STATISTICS IN MEDICINE
Statist. Med. **18**, 2191–2207 (1999)

MODELLING ORDERED CATEGORICAL DATA: RECENT ADVANCES AND FUTURE CHALLENGES

ALAN AGRESTI*

Department of Statistics, University of Florida, Gainesville, Florida 32611-8545, U.S.A.

SUMMARY

This article summarizes recent advances in the modelling of ordered categorical (ordinal) response variables. We begin by reviewing some models for ordinal data introduced in the literature in the past 25 years. We then survey recent extensions of these models and related methodology for special types of applications, such as for repeated measurement and other forms of clustering. We also survey other aspects of ordinal modelling, such as small-sample analyses, power and sample size considerations, and availability of software. Throughout, we suggest problem areas for future research and we highlight challenges for statisticians who deal with ordinal data. Copyright © 1999 John Wiley & Sons, Ltd.

1. INTRODUCTION

This article surveys recent advances in the modelling of ordered categorical (ordinal) response variables. Of the ordinal models introduced in the past 25 years, logit models for cumulative probabilities have been the most popular for applications in medical statistics. Other models that have received attention in the statistics literature include other forms of logit models for multinomial responses and log-linear models. We survey recent extensions of these models. The area of most intense attention has been the modelling of repeated measurement data.

We review recent literature on models for repeated measurement as well as a variety of other topics, such as exact methods for small samples, power and sample size considerations, and the availability of software. With the continuing development of more complex models, an increasingly important but difficult task is communicating to non-statisticians (and to applied statisticians who are not specialists in categorical data analysis) the interpretation of the models and their parameters.

2. MODELS FOR ORDERED CATEGORICAL RESPONSES

Logistic regression models occupy a central place in medical statistics. Not surprisingly, the most popular models for ordinal responses are multi-category generalizations of logistic regression. Here we provide only a brief summary of models described in greater detail in other places.[1–4]

* Correspondence to: Alan Agresti, Department of Statistics, University of Florida, Gainesville, Florida 32611-8545, U.S.A. E-mail: aa@stat.ufl.edu

Contract/grant sponsor: NIH

CCC 0277–6715/99/172191–17$17.50

2.1. Logit Models for Ordinal Responses

Currently, the most popular model for ordinal responses uses logits of cumulative probabilities,[5,6] often called *cumulative logits*. For a c-category response variable Y and a set of predictors $\mathbf{x}$ with corresponding effect parameters β, the model has form

$$\text{logit}[P(Y \leqslant j)] = \alpha_j - \boldsymbol{\beta}'\mathbf{x}, \quad j = 1, \ldots, c-1. \tag{1}$$

(The minus sign in the predictor term makes the sign of each component of $\boldsymbol{\beta}$ have the usual interpretation in terms of whether the effect is positive or negative.) This model applies simultaneously to all $c-1$ cumulative probabilities, and it assumes an identical effect of the predictors for each cumulative probability. It is often referred to as a *proportional odds model*.[7]

This model and related models with alternative link functions such as the probit[7–9] and complementary log-log[7] have several appealing properties. For instance, it is unnecessary to assign scores to the response categories, and if the model holds for a particular set of response categories, it holds with the same effects when the response scale is collapsed in any way.[7] One can motivate the model with identical effect $\boldsymbol{\beta}$ for each j using a regression model for an assumed underlying continuous response.[10] Suppose a continuous response Y^* has mean linearly related to $\mathbf{x}$, and with logistic conditional distribution with constant variance. Then for the categorical variable Y obtained by chopping Y^* into categories, the proportional odds model holds for predictor $\mathbf{x}$, with effects proportional to those in the continuous model; similarly, the probit link applies when the conditional distribution is normal.

The usual sorts of inferences, based on likelihood-ratio, score and Wald statistics, apply to maximum likelihood (ML) estimators of the model parameters. For binary logistic regression, the Wald test (the square of the ratio of the ML estimator to its standard error) can exhibit anomolous behaviour,[11] losing power relative to the likelihood-ratio test when the effect is large. It would be of interest to analyse whether similar results hold for model (1). Software exists for fitting this model, such as PROC LOGISTIC in SAS[12] for the logit and other links, but there are still some surprising gaps (for example, no procedure in SPSS) that may have limited its application is some areas.

Other ordinal logit models utilize single-category probabilities rather than cumulative probabilities. Most important is the *adjacent-categories logit* model,[5,13] which uses logits $\log[P(Y=j)/P(Y=j+1)]$, $j=1, \ldots, c-1$. The cumulative logit and adjacent-categories logit model both imply stochastic orderings of the response distributions for different predictor values. Effects in adjacent-category logit models refer to the effect of a one-unit increase of a predictor on the log odds of response in the lower instead of the higher of any two adjacent categories, whereas the effect in (1) refers to the entire response scale.

With categorical predictors, one can display the data as counts in a contingency table. When the table is not overly sparse, one can test the goodness-of-fit of these models with Pearson or likelihood-ratio statistics. These two models usually fit well in similar situations and provide similar substantive results. When they both fit well, one's choice of model may partly depend on whether one prefers parameter interpretation to refer to particular response categories (in which case the adjacent-categories logit model is natural) or instead to groupings of categories or an underlying continuous variable. When either model with the common effect $\boldsymbol{\beta}$ for each j fits poorly, possible strategies include: (i) trying a link function, such as the log-log, for which the response curve is non-symmetric; (ii) adding additional terms, such as interactions, to the linear predictor; (iii) generalizing the model by adding dispersion parameters[7,14] or permitting separate effects $\boldsymbol{\beta}_j$ for each logit.[15] The models describe location effects, and many applications also have dispersion effects. Although the proportional odds model has been generalized to include dispersion parameters,[7] this generalization has not yet been applied much. In some cases, such as when an ordinal response is measured at several dose levels using litters of mice, lack of fit may reflect overdispersion. This topic has not received much attention for multinomial responses, although recently developed random effects models would be one approach for dealing with it.

When there is lack of fit, another strategy is to try an alternative logit model, such as the *continuation-ratio* logit model, which uses logits $\{\log[P(Y_i = j)/P(Y_i \geqslant j+1)]\}$ or $\{\log[P(Y_i = j+1)/P(Y_i \leqslant j)]\}$. When used with separate effects $\{\boldsymbol{\beta}_j\}$, the multinomial likelihood factors into a product of the binomial likelihoods for the separate logits, which makes such analyses simple with standard software.[16] This model form is also useful when a sequential mechanism determines the response outcome.[17,18]

Another common approach to analyse ordinal response variables assigns scores to categories and uses ordinary regression or ANOVA methods.[19,20] This approach has the advantage of simplicity of interpretation, particularly when it is sufficient to summarize effects in terms of location rather than separate cell probabilities. This is often the case when c is large. One should keep in mind, though, limitations due to a non-normal, bounded response and a tendency for the variance to depend on the mean; for instance, less variability tends to occur when the mean is near the high end or low end of the scale. Preferably, one fits such models assuming a multinomial rather than normal distribution for the response.[19,21]

