Scope for estimation of variances due to sex-linked, maternal and dominance effects in mixed model analyses

Karin Meyer

Animal Genetics and Breeding Unit, University of New England, Armidale
kmeyer@didgeridoo.une.edu.au

Introduction

Recent paper on sexual dimorphism & X-linked variation
- Extended model with autosomal, sex-linked, dominance & maternal effects
- New experimental design proposed
t⇒ full set of covariances among relatives
- ‘Animal model’ analyses advocated
t⇒ no application for X-linked effects so far

Objectives
Examine scope for estimation for model & design suggested

The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity 97:319–328.

Experimental design

Generate many types of relatives
- 3 generations
  - Generation 1: 4 unrelated pairs ⇒ 4 offspring/pair
t⇒ pairs 1 & 2: 1 ♂, 2 ♀; pairs 3 & 4: 4 ♀
  - Generation 2: 4 sires ⇒ each mated to 3 unrelated dams
  - Generation 3: 12 families
⇒ Paternal half-sibs, full-sibs, single & double first cousins
Model

- Standard linear mixed model

\[ y = X\beta + Z(a + s + da + ds) + W(m + c) + e \]

<table>
<thead>
<tr>
<th>Term</th>
<th>Design M.</th>
<th>Variance</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>X</td>
<td>V</td>
<td>observations</td>
</tr>
<tr>
<td>β</td>
<td>X</td>
<td>-</td>
<td>fixed</td>
</tr>
<tr>
<td>a</td>
<td>Z</td>
<td>σ^2_A</td>
<td>additive genetic; autosomal</td>
</tr>
<tr>
<td>s</td>
<td>Z</td>
<td>σ^2_S</td>
<td>additive genetic; sex-linked</td>
</tr>
<tr>
<td>da</td>
<td>Z</td>
<td>σ^2_DA</td>
<td>dominance; autosomal</td>
</tr>
<tr>
<td>ds</td>
<td>Z</td>
<td>σ^2_DS</td>
<td>dominance; sex-linked</td>
</tr>
<tr>
<td>m</td>
<td>W</td>
<td>σ^2_M</td>
<td>maternal genetic; autosomal</td>
</tr>
<tr>
<td>c</td>
<td>W</td>
<td>σ^2_C</td>
<td>maternal perm. environmental</td>
</tr>
<tr>
<td>e</td>
<td>I</td>
<td>σ^2_E</td>
<td>residual</td>
</tr>
</tbody>
</table>

Likelihood

- Unrelated families of equal structure: \( V = I \otimes V_0 \)
- No fixed effects
- Likelihood

\[ \log L = \text{const.} - \frac{d}{2} \left( \log |V_0| + \text{tr} \left( V_0^{-1} M \right) \right) \]
- Hessian matrix

\[ H = -E \left[ \frac{\partial^2 \log L}{\partial \theta_k \partial \theta_m} \right] = \frac{d}{2} \text{tr} \left( V_0^{-1} \frac{\partial V_0}{\partial \theta_k} V_0^{-1} \frac{\partial V_0}{\partial \theta_m} \right) \]

\( d \) degrees of freedom
\( M \) Matrix of mean SQ/CP (across families)
\( \theta_k \) \( k \)-th parameter to be estimated

Profile likelihood

- Vector of parameters \( \theta \)
- Partition into:
  - \( \theta_k \) \( \Rightarrow \) parameter of interest
  - \( \theta_{\neq k} \) \( \Rightarrow \) remaining parameters
- Profile likelihood: \( \log L_P(\theta_k) \)
  - maximize \( \log L(\theta_{\neq k}|\theta_k = t) \)
  - gives likelihood ratio test of \( H_0 : \theta_k = t \)
  - projection of likelihood surface on axis for \( \theta_k \)
  - curvature (2nd derivatives) preserved
Confidence interval

- Hessian matrix:
  - Inverse $H^{-1} = \{h_{km}\}$
  - Asymptotic, lower bound sampling covariances of $\hat{\theta}$
  - Normal approximation: $95\%$ confidence interval for $\hat{\theta}_k$

$$[t_L, t_U] = [\hat{\theta}_k - 1.96\sqrt{h_{kk}}, \hat{\theta}_k + 1.96\sqrt{h_{kk}}]$$

