STATISTICS IN MEDICINE
Statist. Med. 2004; **00**:1–6 *Prepared using* simauth.cls *[Version: 2002/09/18 v1.11]*

Modelling Physicians' Recommendations for Optimal Medical Care by Random Effects Stereotype Regression

Oliver Kuss[1]

[1] *Institute of Medical Epidemiology, Biostatistics, and Informatics, University of Halle-Wittenberg, 06097 Halle (Saale), Germany*

SUMMARY

We show how Anderson's Stereotype regression model can be extended to account for correlated responses by a simple nonlinear parameter restriction on the multinomial logistic model with random effects. We comment on parameter estimation and use a data set on physicians' recommendations and preferences in traumatic brain injury rehabilitation for illustration.

1. Introduction

Many study designs in medical statistics give rise to correlated data. For example, subjects are followed over time, are repeatedly treated under different experimental conditions, or are observed in logical units (e.g. clinics, families, litters). Focusing on regression models, one of the standard tools which adequately accounts for these correlations is the random effects (RE) model, sometimes also called hierarchical or mixed effects model. Comprehensive reviews of the model emphasizing medical applications are given by Brown and Prescott [1](1999), and, emphasizing theory, by Diggle, Heagerty, Liang and Zeger [2]. Random effects models are quite common for continuous responses, and also, despite an enhanced mathematical complexity, for binary responses[3]. Less used have been random effect models for the analysis of discrete non-binary responses, some of the rare examples are Hedeker and Gibbons [4] or Tutz and Hennevogl [5] for ordinal and Hartzel et al.[6] or Lammertyn et al. [7] for nominal responses. To our knowledge, up to now there exists no random effects version of the Stereotype regression model.

The Stereotype regression model was originally proposed by Anderson [8]. He observed that some relevant discrete non-binary responses are not perfectly ordinal in the sense that there is a latent continuous variable which was only observed in discrete and disjunct classes. Rather these responses should be regarded as describing a multidimensional phenomenon where several items determine the grade on an ordinal scale, the most prominent example being the severity of a disease. The Stereotype model results from the standard multinomial model by a simple non-linear parameter restriction. It has less parameters than the multinomial model and models the assumed ordinality of the response in terms of the covariates. Greenland [9] showed that the progression of a disease through various stages is naturally modelled by the Stereotype model, and that the model is valid also under case dependent sampling, as opposed to the

Received 01 January 2004
Revised 01 July 2004

Proportional odds model[10]. A short review of the model with notes on convenient parameter estimation is given by Kuss [11].

In the following we show how the Stereotype model can be extended to account for correlated responses. In Chapter 2 we extend the standard model to a Stereotype model with random effects and comment on accessible methods for parameter estimation. Chapter 3 applies the model to a data set on physicians' recommendations and preferences in traumatic brain injury rehabilitation. Chapter 4 concludes.

2. The Random Effects Stereotype Regression Model

To extend the Stereotype regression model to account for correlated responses we use the fact that the original Stereotype model is derived from the ordinary multinomial logistic regression model by a specific nonlinear parameter restriction. This restriction is simply applied to the multinomial logistic random effects model of Hartzel et al.[6]. The Stereotype model with random effects thus becomes a nonlinear model with random effects and all the well-known theory and estimation methods (see, e.g. Davidian and Giltinan [12]) can be used.

We assume that our data comprises a set of I $(i = 1, \ldots, I)$ independent clusters where the i-th cluster consists of n_i observations. Let Y_{ij} denote the j-th response in cluster i $(j = 1, \ldots, n_i)$, where this response is from one of r $(r = 1, \ldots, R)$ distinct categories and the response probability is $\pi_{ijr} = P(Y_{ij} = r)$. Further, x_{ij} denotes a column vector of p covariates for the j-th observation in the i-th cluster. The model equation is

$$\log\left(\frac{\pi_{ijr}}{\pi_{ijR}}\right) = \theta_r + x'_{ij}\phi_r\beta + u_{ir}, \quad r = 1, \ldots, R-1, \tag{1}$$

where the θ_r are constant terms, and the influences of covariates are assessed through the components of $\beta = (\beta_1, \ldots, \beta_p)'$. The θ_r, the ϕ_r, and the components of β are considered to be fixed effects. For the random effects u_{ir} we assume a multivariate normal distribution with unstructured covariance matrix Σ, that is for $u_i = (u_{i1}, \ldots, u_{i,R-1})'$ we have $u_i \sim N(0, \Sigma)$.

