CORRECTING FOR EFFECTS OF COMMON UNOBSERVED COVARIATES
IN INDIVIDUAL MULTIPLE LOGISTIC REGRESSION MODELS

STEFFEN L. LAURITZEN

TECHNICAL REPORT NO. 72
OCTOBER 1983

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
CORRECTING FOR EFFECTS OF COMMON OBSERVED COVARIATES IN INDIVIDUAL MULTIPLE LOGISTIC REGRESSION MODELS

by

Steffen L. Lauritzen
Aalborg University Centre
and
Stanford University

TECHNICAL REPORT NO. 72
OCTOBER 1983

STUDY ON STATISTICS AND ENVIRONMENTAL FACTORS IN HEALTH (SIMS)

PREPARED UNDER A GRANT TO SIMS FROM
ENVIRONMENTAL PROTECTION AGENCY (EPA)
SLOAN FOUNDATION
NATIONAL SCIENCE FOUNDATION (NSF)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
CORRECTING FOR EFFECTS OF COMMON UNOBSERVED COVARIATES
IN INDIVIDUAL MULTIPLE LOGISTIC REGRESSION MODELS

Steffen L. Lauritzen
Aalborg University Centre
and
Stanford University

ABSTRACT

We suggest including daily panel attack rates as covariates in individual logistic regression models for panel data in studies of the acute health effect of covariates such as pollutant levels. Such a step should give a possibility of correcting for the effect of unobserved covariates that are common to individuals. We give some hints to try to give the procedure a meaning.
1. INTRODUCTION

The present note takes its origin in a particular study of the possible effect of pollution variables on asthma attacks in panel data from the Los Angeles area as described by Korn and Whittemore (1979) and Whittemore and Korn (1980). Both methodological and medical issues are discussed in those papers. A panel of persons (asthma patients) are followed in a period, each of them recording in a diary whether or not they had an asthma attack on any given day. Simultaneously, on any given day, various pollutant levels as well as other covariates are observed.

The analysis presented is based on a logistic multiple regression model for each individual, assessing the influence of the covariates on the health (asthma attacks) of that individual. It is argued (and rightly so) that this "individual" method avoids many biases pertaining to the more traditional method of regressing daily panel attack rates on the covariates in question.

In studies of this type (and in all studies concerned with assessing the influence of covariates on observables) it is a severe problem that the presence of important unobserved covariates may contaminate findings. This difficulty cannot be overcome, but we present an idea in the following which to some extent does correct for the undesired effects of such unobserved covariates. The idea is based on the fact that the daily panel attack rate can be considered as a (crude) measure of the combined influence of all covariates on the attack propensity. Thus by including the daily attack rate as a covariate some correction for this effect is obtained. We shall in the following try to give more precise meaning to this procedure.
2. FITTING LOGISTIC REGRESSIONS WITH DAILY ATTACK RATE AS A COVARIATE

The standard individual logistic regression model described by Korn and Whittemore (1979) is as follows: First we define the function for \( \theta \in ]0,1[ \)

\[
\text{logit } \theta = \log\left( \frac{\theta}{1-\theta} \right).
\]

We then let \( \pi_{it} \) be the conditional probability that person \( i \) has an attack on day \( t \) given the history up to and until time \( t-1 \) and assume

\[
\text{logit } \pi_{it} = \sum_{\nu=1}^{k} \alpha_{i\nu} x_{it}^\nu
\]

where \( x_{it}^\nu \) are covariates that either can be observed directly or are functions of the history of the process up to and until time \( t-1 \). \( \alpha_{i\nu} \) are individual parameters. In the particular study the covariates are pollutant levels, day of week indicators, etc. and, the seemingly most important of them all is the indicator of whether or not the person in question had an attack on the day before.

