Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Animal Nutrition
and Management

Kjell och Märta Beijers Stiftelse
- FRÄMJAR VETENSKAP OCH KULTUR -

Beijerlaboratorierna i Uppsala
The Beijer Laboratories in Uppsala

# Genetic Heteroscedasticity: DHGLM link functions

EAAP Belfast 2016-09-01

Majbritt Felleki, PhD

Researcher at The Beijer Laboratories in Uppsala at Swedish University of Agricultural Sciences

# Heteroscedasticity

- Linear model $y = \mu + e$,

$$y_i = \mu_i + e_i$$

- Heteroscedasticity

$$e_i \sim \mathrm{N}(0, \sigma_i^2)$$

# Genetic heteroscedasticity

- Genetic control of environmental variation

$$y = X\beta + Za + Wp + e,$$

$$e \sim \mathrm{MVN}(0, \Phi),\ \mathrm{diag}\,\Phi = \varphi,$$

$$g(\varphi) = X_d\beta_d + Z_d a_d + W_d p_d$$

# The link function $g$

- San Cristobal-Gaudy *et al.* (1998), the exponential model:

$$g(\varphi) = \log \varphi$$

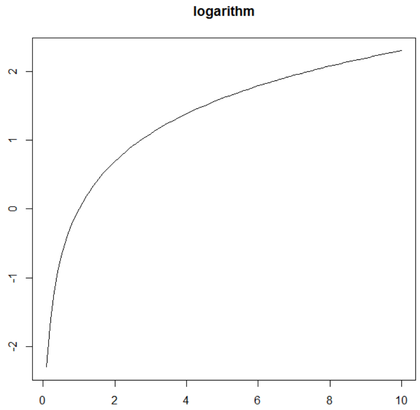
- Garcia *et al.* (2009), the standard deviation model:

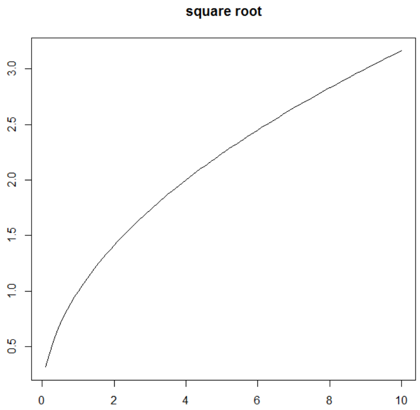
$$g(\varphi) = \sqrt{\varphi}$$

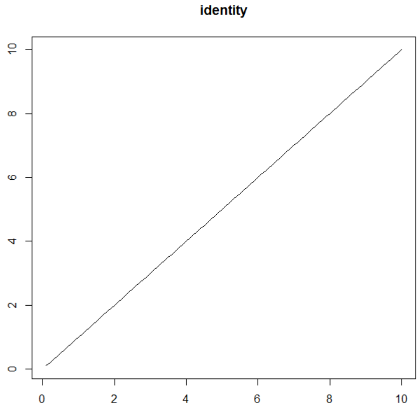
- Hill & Zhang (2004); Mulder *et al.* (2007), the additive model:

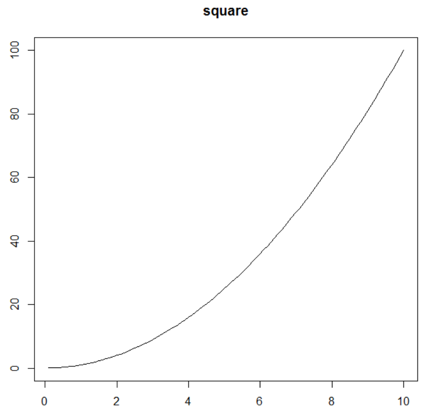
$$g(\varphi) = \varphi$$

# Link functions


logarithm


square root


identity


square

# Aim

- The aim is to implement an validate the use of DHGLM with the link functions

$$g(\varphi) = \log \varphi$$

$$g(\varphi) = \sqrt{\varphi}$$

$$g(\varphi) = \varphi$$

$$g(\varphi) = \varphi^{3/2}$$

$$g(\varphi) = \varphi^2$$

# Pig litter size data

- The implementation was done using the pig litter size data used in Sorensen & Waagepetersen (2003)
- The data includes10 060 records on 4149 sows and a total number of 6437 pigs in the pedigree

# DHGLM Weights

- The DHGLM weights are (mean and residual variance level respectively)

$$\left(g^{-1}(\widehat{y_d})\right)^{-1}$$

$$\left(g'(\hat{\varphi})^2 \frac{2}{1-h} \hat{\varphi}^2\right)^{-1}$$

- The residual variance working variable is

$$y_d = g(\hat{\varphi}) + g'(\hat{\varphi})\left(\frac{e^2}{1-h} - \hat{\varphi}\right)$$