For reasons of identification of parameters we restrict $\theta_R = 0$, $\beta_R = 0$, $u_R = 0$, $\phi_R = 0$, and $\phi_1 = 1$, so that interpretation of parameters is, analogous to the multinomial logistic model, with reference to the R-th category. Note that there are two restrictions ($\phi_R = 0$ and $\phi_1 = 1$) on the ϕ_r.

The model equation of the RE Stereotype regression model is derived from the multinomial logistic random effects model of Hartzel et al.[6] by the non-linear parameter restriction $\beta_r = \phi_r\beta$. The scalars ϕ_r introduce a metric for the common effect of the covariates, where this common covariate effect is assumed constant across the R response categories.

The likelihood contribution of the i-th cluster is

$$l_i(\theta_r, \beta_r, \Sigma) = \int_{-\infty}^{\infty}\left(\prod_{j=1}^{n_i}\left[\frac{\exp(\theta_r + x'_{ij}\phi_r\beta + u_{ir})}{\sum_{q=1}^{R}\exp(\theta_q + x'_{ij}\phi_q\beta + u_{iq})}\right]^{I(Y_{ij}=r)}\right) f_u(u_i, \Sigma)\, du_i, \tag{2}$$

where $f_u(u_i, \Sigma)$ denotes the multivariate normal density, and $I()$ the indicator function. The overall likelihood function is the product of the contributions l_i of the I clusters.

Actual estimation of parameters is complicated by the fact that the likelihood function consists of a product of I integrals which can not be solved in closed form. Thus, numerical

Prepared using simauth.cls

or stochastic integration are viable alternatives. Hartzel et al.[6] suggest adaptive Gaussian quadrature as the preferred method for parameter estimation in non-linear random effects models. As such, the model can be fitted conveniently with, for example, SAS PROC NLMIXED. To check the robustness of this numerical integration procedure we also fitted the model using a MCMC procedure with non-informative priors by using WinBUGS [15].

3. The Motivating Example

The motivation for the derivation of the RE Stereotype model was a data set from a study on physicians' recommendations and preferences in traumatic brain injury (TBI) rehabilitation [16]. In this study, 36 physicians were asked to decide on their preferred rehabilitation setting (out-patient, day-clinic, in-patient) for each of ten typical TBI patients, given to them as case vignettes. Of course, we expect the setting recommendations within the same physician to be correlated. Concerning the three-valued response we recognize that this is not strictly nominal, but has indeed some ordinal flavor, for example, we might think of the "time spent at the rehabilitation unit" as some underlying continuous variable. However, it is not that simple that out-patient, day-clinic, and in-patient rehabilitation only differ by the time that patients spend in clinic. Instead they rather represent different therapeutic concepts and actual treatment in the three settings varies.

Of interest was mainly if we could identify factors (considering physicians and patients) that influence setting preferences. Thus, four covariates, all of them binary, were included in the model, two of them referring to physicians' characteristics (1. Is the physician a neurologist [NEURO] and 2. Is the physician a specialist [SPECIAL]) and two describing the patient (3. Is the time since the TBI event longer than 3 months [TIME] and 4. Is the patient severely or moderately handicapped after the TBI [SEVERITY]). As the reference category of the response we chose the in-patient setting, and compare the day-clinic (DC) and the out-patient(OP) setting to this.

For simple description, in figure 1 the different recommendations are displayed with the physicians as the units on the x-axis and the patients as units on the y-axis. Additionally given is the respective value of the physicians' and the patients' covariates in terms of lines parallel to the axis.

In table I we give the results (estimates and respective standard errors in parentheses) for our data set for the ordinary Stereotype model, the RE multinomial model and the RE Stereotype model. As we expect (and maybe hope as potential patients), physicians' own characteristics do have only small influence on their recommendations, β_{NEURO} and $\beta_{SPECIAL}$ do not deviate significantly from zero in any of the models. Opposed to this, patients' characteristics are highly significant, the β_{TIME} always being larger than zero (rather recommending the OP or DC setting if the TBI happened before more than 3 months), and the β_{TIME} always being smaller then zero (rather less recommending the OP or DC setting if the patient is severely handicapped after the TBI).

