ASReml tutorial

D1 GLM & GLMM

Arthur Gilmour
Iteratively reweighted least squares

Data types in exponential family
  Binary [0,1]
  Binomial [r,n]
  Poisson
  Gamma
  Negative Binomial

Variance is dependent on the mean
Method

- Analyse on a transformed scale using Link function and distribution specific weights

\[ y = X\tau + Zu + D^{-1}(O - E) \]

where \( D \) is derivative of link function

\( O \) is the observed value

and \( E \) is \( g(X\tau + Zu) \), the expected value weights are a function of \( D \) and \( E \)
Method continued

- Apply usual REML to this working variable $y$ updating $E$ and $D$, weights and $y$ each iteration.

- Binomial Logit
  
  \[ Y = X \hat{\tau} + Z \hat{u} \]
  \[ \hat{E} = p = 1/(1 + e^{-Y}) \rightarrow Y = \log(p/(1 - p)) \]
  Derivative is $p(1 - p)$
  Binomial weight is $np(1 - p)$
  Analysis weight is $n/[p(1 - p)]$
Analysis of Deviance

Without random effects, this method is standard GLM (Generalized linear model) and ASReml also calculates the -2L, the maximum likelihood deviance, used to assess improvement of fit from adding an extra term to the model (Analysis of Deviance).
GLMM Disclaimers

- With random effects, the ’Deviance’ is not valid for testing changes in the fixed model,
- the REML LogL reported cannot be used to test changes to the random model because $y$, the working variable changes between models.
GLMM Disclaimers

- The Method does not work for Animal Model variance components: it is biased - especially when number of random effects exceeds N/10.

- In most cases, effectively equivalent results are obtained from analysis on the observed scale.
Families and Links

- BIN [!ID !LOGIT !PROBIT !COMP] !TOTAL \( n \)
- POIS [!ID !LOG !SQRT]
- GAMMA [!ID !LOG !INV] [!phi \( p \)]
- NEGBIN [!ID !LOG !INV] [!PHI \( p \)]

General: !OFFSET \( o \) !DISP !DEV !WORK !RESP !PEAR
Sire model

- Binomial has reasonable base for genetics because the logit [probit] link functions imply underlying residual logistic [normal] distributions with error variance 3.3 [1]

- Can fit the model for Poisson but I am unaware of proper genetic basis for estimating heritability say.
LAMB data

- Foot shape Score

```
#Yr Grp Sex Sire xxx tot l5 l4 ls lr
# 1 1 1 1 18 39 33 6 6 1
# 1 1 0 1 18 50 41 9 2 0
# 1 1 1 2 18 35 30 4 1 0
yr 2 Grp 5 Sex Sire 18
xxx tot L5 !/tot
L4 !/tot LS !/tot LR !/tot
lamb.dat !SKIP 1
L5 !BIN !LOG !TOTAL tot ~ mu ,
Grp Sex Sex.Grp !r Sire
```
D2 ANOVA

Arthur Gilmour
The issues of ANOVA in REML

- Fixed effects in SPARSE part of model
- Types of Sums of Squares
  Incremental and Conditional ($\texttt{!FCON}$)
- Test statistic
- Denominator Degrees of Freedom
  - DDF $-1$ suppressed
  - DDF $1$ numerical derivative
  - DDF $2$ algebraic derivative
OATS - fixed ANOVA

- !FCON
  
  \[ Y \sim \mu \, bl \, var \, bl.wp \, nit \, nit.var \]
  
  ordered to get correct SS and DF.
  
  BUT ASReml uses wrong F ratio because we have not specified that bl and wp are random.

