

# Lab 2: Mixed-effects models

April 7, 2016

# Mixed-effects models

## Laws of probability

### 1. Axiom of conditional probability

$$\Pr(X, Y) = \Pr(Y|X) \Pr(X)$$

- Often easier to specify conditional probabilities than joint probabilities

### 2. Law of total probability

$$\Pr(X) = \int \Pr(X, Y) dY$$

- Used when justifying hierarchical models

# Mixed-effects models

## Bayes rule

- By the Axiom of conditional probability

$$\Pr(\theta|y) \Pr(y) = \Pr(y, \theta) = \Pr(y|\theta) \Pr(\theta)$$

- Therefore

$$\Pr(\theta|y) = \frac{\Pr(y|\theta) \Pr(\theta)}{\Pr(y)}$$

- By the Law of total probability

$$\Pr(y) = \int \Pr(y, \theta) dy = \int \Pr(y|\theta) \Pr(\theta) dy$$

- Therefore

$$\Pr(\theta|y) = \frac{\Pr(y|\theta) \Pr(\theta)}{\int \Pr(y|\theta) \Pr(\theta) dy}$$

- MCMC gives you

$$\Pr(\theta|y) \propto \Pr(y|\theta) \Pr(\theta)$$

# Mixed-effects models

## Empirical Bayes

- By the definition of a likelihood

$$L(\theta; y) = \Pr(y|\theta)$$

- By the Law of total probability

$$\Pr(y|\theta) = \int \Pr(y, \varepsilon|\theta) d\varepsilon$$

- By the Axiom of conditional probability

$$\Pr(y, \varepsilon|\theta) = \Pr(y | \varepsilon, \theta) \Pr(\varepsilon|\theta)$$

- Therefore

$$\Pr(y|\theta) = \int \Pr(y | \varepsilon, \theta) \Pr(\varepsilon|\theta) d\varepsilon$$

# Mixed-effects models

## Generalized linear mixed model

1. Specify distribution for response variable

$$c_i \sim \text{Poisson}(\lambda_i)$$

2. Specify function for expected value

$$g^{-1}(\lambda_i) = x_0 + \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{z}_i^T \boldsymbol{\varepsilon}$$

3. Specify a link function

$$g^{-1}(a) = \log(a) \rightarrow g(a) = \exp(a)$$

4. Specify distribution for random effects

$$\boldsymbol{\varepsilon} \sim \text{Normal}(0, \sigma_{\boldsymbol{\varepsilon}}^2)$$

= General linear model + mixed effect(s)

# Mixed-effects models

How to estimate standard errors?

- Estimate the “Hessian” at the log-marginal likelihood

$$H(\boldsymbol{\theta}; \mathbf{y}) = \begin{bmatrix} \frac{\partial^2 \ln L(\boldsymbol{\theta}; \mathbf{y})}{\partial \theta_1^2} & \frac{\partial^2 \ln L(\boldsymbol{\theta}; \mathbf{y})}{\partial \theta_1 \partial \theta_2} \\ \frac{\partial^2 \ln L(\boldsymbol{\theta}; \mathbf{y})}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 \ln L(\boldsymbol{\theta}; \mathbf{y})}{\partial \theta_2^2} \end{bmatrix}$$

- Calculate its inverse

$$\widehat{\mathbf{V}}(\boldsymbol{\theta}; \mathbf{y}) = \mathbf{H}^{-1}$$

- Extract element and take square root

$$\widehat{\text{SE}}(\theta_i; \mathbf{y}) = \sqrt{\widehat{\mathbf{V}}(\boldsymbol{\theta}; \mathbf{y})_{i,i}}$$

