\{\beta_{kt}\}$ are linear in k. Generalizing Tjur (1982), Agresti (1993) noted that one can estimate $\{\beta_t\}$ in a random effects version of this model using the ordinary ML estimates for the loglinear model

$$\log[\mu_{ab\ldots t}] = a\beta_1 + b\beta_2 + \ldots + t\beta_T + \lambda_{ab\ldots t}\ ,$$

where λ is permutationally invariant. Moreover, these estimates are also conditional ML estimates for the logit model. The loglinear model is a special case of the quasi-symmetry model with linear structure for the main effects. It treats the main effects as variates, with equally-spaced scores, rather than qualitative factors. Each main effect term has a single parameter, rather than the $r-1$ parameters in the more general model.

The complete symmetry model is the special case $\beta_1 = \ldots = \beta_T$. One can test marginal homogeneity using a likelihood-ratio test with $df = T-1$, based on comparing its fit to that of complete symmetry. The ML estimates of $\{\beta_t\}$ have the same order as the sample mean responses (for equally-spaced scores) in the T one-way margins of the r^T table. See Agresti (1993, 1995) and Hatzinger (1994) for examples of the use of GLIM and Agresti (1996, p. 277) for the use of SAS for fitting such models. When $T = 2$, letting $\beta = \beta_2 - \beta_1$, this loglinear model is equivalent to the logit model

$$\log(\pi_{ab}/\pi_{ba}) = \beta(b-a).$$

One can also estimate β using software for logistic regression models, treating $\{n_{ab},\ a < b\}$ as independent binomial variates with sample sizes $\{n_{ab} + n_{ba}\}$.

An alternative model form for ordinal responses uses cumulative logits,

$$\text{logit}[P(Y_{st} \le k)] = \alpha_{ks} - \beta_t.$$

This model has the proportional odds property, for which the T item effects $\{\beta_t\}$ are identical at each k. McCullagh (1977) discussed a related model for $T = 2$. Conditional ML is not available for this model, but Hedeker and Gibbons (1994) presented a random effects approach for a simpler form of the subject term.

For $T = 2$, Agresti and Lang (1993) eliminated the subject parameters by noting a corresponding model for the $r \times r$ marginal table, namely

$$\log\left(\frac{\sum_{a'>a}\sum_{b'\le b}\pi_{a'b'}}{\sum_{a'\le a}\sum_{b'>b}\pi_{a'b'}}\right) + \log\left(\frac{\sum_{a'>b}\sum_{b'\le a}\pi_{a'b'}}{\sum_{a'\le b}\sum_{b'>a}\pi_{a'b'}}\right) = 2(\beta_1 - \beta_2)$$

for all $a \le b$. One can estimate the difference in item parameters by maximizing the likelihood for the $r \times r$ observed table, subject to these constraints for all

combinations of $a \le b$. The special case with no item effect (*i.e.*, constraining the sum of log odds to equal 0 for all $a \le b$) is an alternative characterization of symmetry.

Agresti and Lang (1993) extended this analysis to T items. The general case corresponds to a Rasch model for all $r-1$ binary collapsings of the response, with the same item effects for each collapsing. One can estimate item parameters by fitting a quasi-symmetry model simultaneously to all such collapsings, using the same main effect parameters for each. For $T = 2$, a simple estimate of $\beta = \beta_2 - \beta_1$ uses the fact that for each collapsing, the conditional ML estimate is the log of the ratio of off-main-diagonal counts. A nearly efficient estimator combines these, adding the numerators and adding the denominators before taking their ratio and their logarithm. In terms of the cell counts $\{n_{ij}\}$ in the full $r \times r$ table, the resulting estimate is

$$\tilde{\beta} = \log\{[\sum_{i<j}(j-i)n_{ij}]/[\sum_{i>j}(i-j)n_{ij}]\}.$$

The estimated asymptotic variance of this estimator equals

$$\hat{V}(\tilde{\beta}) = \frac{\sum_{i<j}(j-i)^2 n_{ij}}{[\sum_{i<j}(j-i)n_{ij}]^2} + \frac{\sum_{i>j}(i-j)^2 n_{ij}}{[\sum_{i>j}(i-j)n_{ij}]^2}.$$

See Agresti (1995) for a more detailed discussion of these models, Tutz (1990) for a continuation-ratio ordinal model, and Ten Have and Becker (1995) for a wide variety of loglinear models with quasi-symmetric structure. Hatzinger (1994) provided a more complete overview of the considerable literature about item response models and their connections with GLMs.

5 Comments and Conclusions

For the general multivariate model (1), it would be interesting to analyze whether results for a particular parametric formulation of the random effects vector tend to agree with those for the nonparametric formulation. Assuming parametric structure for the random effects distribution raises other questions, of course. If the specification is correct, do the nonparametric estimates suffer a substantive efficiency loss? If the specification is incorrect, could this introduce much bias? Previous work (e.g., Heckman and Singer, 1984) in a somewhat different context has shown that results may depend on the choice, and this is an advantage of the nonparametric approach (Aitkin 1995).

