ANALYSIS OF REPEATED MEASURES OF ORDINAL DATA .

R.D.MURISON

1. INTRODUCTION.

The pain ratings of data set 2 are ordered categories with 0= pain free and 3 = severe pain. McCULLAGH[4] developed regression models for such data with the cumulative frequencies of the ordered categories being related to explanatory variables by a generalised linear model (GLM). JANSEN[2] shows how the algebra of these models is simplified by the composite link functions of THOMPSON and BAKER [6]. The extension of univariate GLM's to the analysis of longitudinal data was developed by LIANG AND ZEGER[3] using quasi-likelihood generalized estimating equations (GEE).

The purpose of this paper is to give an explicit algorithm for analysing repeated measures of ordinal data using the above theories. The algebra is given in the appendix and data set 2 is used to illustrate the methods. The algorithm may be implemented in any package that handles matrices (eg. GENSTAT , Splus, MATLAB , GAUSS) and data modellers can be flexible and adapt the procedure to their own needs.

2. THE PAIN SEVERITY DATA.

The aim is to compare groups and the data can be summarized by a group $\times$ category table of frequencies such as Table 1. Whilst information on the the individuals has been lost, the groups can still be compared efficiently.

sample		pain rating				total
time	group	0	1	2	3	count
1	treat	0	2	8	6	16
	control	0	3	9	2	14
2	treat	1	9	5	1	16
	control	1	5	8	0	14
3	treat	6	4	3	1	14
	control	5	3	4	0	12
4	treat	6	3	0	1	10
	control	7	3	1	0	11

Table 1. Numbers of patients for pain severity ratings.

3. MODELLING THE VARIANCE.

The important aspect of modelling repeated measures is the representation of the variance-covariance matrix (V). The correlation amongst repeated measures of the latent variable induces correlation (ρ) and overdispersion (σ^2) amongst the counts for each category. Possible marginal variances of the counts from the i th treatment and the j th category at sample time t, y_{ijt} , with fitted values μ_{ijt} , are :-

(i) $var(y_{ijt}) = \sigma^2 \mu_{ijt}$ (scaled Poisson)

(ii) $var(y_{ijt}) = \mu_{ijt} + \sigma^2 \ {\mu_{ijt}}^2$ (negative binomial)

The matrix forms of the variances are explained in the appendix.

Initially, the samples are analysed as if they were independent Poisson variables ($\hat{\rho} = 0, \quad \hat{\sigma}^2 = 1$) and the residuals used to get updated estimates of $\hat{\rho}$ and $\hat{\sigma}^2$. This continues iteratively until convergence and procedures for doing this are given in ZEGER[7] , PRENTICE[5] and BRESLOW[1]. The first assessment of the ratings is not dependent on the subsequent measurements so $\hat{\sigma}^2$ can be set to 1 for the first sample time and estimated for the others.

To some extent, the specification of V is subjective, but minor fluctuations in V do not alter the estimates of treatment effects ($\hat{\beta}$) significantly. LIANG and ZEGER[3] give a robust variance matrix for $\hat{\beta}$ so even if the correlation structure was slightly misspecified, reasonable inference can be made with $\hat{\beta}$ using its asymptotically normal distribution.

4. RESULTS.

An autoregressive correlation with scaled Poisson variance was used for these data. The converged estimates of $\hat{\rho}$ and $\hat{\sigma}^2$ were 0.67 and 1.38 respectively and the converged estimates of the cut points ($\hat{\theta}_i$) and the treatment minus control contrast ($\hat{\beta}_1$) are shown with their standard errors in Table 2. The conclusion is that the slight initial difference between the groups had dissipitated by the second sample time. The fitted values of the distribution function (figure 1) show how the treatment effect changes over the 4 sampling times.

sample time	$\hat{\theta}_1$	*s.e.*	$\hat{\theta}_2$	*s.e.*	$\hat{\theta}_3$	*s.e.*	$\hat{\beta}_1$	*s.e.*
1	−7.02	1.15	−1.25	0.22	1.49	0.34	−0.71	0.37
2	−3.01	0.40	−0.14	0.30	3.12	1.06	0.49	0.36
3	−0.41	0.21	0.73	0.24	3.21	0.94	0.11	0.28
4	0.44	0.27	2.22	0.67	3.00	1.00	0.02	0.41

Table 2. Regression coefficients at each sampling.