Further the attacks of the various persons in the panel on any given day are considered conditionally independent given the covariates and the past, such that persons can be analyzed individually. Assuming that the person \( i \) is a part of the panel from time \( t_i \) until time \( T_i \) and that the covariates do not depend further back than one day, each individual likelihood function becomes

\[
L(\alpha_i) = L(\alpha_{i1}, \ldots, \alpha_{ik})
\]

\[
= \prod_{t=T_i+1}^{T_i} \left[ \sum_{\nu=1}^{k} \frac{e_{i\nu} x_{it}^\nu z_{it}}{1 + e_{i\nu} x_{it}^\nu} \right]
\]

\[
= C_i(\alpha_i) e_{i\nu} \sum_{t=t_i+1}^{T_i} \left[ \sum_{\nu=1}^{k} \alpha_{i\nu} x_{it}^\nu z_{it} \right]
\]
where \((z_{it}, \ldots, z_{iT_i})\) is the attack vector of person \(i\). Standard
maximum likelihood theory gives estimates of \(\alpha_i\) determined by the
equations

\[
\sum_{t=t_i+1}^{T_i} \pi_{it}(\alpha_i) = \sum_{t=t_i+1}^{T_i} x_{it} z_{it} \quad \forall = 1, \ldots, k
\]
\[
i = 1, \ldots, N
\]

where \(N\) is the number of persons in the panel and

\[
\pi_{it}(\alpha_i) = e^{\sum_{\forall=1}^{k} \alpha_{iv} x_{it}} / (1 + e^{\sum_{\forall=1}^{k} \alpha_{iv} x_{it}})
\]

cf. e.g. Cox (1970).

Letting now \(N(t)\) be the number of persons actually in the panel at
time \(t\) and \(S(t)\) the set of persons in the panel, we define daily panel
attack rates as

\[
\lambda_t = \sum_{j \in S(t)} z_{jt} / N(t) = \bar{z}_t
\]
\[
\lambda_t^{(i)} = \sum_{j \in S(t) \setminus \{i\}} z_{jt} / (N(t) - 1) = \bar{z}_t^{(i)}
\]

\(\lambda_t\) is the crude daily panel attack rate and \(\lambda_t^{(i)}\) is that obtained by
excluding person \(i\) from the panel. If \(N(t)\) is large, \(\lambda_t\) and \(\lambda_t^{(i)}\)
are equal for any practical purpose. The "methods" that we shall dis-
cuss are obtained by modifying the equations (2.1) to include an extra
covariate \(x_{it}^{k+1}\) as e.g.

A. \(x_{it}^{k+1} = \lambda_t\cdot -1\)
B. \(x_{it}^{k+1} = \lambda_t\)
C. \(x_{it}^{k+1} = \lambda_t^{(i)}\)
and solve the equations

\[ T_i \sum_{t=t_{i+1}}^T x_{it} \tilde{\pi}_{it}(\alpha_i, \beta_i) = T_i \sum_{t=t_{i+1}}^T x_{it} z_{it} \quad \nu = 1, \ldots, k + 1 \]

\[ i = 1, \ldots, N \]

or

\[ T_i \sum_{t=t_{i+1}}^T x_{it} \tilde{\pi}_{it}(\alpha_i, \beta) = T_i \sum_{t=t_{i+1}}^T x_{it} z_{it} \quad \nu = 1, \ldots, k \]

\[ i = 1, \ldots, N \]

\[ \sum_{i=1}^N \sum_{t=t_{i+1}}^N x_{k+1} \tilde{\pi}_{it}(\alpha_i, \beta) = \sum_{i=1}^N \sum_{t=t_{i+1}}^N x_{k+1} z_{it} \]

where we now let

\[ \tilde{\pi}_{it}(\alpha, \beta) = e^{k \sum_{\nu=1}^k \alpha_{i} x_{\nu}^{i,t} + \beta_{i} x_{it}^{k+1}} / (1 + e^{k \sum_{\nu=1}^k \alpha_{i} x_{\nu}^{i,t} + \beta_{i} x_{it}^{k+1}}) \]

So far we have given no interpretation of the method(s) other than as fitting procedures. It is also clear that other monotone functions of the daily attack rates could be used as well, such as for example \( \text{logit}(\lambda_t) \), etc.

