Estimation of variance components for binary threshold models

Freddy Fikse [1], Lars Rönnegård [1,2]
[1] Dept. of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden
[2] Statistics unit, Dalarna University, Borlänge, Sweden

Aim

Investigate **bias** in **variance components** for **binary threshold models** for two animal breeding examples**.**

Background

- Penalized quasi-likelihood (PQL) common estimation technique for threshold models
- PQL is implemented in widely used software packages
- Severe bias in PQL-estimates of variance components reported in statistical literature

Conclusions

- **MCMC (posterior mode) gave unbiased estimates, except when data contained little information (low h2, low frequency)**

- **PQL both over- and underestimation; depending on data structure**

Methods

Five software packages to estimate variance components:
- DMUAI (uses PQL);
- ASREML (uses PQL);
- GLIMMiX (PQL or Laplace);
- GENSTAT (uses PQL);
- MCMC.

Materials

- Monte Carlo simulation of binary response; frequency levels 5 and 50%.
- Dairy cattle data structure:
    - 90 thousand records, 760 sires.
    - Random: sire, herd-year; Fixed: age at calving, calving month.
- Horse data structure:
    - 1250 records, 33 sires.
    - Random: Sire; Fixed: birth year, gender, region.

Results

- Cattle data structure:
    - PQL always biased;
    - MCMC unbiased.

- Horse data structure:
    - High simulated freq (50%): no significant bias;
    - Low simulated freq (5%): no clear trend.

Bias in sire variance for cattle data structure simulated frequency 50%
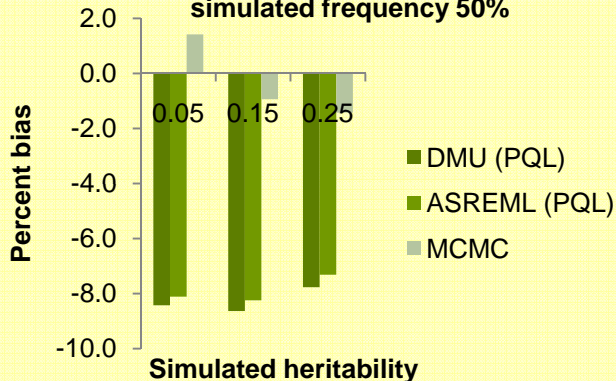
Bias in sire variance for horse data structure simulated frequency 5%