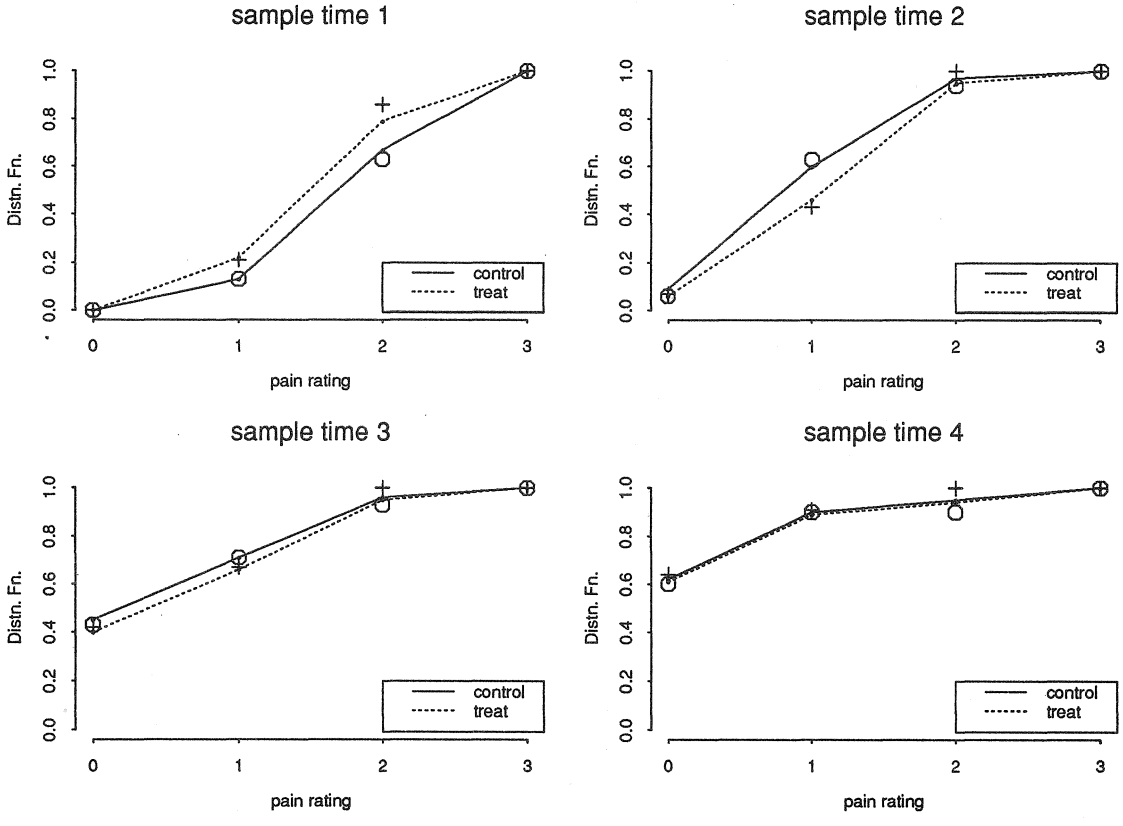

Figure 1. Fitted distribution functions of the treatment and control.

5. DISCUSSION.

The *Splus* language was used to model data set 2. It has an interface to the LINPACK algorithms which can easily handle the high dimension matrices that arise. If the matrices are too big, the modeller may use the block diagonal structure to calculate inverses or use the approximation given by ZEGER [7, p624] whereby the data are first transformed with an autoregressive filter .

6. REFERENCES.

[1] BRESLOW,N.E. (1984). Extra-Poisson Variation in Log-linear Models. *Appl. Stats.* **33**, 38-44.

[2] JANSEN,H. (1988). Using Genstat to Fit Regression Models to Ordinal Data. *Genstat Newsletter* **21**, 28-32.

[3] LIANG,K.Y. and ZEGER,S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

[4] McCULLAGH, P. (1980). Regression Models For Ordinal Data (with discussion). *J.R.Statist.Soc.Ser.B* **42**, 109-142.

[5] PRENTICE,R.L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics* **44**, 1033-1048.

[6] THOMPSON,R. and BAKER,R.J. (1981). Composite link functions in Generalised Linear Models. *Applied Statistics* **30**, 125-131.

[7] ZEGER,S.L. (1988). A regression model for time series of counts *Biometrika* **75**, 621-9.

Centre for Mathematics and its Applications
The Australian National University

7. APPENDIX.

Univariate model.