The estimates obtained by using \( A \) and (2.2) or (2.3) are at least well defined as maximum likelihood estimates in a logistic regression model since the covariate \( \lambda_{t-1} \) is a function of the past history of the process going back in time by only one step. The model corresponding to that method might not be awfully sensible since \( \lambda_{t-1} \) is not as good an estimate of the combined effect of the covariates as \( \lambda_t \), and if drop-outs and drop-ins of patients to and from panel occur in great deal, it is unpleasant that the model seems to depend a lot on the set of persons under study. Thus we might at least suggest that the procedure could have
serious biases if the panel does not remain reasonably constant in the study period.

The most tempting procedures seem to be based on A or B and (2.2) or (2.3), or functions of the attack rates. The methods so derived do not in any exact sense correspond to estimation procedures in well defined statistical models, because the "covariates" are functions of the observations. This makes criticism of the methods based on judging the relevance of models impossible. In the following we shall try to indicate how these methods can be thought of as approximations to procedures derived from models, to repair some of these difficulties. Biases due to the fact that daily attack rates depend on the composition of the panels can occur also in these cases and if the panel is not approximately constant in time periods, the methods seem quite questionable.

Before we proceed to investigate some particular models we can make a few guesses as to the possible good and bad effects of the procedure.

Some correction for unobserved covariates can possibly be obtained. This, however, concerns only such covariates that are common to all individuals and not unobserved covariates that are different for different individuals. In fact, the introduction of the indicator of attack on the previous day can be thought of as a method correcting for unobserved individual covariates.

On the other hand, one might expect that some of the effect of the covariates observed could be hidden by the introduction of the daily attack rate as a covariate. In fact, it is to be expected that only such effects of covariates that are different for different individuals remain significant when daily attack rates are included as covariates, especially when these are very well predicted by the covariates observed. To take
care of this fact, the analysis involving some kind of regression of daily attack rates on covariates, the whole procedure could even be modified by considering residuals from this regression as covariates instead of the crude attack rates themselves. This also gives a possibility of inferring which function of the daily attack rate that is linearly related to the covariates observed.

In any case, it seems to have some interest to investigate the suggested methods, and in the following we shall describe a few models having the same flavor and see how they relate to the methods suggested.

To avoid problems connected with drop-in and drop-out of patients we shall in the following assume that the panel remains constant throughout the period of study and therefore we shall consider the equivalent covariates

\[ z_{it} = \sum_{i=1}^{N} z_{it}^{(i)} \quad z_{t} = z_{t}^{*} - z_{it} \]

instead of \( \lambda_{t} \) and \( \lambda_{t}^{(i)} \). Further we might as well assume \( t_{i} = 0 \) and \( T_{i} = T \). The right hand side of (2.2) then simplifies to (in the case \( \nu = k+1 \))

\[
\sum_{t=1}^{T} z_{t}^* z_{it} = \sum_{j=1}^{N} \sum_{t=1}^{T} z_{it} z_{jt} = s_{i*} = \sum_{j=1}^{N} s_{ij}
\]

where \( s_{ij} = \sum_{t=1}^{T} z_{it} z_{jt} \) is the number of days where person \( i \) and \( j \) both suffered an attack. Similarly the right hand side of the last equation of (2.3) becomes
\[ s_{i} = \sum_{i} \sum_{j} s_{ij}. \]

Or, maybe even simpler

\[ \sum_{i} \sum_{j} \sum_{t} z_{it} z_{jt} = \sum_{t} N(t)^2. \]

The corresponding right hand sides of (2.2) and (2.3) in the case where \( z_{*t}^{(i)} \) is used become respectively

\[ \sum_{t} z_{*t}^{(i)} z_{it} = \sum_{t} \sum_{j} z_{jt} z_{it} - \sum_{t} z_{it} \]

\[ = s_{i} - z_{i}. \]

and

\[ \sum_{t} \sum_{i} z_{*t}^{(i)} z_{it} = \sum_{t} \sum_{i} (z_{*t}^{(i)} - z_{it}) z_{it} \]

\[ = \sum_{t} N(t)^2 - \sum_{i} \sum_{t} z_{it}^2 \]

\[ = \sum_{t} N(t) (N(t) - 1). \]

In all cases we notice that the methods have the flavor of introducing a correlation between persons in the panel. This is the basis for the next two attempts of formulating a model.

