Comparison of the Anderson-Gill approach, poisson and negative binomial regression on simulated recurrent event data with intra-subject correlation

Antje Jahn

20th November 2006
- Recurrent Event Data
- Counting Process Style of Survival Analysis
- Statistical Analysis Methods
- Simulations
- Results
- Conclusions
Recurrent Event Data

A subject can experience the same event more than once. 

Example: Occurrence of otitis media after vaccination

Problems:

- Censored observations
- Events of the same subject are dependent (intra-subject correlation)
- Risk may be affected by previous events
Antje Jahn

Counting Process Style

i is indicating subjects

\[ N_i(t) = \text{number of observed events in } [0,t] \text{ of subject } i \]

\[ C_i = \text{time of censoring} \]

\[ Y_i(t) = 1(t \leq C_i) \]

\[ \alpha_i(t) = \lim_{h \to 0} \frac{P(N_i(t) - N_i(t - h) = 1| \text{sbj } i \text{ at risk at time } t)}{h} \]

in analogy to the hazard rate

\[ \lambda_i(t) = Y_i(t)\alpha_i(t) \text{ the intensity process of } N_i(t) \]

\[ \Rightarrow N_i \text{ is poisson process with rate } \alpha_i(t) \text{ (unstratified approach)} \]
Counting Process Style

Cox proportional hazard model:

\[ \alpha_i(t) = \alpha_0(t) \exp(X_i \beta) \]

Mixed models: \( Z_i \) iid with \( E(Z_i) = 1 \)

\[ \alpha_i(t) = Z_i \alpha_0(t) \exp(X_i \beta) \]
Statistical Analysis: Anderson-Gill approach

Underlying model:

\[ \alpha_i(t) = \alpha_0(t) \exp(\beta X_i) \]

Maximum Likelihood Estimate

\[ l(\beta) = \prod_{i=1}^{n} \prod_{k=1}^{K_i} \left( \frac{\exp(\beta X_i)}{\sum_{l=1}^{n} Y_l(T_{ik}) \exp(\beta X_l)} \right)^{\delta_{ik}} \]

(1)
Robust variance estimation by Grouped Jackknife:

\[ \hat{\beta}_{-i} = \text{ML-Estimator of } \beta \text{ omitting observations of subj } i \]
\[ J_i = \hat{\beta}_{-i} - \hat{\beta} = \text{i-th row of matrix } J \]
\[ \bar{J} = \text{matrix of column means of } J \]

Define the grouped jackknife variance estimator of \( \text{Var}(\hat{\beta}) \) as

\[ \hat{\text{Var}}(\hat{\beta}) = (J - \bar{J})'(J - \bar{J}) \]  
(2)

(Jackknife variance estimators are robust against many departures of the cox proportional hazard assumption, if omitted observations are independent of the included.)
Statistical Analysis: Poisson Regression

Underlying Model: \[ \alpha_i(t) \equiv \alpha_0 \exp(X_i \beta) = \mu_i \]

Poisson process as generalized linear model

\[ f_{N_i}(y; \theta_i, \phi) = \exp\left\{ (y \theta_i - b(\theta_i))/a(\theta_i) + c(y, \phi) \right\} \]

with \( \mu_i = E(N_i) \) \( \theta_i = \log(\mu_i) \) \( a \equiv 1 \) \( b = \exp \)

\( \phi \) is called dispersion parameter and is 1 for the poisson model.

\[ \Rightarrow \text{Var}(N_i) = \exp(\theta_i)\phi = \mu_i\phi = \mu_i \]

Overdispersion \((\phi > 1)\):

\[ \Rightarrow \text{Var}(N_i) > \mu_i \]

\[ \Rightarrow \text{corrected } \hat{\text{Var}}(N_i) = \hat{\phi}\hat{\mu}_i \]

\[ \Rightarrow \text{corrected } \hat{\text{SE}}(\beta) = \hat{\phi}^{1/2}\hat{\text{SE}}(\hat{\beta}) \]
Statistical Analysis: Negative Binomial Regression

Underlying model:

\[ \alpha_i(t) \equiv Z_i \alpha_0 \exp(X_i \beta) = Z_i \mu_i \text{ with } Z_i \sim \Gamma[1/a, a] \]

Then

\[ N_i | Z_i \sim \text{POI}(Z_i \mu_i) \Rightarrow N_i \sim \text{NB}(\mu_i, a) \]

with NB negative binomial distribution, i.e.

\[
\begin{align*}
E(N_i) &= \mu_i \\
\text{Var}(N_i) &= \mu_i + \frac{\mu_i^2}{a}
\end{align*}
\]

Statistical analysis using generalized linear model theory
All methods require independent increments (Poisson process).

Anderson-Gill analysis also fits to non-homogenous Poisson processes.

