Contrasting Marginal and Mixed Effects Models

Recall: two approaches to handling dependence in Generalized Linear Models:

**Marginal models:** based on the *consequences* of dependence on estimating model parameters.

- Target of inference is the *population*.

**Random effect models:** based on the *sources* of dependence.

- Target of inference is the *subject*. 
Special case: Linear models

Model for mean response:

\[ E(Y_i) = X_i \beta \]

Alternative specifications for variance:

- covariance pattern

\[ \text{Cov}(Y_i) = \Sigma_i \]

  e.g. \( AR(1) \) or Toeplitz;
random effects

\[ E(Y_i|b_i) = X_i \beta + Z'_{i,j} b_i \]

with

\[ E(b_i) = 0, \quad \text{Cov}(b_i) = G, \]

whence

\[ \Sigma_i = Z'_{i,j} G Z''_{i,j} + \sigma^2 I_{n_i} \]
Interpretation of $\beta$: in both cases, 

$$E \left( Y_{i,j} \right) = X_{i,j}' \beta = \sum_k X_{i,j,k} \beta_k$$

so

$$\frac{\partial E \left( Y_{i,j} \right)}{\partial X_{i,j,k}} = \beta_k.$$ 

But also, with random effects, 

$$E \left( Y_{i,j} | b_i \right) = X_{i,j}' \beta + Z_{i,j}' b_i$$

so, if the $k^{th}$ covariate is not associated with a random effect, 

$$\frac{\partial E \left( Y_{i,j} | b_i \right)}{\partial X_{i,j,k}} = \beta_k.$$
That is, $\beta_k$ is the rate of change of the mean response as the $k^{th}$ covariate changes, both

- *on average* across subjects (marginally), and

- *for a specific subject* with random effects $b$ (conditionally).

This might be either a rate of change over time, if the covariate is time (within-subject factor), or a treatment effect, if the covariate is a dummy variable (between-subject factor).
Generalized Linear Models

With link function $g(\cdot)$ and inverse link $h(\cdot) = g^{-1}(\cdot)$:

- the marginal model is
  \[
g \left\{ \mathbb{E} \left( Y_{i,j} \right) \right\} = X'_{i,j} \beta,
  \]
  or
  \[
  \mathbb{E} \left( Y_{i,j} \right) = h \left( X'_{i,j} \beta \right);
  \]

- the mixed effects model is
  \[
g \left\{ \mathbb{E} \left( Y_{i,j} \mid b_i \right) \right\} = X'_{i,j} \beta^* + Z'_{i,j} b_i
  \]
  or
  \[
  \mathbb{E} \left( Y_{i,j} \mid b_i \right) = h \left( X'_{i,j} \beta^* + Z'_{i,j} b_i \right).
  \]
In the mixed effects case,

\[
E(Y_{i,j}) = E\left\{E\left(Y_{i,j} \mid b_i \right)\right\} \\
= E\left\{h\left(X_{i,j}'\beta^* + Z_{i,j}'b_i \right)\right\} \\
= \int h\left(X_{i,j}'\beta^* + Z_{i,j}'b_i \right) f_b(b_i) db_i,
\]

and in general

\[
E(Y_{i,j}) \neq h\left(X_{i,j}'\beta \right)
\]

for any \( \beta \).
Logistic regression with random intercept:

\[ \text{logit} \left\{ E \left( Y_{i,j} | b_i \right) \right\} = X_{i,j} \beta^* + b_i \]

so

\[ E \left( Y_{i,j} | b_i \right) = \frac{e^{(X_{i,j} \beta^* + b_i)}}{1 + e^{(X_{i,j} \beta^* + b_i)}} \]