2.2. Models for association with ordinal responses

Logit models for ordinal responses are like ordinary regression models in the sense that they distinguish between response and explanatory variables. *Association models*, on the other hand, are designed to describe association between variables, and they treat those variables symmetrically.[22,23] For cell counts $\{n_{ij}\}$ with expected values $\{\mu_{ij}\}$ in a $r \times c$ contingency table, the models have form[23]

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \sum_{k=1}^{M} \beta_k u_{ik} v_{jk} \tag{2}$$

where $M \leqslant \min(r-1, c-1)$. The saturated model results from $M = \min(r-1, c-1)$. The most useful models are those with $M = 1$; the *linear-by-linear association model*[24] treats the row scores $\{u_{i1}\}$ and the column scores $\{v_{j1}\}$ as fixed monotone constants. The *row effects model* fixes the column scores but treats row scores as parameters, and is also valid when the row variable is nominal. The *column effects model* fixes the row scores but treats column scores as parameters. The *row and column effects* (*RC*) *model* treats both sets of parameters, in which case the model is not log-linear and ML estimation is more difficult.[25,26]

The goodness-of-fit of association models can be checked with ordinary chi-squared statistics. The models with $M = 1$ fit well when there is an underlying bivariate normal distribution.[27] The model with equally-spaced scores for Y relates to logit models for adjacent-category logits,[13] and the models generalize to include covariates.[28,29] Association models naturally describe association in terms of odds ratios for individual cells. For example, the linear-by-linear association model with equally-spaced scores implies a *uniform association* in terms of 'local' odds ratios for sub-tables constructed using adjacent rows and adjacent columns.[22] Alternative association models consider groupings of cells.[30,31] A substantial literature has evolved on applications of these various types of models and on extensions of them.[32–41]

A related literature has developed for correspondence analysis models[42–44] and equivalent canonical correlation models. They have similar structure but use an association term of the form in (2) to model the difference between μ_{ij} and its independence value rather than the difference between $\log \mu_{ij}$ and its independence value. When the association is weak, an approximate relation holds between parameter estimates in association models and canonical correlation models,[23] but otherwise the models refer to different types of ordinal association.[45] Because these models and association models focus primarily on association, they are probably of less use than ordinal logit models in biomedical applications, which usually distinguish between response and explanatory variables.

3. ADVANCES IN MODELLING REPEATED ORDINAL MEASUREMENT DATA

We next review research published within the past ten years on extending these ordinal models. The most active area has been modelling clustered data, such as occur in longitudinal studies and other forms of repeated measurement. Two major types of model for categorical responses differ in terms of whether they have *subject-specific* or *population-averaged* effects. The former refer to conditional distributions at the subject level, whereas the latter refer to marginal distributions, averaged over subjects in the population. The choice of model depends on whether one prefers interpretations to apply at the subject or the population level, the latter being more relevant in epidemiological studies that focus on overall levels of occurrence in a population. Approximate relationships exist between population-averaged and subject-specific parameters under the cumulative logit link,[46] but any particular marginal model need not have any simple and meaningful conditional model that implies it.

Regardless of the choice of model, it is awkward to use ML because of the lack of a natural multivariate family of distributions of categorical responses. For ordinal data, it is often sensible to assume an underlying multivariate normal distribution, which implies a cumulative probit model.[47] A complication for marginal models is that they apply to the marginal distributions of the multivariate response rather than the joint distribution to which the likelihood refers. A weighted least squares approach is simpler to use but limited to categorical predictors with non-sparse data.[48,49]

Most subject-specific models represent subject effects by a random effects term in the model and then assume that repeated responses given that effect are independent. One must integrate out the random effect to obtain the likelihood function. Except in rare cases, this integral does not have closed form and it is necessary to use some approximation, such as numerical integration or some variation of Monte Carlo and EM algorithms.

3.1. Generalized Estimating Equation (GEE) methodology

Rather than attempt to specify fully the joint distribution, one can apply methodology based on generalized estimating equations (GEE). One then specifies models only for marginal distributions and uses a working guess for the correlation structure.[50] This multivariate generalization of quasi-likelihood poses a model for the mean and a variance function. Estimates of model parameters are consistent even if the correlation structure is misspecified. The GEE approach is appealing for categorical data because of not requiring a multivariate distribution, but it has limitations resulting from the lack of a likelihood[51] and its subsequent reliance on Wald methods.

The GEE methodology, originally specified for univariate distributions such as the binomial and Poisson, extends to cumulative logit models[52–57] and cumulative probit models[58] for repeated ordinal responses. An SAS macro is available for the Lipsitz *et al*. approach.[53] For it, let $Y_{ijt} = 1$ if subject i makes response j for tth response. Then for each pair of categories (j, k) one selects a working correlation matrix for the pairs of responses (s, t); for instance, one might choose the exchangeable structure, $\mathrm{corr}(Y_{ijs}, Y_{ikt}) = \rho_{jk}$ for all s and t. Related literature includes applying GEE to the repeated ordinal case with independence estimating equations,[55,56] unstructured correlations,[56] and using a model for global odds ratios.[54,57] More general models allow for dispersion parameters that also depend on covariates.[58]

When marginal models are adopted, the association structure is usually not the primary focus and is regarded as a nuisance. In such cases with ordinal responses, it seems reasonable to use a simple structure for the associations, such as a common local or global odds ratio, rather than to extend much effort modelling it. This has the potential for slight efficiency gain over the independence equations and the more general structures that can have large numbers of parameters to characterize associations, as well as less chance of numerical singularities compared to the latter case. An earlier method related in spirit to GEE methods forms a weighted combination of estimates from separate models fitted to margins of a repeated ordinal response, allowing missing observations and time-dependent covariates.[59] This approach does not allow for simpler working correlation structures, however. When the association structure is itself of interest, a GEE2 approach is available for modelling associations using global odds ratios.[60]

3.2. Maximum likelihood fitting of marginal models

Multivariate logistic models have been defined that have a one-to-one correspondence between joint cell probabilities and parameters of marginal models as well as higher-order parameters of the joint distribution.[61–63] One can then use ML to estimate the model parameters, but the correspondence is awkward to specify for more than a few dimensions. Alternatively, one can treat a marginal model as a set of constraint equations and use methods of maximizing Poisson and multinomial likelihoods subject to constraints.[64,65] In these approaches, it is possible also to model simultaneously the joint distribution or higher-order marginal distributions. For instance, one might use a cumulative logit model for the marginal distributions and a model assuming a common global odds ratio[31,62,63] or a common local odds ratio[65] for the pairwise associations.

Recent computational advances have made ML feasible for relatively large joint distributions, with covariates, both for constrained ML[66,67] and for maximization with respect to joint probabilities expressed in terms of the marginal model parameters and an association model.[68] The latter approach is available in software (MAREG) that can perform either ML or GEE fitting of marginal models for ordinal responses.[68]

3.3. Random effects and mixed models

Random effects in models can represent a variety of situations, including subject heterogeneity, unobserved covariates, and other forms of overdispersion. For binary repeated measurement data, the basic model has a logit link with a linear predictor that contains a random effect having a normal distribution with unknown variance. The model form extends to ordinal logits. For instance, the jth cumulative or adjacent-categories logit for subject (or cluster) s and response t of the multivariate response might have form

$$\text{ordinal logit}_j = \lambda_j + \alpha_s + \boldsymbol{\beta}'\mathbf{x}_t \tag{3}$$

where $\mathbf{x}_t$ is a vector of predictors for that response and $\{\alpha_s\}$ are i.i.d. from a normal distribution.