3. A PAIR POTENTIAL MODEL

An obvious attempt to explain the expressions (2.2) and (2.3) through well defined models is to interpret the probabilities conditionally on the total number of attacks (person i excluded or not). To pursue that further, let us concentrate at a particular point in time \( t_0 \) and suppress the corresponding index to ease notation. Let also
\[ \delta_i = \sum_{v=1}^{k} \alpha_{iv} x_{it0}, \quad N_i = z^{(i)} t_0, \quad z_i = z_{it0} \]

and let us attempt to write

(3.1) \[ \text{logit } P(Z_i = 1|Z_j = z_j, j \neq i) = \delta_i + \beta_i N_i \]

\[ \quad = \delta_i + \beta_i \sum_{j \neq i} z_j. \]

It is by no means clear that this makes sense, i.e. that a probability exists that has these conditional logits. In fact we shall show that this is not the case unless \( \beta_i = \beta \). To see this it shall be convenient for us to change notation and describe the vector of attacks \( \mathbf{z} = (z_1, \ldots, z_N) \) by the subset \( A \) of persons having an attack. We can now specify a joint probability of the attack vector \( \mathbf{z} \) by specifying the probability of any attack set \( A, A \subseteq \{1, \ldots, N\} \). We obtain further by simple calculations

\[
\text{logit } P(Z_i = 1|Z_j = z_j, j \neq i) \\
= \log(P(Z_i = 1|Z_j = z_j, j \neq i)/P(Z_i = 0|Z_j = z_j, j \neq i)) \\
= \log(P(Z_i = 1 \land Z_j = z_j, j \neq i)/P(Z_i = 0 \land Z_j = z_j, j \neq i)) \\
= \log(P(A \cup \{i\])/P(A))
\]

where \( A \) is the subset of \( \{1, \ldots, N\}\{i\} \) describing the attack set of persons when \( i \) is excluded. (3.1) can now be formulated as

(3.2) \[ \log(P(A \cup \{i\})/P(A)) = \delta_i + \beta_i |A| \]

where \( |A| \) denotes the number of elements in \( A \). We then get for a set \( \{1, \ldots, p\} \) that
\[
\log P(\{1, \ldots, p\}) = \sum_{j=1}^{p} \log(P(\{1, \ldots, j\})/P(\{1, \ldots, j-1\})) + \log P(\phi)
\]
\[
= \log P(\phi) + \sum_{j=1}^{p} \delta_j + \sum_{j=1}^{p} \beta_j (j-1).
\]

Using the same argument with the indices permuted in an arbitrary fashion we get
\[
\sum_{j=1}^{p} \beta_j (j-1) = \sum_{j=1}^{p} \beta_{\pi(j)} (j-1)
\]
for any permutation \( \pi \) of \( \{1, \ldots, p\} \), which in turn implies \( \beta_j = \beta \).

If we on the other hand define
\[
(3.3) \quad P(A) = K^{-1}(\delta, \beta) \exp\left\{ \sum_{i \in A} \gamma_i + \beta\left(\frac{|A|}{2}\right)\right\}
\]
where
\[
K(\delta, \beta) = \sum_{A \subseteq \{1, \ldots, N\}} \exp\left\{ \sum_{i \in A} \delta_i + \beta\left(\frac{|A|}{2}\right)\right\}
\]
we do get a perfectly well defined probability and in fact
\[
\log(P(A \cup \{i\})/P(A)) = \delta_i + \beta\left(\frac{|A|+1}{2}\right) - \left(\frac{|A|}{2}\right)
\]
\[
= \delta_i + \beta |A|.
\]