Regarding the values of the model selection criteria we see that the RE Stereotype model is judged superior to the other two models: Compared to the ordinary Stereotype model this means on one hand that it is essential to account for the inherent correlation in the data (which is also confirmed by the elements Σ being significantly different from zero). Compared to the RE multinomial model on the other hand we note that we do not need

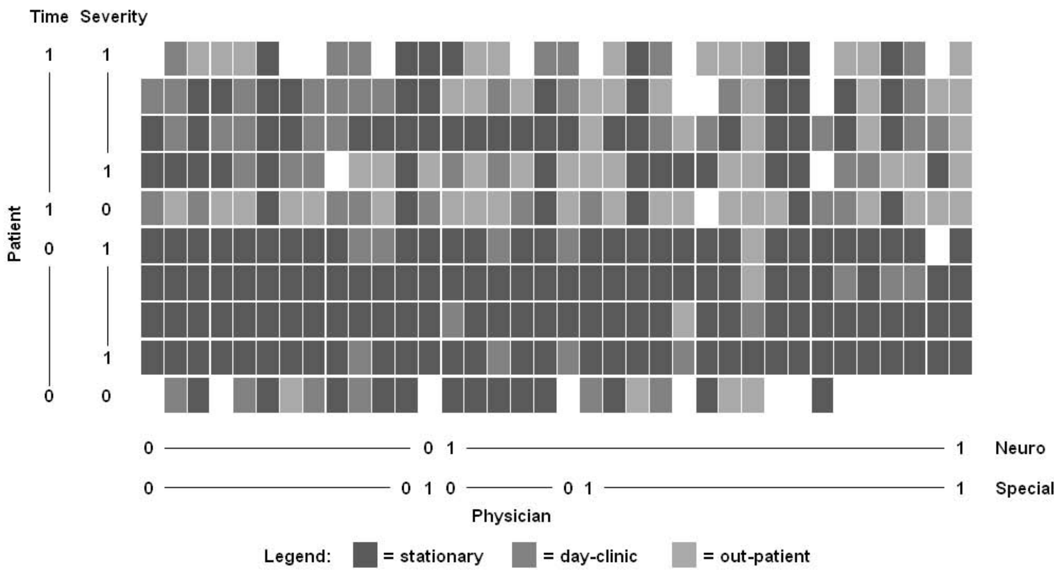

Figure 1. Physicians' recommendations for their preferred rehabilitation setting. Physicians are displayed as units on the x-axis and patients as units on the y-axis. Recommendations are given as rectangles with different shades of gray for the respective pair of physician and patient. Missing recommendations are given as white rectangles. The lines paralleling the axes give the respective values of the four covariates.

the additional information of looking separately at the two response categories. Instead, the RE Stereotype model gives a natural summary of the ordering of response categories and judges the DC category roughly in the middle ($\phi_2 = \phi_{DC} = 0.55$) between the reference category and the OP category (remember that $\phi_1 = \phi_{OP} = 1$ and $\phi_R = \phi_{ST} = 0$ are fixed by definition). The parameters for the comparison of the DC setting with the reference setting, can easily be reconstructed from the results of the RE Stereotype model by inverting the parameter restriction $\beta_r = \phi_r \beta$. Especially, $\beta_{DC,TIME} = \phi_2 \beta_{TIME} = 0.55 * 4.14 = 2.31(0.39)$ and $\beta_{DC,SEVERITY} = \phi_2 \beta_{SEVERITY} = 0.55 * -2.59 = -1.42(0.36)$, where the respective standard errors were estimated via the delta method in an additional SAS PROC NLMIXED ESTIMATE statement. As such, the reconstructed estimates from the RE Stereotype model compare closely to the direct estimates from the RE multinomial model.

4. Conclusion

We showed how Anderson's Stereotype regression model can be extended easily to account for correlated responses. The idea was to impose the nonlinear parameter restriction which relates the ordinary multinomial logistic model to the original Stereotype model to the random effects multinomial model of Hartzel et al. [6]. Proceeding that way, the RE Stereotype model

Table I. Results (estimates and respective standard errors in parentheses) from the ordinary Stereotype model (SAS PROC NLMIXED [11]), the RE multinomial model (SAS PROC NLMIXED [17]) and the RE Stereotype model (SAS PROC NLMIXED and WinBUGS[15] with non-informative priors, 100,000 runs of which 50,000 were discarded as burn-in runs, given is the posterior mean) for the TBI data set.

