This workshop optional lecture next time: The 1.30-3.05 pm (room) usual
(vedcast) 15 Mar 18
why metropolis sampling works, page 1
data science

\[(y_i | \lambda_i) \sim \text{Poisson}(\lambda_i)\]
\[(i = 1, \ldots, n)\]
\[\log \lambda_i = \delta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i\]
\[(\delta_0, \beta_1, \ldots, \beta_k) \sim \text{diffuse}\]
fixed-effects
Poisson regression model (FEPR)

\[(y_i | \lambda_i) \sim \text{Poisson}(\lambda_i)\]
\[\log \lambda_i = \delta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i\]
\[\epsilon_i \sim \text{IID } N(0, \sigma_e)\]

model complexity is somewhere between 3 & (5+2+3).
use data to estimate \(\text{DIC} \quad \text{PD}\) sickness at beginning of experiment variability
\(\text{latent}\) variables random effects
This REPR model is also a lognormal mixture of Poissons; an alternative would be Gamma mixture of Poissons (generalization of Negative Binomial).

The purpose of random effects is to describe within- (E/C) group heterogeneity in # in relation to Poisson.

<table>
<thead>
<tr>
<th>FEPR</th>
<th>1500</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEPR</td>
<td>1419.6</td>
</tr>
</tbody>
</table>

DIC 1

massive difference

.968

parental multi. effect

IHGA is statistically beneficial

FEPR wins
Goal: predict $Y$ for $(x_1, ..., x_k)$

Logistic regression:

$\ell = \log \left( \frac{e^\theta \cdot x}{1 + e^\theta \cdot x} \right)$

$\theta = \text{linear}$

$Y \in \{0, 1\}$

$y \rightarrow Y \rightarrow \logistic$
0. Create a bootstrap copy of on a such
1. Run an algorithm on copy e, obtaining \( \hat{y} \).
2. Collect \( \hat{y} \) together to form an approx. posterior predictive dist.

Linear regression:

\[
Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \epsilon_i \\
\epsilon_i \sim N(0, \sigma^2)
\]

\[
Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}, \\
\epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}
\]
maximized likelihood

\[ \hat{\beta}_{\text{MLE}} = \left( X^T X \right)^{-1} X^T Y \]

\[ \hat{\sigma}^2_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

\[ \hat{y} = X \hat{\beta} \]

Bayesian story:
like MLE but
with prior on \( \beta \)

(see Banerjee notes)

\[ (y_i | p_i; B) \sim \text{Bernoulli}(p_i) \]

\( (i = 1, \ldots, n) \)

\[ \log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} \]

logistic regression

closed-form MLE don't exist; iterative methods needed
model uncertainty

0. Variable selection: optimal subset of \((x_1, \ldots, x_k)\)