Negative binomial approach naturally accommodates intra-subject correlation.
Simulations

- Simulation of event counts for a hypothetical randomized controlled trial
- 50 subjects were randomized, half to treatment group \((x_i = 1)\), half to placebo \((x_i = 0)\).
- 1000 simulations are performed
  - under H0 \((\beta = 0)\) and H1 \((\beta = -0.3)\)
  - for each model...
Simulations: The Models

Homogenous poisson model

\[ \alpha_i(t) = z_i \exp(\beta x_i) \quad (\text{conditional on } Z_i = z_i) \]

Non-homogenous poisson model

\[ \alpha_i(t) = \begin{cases} 
z_i \exp(\beta x_i) & \forall t \in [0, 2] \\
z_i \exp(\beta x_i + \ln(2)) = 2 \cdot z_i \exp(\beta x_i) & \forall t \in ]2, \infty] 
\end{cases} \]

Conditional model

\[ \alpha_{i1}(t) = z_i \exp(\beta x_i) \]
\[ \alpha_{ik}(t) = z_i \exp(\beta x_i + 2) = z_i \exp(2) \cdot \exp(\beta x_i), \quad k \geq 2 \]
Simulations: Correlations and Censoring

Modeling of correlations

\[ Z_i \sim \Gamma(2, \frac{1}{2}) \]

\[ \Rightarrow E(Z_i) = 1, \ Var(Z_i) = \frac{1}{2} \]

Modeling of censoring

\[ C_i \sim \text{MIN}(\text{MAX}(N(4, 2^2), 0.1), 8) \]
Conclusions

Homogenous and non-homogenous poisson models:
- Results of different methods are comparable with acceptable type I error and coverage
- Type I error for AG and Poisson regression can increase to unacceptable levels with higher correlations

Conditional model:
- No method produces unbiased estimates

Time-dependent covariates only with Anderson-Gill approach

Poisson regression applicable for interval-censored data

Poisson regression may be favorable with rare events
## Simulations: Results under H0

<table>
<thead>
<tr>
<th>model</th>
<th>analysis</th>
<th>mean beta</th>
<th>mean SE naive</th>
<th>mean SE robust</th>
<th>type I error naive</th>
<th>type I error robust</th>
</tr>
</thead>
<tbody>
<tr>
<td>homogenous</td>
<td>Anderson-Gill</td>
<td>0,00</td>
<td>0,10</td>
<td>0,18</td>
<td>29,1</td>
<td>6,9</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>0,00</td>
<td>0,18</td>
<td>0,18</td>
<td></td>
<td>6,1</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>0,00</td>
<td>0,10</td>
<td>0,17</td>
<td>30,0</td>
<td>8,3</td>
</tr>
<tr>
<td>non-homog.</td>
<td>Anderson-Gill</td>
<td>0,00</td>
<td>0,08</td>
<td>0,18</td>
<td>36,7</td>
<td>6,4</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>0,00</td>
<td>0,17</td>
<td>0,17</td>
<td></td>
<td>6,2</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>0,00</td>
<td>0,08</td>
<td>0,16</td>
<td>38,2</td>
<td>7,8</td>
</tr>
<tr>
<td>conditional</td>
<td>Anderson-Gill</td>
<td>0,00</td>
<td>0,04</td>
<td>0,18</td>
<td>69,0</td>
<td>7,3</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>0,00</td>
<td>0,22</td>
<td></td>
<td></td>
<td>3,8</td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>0,00</td>
<td>0,04</td>
<td>0,18</td>
<td>68,8</td>
<td>8,0</td>
</tr>
</tbody>
</table>

Antje Jahn
## Simulations: Results under H1

<table>
<thead>
<tr>
<th>model</th>
<th>analysis</th>
<th>mean beta</th>
<th>mean SE naive</th>
<th>mean SE robust</th>
<th>Coverage of 95%-CI naive</th>
<th>Coverage of 95%-CI robust</th>
</tr>
</thead>
<tbody>
<tr>
<td>homogenous</td>
<td>Anderson-Gill</td>
<td>-0.30</td>
<td>0.11</td>
<td>0.19</td>
<td>75.3</td>
<td>93.2</td>
</tr>
<tr>
<td>poisson</td>
<td>neg binomial</td>
<td>-0.30</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>-0.30</td>
<td>0.11</td>
<td>0.18</td>
<td>75.4</td>
<td>92.4</td>
</tr>
<tr>
<td>non-homog.</td>
<td>Anderson-Gill</td>
<td>-0.30</td>
<td>0.09</td>
<td>0.18</td>
<td>63.4</td>
<td>94.5</td>
</tr>
<tr>
<td>poisson</td>
<td>neg binomial</td>
<td>-0.30</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>-0.30</td>
<td>0.09</td>
<td>0.17</td>
<td>62.1</td>
<td>92.6</td>
</tr>
</tbody>
</table>
## Simulations: Results under H1

<table>
<thead>
<tr>
<th>model</th>
<th>analysis</th>
<th>mean beta</th>
<th>mean SE naive</th>
<th>mean SE robust</th>
<th>Coverage of 95%-CI naive</th>
<th>Coverage of 95%-CI robust</th>
</tr>
</thead>
<tbody>
<tr>
<td>homogenous</td>
<td>Anderson-Gill</td>
<td>-0.30</td>
<td>0.11</td>
<td>0.19</td>
<td>75.3</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>-0.30</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>-0.30</td>
<td>0.11</td>
<td>0.18</td>
<td>75.4</td>
<td>92.4</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>-0.30</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>-0.30</td>
<td>0.09</td>
<td>0.17</td>
<td>62.1</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td>neg binomial</td>
<td>-0.36</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poisson</td>
<td>-0.34</td>
<td>0.05</td>
<td>0.18</td>
<td>35.9</td>
<td>92.5</td>
</tr>
</tbody>
</table>

Antje Jahn
Simulations:
Kernel Density Estimation of Standard Error

Homogenous Poisson Model

Antje Jahn
Simulations:
Kernel Density Estimation of Standard Error

Non-homogenous Poisson Model

Antje Jahn
Simulations:
Kernel Density Estimation of Standard Error

Conditional Model

Antje Jahn