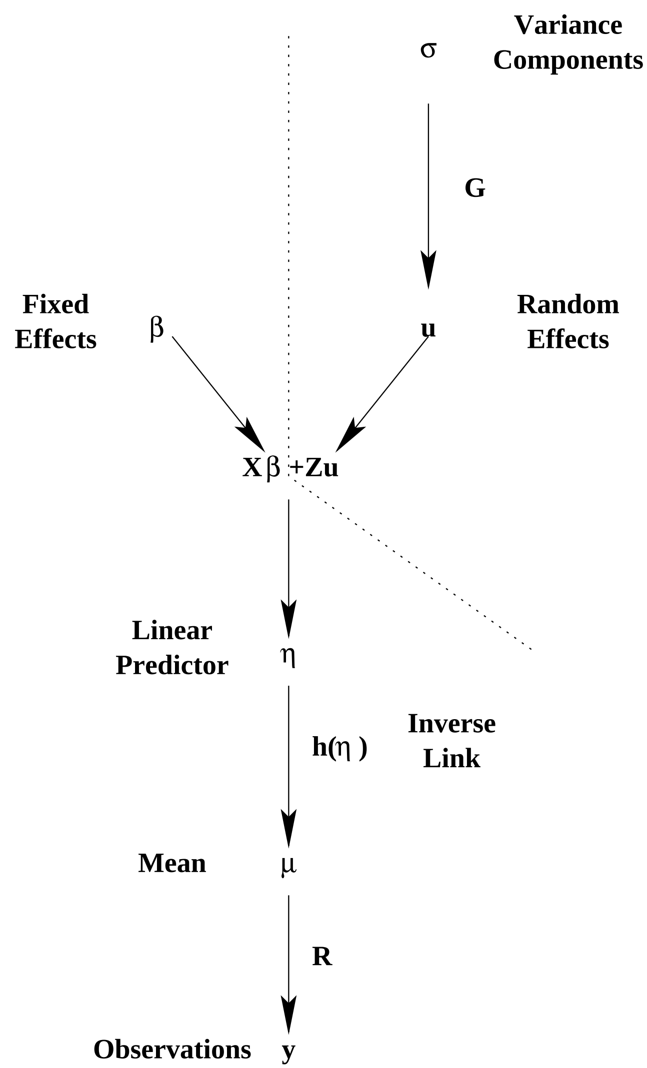
Generalized Linear Mixed Models: Foundations

- Known variance components
- Logistic model
- Marginal likelihood
- Unknown variance components
  - Approximate marginal likelihood
  - MCMC


σ
Variance Components
G
Fixed Effects
β
u
Random Effects
Xβ +Zu
Linear Predictor
η
h(η)
Inverse Link
Mean
μ
R
Observations
y

Density

- Start with $\boldsymbol{y}|\boldsymbol{u}$

$$f_Y(\boldsymbol{y}|\boldsymbol{u};\boldsymbol{\beta},\phi) = \exp\left\{\sum_{i=1}^{N} \ell_i\right\}$$

- and the distribution of $\boldsymbol{u}$

$$f_U(\boldsymbol{u};\boldsymbol{\sigma}) \sim \exp\left\{\frac{-1}{2}(\ln|D| + \boldsymbol{u}'\boldsymbol{D}^{-1}\boldsymbol{u})\right\}$$

- Get the joint density of $\boldsymbol{y}, \boldsymbol{u}$

$$f_{Y,U}(\boldsymbol{y}, \boldsymbol{u}); \boldsymbol{\beta}, \phi, \boldsymbol{\sigma}) = \exp\left\{\sum_{i=1}^{N} \ell_i - \frac{1}{2}(\ln|D| + \boldsymbol{u}'\boldsymbol{D}^{-1}\boldsymbol{u})\right\}$$

- Get the density of $\boldsymbol{y}$

$$f_Y(\boldsymbol{y}; \boldsymbol{\beta}, \phi, \boldsymbol{\sigma}) = \int \dots \int_{\Re^q} \exp\left\{\sum_{i=1}^{N} \ell_i - \frac{1}{2}(\ln|D| + \boldsymbol{u}'\boldsymbol{D}^{-1}\boldsymbol{u})\right\} d\boldsymbol{u}$$

- Difficult to evaluate analytically
  - Numerically
  - Approximately

Penalized Quasi-Likelihood

$$h(\boldsymbol{u}) = \sum_{i=1}^{N} \ell_i - \frac{1}{2}(\ln|D| + \boldsymbol{u}'\boldsymbol{D}^{-1}\boldsymbol{u})$$

- Obtaining the the marginal density of $\boldsymbol{y}$ would be easy if $h(\boldsymbol{u})$ was a quadratic form.

- Second order Taylor's Series expansion of $h(\boldsymbol{u})$ about $\boldsymbol{u}_0$

$$h(\boldsymbol{u}) \doteq h(\boldsymbol{u}_0) + \frac{\partial h}{\partial \boldsymbol{u}'}(\boldsymbol{u} - \boldsymbol{u}_0) + \frac{1}{2}(\boldsymbol{u} - \boldsymbol{u}_0)'\frac{\partial h^2}{\partial \boldsymbol{u} \partial \boldsymbol{u}'}(\boldsymbol{u} - \boldsymbol{u}_0)$$

with the derivatives evaluated at $\boldsymbol{u}_0$.

