Statistical Models for Ordinal Data in A Longitudinal Study Design

Hideki Origasa
Department of Biostatistics
The University of North Carolina
Chapel Hill, N.C. 27514, U.S.A.

SUMMARY

Two general statistical models to extend a model for ordinal data to a longitudinal study design are considered. The first one is to specify a covariance matrix to allow an autocorrelated structure. The second one is to utilize a time-series regression model. This paper presents how these two models work for each of the five different models to deal with ordinal static data. It also discusses advantages and limitations of these models, and provides other possibilities of modifying the models.

1. Introduction

Medical or epidemiologic investigations may often contain ordinal data. For example, severity of a symptom is usually evaluated by an ordered category such as mild, moderate, or severe. If we ignore such an order, then we may lose power to test a hypothesis. Thus, we need a statistical methodology to deal with ordinal categorical data. For extensive references of that, see Agresti (1984), for example.

As well, medical or epidemiologic studies sometimes adopt a design involving a time factor. One of the most typical clinical studies may be the study of natural history of a disease. Although time series analysis is appropriate to a long series of data, it may not be applied to medical data which usually have a small number of occasions, say 3 to 6 at most. Thus, we need a unique method for analyzing such data which is referred to as longitudinal data analysis. A study which produces longitudinal data is called a longitudinal study. We could solve many scientific hypotheses by way of longitudinal studies (Table 1) although they cost much more than other types of study such as a cross-sectional study. For details about the statistical aspects on both design and analysis of longitudinal studies, see Goldstein (1979), Cook and Ware (1983), for example.

Hence, we need to synthesize these two important areas, which has been hardly explored except Koch et al. (1986) and Wei et al. (1985). A schematic structure of ordinal data in longitudinal study design is presented as an example of clinical trials setting (Table 2). This

---

Key words: Longitudinal data analysis; Ordinal categorical data; Time-series regression models; Serial correlation; Covariance structure; Markov property;
Table 1
Scientific hypotheses to be solved by longitudinal studies.

1. persistence
2. structure of individual time paths
3. interaction among events with time
4. stability of multivariate relationship over time
5. detection of a crucial change in structures with time

Table 2
Longitudinal data with trichotomous responses.
(A response has a value of either "1" for mild, "2" for moderate, or "3" for severe as regards the rating of a symptom. A dot "." means a missing.)

<table>
<thead>
<tr>
<th>Group</th>
<th>Observation</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>M</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n_C</td>
<td>2</td>
<td>.</td>
<td>1</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>Drug</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n_D</td>
<td>3</td>
<td>.</td>
<td>1</td>
<td></td>
<td>.</td>
</tr>
</tbody>
</table>

Table 3
Statistical models useful for the ordinal data analysis.

1. Logistic regression model (Walker and Duncan, 1967)
2. Greenland’s model (Greenland, 1985)
3. Loglinear model (Agresti, 1984)
4. Koch’s linear model (Koch et al., 1986)
5. McCullagh’s regression model (McCullagh, 1980)
paper proposes statistical models which are appropriate to analyzing such data, as shown in Table 2. We actually consider an extension of each of the five ordinal data analysis procedures (Table 3) to longitudinal data, using two general modeling schemes to allow a serial correlation within an individual.

2. Models for a longitudinal study design

Many statistical methods assume i.i.d. with normality for an error term. Longitudinal data may not hold the assumption of independence with respect to time since the data at some occasion may correlate with those at the different occasion within an individual. We present a couple of general models for longitudinal data.

Suppose that we have a general multivariate model as

\[ y_i = X_i \beta + \epsilon_i, \]

where \( y_i \) is a \( T \times 1 \) observation vector; \( X_i \) is a \( T \times r \) matrix for covariates; \( \beta \) is an \( r \times 1 \) vector of parameters associated with \( X_i \); and \( \epsilon_i \) is a \( T \times 1 \) vector of errors for each individual \( i \). The covariance of \( \epsilon_i \), \( \Sigma \), is assumed to be \( I_T \) for an independent model. The first idea is to replace the i.i.d. assumption by a correlated intra-individual covariance structure, for example a Markov model. It might be inappropriate to specify more general covariance matrix because of having too many parameters in the covariance matrix. We should specify a particular type of covariance matrix such as a stationary Markov simplex structure (Jöreskog, 1970), that is,

\[
\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix},
\]

where \( (\sigma^2, \rho) \) is a vector of independent parameters.