# Negative variances

- Using links sensible to possible negative predicted values

$$\widetilde{y_d} = \max(\widehat{y_d}, \delta),$$

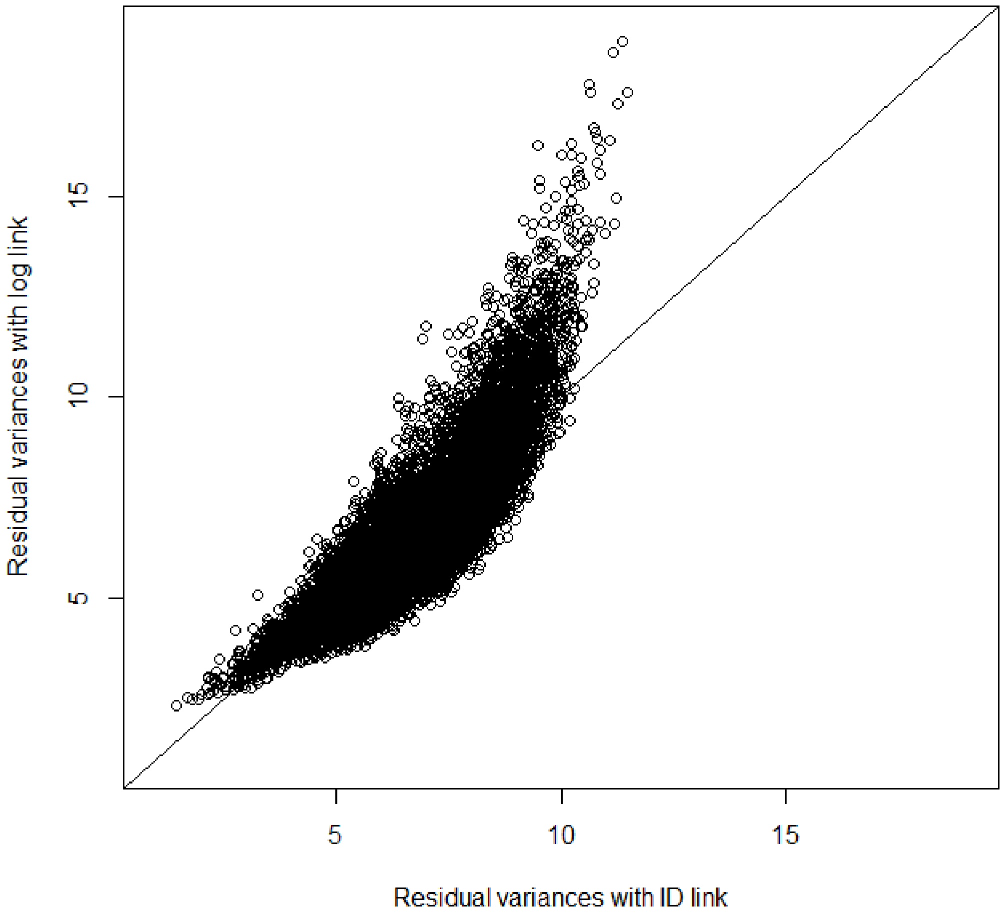
$$\delta = 0.1,\ \varphi \longmapsto \sqrt{\varphi},\ \varphi,\ \text{and}\ \varphi^{3/2}$$

$$\delta = 0.5,\ \varphi \longmapsto \varphi^{2}$$

# Estimates

| Link $g$ | $\sigma_a^2$ | $\rho$ | $\sigma_{a_d}^2$ | $\sigma_p^2$ | $\sigma_{p_d}^2$ | $\sigma_d^2$ | $g^{-1}(\beta_{d_0})$ |
|---|---|---|---|---|---|---|---|
| Log | 1.61 | -0.52 | 0.15 | 0.28 | 0.05 | 0.68 | 5.43 |
| Sq root | 1.60 | -0.51 | 0.20 | 0.40 | 0.00 | 0.79 | 5.56 |
| ID | 1.58 | -0.53 | 3.37 | 0.45 | 0.00 | 0.86 | 5.72 |
| ^(3/2) | 1.57 | -0.55 | 27.3 | 0.51 | 0.00 | 0.90 | 6.01 |
| Square | 1.56 | -0.59 | 136 | 0.58 | 0.00 | 0.93 | 6.32 |

# Comparison residual variances



# Simulation study

- Simulate with log, square root, and identity links
- Analyze with log, square root, and identity links

# Planned simulation study

- 1000 related animals, 5 repeated observations
- 3 sets of parameters (varying the genetic correlation)
- 3 link functions for simulation
- 3 link functions for analysis
- 30 replicates
- Altogether 3*3*3*30=810 runs

# Issues regarding the simulation study

- To translate the values of the parameters for the residual variance level among links

# Conclusion

- DHGLM analysis of heteroscedastic animal models and genetic correlations is implementable using several power link functions
- The residual variances are not identical among models
- Among link translation of the values of the parameters for the residual variance level has to be addressed

SLU

BLU

# Thanks


Kjell och Märta Beijers Stiftelse

- FRÄMJAR VETENSKAP OCH KULTUR -

SLU

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

In cooperation with
Loys Bodin, INRA UMR1388 / INPT ENSAT / INPT ENVT GenPhySe, F-31326 Castanet-Tolosan, France
Arthur Gilmour, Orange, NSW 2800, Australia