Thus one way of interpreting the method given by (2.3) is a pseudo-likelihood estimation procedure patching together likelihood functions based on the conditional distribution of the response of each individual given the past and the responses by other individuals. This then corresponds to the case \( C_{1} \), where the number of individuals among the remaining patients is used as a covariate. It is conceivable that the corresponding exact maximum procedure would perform slightly better.
The joint likelihood would now be equal to (if we let \( \delta_{it} = \sum_{v=1}^{k} \alpha_{iv} x_{it}^v \))

\[
L(\alpha, \beta) = \prod_{t=1}^{T} P(A_t | \text{past})
\]

\[
= e^{\sum_{t=1}^{T} \left( \sum_{v=1}^{k} \alpha_{iv} x_{it}^v z_{it} + \beta \binom{z_{it}}{2} \right)} \prod_{t=1}^{T} K^{-1}(\delta_t, \beta)
\]

\[
= K(\tilde{\alpha}, \beta) e^{\sum_{v=1}^{k} \alpha_{iv} \sum_{t=1}^{T} x_{it}^v z_{it} + \beta \sum_{t=1}^{T} \binom{z_{it}}{2}}
\]

which again leads to an exponential family problem. The actual maximization of the above likelihood can cause some problems in terms of computation time, etc. although the problems seem solvable. Under all circumstances the procedure described by (2.3) seems to be an approximate likelihood method in this model. The model can easily be criticized by claiming that the specific form of the term involving \( z_{it} \) is arbitrary, speculative, and not very well founded. But at least the last remark also then applies to the special case with \( \beta = 0 \), which is exactly the model where persons are considered independent. And the use of other specific functions of \( z_{it} \) than the binomial coefficient. Just that then (2.3) might not be interpretable as an approximate pseudo-likelihood method in that model but the likelihood itself would have to be maximized.

The model just discussed is a special case of the following pair potential model, borrowing a terminology from statistical mechanics and Markov random fields, see e.g. Preston (1973).

A pair potential is a real valued symmetric function of two variables. We denote its values by \( U_{ij} \), \( i=1, \ldots, N; j=1, \ldots, N \). The Gibbs probability with potential \( U \) is then defined as
\[ P_U(A) = K^{-1} \cdot \exp\left( \sum_{i \in A} \sum_{j \in A} U_{ij} \right) \]

where \( K = \sum_{A \subseteq \{1, \ldots, N\}} \exp\left( \sum_{i \in A} \sum_{j \in A} U_{ij} \right) \) is the normalizing constant.

If we now consider the special case

\[ U_{ij} = \frac{1}{2}(\beta_i + \beta_j) \text{ if } i \neq j \]

\[ U_{ii} = \delta_i \]

we get for the joint distribution of the attack set on day \( t \) that

\[ \log P_U(A) = \sum_{i \in A} \delta_i + \frac{1}{2} \sum_{\text{pairs of elements of } A} (\beta_i + \beta_j) - \log K^{-1}. \]

In the special case \( \beta_i \equiv \beta \) we obtain

\[ \log P_U(A) = \sum_{i \in A} \delta_i + \beta \left( \frac{|A|}{2} \right) - \log K^{-1}, \]

i.e. exactly (3.3). The more general model which is perfectly well defined has as logits

\[
\text{logit } P(Z_i = 1|Z_j = z_j, j \neq i) = \log(P(A \cup \{i\})/P(A))
\]

\[ = \delta_i + \frac{1}{2} \beta_i |A| + \frac{1}{2} \sum_{j \in A} \beta_j \]

\[ = \delta_i + \frac{1}{2} \beta_i z^{(i)} + \frac{1}{2} \sum_{j \neq i} \beta_j z_{jt}. \]

This also in an obvious way suggests a modified logistic regression procedure although the simplicity of (2.2) and (2.3) seems lost.