With respect to the parameter estimation, if we add an assumption that \( \epsilon_i \) follows the multivariate normal distribution with mean \( \mathbf{0} \) and covariance matrix \( \Sigma \), then we obtain maximum likelihood and generalized least squares (Aitken) estimate as

\[
\hat{\beta} = (\sum_{i=1}^{N} X_i' \Sigma^{-1} X_i)^{-1} (\sum_{i=1}^{N} X_i' \Sigma^{-1} y_i).
\]

When \( \Sigma \) is unknown, an iterative procedure such as Gauss-Seidel or EM algorithm is needed (Ware and De Gruttola, 1985). The covariance of \( \hat{\beta} \) is given by

\[
\text{Cov} (\hat{\beta}) = (\sum_{i=1}^{N} X_i' \Sigma^{-1} X_i)^{-1},
\]
which turns to be consistent.

Second, we present a time-series regression model. Although it has been developed in the context of econometric theory, it has been applied in the epidemiologic research (Rosner et al., 1985). A conventional time series model requires many occasions; however, it might be sufficient for a time-series regression model to have relatively a few occasions, if we specify a smaller model such as a Markov model in which an observation is regressed by an outcome at the last occasion. Then, for a continuous response, we have

$$y_{it} = \alpha_i + \beta' x_{it} + \rho y_{i,t-1} + \epsilon_{it},$$

where $y_{it}$ is a response for the $i$th individual at time $t$; $\alpha_i$ is an intercept; $x_{it}$ is a vector of covariates; $\beta$ is a linked parameter vector with $x_{it}$; and $\epsilon_{it}$ is independently and normally distributed.

With respect to the parameter estimation, we can obtain maximum likelihood estimates (MLE's) using an iterative algorithm for nonlinear regression. For example, we might use the SAS PROC NLIN (SAS Institute, Inc., 1985) or the BMDP Programs 3R and AR (Dixon, 1983), writing down the derivatives of log-likelihood of the model (2) with respect to the parameters. The property of MLE's not only depends on the way in which the cross-sectional units $N$, or the time series observations $T$ go to infinity, but also depends on the initial conditions (Anderson and Hsiao, 1982).

3. Application of models for ordinal data to longitudinal situation

We first describe an outline of each of the five models for ordinal data. Then we present an application of the two modeling schemes mentioned in the previous section to each model for ordinal data.

3.1 Logistic regression model

Let $y_{it}$ be a dichotomous observation for the $i$th individual at time $t$. Defining $p_{it} = \Pr (y_{it} = 1) = probability of symptom occurrence for the $i$th individual at time $t$, we construct a logistic regression model as

$$\log \frac{p_{it}}{1-p_{it}} = \alpha + \beta' x_{it},$$

where $x_{it}$ is a vector of time-dependent covariates specific to the $i$th individual and $\beta$ is a vector of parameters. We may separate between time-independent and time-dependent covariates by adding $\gamma' z_i$ in (3).

An extension to the $L$ ordinal data is easy to perform. First let us define $y_{it}$ to be integer valued random variable (0 to $L-1$) and also define $w_{itj}$ as
\[ w_{itj} = 1 , \text{ if } y_{it} \text{ has a value more than } j , \]
\[ = 0 , \text{ otherwise.} \]

By letting \( p_{itj} = \Pr (w_{itj} = 1) \), we have a model as
\[
\log \frac{p_{itj}}{1-p_{itj}} = \alpha_j + \beta' \mathbf{x}_{it} , \quad j=0,1,...,L-2 ,
\]
where \( \alpha_j \) is a cut point between the \( j \)th and \( j+1 \)th categories, and so it is not an important parameter (Walker and Duncan, 1967). We are, instead, interested in the parameter \( \beta \).

This type of logistic regression modeling for ordinal data can be generalized to longitudinal data in two ways. Let \( f_{tj} \) be the frequency having the \( j \)th category or higher and \( N \) be the total number of individuals. Then, we assume
\[
\log \frac{f_{tj}}{N - f_{tj}} = \alpha_j + \beta' \bar{x}_t + \epsilon_{tj} , \quad j=0,1,...,L-2 ,
\]
where \( \bar{x}_t \) equals to \( (\sum_i x_{it})/N \). The first way is to replace an i.i.d. assumption for the error term \( \epsilon_{tj} \) by
\[
\text{Cov}(\epsilon_{tj}, \epsilon_{t'j}) = \frac{\sigma^2}{1-\rho} \rho^{t'-t} , \quad t' > t , \quad (4)
\]
which corresponds to a Markov structure although we can specify any other type of covariance matrix. The second way is to construct a Markov regression model as
\[
\text{logit} (p_{itj}) = \alpha_j + \beta' x_{it} + \rho y_{it-1} ,
\]
where the probability having the \( j \)th category or higher for the \( i \)th individual at time \( t \) depends on the outcome at the last occasion, \( y_{it-1} \). The coefficient \( \rho \) should be constant; but it could vary with an individual, or it could depend on time \( t \).