In particular, under the assumption that logit model (1) holds, loglinear model (2) is valid and provides consistent estimates regardless of the true distribution for the random effects. Thus, an informal diagnostic for the parametric approach is to compare estimates under various distributional assumptions to the nonparametric-based loglinear estimates; substantial deviations from the nonparametric estimates indicates a possibly inappropriate choice.

Acknowledgments: This work was partially supported by a grant from the National Institutes of Health.

References

Agresti, A. (1993). Computing Conditional Maximum Likelihood Estimates for Generalized Rasch Models using Simple Loglinear Models with Diagonals Parameters. *Scandinavian Journal of Statistics*, **20**, 63-72.

Agresti, A. (1995). Logit Models and Related Quasi-symmetric Loglinear Models for Comparing Responses to Similar Items in a Survey. *Sociological Methods in Research.* **24**, 68-95.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis.* New York: Wiley.

Agresti, A., and Lang, J. B. (1993). A Proportional Odds Model with Subject-specific Effects for Repeated Ordered Categorical Responses. *Biometrika*, **80**, 527-534.

Aitkin, M. (1995), A General Maximum Likelihood Analysis of Overdispersion in Generalized Linear Models. Unpublished manuscript.

Bock, R. D., and Aitkin, M. (1981). Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm. *Psychometrika*, **46**, 443-59.

Caussinus, H. (1966). Contribution a l'Analyse Statistique des Tableaux de Correlation. *Annales de la Faculte des Sciences de l'Université de Toulouse*, **29**, 77-182.

Coleman, J. S. (1964), *Introduction to Mathematical Sociology.* London: Free Press of Glencoe.

Conaway, M. (1989). Analysis of Repeated Categorical Measurements with Conditional Likelihood Methods. *Journal of the American Statistical Association*, **84**, 53-62.

De Leeuw, J., and Verhelst, N. (1986). Maximum Likelihood Estimation in Generalized Rasch Models. *Journal of Educational Statistics*, **11**, 183-196.

Goodman, L. A. (1974), Exploratory Latent Structure Analysis Using both Identifiable and Unidentifiable Models. *Biometrika*, **61**, 215-231.

Haber, M. (1985), Maximum Likelihood Methods for Linear and Log-linear Models in Categorical Data. *Computational Statistics and Data Analysis*, **3**, 1-10.

Hatzinger, R. (1989). The Rasch Model, some Extensions and their Relation to the Class of Generalized Linear Models. *Statistical Modelling: Proceedings of GLIM89 and the 4th International Workshop on Statistical Modelling.* Lecture Notes in Statistics, **57**, Berlin: Springer.

Hatzinger, R. (1994). A GLM Framework for Item Response Theory Models. Unpublished manuscript.

Heckman, J. J., and Singer, B. (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models of Duration. *Econometrica*, **52**, 271-320.

Hedeker, D., and Gibbons, R. D. (1994). A Random-Effects Ordinal Regression Model for Multilevel Analysis. *Biometrics*, **50**, 933-944.

Kelderman, H. (1984). Loglinear Rasch Model tests. *Psychometrika.* **49**, 223-245.

Lang, J. B., and Agresti, A. (1994). Simultaneously Modeling Joint and Marginal Distributions of Multivariate Categorical Responses. *Journal of the American Statistical Association*, **89**, 625-632.

Lindsay, B., Clogg, C. C., and Grego, J. (1991). Semiparametric Estimation in the Rasch Model and Related Exponential Response Models, Including a Simple Latent Class Model for Item Analysis. *Journal of the American Statistical Association*, **86**, 96-107.

McCullagh, P. (1977). A Logistic Model for Paired Comparisons with Ordered Categorical Data. *Biometrika*, **64**, 449-453.

Neuhaus, J. M., Kalbfleisch, J. D., and Hauck, W. W. (1994), Conditions for Consistent Estimation in Mixed-Effects Models for Binary Matched-Pairs Data. *Canadian Journal of Statistics*, **22**, 139-148.

Rasch, G. (1961). On General Laws and the Meaning of Measurement in Psychology. pp. 321-333 in *Proc. 4th Berkeley Symposum of Mathematical Statistics and Probability*, vol. 4, ed. J. Neyman. Berkeley: Univ. of California Press.

Ten Have, T. R., and Becker, M. P. (1995). Multivariate Contingency Tables and the Analysis of Exchangeability. *Biometrics*, **51**, 1001-1016.

Tjur, T. (1982). A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model. *Scandinavian Journal of Statistics*, **9**, 23-30.

Tutz, G. (1990). Sequential Item Response Models with an Ordered Response. *British Journal of Mathematical and Statistical Psychology*, **43**, 39-55.