The random effects literature for ordinal data so far considers primarily cumulative logit and probit models. Parameter estimation utilizes best linear unbiased prediction of parameters of an underlying continuous model[69] or else an approximation for the likelihood using numerical integration of the random effect[70–73] or the Laplace approximation for such integrals.[74] A FORTRAN program (MIXOR) is available.[71] The log-likelihood has closed form[75,56] for a complementary log-log link with the log of a gamma or inverse Gaussian distribution at the random effects distribution.

Ordinal modelling with random effects should be an active area of research in coming years. Further development for a variety of applications would be useful, for instance to model overdispersion and to allow heterogeneity in association between two variables across strata of a third. A major contribution would be the development of user-friendly software that could use a variety of mixture distributions for the random effect, a variety of link functions, and potentially a multivariate structure for multiple, correlated random effects.

3.4. Comparing marginal distributions

In repeated measurement studies, effects of interest may be either between-subject or within-subject. For categorical responses, methods for the within-subject comparison of marginal distributions have a long history. A common method is based on comparing log-linear models of symmetry and quasi-symmetry.[1] One can also make such comparisons in the context either of marginal models or random effects models. For ordinal responses, for instance, one could make marginal comparisons based on an ordinal logit model such as (3) in which explanatory variables are marginal indicators and their parameters describe these within-subject effects.

For model (3), alternatively one can use a non-parametric approach that makes no assumption about the form of the distribution of the random effect, or one can treat the subject effects as fixed and use conditional maximum likelihood.[77–79] The latter approach generalizes the Rasch item response model. Either approach with an adjacent-categories logit model yields ML estimates that are equivalent to ML estimates for a corresponding quasi-symmetric log-linear model having fixed scores for the ordered response categories.[77,78] Recent applications of generalized Rasch models and corresponding log-linear models include cross-over studies[77,80] and randomized clinical trials with matched-pairs responses. Simple special cases occur for the matched-pairs case, with expected cell counts $\{\mu_{ij} = E(n_{ij})\}$ in a square table for outcome i at the first response and j at the second. The model with a common shift β between margin 2 and margin 1 for each adjacent-category logit, for each subject, has a non-parametric and conditional ML estimate that is the same as the ordinary ML estimate for logit model[77,78]

$$\log(\mu_{ij}/\mu_{ji}) = \beta(j - i).$$

For the corresponding model with cumulative logits (that is, model (3) with indicator $x_1 = 0$ and $x_2 = 1$), simple estimates also exist[81,82] of the effect comparing the margins; one such estimate has form[81]

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

For large multi-dimensional contingency tables, ML can be computationally difficult for marginal comparisons. Recent solutions are motivated by GEE methods[83] or by a randomization approach.[84]

For longitudinal data, Markov chains provide an alternative structure that, unlike models just discussed, takes into account the time ordering. However, this approach seems to have received little attention so far.[85,86]

3.5. Modelling agreement

When the repeated measurement takes the form of ratings by several observers, agreement between pairs of raters or between each rater and a gold standard is usually the primary focus. Traditionally, it has been popular to measure agreement on an ordinal scale using weighted kappa.[87] Recent work has focused on modelling interrater agreement and handling more than two raters, for instance using a latent trait models,[88,89] quasi-symmetric and association models,[90–92] log-linear models for the two-way marginal distributions,[93] random effects,[41] and the area under a receiver operating characteristic (ROC) curve.[58] A cumulative logit model with random effects both for subjects rated and for the observers may be promising for some applications. A recent paper[94] for binary responses uses a conditioning argument for a two-stage agreement analysis in which the first stage focuses on subject-specific agreement and the second stage on marginal agreement; it is of interest to extend this interesting analysis to ordinal responses.

4. OTHER ADVANCES IN MODELLING ORDINAL RESPONSES

Many other topics have received attention in the research literature on ordinal data. These include the following.

4.1. Small-sample inference

Significance tests that take into account the ordering of categories have the potential for substantial power gain over tests that ignore that ordering. Large-sample inference, such as likelihood-ratio tests and confidence intervals for parameters in ordinal models, is well established. Small-sample methods are still under development.

For testing independence in two-way contingency tables, one can construct small-sample exact tests using the generalized hypergeometric distribution that results from conditioning on the row and column totals. Exact tests are available for several statistics, including correlation-type statistics with fixed or mid-rank scores[95,96] and related tests motivated by decision-theoretic considerations.[97,98] The exact conditional approach applies to exponential families with the canonical link. This includes log-linear models for Poisson counts and adjacent-category logit models for multinomial responses, but not cumulative logit models. For two-way tables, the correlation is the sufficient statistic for the association parameter in the linear-by-linear association model, and exact conditional inference (given the row and column totals) is based on that statistic.[95]

For stratified data with an ordinal response, only the case of a binary predictor (for example, two groups) is currently addressed in software,[96] but in principle the exact conditional methodology extends directly to several groups. A FORTRAN program approximates exact score tests of conditional independence for several ordinal log-linear models[99] and provides small-sample analyses for established large-sample tests such as generalized Mantel–Haenszel tests.[100,101] For large tables, exact methods may not be computationally feasible but one can use Monte Carlo to simulate the exact results[99] or use higher-order asymptotic methods such as the saddlepoint to approximate them very well.[102,103]

Open problems for future research include handling models that are more complex than conditional independence and may have non-canonical links for which the conditional approach does not apply. Approximate conditioning methods may provide a useful way of dealing with the conservativeness that sometimes occurs because of extreme discreteness with exact conditional methods. A random effects approach may become a popular way to eliminate nuisance parameters. Markov chain Monte Carlo methods employed for some non-standard log-linear models and for logistic regression[104] may be useful for some analyses.

4.2. Sample size and power

For comparing two groups (for example, two doses of of a drug) with an ordinal response, sample size formulae are available for the proportional odds model.[105] This requires anticipating the c marginal response proportions as well as the size of the effect. Setting $p_j = 1/c$ provides a lower bound for the sample size. The sample size does not depart much from this bound unless a single dominant response category occurs. With equal marginal probabilities, the ratio of the sample size $N(c)$ needed for c categories relative to the sample size $N(2)$ needed for two categories is approximately

$$N(c)/N(2) = 0{\cdot}75/[1 - 1/c^2].$$

Relative to a continuous response ($c = \infty$), using c categories provides efficiency $(1 - 1/c^2)$. The loss of information from collapsing to a binary response is substantial, but little gain results from using more than 4 or 5 categories.