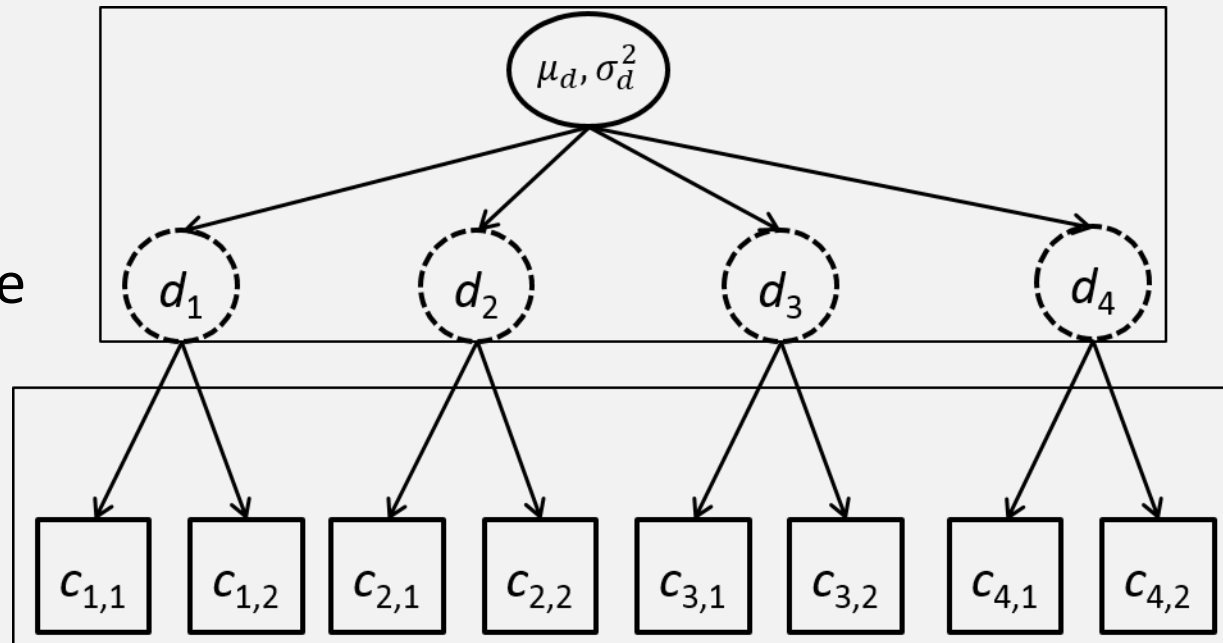
# Mixed-effects models

Example – Hierarchical count samples

$$\log(d_j) \sim \text{Normal}(\mu_d, \sigma_d^2)$$

$$c_{i,j} \sim \text{Poisson}(d_j)$$

- Counts
  - 4 sites
  - 2 observations/site
  - 3 fixed effects
  - 4 random effects



# *Mixed-effects models*

- Simulating data
  - [See R code]



# Mixed-effects models

## Fit using R

- Using *lme4* package
- *formula*: way to specify model

### 1. Linear model – *lm(formula= ... )*

- $\text{Count} \sim 0 + \text{factor}(\text{Site})$
- “Count” – response variable
- “0” – Don’t include intercept
- “factor(Site)” – Include a fixed effect for each site

### 2. Linear mixed model – *lm(formula = ... | ... )*

- $\text{Count} \sim ( 1 \mid \text{factor}(\text{Site}))$
- “( 1 | factor(Site) )” – Include a random effect for each site

# *Mixed-effects models*

**Fit using R**

– [See R code]

# Mixed-effects models

## Fit using TMB

### Steps during optimization

1. Write joint log-likelihood  $\Pr(y, \varepsilon | \theta)$  in CPP file

$$f(\theta, \varepsilon) = \log(\Pr(y | \theta_1, \varepsilon) \Pr(\varepsilon | \theta_2))$$

2. Choose initial values for fixed  $\theta_0$  and random  $\varepsilon_0$
3. “Inner optimization” – Optimize random effects with  $\theta_0$  held constant

$$\hat{\varepsilon} = \operatorname{argmax}_{\varepsilon} (f(\theta_0, \varepsilon))$$

4. Calculate Laplace approx. for marginal likelihood of fixed effects

$$\ln L(\theta_0; y) \cong f(\theta_0, \hat{\varepsilon}) - \frac{1}{2} \log(|\mathbf{H}|)$$

- TMB also provides the gradient of the penalized likelihood with respect to fixed effects
5. “Outer optimization” – Repeat steps 2-3
- Outer optimization is done in R using the function value and gradient provided by TMB

# *Mixed-effects models*

**Fit using TMB**

[See R code]

# Mixed-effects models

- Benefits of using linear mixed models
  - Separate estimate of measurement and between-site variability
  - Include covariates for either one
  - Improved precision
  - “Shrinkage”
- Draw-backs
  - Biased if random effects aren’t “exchangeable”

# Mixed-effects models

## Restricted maximum likelihood models (REML)

- Maximum likelihood (ML) estimates of variance parameters are biased

- ML estimate  $\hat{\sigma}_{ML}^2$

$$\hat{\sigma}_{ML}^2 = \frac{1}{n_i} \sum_{i=1}^{n_i} (y - \hat{\mu}_i)^2$$

- Expectation

$$\hat{\sigma}_{unbiased}^2 = \frac{1}{n_i - 1} \sum_{i=1}^{n_i} (y - \hat{\mu}_i)^2$$

- Same problem arises for variance estimates of random effects
- REML gives unbiased estimates of random-effect variances
  - Also sometimes helps convergence
  - Important when log-likelihood function is correlated with respect to random and fixed effects

# *Mixed-effects models*

[Explore “map” argument to TMB]