and

\[
E \left( Y_{i,j} \right) = E \left\{ \frac{e^{(X_{i,j} \beta^* + b_i)}}{1 + e^{(X_{i,j} \beta^* + b_i)}} \right\} \\
= \int_{-\infty}^{\infty} \frac{e^{(X_{i,j} \beta^* + b_i)}}{1 + e^{(X_{i,j} \beta^* + b_i)}} \times \frac{1}{\sqrt{2\pi \sigma_b^2}} e^{-\frac{1}{2} \frac{b_i^2}{\sigma_b^2}} db_i.
\]
No closed-form solution, but

\[
\text{logit} \left\{ \mathbb{E} \left( Y_{i,j} \right) \right\} \approx \frac{X'_{i,j} \beta^*}{\sqrt{1 + k^2 \sigma_b^2}}
\]

where \( k = \frac{16\sqrt{3}}{15\pi} = 0.588 \) and \( k^2 = 0.346 \), so

\[
\beta \approx \frac{\beta^*}{\sqrt{1 + 0.346 \sigma_b^2}}.
\]

If \( \sigma_b^2 = 8 \), then \( \beta \approx \beta^*/2 \)
Example: one covariate, sample of size 13 with $\beta_1^* = -1.5$, $\beta_2^* = 0.75$, $g_{1,1} = \sigma_b^2 = 4$; population average is shown in red, and the approximation in blue.
Case study: abnormal ECG in a drug trial.

options linesize = 80 pagesize = 21 nodate;

data ecg;
  retain id 0;
  infile 'ecg.txt' firstobs = 32;
  input seq r1 r2 count;
  if seq = 1
    then /* Placebo followed by drug */
      do i = 1 to count;
        id = id + 1;
        trt = 0; y = r1; period = 0; output;
        trt = 1; y = r2; period = 1; output;
      end;
    else /* Drug followed by placebo */
      do i = 1 to count;
        id = id + 1;
        trt = 1; y = r1; period = 0; output;
        trt = 0; y = r2; period = 1; output;
      end;
  run;
title1 'Marginal Logistic Regression Model';
title2 'Crossover Trial on Cerebrovascular Deficiency';

proc genmod descending;
   class id;
   model y = trt period / d=bin;
   repeated subject=id / logor=fullclust;
run;

title1 'Mixed Effects Logistic Regression Model (Random Intercept)';
title2 'Crossover Trial on Cerebrovascular Deficiency';

proc nlmixed qpoints=100;
   /* Initial values from GEE output: */
   parms beta1=-1.2433 beta2=.5689 beta3=.2951 g11=1;
   eta=beta1 + beta2*trt + beta3*period + b;
   p=exp(eta)/(1 + exp(eta));
   model y ~ binary(p);
   random b ~ normal(0,g11) subject=id;
run;
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Model Information

Data Set WORK.ECG
Distribution Binomial
Link Function Logit
Dependent Variable y

Number of Observations Read 134
Number of Observations Used 134
Number of Events 42
Number of Trials 134
Marginal Logistic Regression Model  
Crossover Trial on Cerebrovascular Deficiency  

The GENMOD Procedure  

Class Level Information  

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>67</td>
<td>1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67</td>
</tr>
</tbody>
</table>
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Response Profile

<table>
<thead>
<tr>
<th>Ordered Value</th>
<th>Total y</th>
<th>Total Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>92</td>
</tr>
</tbody>
</table>

PROC GENMOD is modeling the probability that $y='1'$.
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Parameter Information

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prm1</td>
<td>Intercept</td>
</tr>
<tr>
<td>Prm2</td>
<td>trt</td>
</tr>
<tr>
<td>Prm3</td>
<td>period</td>
</tr>
</tbody>
</table>

Criteria For Assessing Goodness Of Fit

<table>
<thead>
<tr>
<th>Criterion</th>
<th>DF</th>
<th>Value</th>
<th>Value/DF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviance</td>
<td>131</td>
<td>163.8863</td>
<td>1.2510</td>
</tr>
<tr>
<td>Scaled Deviance</td>
<td>131</td>
<td>163.8863</td>
<td>1.2510</td>
</tr>
<tr>
<td>Pearson Chi-Square</td>
<td>131</td>
<td>133.5123</td>
<td>1.0192</td>
</tr>
</tbody>
</table>
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Criteria For Assessing Goodness Of Fit