- Select $\boldsymbol{u}_0$ such that

$$\frac{\partial h}{\partial \boldsymbol{u}'}|_{\boldsymbol{u}=\boldsymbol{u}_0} = 0$$

$$\frac{\partial h}{\partial \boldsymbol{u}'} = \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) - \boldsymbol{D}^{-1}\boldsymbol{u}$$

- Which yields

$$h(\boldsymbol{u}) \doteq h(\boldsymbol{u}_0) + \frac{1}{2}(\boldsymbol{u}-\boldsymbol{u}_0)'\frac{\partial h^2}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}(\boldsymbol{u}-\boldsymbol{u}_0)$$

- The second partial

$$\frac{\partial h^2}{\partial \boldsymbol{u}\partial \boldsymbol{u}'} = -\boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z} - \boldsymbol{D}^{-1}$$

- Which yields

$$h(\boldsymbol{u}) \doteq \sum_{i=1}^{N} \ell_i(\boldsymbol{u}_0) - \frac{1}{2}(\ln|D| + \boldsymbol{u}_0'\boldsymbol{D}^{-1}\boldsymbol{u}_0) - \frac{1}{2}(\boldsymbol{u}-\boldsymbol{u}_0)'[\boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z} + \boldsymbol{D}^{-1}](\boldsymbol{u}-\boldsymbol{u}_0)$$

- Integrating yields

$$\ell \doteq \sum_{i=1}^{N} \ell_i(\boldsymbol{u}_0) - \frac{1}{2}(\ln|D| + \boldsymbol{u}_0'\boldsymbol{D}^{-1}\boldsymbol{u}_0) - \frac{1}{2}\ln|\boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z} + \boldsymbol{D}^{-1}|$$

- Taking the partial with respect $\boldsymbol{\beta}$

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} \doteq \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$$

- Approximate MLE by solving

$$\begin{aligned} \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) &= \mathbf{0} \\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) - \boldsymbol{D}^{-1}\boldsymbol{u} &= \mathbf{0} \end{aligned}$$

- Using Fisher scoring

$$\begin{pmatrix} \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z} + \boldsymbol{D}^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{u}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{\eta}) \\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{\eta}) \end{pmatrix}$$

- Has the advantage of being simple to implement.

- Depends on how closely the conditional distribution of $\boldsymbol{y}$ given $\boldsymbol{u}$ approximates a normal distribution.

  – Work well **when** the central limit theorem applies

Estimation of variance components

- Recall that REML in the linear mixed model case involved setting a set a quadratic forms equal to their expectation and solving.

$$\widehat{\boldsymbol{u}}_i' \widehat{\boldsymbol{u}}_i$$

- Using the same set of quadratic forms and their approximate expectation:

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{u}} \end{pmatrix} = \boldsymbol{C} \begin{pmatrix} \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1} \\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1} \end{pmatrix} (\boldsymbol{y} - \widehat{\boldsymbol{\mu}} + \boldsymbol{H}\widehat{\boldsymbol{\eta}})$$

$$(\boldsymbol{y} - \boldsymbol{\mu} + \boldsymbol{H}\boldsymbol{\eta}) \dot{\sim} (\boldsymbol{H}\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{R} + \boldsymbol{H}\boldsymbol{Z}\boldsymbol{D}\boldsymbol{Z}'\boldsymbol{H}')$$

$$\mathrm{E} \begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{u}} \end{pmatrix} \doteq \boldsymbol{C} \begin{pmatrix} \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X}\boldsymbol{\beta} \\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X}\boldsymbol{\beta} \end{pmatrix}$$

$$\mathrm{E}(\widehat{\boldsymbol{\beta}}) \doteq [\boldsymbol{C}^{XX}\boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} + \boldsymbol{C}^{XZ}\boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X}]\boldsymbol{\beta}$$

$$\mathrm{E}(\widehat{\boldsymbol{u}}) \doteq \boldsymbol{0}$$

$$\operatorname{var}\begin{pmatrix}\widehat{\boldsymbol{\beta}}\\ \widehat{\boldsymbol{u}}\end{pmatrix} \doteq \boldsymbol{C}\begin{pmatrix}\boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z}\\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z}\end{pmatrix}\boldsymbol{C}$$

$$+\boldsymbol{C}\begin{pmatrix}\boldsymbol{X}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z}\\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z}\end{pmatrix}\boldsymbol{D}\begin{pmatrix}\boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{X}\\ \boldsymbol{Z}'\boldsymbol{H}'\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{Z}\end{pmatrix}\boldsymbol{C}$$

$$= [\boldsymbol{C} - \boldsymbol{C}\boldsymbol{D}^{-1}\boldsymbol{C}] + [\boldsymbol{D} - \boldsymbol{C} - \boldsymbol{C} + \boldsymbol{C}\boldsymbol{D}^{-1}\boldsymbol{C}]$$

$$\widehat{\boldsymbol{u}} \dot{\sim} (\boldsymbol{0}, \boldsymbol{D} - \boldsymbol{C}^{ZZ})$$

Which is identical to the linear mixed model case.

