

Hierarchical Generalized Linear Models with Random Effects and Variance Heterogeneity for GWAS and Beyond

Lars Rönnegård
SLU & Dalarna University, Sweden

Hierarchical Generalized Linear Models

This poster is a review of

- the HGLM framework
- their implementations in the R software
- the variety of possible applications in genetics.

In R, the *glm* function has been extended in the two packages **dglm** (by GK Smyth) and **hglm**. These extensions give modelling flexibility, especially useful in quantitative genetics (Figure 1).

The HGLM framework extends generalized linear models to include random effects and variance heterogeneity

Double hierarchical generalized linear models (DHGLM)

- includes also random effects in the residual variance
- a fast implementation is available in ASReml.

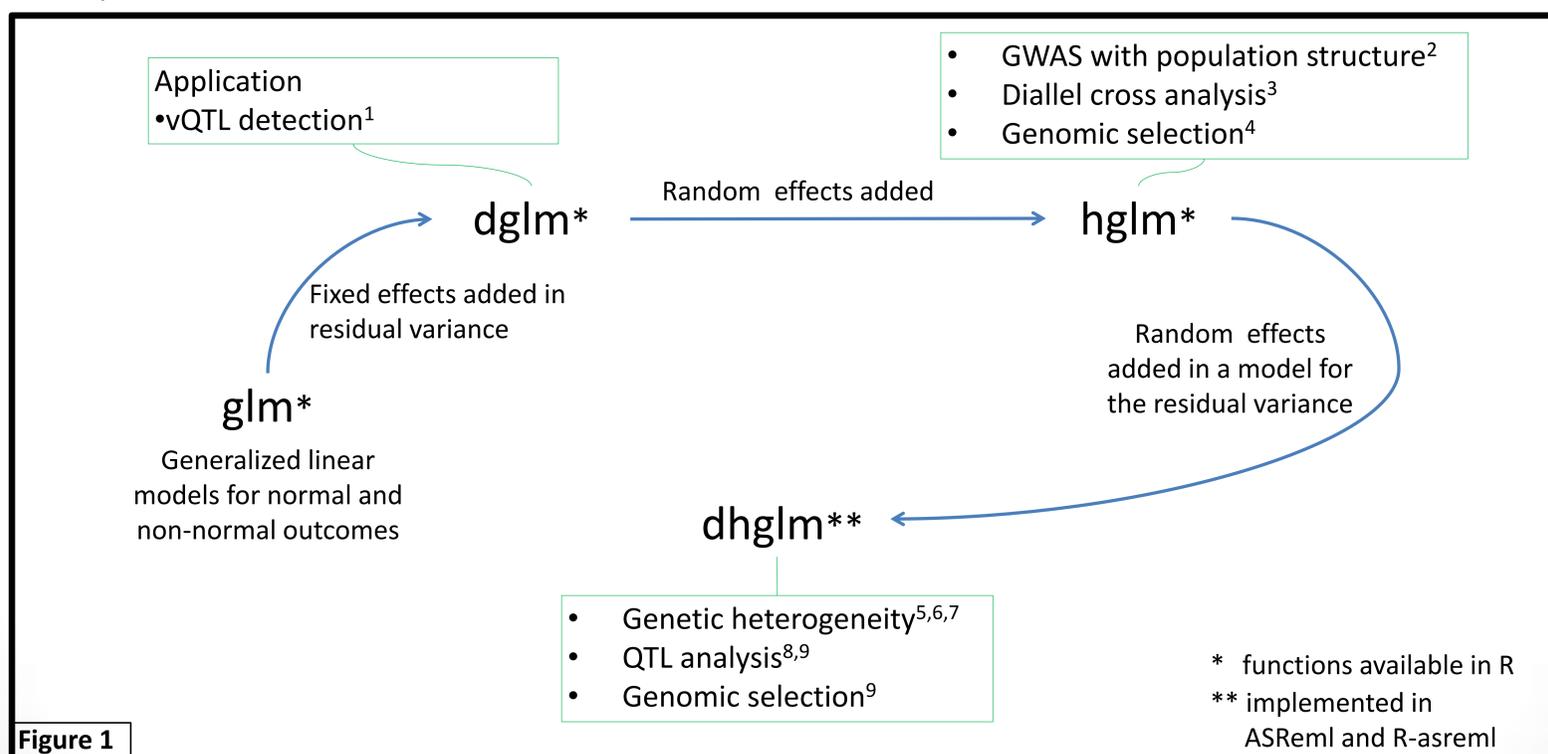
The h-likelihood: a third school in statistics

The two dominating schools in statistics are the frequentist and Bayesian schools.

A third, and computationally attractive, alternative is Lee & Nelder's hierarchical (h-) likelihood inference, where

- GLMs with random effects can be analyzed
- estimates can be computed using interconnected GLMs.
- consists of three objects: data, fixed parameters and unobserved random effects (unobservables).

This is different from the traditional Bayesian models that consist of the two objects, the data and unobservables, while frequentist's models consist of the data and parameters.



References

1. Rönnegård & Valdar (2011) Detecting major genetic loci controlling phenotypic variability in experimental crosses. **Genetics** 188:435-447.

2. Rönnegård, Shen & Alam (2010) hglm: A package for fitting hierarchical generalized linear models. **The R Journal** 2:20-28. (used in the *GenABEL* package for GWAS with population structure)

3. Lenarcic et al. (2012) **Genetics** 190:413-435.

4. Shen, Alam, Fikse, Rönnegård (2012) Fast generalized ridge regression for models including heteroscedastic effects in quantitative genetics. (manuscript)

5. Rönnegård, Felleki, Fikse, Mulder & Strandberg (2010) Genetic heterogeneity of residual variance - estimation of variance components using double hierarchical generalized linear models. **Genetics Selection Evolution** 42:8.

6. Rönnegård, Felleki, Fikse, Mulder & Strandberg (2012) Variance component and breeding value estimation for environmental sensitivity in Swedish Holstein dairy cattle (submitted)

7. Felleki, Lee, Lee, Gilmour & Rönnegård (2012) Estimation of breeding values for mean and dispersion, their variance and correlation using double hierarchical generalized linear models. (manuscript)

8. Rönnegård & Lee (2010) Hierarchical generalized linear models have a great potential in genetics and animal breeding. In proceedings: WCGALP, Leipzig, Germany.

9. Shen, Rönnegård & Carlborg (2011). Hierarchical likelihood opens a new way of estimating genetic values using genome-wide dense marker maps. **BMC Proceedings** 5(Suppl 3):S14.