For stratified data, one may need a somewhat increased sample size to preserve the desired power.[105] However, the variation among strata in the category probabilities has to be substantial before sample size is greatly affected. A somewhat different approach to sample size determination evaluates the exact conditional distribution with a network algorithm by simulation.[106]

4.3. Choice of scores and categories

Some methods for ordinal data, such as association models, require assigning scores to response categories. Various factors are relevant in choosing scores,[107–110] but it usually makes more sense to select scores that seem meaningful for the categories rather than to use automatic methods such as mid-rank scoring. Mid-rank scores need not provide reasonable scalings, since scores for neighbouring categories having relatively few observations are necessarily close. For highly unbalanced data, such as one response category has much greater frequency than the others, results may depend strongly on the choice of scores.[108]

In medical research, continuous variables are often converted to ordered categorical variables by grouping values. Grouping introduces an extreme form of measurement error with an associated loss of power[111] that, as mentioned above, can be severe in the binary case.

4.4. Order-restricted inference

Occasionally, one may want to account for the ordering but make weaker assumptions than ordinary models about structural forms of relationships. For instance, one might conduct an order-restricted inference that assumes only a stochastic ordering of response distributions. Likelihood ratio tests exist for comparing two multinomial distributions against the alternative of a stochastic ordering[112,113] or a narrower alternative of non-negative local log-odds ratios.[114] The large-sample distribution of the test statistics is chi-bar squared, the distribution of a weighted average of chi-squared variates with differing degrees of freedom. Other order-restricted approaches for stochastic ordering alternatives are motivated by decision-theoretic considerations[115,116] Recent evidence shows possibly anomalous behaviour by likelihood-ratio tests for the stochastic ordering alternative.[116,117]

Results for order-restricted comparisons of several multinomial distributions are incomplete, although tests have been suggested for local odds ratios,[114,117,118] cumulative odds ratios,[117,119] and continuation ratios.[120] It is possible to simulate exact conditional distributions of likelihood-ratio tests for various ordered alternatives.[117] Other work has focused on order-restricted inference in the context of models. For instance, for an association model having a parameter for each level of an ordinal response or predictor, one can fit the model subject to an ordering constraint on the parameter estimates.[121,122] Little attention has yet been paid to order-restricted inference in repeated measures problems.[123]

4.5. Goodness-of-fit

For multivariate categorical response models, it is possible to partition goodness-of-fit statistics.[124] For sparse data or continuous predictors, chi-squared fit statistics are inappropriate. Alternatives include a generalization[125] of the Hosmer–Lemeshow statistic for binary logistic regression, which compares observed to fitted counts for a partition of the possible response (for example, cumulative logit) values. A second way[15,126] tests the proportional odds assumption for model (1), and a third way checks the choice of link function. It would be worthwhile to develop residual analyses and other diagnostics that exploit the ordinal nature of the response, as well as develop and evaluate indices such as AIC to compare the fits of distinctly different forms of models.

4.6. Latent variable models

Most of the latent class modelling literature treats the observed categorical variables as nominal scale. For ordinal variables, it normally makes sense to have a continuous or ordinal latent variable. In the context of joint log-linear modelling of the observed and latent variables, one could assume linear-by-linear structure between observed and latent variables.[128,129] Alternatively, various types of logits, such as cumulative logits or cumulative probits,[130] could be applied to the observed responses, or one could use an ordinal latent variable even if observed variables are nominal.[131,152] Recent applications of latent variable models include modelling rater agreement,[128] cross-over trials,[80] and household fertility.

4.7. Bayesian inference

With recent advances in computational methods, Bayesian approaches with ordinal models are ripe for development. A recent paper uses Gibbs sampling with a normal or t prior for the association parameter in a linear-by-linear association model.[133] Another paper uses models for cumulative probabilities with prior information about the choice of link.[134] The cumulative probit form of model was applied to modelling rater agreement.[88] An alternative, order-restricted, approach conducts inference solely under the assumption that several multinomial distributions are stochastically ordered.[119] A serious challenge in applying Bayesian methods with multivariate categorical data is the large number of parameters, most of which may be nuisance parameters.

4.8. Smoothing ordinal data

The Bayesian approach is natural for smoothing data, for instance to eliminate sampling zeros or to estimate cell probabilities in contingency tables without assuming parametric models.[135] Alternatively, for ordinal data one can achieve this aim using kernel methods,[136] penalized likelihood methods,[137] or local likelihood estimation. For continuous and binary data, generalized additive models provide a smoothing that reflects a sampling model and a link function yet does not require a linear structural form. This approach could also be useful for ordinal responses, for instance by formulating a generalized additive model of proportional odds form.

4.9. Paired preference modelling

The Bradley–Terry model describes outcomes of pairwise competitions of a set of items (for example, tennis players, types of wines). Each item has a parameter, and the logit of the probability of preference of item i over j equals the difference between the parameters for those items. This model extends to ordinal variables,[138–140] for instance using compulative logits. For future work, a random effects approach seems natural here to reflect dependence among repeated comparisons by the same subjects.

4.10. Missing data

Missing data are an all too common problem, especially with longitudinal designs.[141] The relative paucity of literature about handling them with ordinal models includes a score test of independence in two-way tables with extensions for stratified data,[142] a comparison of likelihood-based and GEE methods for repeated responses with missing data,[52,55] handling non-random drop-out,[143] and Bayesian tobit modelling in studies with longitudinal ordinal data.[144] This is a promising and important area for future research.

4.11. Other areas

Areas not mentioned above on which work has appeared in recent years include diagnostics summarizing higher order effects,[145] odds ratio estimates for highly sparse situations in which ML estimates are inconsistent,[146] joint modelling of ordinal and continuous responses,[147] decompositions of chi-squared statistics for ordered alternatives,[148–150] a two-sample permutation test comparing two groups of clusters,[151] transitional models (pp. 201–203),[152] modelling overdispersion for ordinal responses,[153] interval censoring with an ordinal response,[154] and R-squared measures for ordinal models.[155]

5. OTHER CONSIDERATIONS

5.1. Software

New methods are rarely used in practice unless accompanied by user-friendly software. Of particular benefit would be a program that can handle a variety of strategies for multivariate ordinal logit models, including ML fitting of marginal models, GEE methods, and mixed models, all for a variety of link functions. Even with binary data, such goals currently require a variety of software, and more basic software needs exists that are not nearly as ambitious. For instance, some major statistical packages do not yet contain procedures for fitting univariate ordinal logit models or other models for multinomial responses.

5.2. Model interpretation

The tremendous improvements in computing power in recent years have fuelled the development of ever more complex statistical methodology. A major challenge for research statisticians is to explain this methodology to statisticians who are not specialists in their area and to scientists who could benefit from the methods. Statisticians should not underestimate this challenge. For instance, to what extent are the advances of the past 25 years for analysing ordinal data used in practice? One of the most useful advances is regression modelling using cumulative logits. To what extent has this model been adopted by consulting statisticians and by quantitative methodologists in other disciplines? My limited experience suggests it is gradually becoming better known among statisticians but that it still finds little use in many areas (such as the social sciences) in which ordinal responses are common.

This challenge is becoming greater because of the complexity of much of the newly developing methodology. Whereas it may not be crucial for a scientist to understand the technical details of how to produce the parameter estimates and their standard errors, it is important for the scientist to know what the estimate means. How can one explain differences among the various potential approaches and when they may be appropriate? Even basic distinctions such as (population-averaged/subject-specific) and (marginal/conditional) are confusing to many. In terms of actual applied impact, this may be a bigger challenge than any mentioned in this paper regarding the development of new methodology.