- Degrees Freedom Stratum Variances
  
  45.00  177.083  1.0

  Source  Modterms  Gam  Component  C/SE  %C
  Variance  72  45  1.0  177.083  4.74  0P
Fixed ANOVA

<table>
<thead>
<tr>
<th>AnOVar</th>
<th>Num</th>
<th>DenDF</th>
<th>F-inc</th>
<th>F-con</th>
<th>M</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu</td>
<td>1</td>
<td>45.0</td>
<td>4395.31</td>
<td>NA</td>
<td>.</td>
<td>NA</td>
</tr>
<tr>
<td>blocks</td>
<td>5</td>
<td>45.0</td>
<td>17.93</td>
<td>17.93</td>
<td>A</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>variety</td>
<td>2</td>
<td>45.0</td>
<td>5.04</td>
<td>5.04</td>
<td>a</td>
<td>0.011</td>
</tr>
<tr>
<td>bl.wp</td>
<td>10</td>
<td>45.0</td>
<td>3.40</td>
<td>3.40</td>
<td>B</td>
<td>0.002</td>
</tr>
<tr>
<td>nitrogen</td>
<td>3</td>
<td>45.0</td>
<td>37.69</td>
<td>37.69</td>
<td>A</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>nit.var</td>
<td>6</td>
<td>45.0</td>
<td>0.30</td>
<td>0.30</td>
<td>B</td>
<td>0.932</td>
</tr>
</tbody>
</table>
Proper tests

- **mu**: $F = 245.14 = 4395.31 / 17.31$ with 1,5 DF
- **var**: $F = 1.49 = 5.04 / 3.40$ with 2,10 DF

<table>
<thead>
<tr>
<th>Source</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
<th>Value5</th>
<th>Value6</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu</td>
<td>45.0</td>
<td>4395.31</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>blocks</td>
<td>45.0</td>
<td>17.93</td>
<td>17.93</td>
<td>A</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>variety</td>
<td>45.0</td>
<td>5.04</td>
<td>5.04</td>
<td>a</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bl.wp</td>
<td>45.0</td>
<td>3.40</td>
<td>3.40</td>
<td>B</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
OATS - REML analysis

!FCON

Y ~ mu nit var nit.var !r bl bl.wp

Degrees Freedom and Stratum Variances

<table>
<thead>
<tr>
<th>Degrees Freedom</th>
<th>Variance</th>
<th>12.0</th>
<th>4.0</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.00</td>
<td>3175.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.00</td>
<td>601.331</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.00</td>
<td>177.083</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source Modterms Gamma Component /SE %C

<table>
<thead>
<tr>
<th>Source</th>
<th>Modterms</th>
<th>Gamma</th>
<th>Component</th>
<th>/SE</th>
<th>%C</th>
</tr>
</thead>
<tbody>
<tr>
<td>blocks</td>
<td>6</td>
<td>6</td>
<td>1.21</td>
<td>214.477</td>
<td>1.27</td>
</tr>
<tr>
<td>bl.wp</td>
<td>18</td>
<td>18</td>
<td>0.599</td>
<td>106.062</td>
<td>1.56</td>
</tr>
<tr>
<td>Variance</td>
<td>72</td>
<td>60</td>
<td>1.00</td>
<td>177.083</td>
<td>4.74</td>
</tr>
</tbody>
</table>
Proper F-inc value

F-con is here same as F-inc because a balanced design.

<table>
<thead>
<tr>
<th>AnOVar</th>
<th>Num</th>
<th>DenDF</th>
<th>F-inc</th>
<th>F-con</th>
<th>M</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu</td>
<td>1</td>
<td>5.0</td>
<td>245.14</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>nitrogen</td>
<td>3</td>
<td>45.0</td>
<td>37.69</td>
<td>37.69</td>
<td>A</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>variety</td>
<td>2</td>
<td>10.0</td>
<td>1.49</td>
<td>1.49</td>
<td>A</td>
<td>0.272</td>
</tr>
<tr>
<td>nit.var</td>
<td>6</td>
<td>45.0</td>
<td>0.30</td>
<td>0.30</td>
<td>B</td>
<td>0.932</td>
</tr>
</tbody>
</table>
Wald tests

- F-inc based on Sum of Squares explained by adding the term in the order specified divided by NumDF and $\sigma^2$

- In mixed model, not easy to calculate DenDF. Not available in ASReml 1.1
  - Not available for large models in ASReml 2
  - Not necessarily the same for F-inc and F-con
Main effects before interactions:
- If you specify an interaction first, the main effects will typically disappear
- If you put an interaction in the sparse section, it will sweep out mu, and higher order terms.
In practice