<table>
<thead>
<tr>
<th>Criterion</th>
<th>DF</th>
<th>Value</th>
<th>Value/DF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaled Pearson X2</td>
<td>131</td>
<td>133.5123</td>
<td>1.0192</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td></td>
<td>-81.9432</td>
<td></td>
</tr>
</tbody>
</table>

Algorithm converged.
### Marginal Logistic Regression Model

Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

### Analysis Of Initial Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Error</th>
<th>Wald 95% Confidence Limits</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-1.2186</td>
<td>0.3441</td>
<td>-1.8931 -0.5440</td>
<td>12.54</td>
<td>0.0004</td>
</tr>
<tr>
<td>trt</td>
<td>1</td>
<td>0.5582</td>
<td>0.3784</td>
<td>-0.1835 1.2998</td>
<td>2.18</td>
<td>0.1402</td>
</tr>
<tr>
<td>period</td>
<td>1</td>
<td>0.2743</td>
<td>0.3768</td>
<td>-0.4642 1.0129</td>
<td>0.53</td>
<td>0.4666</td>
</tr>
<tr>
<td>Scale</td>
<td>0</td>
<td>1.0000</td>
<td>0.0000</td>
<td>1.0000 1.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** The scale parameter was held fixed.
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

GEE Model Information

<table>
<thead>
<tr>
<th>Log Odds Ratio Structure</th>
<th>Fully Parameterized Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Effect</td>
<td>id (67 levels)</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>67</td>
</tr>
<tr>
<td>Correlation Matrix Dimension</td>
<td>2</td>
</tr>
<tr>
<td>Maximum Cluster Size</td>
<td>2</td>
</tr>
<tr>
<td>Minimum Cluster Size</td>
<td>2</td>
</tr>
</tbody>
</table>
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Log Odds Ratio
Parameter Information

Parameter | Group
---|---
Alpha1 | (1, 2)

Algorithm converged.
Marginal Logistic Regression Model
Crossover Trial on Cerebrovascular Deficiency

The GENMOD Procedure

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

| Parameter | Estimate | Error  | Lower Limit | Upper Limit | Z Value | Pr > |Z| |
|-----------|----------|--------|-------------|-------------|---------|-------|--|
| Intercept | -1.2433  | 0.2999 | -1.8311     | -0.6556     | -4.15   | < .0001 |
| trt       | 0.5689   | 0.2335 | 0.1112      | 1.0266      | 2.44    | 0.0148 |
| period    | 0.2951   | 0.2319 | -0.1593     | 0.7496      | 1.27    | 0.2030 |
| Alpha1    | 3.5617   | 0.8148 | 1.9647      | 5.1587      | 4.37    | < .0001 |
Mixed Effects Logistic Regression Model (Random Intercept) Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

Specifications

- Data Set: WORK.ECG
- Dependent Variable: y
- Distribution for Dependent Variable: Binary
- Random Effects: b
- Distribution for Random Effects: Normal
- Subject Variable: id
- Optimization Technique: Dual Quasi-Newton
- Integration Method: Adaptive Gaussian Quadrature
Mixed Effects Logistic Regression Model (Random Intercept)  
Crossover Trial on Cerebrovascular Deficiency  

The NLMIXED Procedure  

Dimensions  

Observations Used 134  
Observations Not Used 0  
Total Observations 134  
Subjects 67  
Max Obs Per Subject 2  
Parameters 4  
Quadrature Points 100
Mixed Effects Logistic Regression Model (Random Intercept) Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

Parameters

<table>
<thead>
<tr>
<th>beta1</th>
<th>beta2</th>
<th>beta3</th>
<th>g11</th>
<th>NegLogLike</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.2433</td>
<td>0.5689</td>
<td>0.2951</td>
<td>1</td>
<td>76.6705695</td>
</tr>
</tbody>
</table>