The details of this section may be found in McCULLAGH[4] , THOMPSON and BAKER[6] and JANSEN[2].

The probabilities of each group $\times$ category are given by $\pi_{ij} = F(\theta_j - \beta_i) - F(\theta_{j-1} - \beta_i)$ where θ_j is a cut point and β_i is the effect for treatment i . The parameters θ and β are combined to form the linear predictor $\eta_{ij} = \theta_j - \beta_i$ and the logistic distribution may be assumed as the latent distribution, ie. $F(\eta_{ij}) = (1 + \exp(-\eta_{ij}))^{-1}$.

At the first sampling, the model for the observed counts y and their fitted values μ is $\mu_1 = N_1 \cdot \{C \cdot F(\eta_1) + \psi\}$ where ,

$$\eta_1 = \begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{13} \\ \eta_{21} \\ \eta_{22} \\ \eta_{23} \end{pmatrix} = \overset{X}{\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}} \overset{\beta^*}{\begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ -\beta_1 \end{pmatrix}}, \quad \mu_1 = \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{14} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \\ \mu_{24} \end{pmatrix}, \quad \psi = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$N_1 = \begin{pmatrix} 16 & & & & & & & \\ & 16 & & & & & & \\ & & 16 & & & & & \\ & & & 16 & & & & \\ & & & & 14 & & & \\ & & & & & 14 & & \\ & & & & & & 14 & \\ & & & & & & & 14 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & & & & \\ -1 & 1 & 0 & & & 0 & \\ 0 & -1 & 1 & & & & \\ 0 & 0 & -1 & & & & \\ & & & 1 & 0 & 0 & \\ & 0 & & -1 & 1 & 0 & \\ & & & 0 & -1 & 1 & \\ & & & 0 & 0 & -1 & \end{pmatrix}$$

The parameter estimates are obtained by the iterative solution of the GLM by :- $(X^{*\prime}WX^*)\beta^* = X^{*\prime}WZ^*$ where $X^* = NCHX$, $Z^* = NCH\eta_1 + (y - \mu)$, $W^{-1} = V = \text{diag}(\mu_i)$ and $H = \text{diag}\left(\frac{\partial F(\eta)}{\partial \eta}\right) = \text{diag}\left(e^\eta \cdot (1 + e^\eta)^{-2}\right)$.

Repeated Measures Model.

This section is an adaption of the theory that LIANG and ZEGER[3] have given for a logistic model.

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & N_3 & \\ & & & N_4 \end{pmatrix} \left\{ \begin{pmatrix} C & & & \\ & C & & \\ & & C & \\ & & & C \end{pmatrix} \begin{pmatrix} F(\eta_1) \\ F(\eta_2) \\ F(\eta_3) \\ F(\eta_4) \end{pmatrix} + \begin{pmatrix} \psi \\ \psi \\ \psi \\ \psi \end{pmatrix} \right\}$$

Using the previous notation ,the generalised estimating equation (G.E.E.) is :-

$$\frac{\partial \mu}{\partial \beta^*}' \cdot W \cdot (\mathbf{y} - \mu) = 0$$

$$\text{ie. } (NCHX)' \cdot W \cdot (\mathbf{y} - \mu) = 0$$

The treatment effects and the cut points are found by iterative weighted least squares , $\hat{\beta}^* = (X^{*\prime}WX^*)^{-1}X^{*\prime}WZ^*$, with W,H, and Z^* being updated each iteration.

The robust variance-covariance matrix of the regression parameters is :-

$V_G(\beta^*) = (X^{*\prime}WX^*)^{-1}(X^*WSS'WX^*)(X^{*\prime}WX^*)^{-1}$ where SS' is the error sum of squares. For a sample size of N, $N^{1/2}(\hat{\beta} - \beta) \sim \text{Normal}\,(0, V_G(\beta))$ In this case, W is not a diagonal matrix but has off diagonal elements to account for the correlation amongst the samplings.