### 3.2 Greenland’s model

Greenland (1985) proposed a general statistical modeling scheme for ordinal data analysis which covers an ordinary logistic regression model as a special case of that.

Suppose that a response \( (R) \) has \( L \) ranked values. Let the conditional probability that a response has the \( j \)th category given covariates \( \mathbf{x} \) denote as
\[
\pi_j (\mathbf{x}) = \Pr (R=r_j \mid \mathbf{x}) \]
\[ = \Pr (j; \theta) , \quad j=0,1,...,L-1 . \]
The parameter \( \theta \) can be modeled as

\[
\theta = g(x; \beta),
\]

where \( \beta \) is a vector of parameters linking with covariates \( x \), and \( g \) is a function to be specified. He obtained an impressive example as follows: First let

\[
\Pr(j; \theta) = \frac{\theta_j}{\sum_{l=0}^{L-1} \theta_l / l!},
\]

and

\[
\theta = g(x; \beta) = \exp(\beta'x). \tag{6}
\]

Then, we obtain

\[
\pi_j(x) = \frac{\exp(\beta'x_j) / j!}{\sum_{l=0}^{L-1} \exp(\beta'x_l) / l!},
\]

by combining (5) and (6). To be interesting, if \( L=2 \), it exactly corresponds to an ordinary logistic regression model. He provided more general approach where he modeled \( \pi_j(x) \) directly, that is,

\[
\log\left( \frac{\pi_j(x)}{\pi_0(x)} \right) = w_j(x) = a_j + \beta'x, \quad j=0,1,\ldots,L-1,
\]

which led to

\[
\pi_j(x) = \frac{\exp(a_j + \beta'x_j)}{\sum_{l=0}^{L-1} \exp(a_l + \beta'x_l)}, \tag{7}
\]

Although it assumed a uniform trend in terms of exponential function for adjacent categories, we can specify a non-uniform trend using scores \( (y_j's) \) as

\[
\pi_j(x) = \frac{\exp(a_j + \beta'x_j y_j)}{\sum_{l=0}^{L-1} \exp(a_l + \beta'x_l y_l)},
\]

where \( y_j \) is a score for the \( j \)th category.

We extend the model (7) to a longitudinal study design as the case of logistic regression model. First, we could suppose an autocorrelated structure for the error term instead of an independent structure. Let the frequencies having the \( j \)th category at time \( t \) and
the first category at time \( t \) to be \( f_{jt} \) and \( f_{0t} \), respectively. The model is written as

\[
\frac{f_{jt}}{f_{0t}} = a_j + \beta'x_j + \epsilon_{jt}, \quad j=0,1,\ldots,L-1.
\]

Then, we can specify a similar expression to (4) for the covariance structure of errors. Second, we could construct a Markov regression model as

\[
\log \left\{ \frac{\pi_{ijt}(x)}{\pi_{i0t}(x)} \right\} = a_j + \beta'x_j + \rho \gamma_i, t-1, \quad j=0,1,\ldots,L-1,
\]

where \( \pi_{ijt}(x) \) is the probability such that the \( i \)th individual has the \( j \)th category at time \( t \), given the covariates, \( x \). The \( \gamma_i, t-1 \) is a realization or an outcome of the response at time \( t-1 \).

### 3.3 Loglinear model

Loglinear model is one of the most classical model for categorical data analysis. Suppose that we have longitudinal ordinal data by treatment group, as in Table 4. First, if we pick up an 2×3 table at a specific time point from Table 4, we can construct a model as

\[
\log m_{ij} = \mu + \lambda_i^X + \lambda_j^Y,
\]

where \( m_{ij} \) is the expected value of frequency in a \((i,j)\) cell; \( \mu \) is an overall mean; \( \lambda_i^X \) is a latent variable for treatment effect; and \( \lambda_j^Y \) is one for symptom for \( i=1(\text{Control}), 2(\text{Drug}) \) and \( j=0(\text{mild}), 1(\text{moderate}), 2(\text{severe}) \). It does not assume any ordering for the symptom. We should change a model for ordinal trichotomous data.