- The estimating equations become

$$.5\left[\operatorname{Diag}\left(\frac{1}{\sigma_i^4}\{q_i - 2\operatorname{tr}(\boldsymbol{C}^{ii})/\sigma_i^2\}\right) + \left\{\frac{1}{\sigma_i^4\sigma_j^4}\operatorname{tr}(\boldsymbol{C}^{ij}\boldsymbol{C}^{ji})\right\}\right]\widehat{\boldsymbol{\sigma}} = .5\left\{\frac{\widehat{\boldsymbol{u}}_i'\widehat{\boldsymbol{u}}_i}{\sigma_i^4}\right\}$$

Germination Rate

- Data[5] from an experiment examining the germination rates
- Two types of seeds ($O.\ aegyptiaco\ 75\ and\ 73$)
- Two types of root extracts (bean and cucumber).
- Dependent variable: Number of seeds that germinated r_{ijk} out of the the total number present in a dish n_{ijk}. table.

[5] Source: Crowder, M. (1978) *Beta-binomial Anova for proportions*. Appl. Statist. **27**, 34–37 (Table 3)

O. aegyptiaco 75				*O. aegyptiaco 73*			
Bean		Cucumber		Bean		Cucumber	
r_{11k}	n_{11k}	r_{12k}	n_{12k}	r_{21k}	n_{21k}	r_{22k}	n_{22k}
10	39	5	6	8	16	3	12
23	62	53	74	10	30	22	41
23	81	55	72	8	28	15	30
17	39	46	79	0	4	3	7
		10	13				

$$y_{ijk} = r_{ijk}/n_{ijk}|\boldsymbol{u} \sim \frac{1}{n_{ijk}}\,\mathrm{Bin}(n_{ijk}, \mu_{ijk})$$

$$\mathrm{E}(y_{ijk}|\boldsymbol{u}\boldsymbol{u}) = \mu_{ijk}$$

$$\mathrm{var}(y_{ijk}|\boldsymbol{u}) = \frac{\mu_{ijk}(1-\mu_{ijk})}{n_{ijk}}$$

$$\ell(\boldsymbol{\mu}; \boldsymbol{y}|\boldsymbol{u}) = c + \sum_{i=1}^{2}\sum_{j=1}^{7} n_{ijk}[y_{ijk}\ln(\mu_{ijk}) + (1-y_{ijk})\ln(1-\mu_{ijk})]$$

The linear predictor is

$$\eta_{ij} = \mu + T_i + E_j + (TE)_{ij} + D_{ijk}$$

- μ is the intercept
- T_i is the main effect of seed type i
- E_j is the main effect of root extract j,
- D_{ijk}: is a random dish effect iid $\mathrm{N}(0, \sigma_D^2)$.

ASREML:

```
Germination Example
  type !A
  extract !A
  r
  n
  y
  plate !I
germ.dat !MAXIT=100
r !BINOMIAL !TOTAL n ~ type*extract !r plate
0 0 1
plate 1
0 0 I 1
predict type extract !TDIFF
predict type !TDIFF
```

Results:

```
  1 LogL=-6.22355     S2=  1.0000         13 df    1.000      1.000
  2 LogL=-5.92972     S2=  1.0000         13 df    1.000     0.9524
  3 LogL=-3.76174     S2=  1.0000         13 df    :   1 components constrained
  4 LogL=-.878734     S2=  1.0000         13 df    1.000     0.3252E-01
  5 LogL=-.811204     S2=  1.0000         13 df    1.000     0.4485E-01
  6 LogL=-.787669     S2=  1.0000         13 df    1.000     0.5803E-01
  7 LogL=-.787653     S2=  1.0000         13 df    1.000     0.5904E-01
  8 LogL=-.787651     S2=  1.0000         13 df    1.000     0.5901E-01
 Final parameter values                       1.0000    0.59009E-01
 Deviance from GLM fit              13       10.55
 Variance heterogenity factor [Deviance/DF]     0.81
```

```
Degrees of Freedom and Stratum Variances
          1.35   0.590093E-01     1.0

Source                Model  terms     Gamma     Component    Comp/SE   % C
plate               identity    17  0.590093E-01  0.590093E-01   0.82   0 U

Analysis of Variance              NumDF     DenDF    F_inc              Prob
  1 type                              2      10.8     2.35             0.141
  2 extract                           1       9.7    30.96             <.001
  7 type.extract                      1      13.0     4.54             0.053

                    Estimate       Standard Error    T-value     T-prev
  7 type.extract
                   4  -0.978362       0.460229         -2.13
  2 extract
    Cucumber            1.54346       0.267891          5.76
  1 type
    O_aeg_75          -0.695085       0.189894         -3.66
    O_aeg_73          -0.696893       0.276081         -2.52     -0.01
  6 plate                              17 effects fitted
```