If we return to the simple case, i.e. where

\[ U_{ii} = \delta_i, \quad U_{ij} = \beta, \quad i \neq j \]
we note that the distribution of the number of persons having an
attack is given as

$$P(|A| = s) = K^{-1} e^{\beta(s)} \gamma_s(e^\delta)$$

where $e^\delta = (e_1^\delta, \ldots, e_N^\delta)$ and $\gamma_s$ is the elementary symmetric
function

$$\gamma_s(\lambda_1, \ldots, \lambda_N) = \sum_{x_1 + \ldots + x_N = s} \frac{x_1^{\lambda_1} x_2^{\lambda_2} \ldots x_N^{\lambda_N}}{x_i \in \{0,1\}}$$

If we then continue to calculate the conditional distribution of the
attack set given the total number of attacks we obtain

$$(3.4) \quad P(A|N_t=s) = \begin{cases} 0 & \text{if } |A| \neq s \\ \sum_{\delta_i} e^{\delta_i} \gamma_s(e^\delta) & \text{if } |A| = s \end{cases}$$

and observe the interesting fact that this does not depend on $\beta$! This
phenomenon and the distributions above are well known from the Rasch
model for intelligence tests, see Rasch (1960, 1980). It gives us the
potential possibility of letting $\beta$ depend on $t$ and leads to yet
another "explanation" of the regression methods discussed in Section 2.

4. A RANDOM ENVIRONMENT MODEL

Consider a real-valued stochastic process $\theta = \{\theta_t, t=0, \pm 1, \pm 2, \ldots\}$ supposed to describe the impact of the environment on the
attack pattern in the panel at time $t$ and assume the model
\[ P(A_t | A_0, \ldots, A_{t-1}, \theta) = K^{-1}(\theta_t, \delta_t) \exp\{ \sum_{i \in A_t} \delta_{it} + \theta_t | A_t | \} . \]

This clearly corresponds to assuming that the attacks of different persons at time \( t \) are conditionally independent given the past and the "environment" at time \( t \) with corresponding conditional logits

\[ \text{logit} \, \pi_{it}(\alpha, \theta) = \delta_{it} + \theta_t \]

\[ = \sum_{\nu=1}^{k} \alpha_{i\nu} x_{it}^{\nu} + \theta_t . \]

In other words, there is just at each point in time a shift \( \theta_t \) of the logit of all person probabilities due to unobserved environmental factors.

One could assume a specific distribution for \( \theta_t \), e.g. of the type of a regression on observed covariates with stationarily autocorrelated errors, but also the completely nonparametric specification of the distribution of \( \theta_t \) has interest, in which case \( \theta_t \) can be considered just an unknown sequence of nuisance parameters and "varying environment" might then be a more appropriate name of the model. In the case where the only observed covariate is the constant \( x_{it}^{\nu} \equiv 1 \) we obtain

\[ \text{logit} \, \pi_{it}(\alpha, \theta) = \alpha_i + \theta_t \]

which exactly is the Rasch model.

Calculating the conditional distribution of an attack set given the number of persons having the attack, we obtain

\[
(4.1) \quad P(A_t | N_t = s) = \begin{cases} 0 & \text{if } |A_t| = s \\ \sum_{i \in A_t} \delta_i / \gamma_s(e^\delta) & \text{otherwise} \end{cases}
\]
which does not depend on $\theta$ and exactly is equal to (3.4).

Rasch (1960, 1980) argued that the nuisance parameters $(\theta_t)$ will destroy the performance of the maximum likelihood estimates and suggested a conditional procedure based on (4.1) to estimate the "row effects."

From (4.1) we obtain (conditionally on the past)

$$P(Z_{it} = 1|Z_{*t} = s) = e^{\delta_{it}} \gamma_{s-1}(e_t/e_i)^{\delta_{it}}/\gamma_s(e_t)$$

where $\gamma_s(\lambda_i)$ is the elementary symmetric function of the $\lambda$-vector obtained by omitting $\lambda_i$ and as usual $\delta_{it} = \sum_{v=1}^{k} \alpha_{1v} \lambda_{vt}$.