ACKNOWLEDGEMENTS

This work was partially supported by an NIH grant.

REFERENCES

1. Agresti, A. *An Introduction to Categorical Data Analysis*, Wiley, 1996.
2. Barnhart, H. X. and Sampson, A. R. 'Overview of multinomial models for ordinal data', *Communications in Statistics – A. Theory and Methods*, **23**, 3395–3416 (1994).
3. Clogg, C. C. and Shihadeh, E. S. *Statistical Models for Ordinal Variables*, Sage Publications, Thousand Oaks, CA, 1994.
4. Etzioni, R. D., Fienberg, S. E., Gilula, Z. and Haberman, S.J. 'Statistical models for the analysis of ordered categorical data in public health and medical research', *Statistical Methods in Medical Research* **3**, 179–204 (1994).
5. Simon, G. 'Alternative analyses for the singly-ordered contingency table', *Journal of the American Statistical Association*, **69**, 971–976 (1974).
6. Williams, O. D. and Grizzle, J. E. 'Analysis of contingency tables having ordered response categories', *Journal of the American Statistical Association*, **67**, 55–63 (1972).
7. McCullagh, P. 'Regression models for ordinal data' (with discussion), *Journal of the Royal Statistical Society*, *Series*, *B*, **42**, 109–142 (1980).
8. Aitchison, J. and Silvey, S. D. 'The generalization of probit analysis to the case of multiple responses', *Biometrika*, **44**, 131–140 (1957).

9. McKelvey, R. D. and Zovoina, W. 'A statistical model for the analysis of ordinal level dependent variables', *Journal of Mathematical Sociology*, **4**, 103–120 (1975).
10. Anderson, J. A. and Philips, P. R. 'Regression, discrimination and measurement models for ordered categorical variables', *Applied Statistics*, **30**, 22–31 (1981).
11. Hauck, W. W. and Donner, A. 'Wald's test as applied to hypotheses in logit analysis', *Journal of the American Statistical Association*, **72**, 851–853 (1977).
12. Stokes, M. E., Davis, C. S. and Koch, G. G. *Categorical Data Analysis Using the SAS System*, SAS Institute, Cary, NC, 1995.
13. Goodman, L. A. 'The analysis of dependence in cross-classifications having ordered categories, using log-linear models for frequencies and log-linear models for odds', *Biometrica*, **39**, 149–160 (1983).
14. Cox, C. 'Location scale cumulative odds models for ordinal data: A generalized non-linear model approach', *Statistics in Medicine*, **14**, 1191–1203 (1995).
15. Brant, R. 'Assessing proportionality in the proportional odds model for ordinal logistic regression', *Biometrics*, **46**, 1171–1178 (1990).
16. Lindsey, J. K., Jones, B. and Ebbutt, A. F. 'Simple models for repeated ordinal responses with an application to a seasonal rhinitis clinical trial', *Statistics in Medicine*, **16**, 2873–2882 (1997).
17. Tutz, G. 'Sequential item response models with an ordered response', *British Journal of Mathematical and Statistical Psychology*, **43**, 39–55 (1990).
18. Tutz, G. 'Sequential models in categorical regression', *Computational Statistics & Data Analysis*, **11**, 275–295 (1991).
19. Grizzle, J. E., Starmer, C. F. and Koch, G. G. 'Analysis of categorical data by linear models', *Biometrics*, **25**, 489–504 (1969).
20. Heeren, T. and D'Agostino, R. 'Robustness of the two independent samples *t*-test when applied to ordinal scaled data', *Statistics in Medicine*, **6**, 79–90 (1987).
21. Lipsitz, S. R. 'Methods for estimating the parameters of a linear model for ordered categorical data', *Biometrics*, **48**, 271–281 (1992).
22. Goodman, L. A. 'Simple models for the analysis of association in cross-classifications having ordered categories', *Journal of the American Statistical Association*, **74**, 537–552 (1979).
23. Goodman, L. A. 'The analysis of cross-classified data having ordered and/or the unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries', *Annals of Statistics*, **13**, 10–69 (1985).
24. Haberman, S. J. 'Log-linear models for frequency tables with ordered classifications', *Biometrics*, **36**, 589–600 (1974).
25. Haberman, S. J. 'Tests for independence in two-way contingency tables based on canonical correlation and on linear-by-linear interaction', *Annals of Statistics*, **9**, 1178–1186 (1981).
26. Haberman, S. J. 'Computation of maximum likelihood estimates in association models', *Journal of the American Statistical Association*, **90**, 1438–1446 (1995).
27. Becker, M. P. 'On the bivariate normal distribution and association models for ordinal categorical data', *Statistics & Probability Letters*, **8**, 435–440 (1989).
28. Becker, M. P. 'Models for the analysis of association in multivariate contingency tables', *Journal of the American Statistical Association*, **84**, 1014–1019 (1989).
29. Becker, M. P. and Clogg, C. C. 'Analysis of sets of two-way contingency tables using association models', *Journal of the American Statistical Association*, **84**, 142–151 (1989).
30. Dale, J. R. 'Global cross-ratio models for bivariate, discrete, ordered responses', *Biometrics*, **42**, 909–917 (1986).
31. Molenberghs, G. and Lesaffre, E. 'Marginal modeling of correlated ordinal data using a multivariate Plackett distribution', *Journal of the American Statistical Association*, **89**, 633–644 (1994).
32. Ashby, D., West, C. R. and Ames, D. 'The ordered logistic regression model in psychiatry: rising prevalence of dementia in old people's homes', *Statistics in Medicine*, **8**, 1317–1326 (1989).
33. Gange, S. J., Linton, K. L. P., Scott, A. J., DeMets, D. L. and Klein, R. 'A comparison of methods for correlated ordinal measures with opthalmic applications', *Statistics in Medicine*, **14**, 1961–1974 (1995).
34. Greenland, S. 'Alternative models for ordinal logistic regression', *Statistics in Medicine*, **13**, 1665–1677 (1994).
35. Hastie, T. J., Botha, J. L. and Schnitzler, C. M. 'Regression with an ordered categorical response', *Statistics in Medicine*, **8**, 785–794 (1989).

