Poisson Regression

STA 312 Fall 2012

See last slide for copyright information
Regression: Outcomes are Counts

- Poisson process model roughly applies
- Examples: Relationship of explanatory variables to
  - Number of children
  - Number of typos in a short document
  - Number of workplace accidents in a short time period
  - Number of marriages

- For large $\lambda$, CLT says a normality assumption is okay, but not constant variance
Linear Model for log $\lambda$

- $\log \lambda = \beta_0 + \beta_1 x_1 + ... + \beta_{p-1} x_{p-1}$
- Implicitly for $i = 1, ..., n$
- Everybody in the sample has a different $\lambda = \lambda_i$
- Take exponential function of both sides
- Substitute into Poisson likelihood
- Maximum likelihood as usual
- Likelihood ratio tests, etc.
\[
\log \lambda = \beta_0 + \beta_1 x_1 + \ldots + \beta_{p-1} x_{p-1}
\]

- Increase \( x_k \) with everything else held constant, and
  - \( \log \lambda \) increases by \( \beta_k \)
  - \( \lambda \) is multiplied by \( e^{\beta_k} \)
Back to the job study: n=200 Students

- 106 employed in a job related to field of study
- 74 employed in a job unrelated to field of study
- 20 unemployed
- Could be independent Poisson processes

Conditionally on the total number of students, multinomial with

- \( \pi_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \)
- \( \pi_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \)
- \( \pi_3 = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \)
Poisson regression with dummy variables

<table>
<thead>
<tr>
<th>Job Status</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$\log \lambda = \beta_0 + \beta_1 d_1 + \beta_2 d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Related</td>
<td>0</td>
<td>0</td>
<td>$\beta_0$ = $\log \lambda_1$</td>
</tr>
<tr>
<td>Unrelated</td>
<td>1</td>
<td>0</td>
<td>$\beta_0 + \beta_1$ = $\log \lambda_2$</td>
</tr>
<tr>
<td>Unemployed</td>
<td>0</td>
<td>1</td>
<td>$\beta_0 + \beta_2$ = $\log \lambda_3$</td>
</tr>
</tbody>
</table>

On average, we expect $e^{\beta_2}$ times as many unemployed students as students with jobs related to their fields of study.
Copyright Information

This slide show was prepared by Jerry Brunner, Department of Statistics, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. These Powerpoint slides will be available from the course website: 
http://www.utstat.toronto.edu/brunner/oldclass/312f12