Interpret the following ANOVA

<table>
<thead>
<tr>
<th>Src</th>
<th>NDF</th>
<th>DDF</th>
<th>F-inc</th>
<th>F-con</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1</td>
<td>45</td>
<td>1.</td>
<td>21.</td>
</tr>
<tr>
<td>X2</td>
<td>1</td>
<td>45</td>
<td>23.</td>
<td>1.</td>
</tr>
<tr>
<td>X3</td>
<td>1</td>
<td>45</td>
<td>10.</td>
<td>10.</td>
</tr>
</tbody>
</table>
In practice

- Interpret the following ANOVA

<table>
<thead>
<tr>
<th>Src</th>
<th>NDF</th>
<th>DDF</th>
<th>F-inc</th>
<th>F-con</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1</td>
<td>45</td>
<td>1.21</td>
<td>21.</td>
</tr>
<tr>
<td>X2</td>
<td>1</td>
<td>45</td>
<td>23.1</td>
<td>1.</td>
</tr>
<tr>
<td>X3</td>
<td>1</td>
<td>45</td>
<td>10.</td>
<td>10.</td>
</tr>
</tbody>
</table>

- X1+X2+X3 explains 34.
  - X1+X2 explains 24
  - X1+X3 explains 33
  - X2+X3 explains 13
In practice

Interpret the following ANOVA

<table>
<thead>
<tr>
<th>Src</th>
<th>NDF</th>
<th>DDF</th>
<th>F-inc</th>
<th>F-con</th>
<th>M</th>
<th>%P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>2</td>
<td>45</td>
<td>11.</td>
<td>21.</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>11</td>
<td>45</td>
<td>8.</td>
<td>1.</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Var</td>
<td>12</td>
<td>45</td>
<td>10.</td>
<td>10.</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Var.Site</td>
<td>99</td>
<td>45</td>
<td>5.</td>
<td>5.</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>
In practice

- Interpret the following ANOVA

<table>
<thead>
<tr>
<th>Src</th>
<th>NDF</th>
<th>DDF</th>
<th>F-inc</th>
<th>F-con</th>
<th>M</th>
<th>%P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>2</td>
<td>45</td>
<td>11.</td>
<td>21.</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>11</td>
<td>45</td>
<td>8.</td>
<td>1.</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Var</td>
<td>12</td>
<td>45</td>
<td>10.</td>
<td>10.</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Var.Site</td>
<td>99</td>
<td>45</td>
<td>5.</td>
<td>5.</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

- Sites are nested in Region so Region may not be tested after Site (Var.Site).
## Conditional F-tests

<table>
<thead>
<tr>
<th>Term</th>
<th>F-inc</th>
<th>F-con</th>
<th>Mcode</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B,C,B.C</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>A.B</td>
<td>A.B</td>
<td>A,B</td>
<td>A.B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>A,B,A.B</td>
<td>C</td>
</tr>
</tbody>
</table>
More complicated example

<table>
<thead>
<tr>
<th>AnOVar</th>
<th>NumDF</th>
<th>DenDF</th>
<th>F-inc</th>
<th>F-con</th>
<th>M</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>FD0</td>
<td>1</td>
<td>220.0</td>
<td>36.81</td>
<td>34.30</td>
<td>C</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>GreenDM</td>
<td>1</td>
<td>10.7</td>
<td>9.52</td>
<td>8.59</td>
<td>A</td>
<td>0.014</td>
</tr>
<tr>
<td>Pos</td>
<td>1</td>
<td>10.0</td>
<td>93.09</td>
<td>93.03</td>
<td>A</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Green.Pos</td>
<td>1</td>
<td>10.0</td>
<td>0.57</td>
<td>0.04</td>
<td>B</td>
<td>0.847</td>
</tr>
<tr>
<td>HM</td>
<td>1</td>
<td>9.5</td>
<td>0.12</td>
<td>3.39</td>
<td>B</td>
<td>0.095</td>
</tr>
<tr>
<td>Pos.HM</td>
<td>1</td>
<td>10.0</td>
<td>9.40</td>
<td>9.88</td>
<td>C</td>
<td>0.010</td>
</tr>
<tr>
<td>SLen</td>
<td>1</td>
<td>420.8</td>
<td>0.94</td>
<td>0.82</td>
<td>A</td>
<td>0.368</td>
</tr>
<tr>
<td>SL.Pos</td>
<td>1</td>
<td>1004.2</td>
<td>1.66</td>
<td>1.72</td>
<td>B</td>
<td>0.190</td>
</tr>
<tr>
<td>SL.Green</td>
<td>1</td>
<td>396.3</td>
<td>4.18</td>
<td>6.18</td>
<td>B</td>
<td>0.014</td>
</tr>
<tr>
<td>SL.Pos.Green</td>
<td>1</td>
<td>1275.0</td>
<td>1.55</td>
<td>11.52</td>
<td>C</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
### More complicated example