	Stereotype Model	RE Multinomial Model		RE Stereotype Model (NLMIXED)	RE Stereotype Model (WinBUGS)
		Fixed effects			
		OP	DC		
$\hat{\beta}_{NEURO}$	0.89 (0.46)	1.33 (0.92)	0.25 (0.73)	1.36 (0.93)	1.37 (0.98)
$\hat{\beta}_{SPECIAL}$	0.19 (0.43)	0.28 (0.86)	-0.46 (0.71)	0.40 (0.88)	0.33 (0.91)
$\hat{\beta}_{TIME}$	3.26 (0.45)	4.14 (0.59)	2.36 (0.40)	4.23 (0.58)	4.31 (0.60)
$\hat{\beta}_{SEVERITY}$	-2.00 (0.43)	-2.59 (0.53)	-1.60 (0.47)	-2.60 (0.53)	-2.63 (0.53)
$\hat{\phi}_{DC}$	0.50 (0.10)	–	–	0.55 (0.09)	0.55 (0.09)
		Random effects			
$\hat{\sigma}^2_{OP}$	–	2.61 (1.20)	–	2.62 (1.20)	3.23 (1.56)
$\hat{\sigma}^2_{DC}$	–	–	1.65 (0.80)	2.01 (0.96)	2.32 (1.14)
$\hat{\sigma}_{OPDC}$	–	1.89 (0.86)		1.94 (0.93)	2.18 (1.12)
		Model selection criteria			
AIC	517.3	499.3		487.5	–
BIC	543.9	516.7		503.3	–

becomes a nonlinear random effects model and standard theory and estimation methods apply. We used SAS PROC NLMIXED and WinBUGS for parameter estimation and results were seen to be comparable.

In terms of our motivating example we were able to identify factors which influence physicians' preferences on optimal rehabilitation setting in TBI patients. We learned that we had to account for the inherent correlation in the data but did not need the additional complexity of the RE multinomial model. Moreover, we got information about the ordering of the response categories.

In the future we are mainly interested in additional estimation techniques to judge robustness of our results, where nonparametric ML methods might be a promising candidate. For example, it is straightforward to relax the assumption of normality of the random effects distribution, Davidian/Gallant [13] and Vermunt [14] gave estimation algorithms for more general random effects distributions in non-linear mixed models.

5. Acknowledgement

We are grateful to Uwe Hasenbein and Prof. C.-W. Wallesch (Institute of Neurological and Neurosurgical Rehabilitation, Magdeburg, Germany) for providing us with the neurological background and for letting us use their data.

REFERENCES

1. Brown, H., and Prescott, R. (1999). Applied Mixed Models in Medicine. New York: John Wiley & Sons.
2. Diggle, P.J., Haegerty, P, Liang, K.-Y., and Zeger, S.L. (2002). *Analysis of Longitudinal Data*. London: Chapman & Hall.
3. Molenberghs, G., Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. New York, NY: Springer.
4. Hedeker, D., and Gibbons, R.D. (1994). A Random Effects Ordinal Regression Model for Multilevel Analysis. *Biometrics*, **50**, 933–944.
5. Tutz, G., and Hennevogl, W. (1996). Random Effects in Ordinal Regression Models. *Computational Statistics and Data Analysis*, **22**, 537–557.
6. Hartzel, J., Agresti, A., and Caffo, B. (2001). Multinomial Logit Random Effects Models. *Statistical Modelling*, **1**, 81–102.
7. Lammertyn, J., De Ketelaere, B., Marquenie,D., Molenberghs, G., and Nicola, B. M. (2003). Mixed models for multicategorical repeated response: modelling the time effect of physical treatments on strawberry sepal quality. *Postharvest Biology and Technology*, **30**, 195–207.
8. Anderson, J.A. (1984). Regression and Ordered Categorical Variables. *Journal of the Royal Statistical Society, Series B*, **46**, 1–30.
9. Greenland, S. (1994). Alternative models for ordinal logistic regression. *Statistics in Medicine*, **13**, 1665-1677.
10. McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society, Series B*, **42**, 109–142.
11. Kuss, O. (2006). On the estimation of the stereotype regression model. *Computational Statistics and Data Analysis*, **50**, 1877–1890.
12. Davidian, M., and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. London: Chapman & Hall.
13. Davidian, M., and Gallant, A.R. (1993). The nonlinear mixed effects model with a smooth random effects density. *Biometrika*. **80**, 475–488.
14. Vermunt, J.K. (2004). An EM algorithm for the estimation of parametric and nonparametric hierarchical nonlinear models. *Statistics Neerlandica*, **58**, 220-233.
15. Spiegelhalter, D. J., Thomas, A., Best, N., and Lunn, D. (2003). WinBUGS, Version 1.4 . MRC Biostatistics Unit, Cambridge, UK.
16. Hasenbein, U., Kuss, O., Bäumer, M., Schert, C., Schneider, H., and Wallesch, C.W. (2003). Physicians' preferences and expectations in traumatic brain injury rehabilitation - results of a case-based questionnaire survey. *Disability and Rehabilitation*, **25**, 136–142.
17. Kuss, O., and McLerran D. (2007). A Note on the Estimation of the Multinomial Logistic Model with Correlated Responses in SAS. *Computer Methods and Programs in Biomedicine*, in press.