Iteration History

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>75.877218</td>
<td>0.793351</td>
<td>3.143617</td>
<td>-31.3385</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>71.2225189</td>
<td>4.654699</td>
<td>0.769971</td>
<td>-36.6519</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>70.9616922</td>
<td>0.260827</td>
<td>0.959227</td>
<td>-1.3592</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>69.3279449</td>
<td>1.633747</td>
<td>0.692069</td>
<td>-1.90957</td>
</tr>
</tbody>
</table>
Mixed Effects Logistic Regression Model (Random Intercept)  
Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

Iteration History

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>69.0008522</td>
<td>0.327093</td>
<td>0.679054</td>
<td>-0.73402</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>68.5033645</td>
<td>0.497488</td>
<td>0.630885</td>
<td>-0.57466</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>68.2882471</td>
<td>0.215117</td>
<td>0.611136</td>
<td>-0.17656</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>68.1688967</td>
<td>0.11935</td>
<td>0.296866</td>
<td>-0.0794</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>68.1357412</td>
<td>0.033156</td>
<td>0.085458</td>
<td>-0.01492</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>68.1309946</td>
<td>0.004747</td>
<td>0.021189</td>
<td>-0.00549</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>68.1302223</td>
<td>0.000772</td>
<td>0.013186</td>
<td>-0.00148</td>
</tr>
<tr>
<td>12</td>
<td>23</td>
<td>68.1301696</td>
<td>0.000053</td>
<td>0.000147</td>
<td>-0.0001</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>68.1301694</td>
<td>2.006E-7</td>
<td>0.000046</td>
<td>-3.62E-7</td>
</tr>
</tbody>
</table>
Mixed Effects Logistic Regression Model (Random Intercept)  
Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood 136.3
AIC (smaller is better) 144.3
AICC (smaller is better) 144.6
BIC (smaller is better) 153.1
Mixed Effects Logistic Regression Model (Random Intercept)  
Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

Parameter Estimates

| Parameter | Estimate | Error  | DF  | t Value | Pr > |t|  | Alpha | Lower   |
|-----------|----------|--------|-----|---------|------|----|-------|--------|
| beta1     | -4.0816  | 1.6711 | 66  | -2.44   | 0.0173 | 0.05 | -7.4180 |
| beta2     | 1.8631   | 0.9269 | 66  | 2.01    | 0.0485 | 0.05 | 0.01241 |
| beta3     | 1.0375   | 0.8189 | 66  | 1.27    | 0.2096 | 0.05 | -0.5974 |
| g11       | 24.4355  | 18.8489| 66  | 1.30    | 0.1994 | 0.05 | -13.1974 |
Mixed Effects Logistic Regression Model (Random Intercept)  
Crossover Trial on Cerebrovascular Deficiency

The NLMIXED Procedure

Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Upper</th>
<th>Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta1</td>
<td>-0.7452</td>
<td>-8.31E-6</td>
</tr>
<tr>
<td>beta2</td>
<td>3.7137</td>
<td>0.000028</td>
</tr>
<tr>
<td>beta3</td>
<td>2.6725</td>
<td>-0.00005</td>
</tr>
<tr>
<td>g11</td>
<td>62.0685</td>
<td>-1.34E-7</td>
</tr>
</tbody>
</table>
Notes

- The treatment effect is significant in both analyses, but much larger in the mixed effects model than in the marginal model.

- The \texttt{proc nlmixed} fit has a very large $\hat{g}_{1,1}$ of 24.4, but with a large standard error and a non-significant $t$-statistic.

- But the \texttt{proc nlmixed} output gives $\text{AIC} = 144.3$ and the “independence” part of the \texttt{proc genmod} output gives Log Likelihood $= -81.9432$, whence

\[
\text{AIC} = (-2) \times (-81.9432) + 2 \times 3 = 169.9,
\]

showing that including the variance parameter leads to a much better fit.