<table>
<thead>
<tr>
<th>Description</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
<th>Value5</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>at (Group, 4)</td>
<td>11.2</td>
<td>5.02</td>
<td>0.02</td>
<td>A</td>
<td>0.878</td>
<td></td>
</tr>
<tr>
<td>at (Group, 1)</td>
<td>9.4</td>
<td>5.98</td>
<td>0.00</td>
<td>A</td>
<td>0.978</td>
<td></td>
</tr>
<tr>
<td>at (G, 4).Pos</td>
<td>10.1</td>
<td>0.22</td>
<td>0.60</td>
<td>B</td>
<td>0.456</td>
<td></td>
</tr>
<tr>
<td>at (G, 1).Pos</td>
<td>9.9</td>
<td>0.99</td>
<td>0.15</td>
<td>B</td>
<td>0.703</td>
<td></td>
</tr>
<tr>
<td>at (G, 4).Pos.SL</td>
<td>1786.9</td>
<td>3.20</td>
<td>18.67</td>
<td>C</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>at (G, 1).Pos.SL</td>
<td>1125.4</td>
<td>4.69</td>
<td>10.95</td>
<td>C</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Clover</td>
<td>9.8</td>
<td>0.71</td>
<td>0.73</td>
<td>A</td>
<td>0.414</td>
<td></td>
</tr>
<tr>
<td>Clov.SL</td>
<td>535.4</td>
<td>1.07</td>
<td>1.07</td>
<td>B</td>
<td>0.302</td>
<td></td>
</tr>
<tr>
<td>Clov.Pos</td>
<td>10.0</td>
<td>0.57</td>
<td>0.57</td>
<td>B</td>
<td>0.467</td>
<td></td>
</tr>
<tr>
<td>Clov.SL.Pos</td>
<td>1269.5</td>
<td>16.06</td>
<td>16.06</td>
<td>C</td>
<td>&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>
ASReml tutorial

D3 Prediction in Linear Mixed Models

Arthur Gilmour
Prediction

- This is the process of calculating linear combinations of the effects to summarise aspects of the analysis. e.g.
  - plotting a fitted spline curve
  - creating treatment means

- Details available in Welham et al and Gilmour et al. with coauthors B Cullis, B Gogel & R Thompson
Plan

- Introduction - The linear mixed model
- Prediction in Large problems
- Estimability
- Random effects
- Specification
**linear mixed model**

\[ y = X\tau + Zu + e; \]
\[ y \sim N(X\tau, R + ZGZ^T); \]
\[ u \sim N(0, G); e \sim N(0, R) \]
Mixed model equations

\[
\begin{pmatrix}
X^T R^{-1} X & X^T R^{-1} Z \\
Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1}
\end{pmatrix}
\begin{pmatrix}
\tau \\
u
\end{pmatrix}
= 
\begin{pmatrix}
X^T R^{-1} y \\
Z^T R^{-1} y
\end{pmatrix}
\]
rewritten as \( C\beta = W^T R^{-1} y \)
giving \( \hat{\beta} = C^{-1} W^T R^{-1} y \)
prediction \( \pi = D\hat{\beta} \)
with variance \( D C^{-1} D^T \)
Large Models

- ASReml fits large models; avoids forming all elements of $C^{-1}$
- Absorbing $C$ in
  \[
  \begin{pmatrix}
    y^T R^{-1} y & 0 & y^T R^{-1} W \\
    0 & 0 & D \\
    W^T R^{-1} y & D^T & C
  \end{pmatrix}
  \]
gives

\[
  \begin{pmatrix}
    y^T P y & -\pi^T \\
    -\pi & -DC^{-1} D^T
  \end{pmatrix}
  \]