36. Holtbrügge, W. and Schumacher, M. 'A comparison of regression models for the analysis of ordered categorical data', *Applied Statistics*, **40**, 249–259 (1991).
37. Qu, Y., Piedmonte, M. R. and Medendorp, S. V. 'Latent variable models for clustered ordinal data', *Biometrics*, **51**, 268–275 (1995).
38. Rosner, B. and Glynn, R. J. 'Multivariate methods for clustered ordinal data with applications for survival analysis', *Statistics in Medicine*, **16**, 357–372 (1997).
39. Sall, J. 'A monotone regression smoother based on ordinal cumulative logistic regression', *ASA Proceedings of Statistical Computing Section*, 276–281 (1991).
40. Tamara, R. N., Tanaka, Y., Satoh, K. and Takahashi, M. 'Using the proportional odds model to assess the relationship between a multi-item and a global item efficacy scale in a psychiatric clinical trial', *Journal of Biopharmaceutical Statistics*, **6**, 127–137 (1996).
41. Williamson, J. M. and Manatunga, A. K. 'Assessing interrater agreement from dependent data', *Biometrics*, **53**, 707–714 (1997).
42. Gilula, Z. and Haberman, S. J. 'The analysis of multivariate contingency tables by restricted canonical and restricted association models', *Journal of the American Statistical Association*, **83**, 760–771 (1988).
43. Gilula, Z. and Ritov, Y. 'Inferential ordinal correspondence analysis: Motivation, derivation and limitations', *International Statistical Review*, **58**, 99–108 (1990).
44. Goodman, L. A., 'Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables', *International Statistical Review*, **54**, 291–309 (1986).
45. Gilula, Z., Krieger, A. M. and Ritov, Y. 'Ordinal association in contingency tables: Some interpretive aspects', *Journal of the American Statistical Association*, **83**, 540–545 (1988).
46. Ten Have, T. R., Landis, J. R. and Hartzel, J. 'Population-averaged and cluster-specific models for clustered ordinal response data', *Statistics in Medicine*, **15**, 2573–2588 (1996).
47. Kim, K. 'A bivariate cumulative probit regression model for ordered categorical data', *Statistics in Medicine*, **14**, 1341–1352 (1995).
48. Agresti, A. 'A survey of models for repeated ordered categorical response data', *Statistics in Medicine*, **8**, 1209–1224 (1989).
49. Koch, G. G., Landis, J. R., Freeman, J. L., Freeman, D. H. and Lehnen, R. G. 'A general methodology for the analysis of experiments with repeated measurement of categorical data', *Biometrics*, **33**, 133–158 (1997).
50. Liang, K. Y., Zeger, S. L. and Qaqish, B. F. 'Multivariate regression analyses for categorical data (with discussion)', *Journal of the Royal Statistical Society*, *Series B*, **54**, 3–40 (1992).
51. Lindsey, J. K. *Models for Repeated Measurements*, Clarendon Press, Oxford, 1993.
52. Kenward, M. G., Lesaffre, E. and Molenberghs, G. 'An application of maximum likelihood and generalized estimating equations to the analysis of ordinal data from a longitudinal study with cases missing at random', *Biometrics*, **50**, 945–953 (1994).
53. Lipsitz, S. R., Kim, K. and Zhao, L. 'Analysis of repeated categorical data using generalized estimating equations', *Statistics in Medicine*, **13**, 1149–1163 (1994).
54. Lumley, T. 'Generalized estimating equations for ordinal data: A note on working correlation structures', *Biometrics*, **52**, 354–361 (1996).
55. Mark, S. D. and Gail, M. H. 'A comparison of likelihood-based and marginal estimating equation methods for analysing repeated ordered categorical responses with missing data', *Statistics in Medicine*, **13**, 479–493 (1994).
55. Liang, K. Y., Zeger, S. L. and Qaqish, B. F. 'Multivariate regression analyses for categorical data (with discussion)', *Journal of Royal Statistical Society*, *Series B*, **543**, 3–40 (1992).
56. Miller, M. E., Davis, C. S. and Landis, J. R. 'The analysis of longitudinal polytomous data: generealized estimating equations and connections with weighted least squares', *Biometrics*, **49**, 1033–1044 (1993).
57. Williamson, J. M., Kim, K. and Lipsitz, S. R. 'Analyzing bivariate ordinal data using a global odds ratio', *Journal of the American Statistical Association*, **90**, 1432–1437 (1995).
58. Toledano, A. Y. and Gatsonis, C. 'Ordinal regression methodology for ROC curves derived from correlated data', *Statistics in Medicine*, **15**, 1807–1826 (1996).
59. Stram, D., Wei, L. J. and Ware, J. 'Analysis of repeated ordered categorical outcomes with possibly missing observations and time-dependent covariates', *Journal of the American Statistical Association*, **83**, 631–637 (1998).

60. Heagerty, P. J. and Zeger, S. L. 'Marginal regression models for clustered ordinal measurements', *Journal of the American Statistical Association*, **91**, 1024–1036 (1996).
61. Fitzmaurice, G. and Laird, N. 'A likelihood-based method for analysing longitudinal binary responses', *Biometrika*, **80**, 141–151 (1993).
62. Glonek, G. F. V. 'A class of regression models for multivariate categorical responses', *Biometrika*, **83**, 115–128 (1996).
63. McCullagh, P. and Glonek, G. F. V. 'Multivariate logistic models', *Journal of the Royal Statistical Society*, *Series B*, **57**, 533–546 (1995).
64. Lang, J. B. 'Maximum likelihood methods for a general class of log-linear models', *Annals of Statistics*, **24**, 726–752 (1996).
65. Lang, J. B. and Agresti, A. 'Simultaneously modeling joint and marginal distributions of multivariate categorical responses', *Journal of the American Statistical Association*, **89**, 625–632 (1994).
66. Bergsma, W. P. 'Marginal models for categorical data', Ph.D. thesis, Tilburg University, Netherlands, 1997.
67. Lang, J. B., McDonald, J. W. and Smith, P. W. F. 'Association marginal modeling of multivariate polytomous responses: A maximum likelihood approach for large, sparse tables', Technical report 263, University of Iowa Department of Statistics and Actuarial Science, 1997.
68. Heumann, C. 'Likelihoodbasierte marginale regressionsmodelle fur korrelierte kategoriale daten', PhD thesis, Ludwig–Maximilians–Universitat, Munchen, 1997.
69. Harville, D. A. and Mee, R. W. 'A mixed-model procedure for analyzing ordered categorical data', *Biometrics*, **40**, 393–408 (1984).
70. Ezzet, F. and Whitehead, J. 'A random effects model for ordinal responses from a crossover trial', *Statistics in Medicine*, **10**, 901–907 (1991).
71. Hedeker, D. R. and Gibbons, R. D. 'A random-effects ordinal regression model for multilevel analysis', *Biometrics*, **50**, 933–944 (1994).
72. Jansen, J. 'On the statistical analysis of ordinal data when extravariation is present', *Applied Statistics*, **39**, 75–84 (1990).
73. Tutz, G. and Hennevogl, W. 'Random effects in ordinal regression models', *Computational Statistics & Data Analysis*, **22**, 537–557 (1996).
74. Sheiner, L. B., Beal, S. L. and Dunne, A. 'Analysis of nonrandomly censored ordered categorical longitudinal data from analgesic trials', *Journal of the American Statistical Association*, **92**, 1235–1244 (1997).
75. Crouchley, R. 'A random-effects model for ordered categorical data', *Journal of the American Statistical Association*, **90**, 489–498 (1995).
76. Ten Have, T. R. 'A mixed effects model for longitudinal ordinal response data including correlated discrete failure times with ordinal responses', *Biometrics*, **52**, 473–491 (1996).
77. Agresti, A. 'Distribution-free fitting of logit models with random effects for repeated categorical responses', *Statistics in Medicine*, **12**, 1969–1987 (1993).
78. Agresti, A. 'Computing conditional maximum likelihood estimates for generalized Rasch models using simple loglinear models with diagonals parameters', *Scandinavian Journal of Statistics*, **20**, 63–72 (1993).
79. Conaway, M. R. 'Analysis of repeated categorical measurements with conditional likelihood methods', *Journal of the American Statistical Association*, **84**, 53–62 (1989).
80. Kenward, M. G. and Jones, B. 'The analysis of categorical data from cross-over trials using a latent variable model', *Statistics in Medicine*, **10**, 1607–1619 (1991).
81. Agresti, A. and Lang, J. B. 'A proportional odds model with subject-specific effects for repeated ordered categorical responses', *Biometrika*, **80**, 527–534 (1993).
82. McCullagh, P. 'A logistic model for paired comparisons with ordered categorical data', *Biometrika*, **64**, 449–453 (1977).
83. Agresti, A., Lipsitz, S. and Lang, J. 'Comparing marginal distributions of large, sparse contingency tables', *Computational Statistics & Data Analysis*, **14**, 55–73 (1992).
84. Landis, J. R., Miller, M. E., Davis, C. S. and Koch, G. G. 'Some general methods for the analysis of categorical data in longitudinal studies', *Statistics in Medicine*, **7**, 109–137 (1988).
85. Albert, P. S. 'A Markov model for sequences of ordinal data from a relapsing-remitting disease', *Biometrics*, **50**, 51–60 (1994).