### Table 4

Longitudinal ordinal data reformed into a series of contingency tables. The "C" and "D" denote the control and drug groups, respectively.

<table>
<thead>
<tr>
<th>severity of symptom</th>
<th>mild</th>
<th>moderate</th>
<th>severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>C ( a_{111} ) ( a_{112} ) ( a_{113} ) ( a_{11} )</td>
<td>( a_T11 ) ( a_T12 ) ( a_T13 ) ( a_T1 )</td>
<td>( a_{T11} ) ( a_{T12} ) ( a_{T13} ) ( a_{T1} )</td>
<td></td>
</tr>
<tr>
<td>( a_{121} ) ( a_{122} ) ( a_{123} ) ( a_{12} )</td>
<td>( a_{T21} ) ( a_{T22} ) ( a_{T23} ) ( a_{T2} )</td>
<td>( a_{T21} ) ( a_{T22} ) ( a_{T23} ) ( a_{T2} )</td>
<td></td>
</tr>
<tr>
<td>( a_{11} ) ( a_{12} ) ( a_{13} ) ( a_{1} )</td>
<td>( a_{T1} ) ( a_{T2} ) ( a_{T3} ) ( a_{T} )</td>
<td>( a_{T1} ) ( a_{T2} ) ( a_{T3} ) ( a_{T} )</td>
<td></td>
</tr>
</tbody>
</table>

-7-
The model is expressed as

\[
\log m_{ij} = \mu + \lambda_i^X + \lambda_j^Y + \tau_i^X (\nu_j - \bar{\nu}) , \quad j=0,1,2,
\]

where \( \sum \lambda_i^X = \sum \lambda_j^Y = \sum \tau_i^X = 0 \), and the last term is an interaction term which includes an adjustment for ordering with a monotone increase as the later category (Agresti, 1984). With respect to an extension to longitudinal data, we can utilize the same idea as before. By adding an indice \( k \) for time factor,

\[
\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \tau_i^X (\nu_j - \bar{\nu}) ,
\]

if we assume no interactions except the one between treatment and symptom. If we replace \( m_{ijk} \) by an observed frequency \( a_{ijk} \), then

\[
\log a_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \tau_i^X (\nu_j - \bar{\nu}) + \epsilon_{ijk} ,
\]

where \( \epsilon_{ijk} \) is often assumed to be independently and normally distributed with mean 0 and a common variance \( \sigma^2 \). For longitudinal data, we may assume

\[
\text{Corr} (\epsilon_{ijk}, \epsilon_{ijk'}) = \rho^{|k-k'|} , \quad k'>k,
\]

where \( \rho \) is a correlation coefficient for a unit change of time. Another way is to construct a Markov regression model as

\[
\log a_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \tau_i^X (\nu_j - \bar{\nu}) + \rho a_{ij,k-1} + \epsilon'_{ijk} ,
\]

where \( \epsilon'_{ijk} \) is independently and normally distributed with mean 0 and common variance \( \sigma^2 \).

### 3.4 Koch’s model

Grizzle et al. (1969) developed a general linear model for categorical data analysis. Koch et al. (1977) extended it to longitudinal data. They construct a profile which is a vector of responses over the occasions for each individual and then enumerate the frequency by profile. They suppose a linear model for the vector of frequencies. It may be useful especially for relatively small number of occasions because the number of distinct profiles is \( L^d \) where \( d \) is the number of occasions and \( L \) is the number of categories. It does not consider any autoregressive structure for the components of a vector of response profiles. Each of \( L^d \) different profiles is considered as separate although they could be correlated each other.

For ordinal data, Koch et al. (1986) used to apply an equal adjacent odds-ratio technique, scoring one, and nonparametric rank measure of association, any of which is specified in the design matrix in an appropriate way. Let \( F \) be a vector of frequencies or any
function of them by the \( L^d \) profiles. The model is expressed as

\[
F = X\beta + E,
\]

where \( X \) is a design matrix where an appropriate specification for ordinal data is to be provided. They apply a weighted least squares (WLS) procedure for estimating the parameter \( \beta \). As regards an extension of Koch's model to ordinal responses, an equal adjacent odds-ratio model is valid under a very special assumption. Scoring method is also arbitrary to perform. Thus, it may be recommended to use nonparametric rank measures of association between distinct profiles which Koch et al. (1986) also mentioned. First, we have to define an appropriate measure of association and then specify the covariance matrix of \( E \) involving an autocorrelation. Thus, we can utilize a WLS procedure for parameter estimation.

