Today

Goal

Find maximum likelihood estimates for parameters in Generalized Linear Models.

Generalized Linear Model

1. Likelihood; $f(y;\theta)$, member of the exponential family.
2. Link function; $g(\mu_i) = x_i\beta$, where $\mu_i = E(Y_i)$ and $g()$ is monotone and differentiable.
3. Linear component of explanatory variables; $g(\mu) = X\beta$

TO DO:

- Linguist example, ML
- Newton-Raphson method
- Method of scoring for GLM
  - Univariate
  - Multivariate $\Rightarrow$ iterative weighted least square
- Examples

- Inspired by Ch 3.5.2.
- Interested in languages that descend from the same historical languages.
  - Norwegian and Swedish from Norse.
  - Modern French and Spanish from Latin.
- Languages that are separated by time t.
- Probability that a particular meaning has cognate words, $\exp(-\lambda t)$.
- Data: A linguist (*Clue*) judges if N different meanings are cognate:

Meaning	**Norwegian**	**Swedish**	**Cognate**
Laugh	Le	Skratta	No
House	Hus	Hus	Yes

Similar data for Spanish and French.

Model

1. Probability function, $Y_i \sim f(y, \theta_i)$
2. Link function, $g(\mu_i) = x_i^T \beta$
3. Explanatory variables and parameters

Model

1. Probability function, $Y_i \sim f(y, \theta_i)$
   - $Y_i \sim bin(1, p_i)$ with $p_i = \exp(-\lambda t_i) = \theta_i$
2. Link function, $g(\mu_i) = x_i^T \beta$
3. Explanatory variables and parameters

Model

1. Probability function, $Y_i \sim f(y, \theta_i)$
    - $Y_i \sim bin(1, p_i)$ with $p_i = \exp(-\lambda t_i) = \theta_i$
2. Link function, $g(\mu_i) = x_i^T \beta$

3. Explanatory variables and parameters

   Parameter: $\beta = \lambda$

   Explanatory variable: t_i, time since separation.
   $X = [t_{ns}, t_{ns}, \ldots t_{ns}, t_{sf}, \ldots, t_{sp}]$. t-s are assumed known.

Model

1. Probability function, $Y_i \sim f(y, \theta_i)$
   - $Y_i \sim bin(1, p_i)$ with $p_i = \exp(-\lambda t_i) = \theta_i$
2. Link function, $g(\mu_i) = x_i^T \beta$
   - $\mu_i = E(Y_i) = p_i = \exp(-\lambda t_i)$
   - $g(\mu_i) = log(\mu_i)$
3. Explanatory variables and parameters

   Parameter: $\beta = \lambda$

   Explanatory variable: t_i, time since separation. $X = [t_{ns}, t_{ns}, \ldots t_{ns}, t_{sf}, \ldots, t_{sp}]$. t-s are assumed known.

Exponential family

$f(y;\theta)$ belongs to the exponential family if

$$f(y;\theta) = \exp[a(y)b(\theta)) + c(\theta) + d(y)]$$

Score statistics

Let $l(\theta; y_i)$ be log-likelihood function. Then the score statistic is:

$$U(\theta; y) = \frac{\partial l(\theta; y_i)}{\partial \theta}$$

Information

Let $U = U(\theta; Y)$ be the score statistic. Then the information is

$$\Im = Var(U)$$

If Y_i has pdf from exponential family:

- $\Im = E(U^2) = -E(\frac{\partial U}{\partial \theta}) = -E(\frac{\partial^2 l_i(\theta; y_i)}{\partial \theta^2})$

Model, historical linguistics

1. Probability function, $Y_i \sim f(y, \theta_i)$
   - $Y_i \sim bin(1, p_i)$ with $p_i = \exp(-\lambda t_i) = \theta_i$
2. Link function, $g(\mu_i) = x_i^T \beta$
   - $\mu_i = E(Y_i) = p_i = \exp(-\lambda t_i)$
   - $g(\mu_i) = log(\mu_i)$
3. Explanatory variables and parameters

Parameter: $\beta = \lambda$

Explanatory variable: t_i, time since separation. $X = [t_{ns}, t_{ns}, \ldots t_{ns}, t_{sf}, \ldots, t_{sp}]$. t-s are assumed known.

Model, historical linguistics

1. Probability function, $Y_i \sim f(y, \theta_i)$
   - $Y_i \sim bin(1, p_i)$ with $p_i = \exp(-\lambda t_i) = \theta_i$
2. Link function, $g(\mu_i) = x_i^T \beta$
   - $\mu_i = E(Y_i) = p_i = \exp(-\lambda t_i)$
   - $g(\mu_i) = log(\mu_i)$
3. Explanatory variables and parameters

   Parameter: $\beta = \lambda$

   Explanatory variable: t_i, time since separation. $X = [t_{ns}, t_{ns}, \dots t_{ns}, t_{sf}, \dots, t_{sp}]$. t-s are assumed known.

- $Var(Y_i) = p_i(1 - p_i) = \exp(-\lambda t_i)(1 - \exp(-\lambda t_i))$
- Different variance $\Rightarrow$ weights
- a non-linear link.

Leukemia (approx exercise 4.1)

Data: Survival (y_i) in weeks and log initial number of white blood cell count (x_i) for 17 patients.

Model:
- $Y_i \sim gamma(\theta_1, \theta_2)$
- $\mu = E(Y_i) = \theta_1\theta_2 = \exp(\beta_0 + \beta_1 x)$
- $g(\mu) = log(\mu) = X\beta$