86. Chuang, C. and Francom, S. F. 'A structural model for the Markov chain of change in vote intention', *Communications in Statistics – A. Theory and Methods*, **15**, 3475–3487 (1986).
87. Fleiss, J. L. Cohen, J. and Everitt, B. S. 'Large-sample standard errors of kappa and weighted kappa', *Psychological Bulletin*, **72**, 3223–3327 (1969).
88. Johnson, V. E. 'On Bayesian analysis of multirater ordinal data: An application to automated essay grading', *Journal of the American Statistical Association*, **91**, 42–51 (1996).
89. Uebersax, J. S. and Grove, W. M. 'A latent trait finite mixture model for the analysis of rating agreement', *Biometrics*, **49**, 823–835 (1993).
90. Agresti, A. 'A model for agreement between ratings on an ordinal scale', *Biometrics*, **44**, 539–548 (1988).
91. Becker, M. P. 'Using association models to analyse agreement data: two examples', *Statistics in Medicine*, **8**, 1199–1207 (1989).
92. Becker, M. P. 'Quasisymmetric models for the analysis of square contingency tables', *Journal of the Royal Statistical Society*, *Series B*, **52**, 369–378 (1990).
93. Becker, M. P. and Agresti, A. 'Log-linear modelling of pairwise interobserver agreement on a categorical scale', *Statistics in Medicine*, **11**, 101–114 (1992).
94. Cook, R. J. and Farewell, V. T. 'Conditional inference for subject-specific and marginal agreement: two families of agreement measures', *Canadian Journal of Statistics*, **23**, 333–344 (1995).
95. Agresti, A., Mehta, C. R. and Patel, N. R. 'Exact inference for contingency tables with ordered categories', *Journal of the American Statistical Association*, **85**, 453–458 (1990).
96. StatXact. *StatXact3 for Windows*: *Statistical Software for Exact Nonparametric Inference*, *User Manual*, Cytel Software, 1995.
97. Cohen, A. and Sackrowitz, H. B. 'Tests for independence in contingency tables with ordered categories', *Journal of Multivariate Analysis*, **36**, 56–67 (1991).
98. Cohen, A. and Sackrowitz, H. B. 'Improved tests for comparing treatments against a control and other one-sided problems', *Journal of the American Statistical Association*, **87**, 1137–1144 (1992).
99. Kim, D. and Agresti, A. 'Nearly exact tests of conditional independence and marginal homogeneity for sparse contingency tables', *Computational Statistics and Data Analysis*, **24**, 89–104 (1997).
100. Landis, J. R., Heyman, E. R. and Koch, G. G. 'Average partial association in three-way contingency tables: A review and discussion of alternative tests', *International Statistical Review*, **46**, 237–254 (1978).
101. Mantel, N. 'Chi-square tests with one degree of freedom: extensions of the Mantel–Haenszel procedure', *Journal of the American Statistical Association*, **58**, 690–700 (1963).
102. Agresti, A., Lang, J. and Mehta, C. 'Some empirical comparisons of exact, modified exact, and higher-order asymptotic tests of independence for ordered categorical variables', *Communications in Statistics*, *Simulation and Computation*, **22**, 1–18 (1993).
103. Pierce, D. A. and Peters, D. 'Practical use of higher-order asymptotics for multiparameter exponential families' (with discussion), *Journal of Royal Statistical Society*, *Series B*, **54**, 701–737 (1992).
104. Forster, J. J., McDonald, J. W. and Smith, P. W. F. 'Monte Carlo exact conditional tests for log-linear and logistic models', *Journal of the Royal Statistical Society*, *Series B*, **58**, 445–453 (1996).
105. Whitehead, J. 'Sample size calculations for ordered categorical data', *Statistics in Medicine*, **12**, 2257–2271 (1993).
106. Hilton, J. F. and Mehta, C. R. 'Power and sample size calculations for exact conditional tests with ordered categorical data', *Biometrics*, **49**, 609–616 (1993).
107. Gastwirth, J. L. and Podgor, M. J. 'Scores', *Encyclopedia of Biostatistics*, Wiley, New York, 1998.
108. Graubard, B. I. and Korn, E. L. 'Choice of column scores for testing independence in ordered $2 \times K$ contingency tables', *Biometrics*, **43**, 471–476 (1987).
109. Kimeldorf, G., Sampson, A. R. and Whitaker, L. R. 'Min and max scorings for two-sample ordinal data', *Journal of the American Statistical Association*, **87**, 241–247 (1992).
110. Rahlfs, V. W. and Zimmerman, H. 'Scores: Ordinal data with few categories – how should they be analyzed?', *Drug Information Journal*, **27**, 1227–1240 (1993).
111. Altman, D. G. 'Categorizing continuous variables', in *Encyclopedia of Biostatistics*, Wiley, Chichester, 1998.
112. Grove, D. M. 'A test of independence against a class of ordered alternatives in a $2 \times c$ contingency table', *Journal of the American Statistical Association*, **75**, 454–459 (1980).
113. Robertson, T. and Wright, F. T. 'Likelihood-ratio tests for and against a stochastic ordering between multinomial populations', *Annals of Statistics*, **9**, 1248–1257 (1981).