- Done in final iteration
Estimability

- No data. Height of males, females and angels
- Overparameterization: model $\mu + \alpha_i$
  $\alpha_i$ is non-estimable - infinitely many solutions

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>-2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
## Estimability

- **Over-modelling** \( \mu + \alpha_i + \beta_j \)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>10</td>
<td>14</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>15</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>17</td>
<td>18</td>
</tr>
</tbody>
</table>
Averaging over incomplete tables

- No problem if table is all estimable even though some cells have no actual data i.e. interaction is omitted or is random.

- Sometimes sensible to average only over cells containing data. E.g. variety.year.location table: form variety table averaging over the experiments which form an incomplete year.location table.
During absorption process, nonestimability is present when the row of $D$ is not zero but the row of $C$ is zero.

\[
\begin{array}{c|ccc|c}
-25/5 & -1/5 \\
* & 0 & -1/0 \\
-30/6 & 0 & 0 & -1/5 \\
\hline
25 & 1 & 0 & 0 & 5 \\
0 & 0 & 1 & 0 & 0 \\
30 & 0 & 0 & 1 & 0 & 0 & 5 \\
\end{array}
\]
Random effects

- The residual term is usually ignored. But kriging is prediction based on a correlated residual.

- Other random terms might be
  - error terms (usually ignored) or
  - treatment terms - may be predicted, averaged (conditional) or ignored (marginal).

- Averaging over random lowers SE, may not affect SED
Specification

- Mixed model consists of model terms derived from factors and covariates.
  - Classify set - defines table to predict
  - Average set - usually other fixed variables
  - Ignored set - usually other random variables

- Set specific levels to be predicted
- Whether to average over cells with data
- Specific in/ex-clusion of model terms
- Whether to print nonestimable solutions
Growth of 5 Orange trees

Title: orange.  
Y = circumference  
X = age  
Y-axis: 30.0000 214.0000  
X-axis: 118.0000 1582.0000
Orange tree model

- Fixed - mu season day
- Random - Tree Tree.day spl(day)
Conclusion A

- Prediction from a linear model is essential for reporting results
- User must be advised if predictions are non-estimable functions
- Algorithm must be able to handle large and complex models
- Algorithm must give control of how to handle random terms
Predict directive

- basic PREDICT syntax
- Where to place the PREDICT directive
- additional qualifiers
- other issues - spline interpolation
Basic syntax

- PREDICT classify set [qualifiers]
  !present present set,
  !average factor [weights]
- Each factor name in the classify and present sets may be followed by list of actual values to predict
Where it goes

- Immediately after the linear model (before the Variance header line)
  - or after all R and G structure lines

- There can be many PREDICT statements.
  - Parsed at start
  - Design formed before first iteration
  - Prediction done in final iteration
Basic concepts

- A hyper table defined by the factors in the model possibly ignoring some which only appear in random terms e.g. breed year animal sex

- Collapsed by averaging dimensions to produce the predict table.

- E.g \( y \sim \mu \text{ variety } !r \text{ block predict variety} \)
Hypertable control

- !PARALLEL a b c
  - In model need to link factor and covariate versions of same data with lin() or spl()
- data line qualifiers !PPOINTS and !PVALS for covariates (especially splines)
- specify particular levels of classify and present factors.
  - Predict breed sex 1
Fine control - ignoring model terms

- Default - ignore model terms involving factors which only occur in random model terms (e.g. animal but not spl())

- Control
  - !USE model_terms
  - !ONLYUSE model_terms
  - !IGNORE model_terms
  - !EXCEPT model_terms
Output control

- !PRINTALL
- !SED, !VPV
  - increased memory required for V matrix
- Backtransformation !LOGIT !PROBIT
  !INVERSE !SQRT !COMPLOGLOG !LOGN
  !POWER p Deprecated qualifier
- !FITMARGIN if classify set is two-way table causes marginal tables to be produced.
General

- !FINAL command line option does one !CONTINUE iteration making it easy to do modelling then add predict statements at