Finally notice that both Koch's approach and loglinear model construct a model for marginal probabilities in the contingency table, and also all the covariates must be categorical or categorized in some way.

3.5 McCullagh's model

McCullagh (1980) proposed a general model for ordinal data analysis. Let \( Y_j(x) = \Pr( R \leq j ) \) where \( R \) stands for a random variable with ordered rank. The model is written as

\[
\text{link} ( Y_j(x) ) = \theta_j - \beta^t x,
\]

where \text{link} is any monotone increasing function from \((0,1)\) onto \((-\infty,\infty)\), for example a logistic function, that is,

\[
\frac{\gamma_j(x)}{1 - \gamma_j(x)} = \theta_j - \beta^t x,
\]

where \( x \) is a vector of covariates (explanatory variables) of interest and \( \beta \) is a vector of parameters linked with \( x \). We are more interested in the parameter \( \beta \) than \( \theta \) since \( \theta \) is a cut point between the \( j \)th and the \((j+1)\)th categories.

Wei et al. (1985) tried to extend it to longitudinal clinical trials data. They supposed a model as

\[
\zeta(\gamma_{1jk}) = \theta_{jk} + \beta_k, \quad \zeta(\gamma_{2jk}) = \theta_{jk},
\]

(8)

where \( \gamma_{1jk} = \sum_{j' \leq j} p_{ij'k} \). They actually constructed a model with the different parameter for time \( k \) (\( k=1,\ldots,T \)) and category \( j \). In the model (8), \( \beta_k \) is an increment for treatment 1. The parameter, \( \theta_{jk} \) is a nuisance parameter which represents a baseline of treatment 2 for the \( j \)th category and \( k \)th time point. However, the baseline
parameter, \( \theta_{jk} \), may be related to the baseline parameter of the \( l \)th
\( (l=0,1,\ldots,L-1) \) category and the \( (k-1) \)th time point. For example, the
severity of a symptom might depend on the severity on the previous
time. Thus, for any \( k \), we propose an additional parameterization as

\[
\theta_{jk} = \sum_{l=0}^{L-1} p_{lj} \theta_{l,k-1}, \quad j=0,1,\ldots,L-1,
\]

where \( p_{lj} \) is the probability of having severity \( j \) at time \( k \), conditional
on having severity \( l \) at time \( k-1 \).

When we take a logistic function for \( \zeta(\cdot) \) in (8),

\[
\log \left\{ \frac{\gamma_{1jk}}{1 + \gamma_{1jk}} \right\} - \log \left\{ \frac{\gamma_{2jk}}{1 + \gamma_{2jk}} \right\} = \beta_k. \quad (10)
\]

Clayton (1974) modified the model (10) using a weight \( (w_j) \) for
category \( j \) as

\[
\hat{\beta}_k = \sum_j w_j [\log \left\{ \frac{\gamma_{1jk}}{1 + \gamma_{1jk}} \right\} - \log \left\{ \frac{\gamma_{2jk}}{1 + \gamma_{2jk}} \right\}] / \sum_j w_j.
\]

If we specify these weights in an appropriate way, it could be applied
as one of more general models.

4. Discussions

As Wei et al. (1985) mentioned, there are two distinct approaches
of modeling longitudinal ordinal data. The first one is to model the
marginal probabilities which has been extensively described in this
paper. The second one is to model the probabilities for transitions
between the degrees of symptom where we can use a Markov chain or a
Markov process, for example. Let us consider a use of Markov chain
model for longitudinal ordinal data. This one is different from those
mentioned before. We first assume a time series structure, for
example, Markov chain model with \( L \) ordered categories. Then, we
define a transition probability matrix as

\[
P = \begin{pmatrix}
0 & 1 & 2 & \ldots & L-1 \\
1 & 2 & \ldots & L-1 & \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
L-1 & \vdots & \ddots & \vdots & \vdots \\
\end{pmatrix}
\]