(i) $W^{-1} = V = (\Sigma M)^{1/2} \cdot R(\rho) \cdot (\Sigma M)^{1/2}$ or

(ii) $W^{-1} = V = M + (\Sigma M)^{1/2} \cdot R(\rho) \cdot (\Sigma M)^{1/2}$

$$M = \text{diag}(\overbrace{\mu_{11}\dots\mu_{24}}^{t_1}\ \overbrace{\mu_{11}\dots\mu_{24}}^{t_2}\ \overbrace{\mu_{11}\dots\mu_{24}}^{t_3}\ \overbrace{\mu_{11}\dots\mu_{24}}^{t_4})$$

$$\Sigma = \text{diag}(\ 1\dots1,\quad \sigma^2\dots\sigma^2\quad \sigma^2\dots\sigma^2\quad \sigma^2\dots\sigma^2)$$

An autocorrelation structure amongst the samples would be represented as :-

$$R(\rho) = \begin{array}{c} \\ t_1 \\ \\ \\ t_2 \\ \\ \\ t_3 \\ \\ \\ t_4 \\ \\ \end{array} \begin{array}{c} \overbrace{y_{11}\dots y_{24}}^{t_1}\ \overbrace{y_{11}\dots y_{24}}^{t_2}\ \overbrace{y_{11}\dots y_{24}}^{t_3}\ \overbrace{y_{11}\dots y_{24}}^{t_4} \\ \begin{array}{c} y_{11} \\ \vdots \\ y_{24} \\ y_{11} \\ \vdots \\ y_{24} \\ y_{11} \\ \vdots \\ y_{24} \\ y_{11} \\ \vdots \\ y_{24} \end{array} \begin{pmatrix} 1 & \ddots & 1 & \rho & \ddots & \rho & \rho^2 & \ddots & \rho^2 & \rho^3 & \ddots & \rho^3 \\ \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots \\ 1 & \ddots & 1 & \rho & \ddots & \rho & \rho^2 & \ddots & \rho^2 & \rho^3 & \ddots & \rho^3 \\ \rho & \ddots & \rho & 1 & \ddots & 1 & \rho & \ddots & \rho & \rho^2 & \ddots & \rho^2 \\ \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots \\ \rho & \ddots & \rho & 1 & \ddots & 1 & \rho & \ddots & \rho & \rho^2 & \ddots & \rho^2 \\ \rho^2 & \ddots & \rho^2 & \rho & \ddots & \rho & 1 & \ddots & 1 & \rho & \ddots & \rho \\ \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots \\ \rho^2 & \ddots & \rho^2 & \rho & \ddots & \rho & 1 & \ddots & 1 & \rho & \ddots & \rho \\ \rho^3 & \ddots & \rho^3 & \rho^2 & \ddots & \rho^2 & \rho & \ddots & \rho & 1 & \ddots & 1 \\ \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots & \ddots & & \ddots \\ \rho^3 & \ddots & \rho^3 & \rho^2 & \ddots & \rho^2 & \rho & \ddots & \rho & 1 & \ddots & 1 \end{pmatrix} \end{array}$$

Estimating $\hat{\sigma}^2$ and $\hat{\rho}$ from moments.

The papers by LIANG and ZEGER[3] , PRENTICE[5] and BRESLOW[1] provide the details for this section.

After fitting the model with initial estimates of ρ and σ^2 , the standardised residuals are given by $\hat{r}_{ijt} = (y_{ijt} - \mu_{ijt})/(\mu_{ijt})^{1/2}$. Updated estimates are found by :-

(i) $\hat{\sigma}^2 = \sum_{t=1}^{4}\sum_{j=1}^{4}\sum_{i=1}^{2}(r_{ijt})^2/(N-p)$ (scaled Poisson) or

(ii) $\hat{\sigma}^2 = \sum_t\sum_j\sum_i(y_{ijt}-\hat{\mu}_{ijt})^2/\left(\hat{\mu}_{ijt}(\hat{\mu}_{ijt}+\sigma^{-2})\right)$ (negative binomial)

and the correlation by $\hat{\rho} = \sum_t\sum_j\sum_i r_{ijt}\cdot r_{ij,t+1}\cdot\sigma^{-2}$

Starting Values.

The original data are transformed to adjust the zeroes.

$$y_{ij} = \frac{N(y_{ij}+\frac{1}{2})}{N+\frac{1}{2}}$$

$$\pi_{ij0} = \frac{y_{ij}}{N}$$

$$F_0(\eta_{ij}) = C^-(\pi_{ij0}-\psi) \text{ where } C^- \text{ is a generalised inverse of } C$$

$$\eta_{ij0} = \log\left(F_0(\eta_{ij})/\left(1-F_0(\eta_{ij})\right)\right)$$

$$\beta_0{}^* = (X'X)^{-1}X'\eta_{ij0}$$