Poisson Regression

Poisson Process

- Events happening randomly in space or time
- Independent increments
- For a small region or interval,
  - Chance of 2 or more events is negligible
  - Chance of an event roughly proportional to the size of the region or interval
- Then (solve a system of differential equations), the probability of observing *x* events in a region of size *t* is

$$\frac{e^{-\lambda t}(\lambda t)^x}{x!} \text{ for } x = 0, 1, \ldots$$

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 λ a normality assumption is okay, but not constant variance

Linear Model for log λ

- $\log \lambda = \beta_0 + \beta_1 x_1 + \ldots + \beta_{p-1} x_{p-1}$
- Implicitly for $i = 1, \ldots N$
- 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, Wald tests, etc.