- Profile log $L$
  - Find values $t_L, t_U$ for
  $$\log L_P - \log L_m = -1.96^2 \chi^2_{1.95\%}$$
  $$= -1.92$$

Simulation study

- Design suggested by Fairbairn & Roff (2006)
  - 4 offspring/mating in generation 3 ($2\sigma$, $2\varphi$)
  - 72 records per family + 8 dams w/out records
  - 200 unrelated families

- Sample $M$ from Wishart distribution
- Population values
  - $\sigma_A^2 = 400$, $\sigma_S^2 = 100$, $\sigma_M^2 = 120$, $\sigma_{AM} = -30$, $\sigma_C^2 = 150$, $\sigma_{DA}^2 = 60$, $\sigma_{DS}^2 = 20$ and $\sigma_E^2 = 600$.

- Consider full model & models with subsets of effects
- 50,000 replicates per analysis
- REML estimation using Method of Scoring

Profile likelihood

- Constrain parameters
  - Variance components $\geq 10^{-8}$
  - $\sigma_{AM}/\sqrt{\sigma_A^2 \sigma_M^2} \in [-1, 1]$

- Obtain $A$ and $S$ from pedigree (tabular method)

- Construct $D_A$ and $D_S$ from coefficients in $E[Cov(\text{relatives})]$ given by F & R (2006).

- Empirical 95% confidence limits
  - Values truncating top & bottom 2.5% of estimates
Confidence intervals

<table>
<thead>
<tr>
<th>Variance component</th>
<th>( \sigma_A^2 )</th>
<th>( \sigma_S^2 )</th>
<th>( \sigma_M^2 )</th>
<th>( \sigma_{AM}^2 )</th>
<th>( \sigma_C^2 )</th>
<th>( \sigma_{DA}^2 )</th>
<th>( \sigma_{DS}^2 )</th>
<th>( \sigma_E^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hessian</td>
<td>-29</td>
<td>28</td>
<td>-63</td>
<td>-10</td>
<td>-3</td>
<td>6</td>
<td>-15</td>
<td></td>
</tr>
<tr>
<td>Profile log L</td>
<td>-30</td>
<td>-9</td>
<td>-5</td>
<td>6</td>
<td>2</td>
<td>-21</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>28</td>
<td>-9</td>
<td>-65</td>
<td>-33</td>
<td>-16</td>
<td>1</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

- Below diagonal: values from Hessian matrix
- Above diagonal: empirical values from simulation

Sampling correlations \((\times 100)\)

<table>
<thead>
<tr>
<th>( \sigma_A^2 )</th>
<th>( \sigma_S^2 )</th>
<th>( \sigma_M^2 )</th>
<th>( \sigma_{AM}^2 )</th>
<th>( \sigma_C^2 )</th>
<th>( \sigma_{DA}^2 )</th>
<th>( \sigma_{DS}^2 )</th>
<th>( \sigma_E^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHS</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
<td>1</td>
<td>1</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>FS</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1</td>
<td>1</td>
<td>1/4</td>
</tr>
<tr>
<td>SFC</td>
<td>1/8</td>
<td>3/8</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/4</td>
</tr>
<tr>
<td>DFC</td>
<td>1/4</td>
<td>3/8</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/16</td>
</tr>
</tbody>
</table>

- Design suggested provides many types of covariances between relatives
- But: unlikely to support accurate estimation of all 8 \((co)\)variance components
- ‘Animal model’ analysis not a magic bullet
- Inspection of \(E[Cov(\text{relatives})]\) \(\Rightarrow\) incomplete information
- Profile (log) likelihood calculations recommended
- Provide additional insights
- Good agreement with simulation results
- Straightforward & computationally undemanding
- Should be part of planning experiments

Conclusions

<table>
<thead>
<tr>
<th>Cov(\text{relatives})</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2_A )</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>PHS</td>
</tr>
<tr>
<td>FS</td>
</tr>
<tr>
<td>SFC</td>
</tr>
<tr>
<td>DFC</td>
</tr>
</tbody>
</table>

- SFC: Dams full sisters
- DFC: Sires full brothers, dams full sisters