114. Dykstra, R., Kochar, S. and Robertson, T. 'Inference for likelihood ratio ordering in the two-sample problem', *Journal of the American Statistical Association*, **90**, 1034–1040 (1995).
115. Berger, V. and Sackrowitz, H. 'Improving tests for superior treatment in contingency tables', *Journal of the American Statistical Association*, **92**, 700–705 (1997).
116. Cohen, A. and Sackrowitz, H. 'Testing whether treatment is better than control with ordered categorical data', Technical Report 96–019, Department of Statistics, Rutgers University, 1996.
117. Agresti, A. and Coull, B. C. 'Order-restricted inference for monotone trend alternatives in contingency tables', *Computational Statistics & Data Analysis*, **28**, 139–155 (1998).
118. Patefield, W. M. 'Exact tests for trends in ordered contingency tables', *Applied Statistics*, **31**, 32–43 (1982).
119. Evans, M., Gilula, Z., Guttman, I. and Swart, T. 'Bayesian analysis of stochastically ordered distributions of categorical variables', *Journal of the American Statistical Association*, **92**, 208–214 (1997).
120. Grove, D. M. 'Positive association in a two-way contingency table: Likelihood ratio tests', *Communications in Statistics, A: Theory and Methods*, **13**, 931–945 (1984).
121. Agresti, A., Chuang, C. and Kezouh, A. 'Order-restricted score parameters in association models for contingency tables', *Journal of the American Statistical Association*, **82**, 619–623 (1987).
122. Ritov, Y. and Gilula, Z. 'The order-restricted RC model for ordered contingency tables: Estimation and testing for fit', *Annals of Statistics*, **19**, 2090–2101 (1991).
123. Albert, P. S., Hunsberger, S. A. and Biro, F. M. 'Modeling repeated measures with monotonic ordinal responses and misclassification, with application to studying maturation', *Journal of the American Statistical Association*, **92**, 1304–1311 (1997).
124. Lang, J. B. 'On the partitioning of goodness-of-fit statistics for multivariate categorical response models', *Journal of the American Statistical Association*, **91**, 1017–1023 (1996).
125. Lipsitz, S. R., Fitzmaurice, M. G. and Molenberghs, G. 'Goodness-of-fit tests for ordinal response regression models', *Applied Statistics*, **45**, 175–190 (1996).
126. Peterson B. and Harrell, F. E., Jr. 'Partial proportional odds models for ordinal response variables', *Applied Statistics*, **39**, 205–217 (1990).
127. Genter, F. C. and Farewell, V. T. 'Goodness-of-link testing in ordinal regression models', *Canadian Journal of Statistics*, **13**, 37–44 (1985).
128. Agresti, A. and Lang, J. 'Quasi-symmetric latent class models, with application to rater agreement', *Biometrics*, **49**, 131–140 (1993).
129. Heinen, T. *Discrete Latent Variable Models*, Tilburg University Press, 1993.
130. Sobel, M. E. and Arminger, G. 'Modeling household fertility decisions: A nonlinear simultaneous probit model', *Journal of the American Statistical Association*, **87**, 38–47 (1992).
131. Croon, M. 'Latent class analysis with ordered latent classes', *British Journal of Mathematical and Statistical Psychology*, **43**, 171–192 (1990).
131. Böckenholt, U. and Böckenholt, I. 'Constrained latent class analysis: Simultaneous classification and scaling of discrete choice data', *Psychometrika*, **56**, 699–716 (1991).
132. Croon, M. 'Latent class analysis with ordeed latent classes', *British Journal of Mathematical and Statistical Psychology*, **43**, 171–192 (1990).
133. Brink A. and Smith, A. F. M. 'Bayesian modelling of the association in contingency tables', in *Statistical Modelling: Proceedings of the 11th International Workshop on Statistical Modelling,* Orvieto, Italy, 1996.
134. Lang, J. B. 'Bayesian ordinal and binary regression models with a parametric family of mixture links', Technical report 267, University of Iowa Department of Statistics and Actuarial Science, 1997.
135. Agresti, A. and Chuang, C. 'Model-based Bayesian methods for estimating cell proportions in cross-classification tables having ordered categories', *Computational Statistics & Data Analysis*, **7**, 245–258 (1989).
136. Titterington, D. M. and Bowman, A. W. 'A comparative study of smoothing procedures for ordered categorical data', *Journal of Statistical Computation and Simulation*, **21**, 291–312 (1985).
137. Simonoff, J. S. 'Probability estimation via smoothing in sparse contingency tables with ordered categories', *Statistics & Probability Letters*, **5**, 55–63 (1987).
138. Agresti, A. 'Analysis of ordinal paired comparison data', *Applied Statistics*, **41**, 287–297 (1992).
139. Böckenholt, U. and Dillon, W. R. 'Modeling within-subject dependencies in ordinal paired comparison data', *Psychometrika*, **62**, 411–434 (1997).

140. Fahrmeir, L. and Tutz, G. 'Dynamic stochastic models for time-dependent ordered paired comparison systems', *Journal of the American Statistical Association*, **89**, 1438–1449 (1994).
141. Lesaffre, E., Molenberghs, G. and DeWulf, L. 'Effects of dropouts in a longitudinal study: An application of a repeated ordinal model', *Statistics in Medicine*, **15**, 1123–1141 (1996).
142. Lipsitz, S. R. and Fitzmaurice, G. M. 'The score test for independence in $R \times C$ contingency tables with missing data', *Biometrics*, **52**, 751–762 (1996).
143. Molenberghs, G., Kenward, M. G. and Lesaffre, E. 'The analysis of longitudinal ordinal data with nonrandom drop-out', *Biometrika*, **84**, 33–44 (1997).
144. Cowles, M. K., Carlin, B. P. and Connett, J. E. 'Bayesian tobit modeling of longitudinal ordinal clinical trial compliance data with nonignorable missingness', *Journal of the American Statistical Association*, **91**, 86–98 (1996).
145. Simonoff, J. S. and Tsai, C.-L. 'Higher order effects in log-linear and log-non-linear models for contingency tables with ordered categories', *Applied Statistics*, **40**, 449–458 (1991).
146. Liu, I. M. and Agresti, A. 'Mantel–Haenszel-type inference for cumulative odds ratios', *Biometrics*, **52**, 1222–1234 (1996).
147. Catalano, P. J. 'Bivariate modelling of clustered continuous and ordered categorical outcomes', *Statistics in Medicine*, **16**, 883–900 (1997).
148. Best, D. J. and Rayner, J. C. W. 'Nonparametric analysis for doubly ordered two-way contingency tables', *Biometrics*, **52**, 1153–1156 (1996).
149. Nair, V. N. 'Testing in industrial experiments with ordered categorical data', *Technometrics*, **28**, 283–291 (1986).
150. Nair, V. N. 'Chi-squared-type tests for ordered alternatives in contingency tables', *Journal of the American Statistical Association*, **82**, 283–291 (1987).
151. Fay, M. P. and Gennings, C. 'A Non-parametric two-sample tests for repeated ordinal responses', *Statistics in Medicine*, **15**, 429–442 (1996).
152. Diggle, P. J., Liang, K.-Y. and Zeger, S. L. *Analysis of Longitudinal Data*, Clerendon Press, Oxford 1994.
153. Wang, Y.-G. 'A quasi-likelihood approach for ordered categorical data with overdispersion', *Biometrics*, **52**, 1252–1258 (1996).
154. Simpson, D. G., Carroll, R. J., Zohu, H. and Guth, D. J. 'Interval censoring and marginal analysis in ordinal regression', *Journal of Agricultural, Biological, and Environmental Statistics*, **1**, 354–376 (1996).
155. Agresti, A. 'Applying R^2-type measures to ordered categorical data', *Technometrics*, **28**, 133–138 (1986).