Further, the conditional logits become

$$\logit P(Z_{it} = 1|Z_{*t} = s)$$

$$= \log(P(Z_{it} = 1 \land Z_{t} = s-1)/P(Z_{it} = 0 \land Z_{t} = s))$$

$$= \delta_{it} + \log \frac{\gamma_{s-1}(e_t/e_i)^{\delta_{it}}}{\gamma_s(e_t/e_i)}$$

In the special case where $\delta_{it} = \delta_t$ we obtain

$$\logit P(Z_{it} = 1|Z_{*t} = s) = \logit \frac{s}{n}$$

$$= \log \frac{s}{n-s}$$

which seems to suggest that to approximate the conditional maximum likelihood procedure by introducing an extra covariate should use the logit of the attack rate rather than the attack rate itself. A detailed investigation of this demands approximate expressions for the elementary symmetric functions $\gamma_s(\lambda)$. Such expressions have been given by Johansen (1966). See also Martin-Löf (1970).
In fact we have approximately (for small $\delta_{it}$) that

$$
\gamma_s(e^{\delta_{it}}) = \left(\begin{array}{c} N-1 \\ s \end{array}\right) \frac{1}{N-1} \sum_{j \neq i} \delta_{jt}
$$

whereby

$$
\text{logit } P(Z_{it} = 1 | Z_{*t} = s_t) \approx \delta_{it} - \frac{1}{N-1} \sum_{j \neq i} \delta_{jt} + \log \frac{s_t}{N-s_t}
$$

$$
= \delta_{it} (1 + \frac{1}{N-1}) - \frac{N}{N-1} \delta_{*t} + \text{logit } \lambda_t
$$

implying that the conditional procedure approximately corresponds to including $\text{logit } \lambda_t$ as a covariate with coefficient equal to one.

Simultaneously we see that if individuals are equal ($\delta_{it} = \delta_{*t}$), the effects of observed covariates "disappear" and have to be detected from an analysis of $\lambda_t$.

An exact conditional procedure is now obtained from standard exponential family theory cf. e.g. Barndorff-Nielsen (1978) or Martin-Löf (1970) by solving the system of equations

$$(4.2) \quad \sum_{t=1}^{T} x_{it} \hat{\pi}_{it}(\alpha) = \sum_{t=1}^{T} z_{it} x_{it}, \quad \nu = 1, \ldots, k$$

$$i = 1, \ldots, N$$

where now $\hat{\pi}_{it}(\alpha)$ is given by

$$
\text{logit } \hat{\pi}_{it}(\alpha) = \sum_{\nu=1}^{k} \alpha_{i\nu} x_{it} + \log \frac{1}{\gamma_s(e^{\delta_{*t} \delta_{it}})}
$$

The "only" problem with this procedure is that it is computationally problematic since evaluation of the elementary symmetric functions is rather time consuming.
Given $\tilde{\alpha}$ and the past, the distribution of $z_{t*}$ is given as

$$(4.3) \quad P(z_{t*} = s) = e^{\theta_t s} \gamma_s(e^{\delta_t}) \prod_{i=1}^{N} \frac{\delta_{it} + \theta_t}{1 + e^{\delta_{it} + \theta_t}}^{-1}$$

and the following procedure seems now to have potential possibilities:

1) Estimate $\tilde{\alpha}$ (and thereby $\tilde{\delta}_t$) by solving equations (4.2) to obtain values $\hat{\delta}_t$.

2) Work on (4.3) as if $\delta_{it}$ were known and equal to $\hat{\delta}_{it}$ to obtain estimates of $\theta_t$ by maximum likelihood, i.e. determine $\hat{\theta}_t$ by the equation

$$\sum_{s=0}^{N} \hat{\theta}_t s \gamma_s(e^{\hat{\delta}_t}) \prod_{i=1}^{N} \frac{\delta_{it} + \hat{\theta}_t}{1 + e^{\delta_{it} + \hat{\theta}_t}}^{-1} = z_{t*}.$$ 

3) Investigate now the relation between $\hat{\delta}_t$ and the covariates that are not individually different to obtain information about the overall effect of the observed covariates on the attack frequency of individuals, possibly by regression methods.

5. **FINAL COMMENTS**

Note that all of the methods sketched give a potential possibility of testing whether or not common unobserved covariates have an influence on the conclusions by introducing extra parameters and testing these to be zero. All methods should be used with special care if the panel is highly variable over time. We hope to have argued that methods of the above type at least deserve some further study and note that in principle they can be extended to other generalized linear models than the logistic regression, in particular to regressions in the normal distribution where one can hope for more explicit expressions for the terms involved.
REFERENCES