For the binary case, \( L=2 \), let us define \( p_{ab} = \Pr (Y_{t+1} = b / Y_t = a) \),
for any time \( t \) and \( a,b=0 \) or \( 1 \). We may specify a model as

\[
p_{01} = \frac{e^\alpha}{1 + e^\alpha}, \quad p_{00} = 1 - p_{01}, \quad (11)
\]

and
where \((\alpha, \Delta)\) is a vector of parameters. Muenz and Rubinstein (1985) proposed a similar logistic model which allowed to include time-independent covariates. Returning to an ordinal case, let us define a

\[
\gamma_{ij} = \sum_{j=1}^{\infty} p_{ij}, \quad \delta_{ij} = 1 - \gamma_{ij} = \sum_{j=1}^{\infty} p_{ij},
\]

for \(j=0,1,\ldots,L-1\) and \(i=0,1,\ldots,L-1\). Since \(\gamma_{ij} + \delta_{ij} = 1\), we consider a model only for \(\gamma_{ij}\). Suppose a parameterization as

\[
\logit(\gamma_{ij}) = \log \left( \frac{\gamma_{ij}}{1 - \gamma_{ij}} \right) = \theta_{ij}, \quad j=0,1,\ldots,L-1.
\]

From a simple algebra, we obtain

\[
\gamma_{ij} = \frac{e^{\theta_{ij}}}{1 + e^{\theta_{ij}}},
\]

It actually includes a binary case, (11) and (12), by letting \(\theta_{01} = \alpha\), \(\theta_{10} = \alpha + \Delta\). We have to construct a model for \(\gamma_{ij}\). For a binary response, we need at least two parameters since only one parameter is needed for each \(i\) \((i=0,1)\). For a trichotomous case, at least two parameters are needed for each \(i\) \((i=0,1,2)\); so we need at least six parameters, and so forth.

Let us mention some comments on the advantages and/or limitations for each model. First, both logistic regression model and Greenland's model which is thought as more general one than the former one work quite well and are also interpretable. Loglinear and Koch's WLS procedure have the same limitation such that all the covariates should be categorical and time-independent. Koch et al. (1986) proposed four different ideas for ordinal data analysis, that is, equal adjacent odds-ratio model, scoring, nonparametric rank measure of association between treatment group and response, and cumulative logit model which is the same as McCullagh's regression model. The first two procedures hold a restrictive assumption for ordinal categories. For example, scores for ordinal categories must be arbitrarily assigned and the assumption of equal odds-ratio for adjacent categories may not be true. On the other hand, we can specify the most appropriate model for each problem because their WLS procedure is very general to apply. Although McCullagh's regression model works well, Fienberg criticized that loglinear or logit model would be more useful; and Koch also mentioned that his WLS procedure could be more general in the discussions of the paper by McCullagh (1980).

With respect to parameter estimation based on the model (8), McCullagh (1980) has shown an iterative reweighted least squares
method which essentially corresponds to the maximum likelihood estimation. However, Williams and Grizzle (1972) utilized a WLS, and Koch et al. (1985) proposed a two-stage procedure using both MLE and WLS. At the first stage, they obtained the MLE's of $\hat{\beta}_1, \ldots, \hat{\beta}_T$ in the expression (8) and their covariance matrix. Then, they fitted $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_T)'$ to a linear model

$$E(\hat{\beta}) = \mathbf{Z}\gamma,$$

where they utilized a WLS procedure to estimate $\gamma$. Thus, we need to compare these estimation procedures from the viewpoints of computation time and statistical precision for estimation by simulation.

Although Wei et al. (1985) have shown an extension of McCullagh’s regression model to longitudinal data, it does not take a serial correlation into account. We have revised it using an additional parameterization as (9). The assumption (9) includes a set of additional parameters, $\rho_{lj}$, for $l, j=0,1,\ldots,L-1$, which gives a problem of too many parameters in the model. Thus, we may need a simpler model for the baseline parameter $(\theta_{jk})$ in the expression (9) to consider an autocorrelation within an individual. If we assume an equally adjacent scheme over the severities of symptom, then we can write as

$$\theta_{jk} = \sum_{l=0}^{L-1} \rho^{L-|l-j|} \theta_{l,k-1},$$

where $\rho$ is the probability that severity changes by just one grade for a unit of time. This parameterization for a serial correlation can be combined into the model (8) for ordinal data.

Finally, we would urge to make a further research on a computational procedure for statistical inference since we have just proposed several statistical models for longitudinal ordinal data. We also need to investigate statistical property of the estimates, in either a large sample or a small sample from the practical viewpoint.

ACKNOWLEDGEMENT

The author is grateful to the referees for a careful reading of the first draft.

REFERENCES


[Received February, 1987]