

DATA ANALYSIS USING HIERARCHICAL GENERALIZED LINEAR MODELS WITH R

Youngjo Lee, Lars Rönnegård
& Maengseok Noh

Poisson

$$L = \prod_{i=1}^n \frac{\mu^{y_i} e^{-\mu}}{y_i!}$$

$$\log(\mu) = X\beta$$

Binomial

$$L = \prod_{i=1}^n \binom{n}{y_i} \mu^{y_i} (1 - \mu)^{n - y_i}$$

$$\text{logit}(\mu) = X\beta$$

Gamma

$$L = \prod_{i=1}^n \frac{1}{\Gamma(k)} \theta^k y_i^{k-1} e^{-\frac{y_i}{\theta}}$$

$$k\theta \equiv \mu; \quad 1/\mu = X\beta$$

Gaussian

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2}$$

$$\mu = X\beta$$

GLM

Common estimation algorithm using iterative regression

- Fast and easy to implement
- ***Linear regression model checking tools!!!***



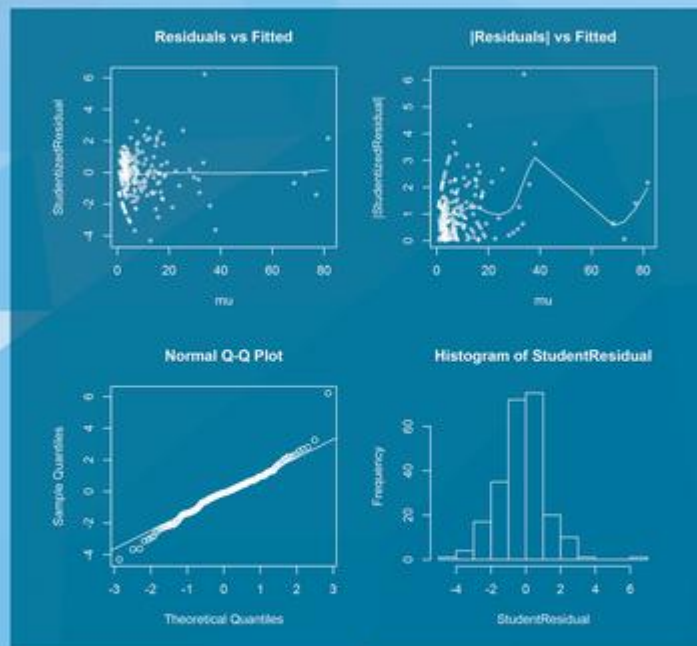
Lee Y. & Nelder J. A. (1996) "Hierarchical generalized linear models" JRSS B 619-678

GLM approach for fitting

- Linear mixed models
- Generalized linear mixed models (Laplace approximation)
- Mixed models with non-Gaussian random effects
- Above models + dispersion modelling

DATA ANALYSIS USING HIERARCHICAL GENERALIZED LINEAR MODELS WITH R

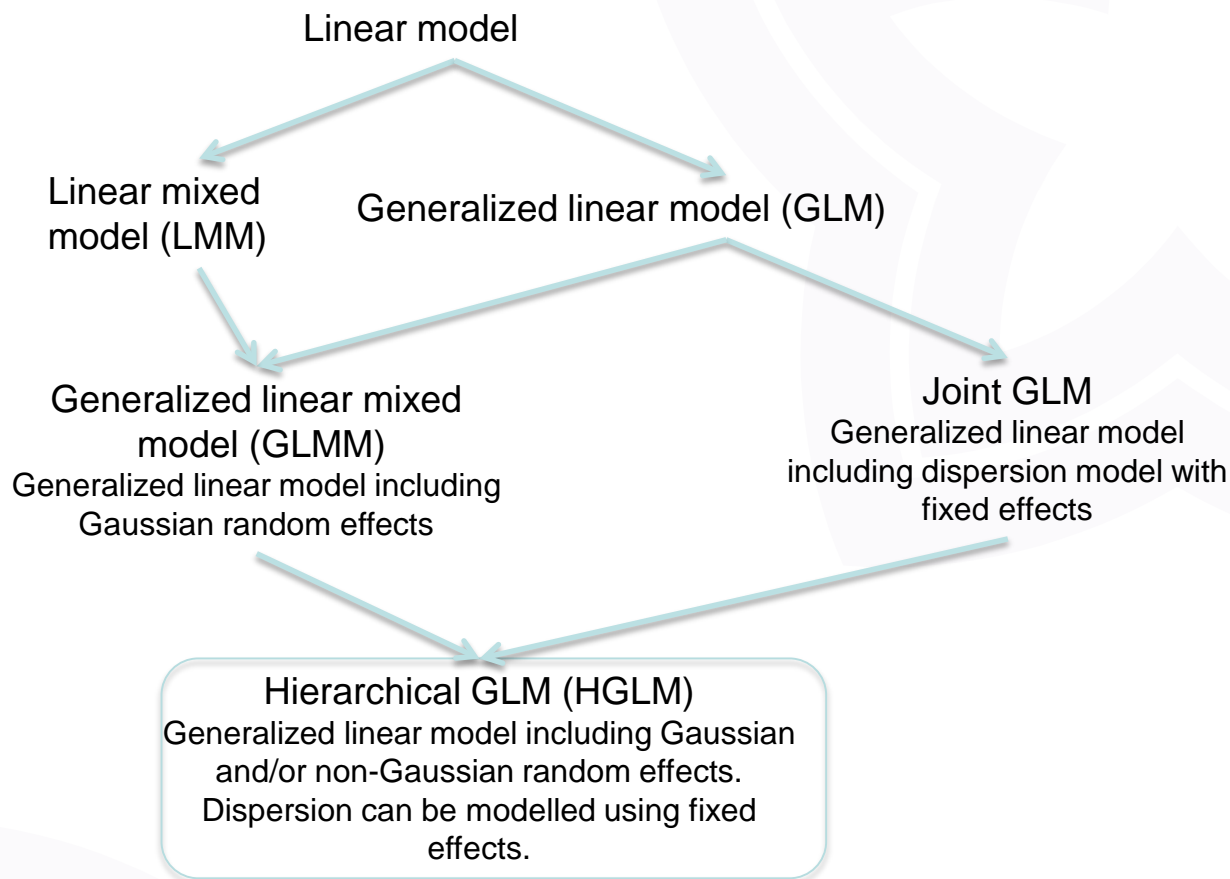
Coming out July 2017

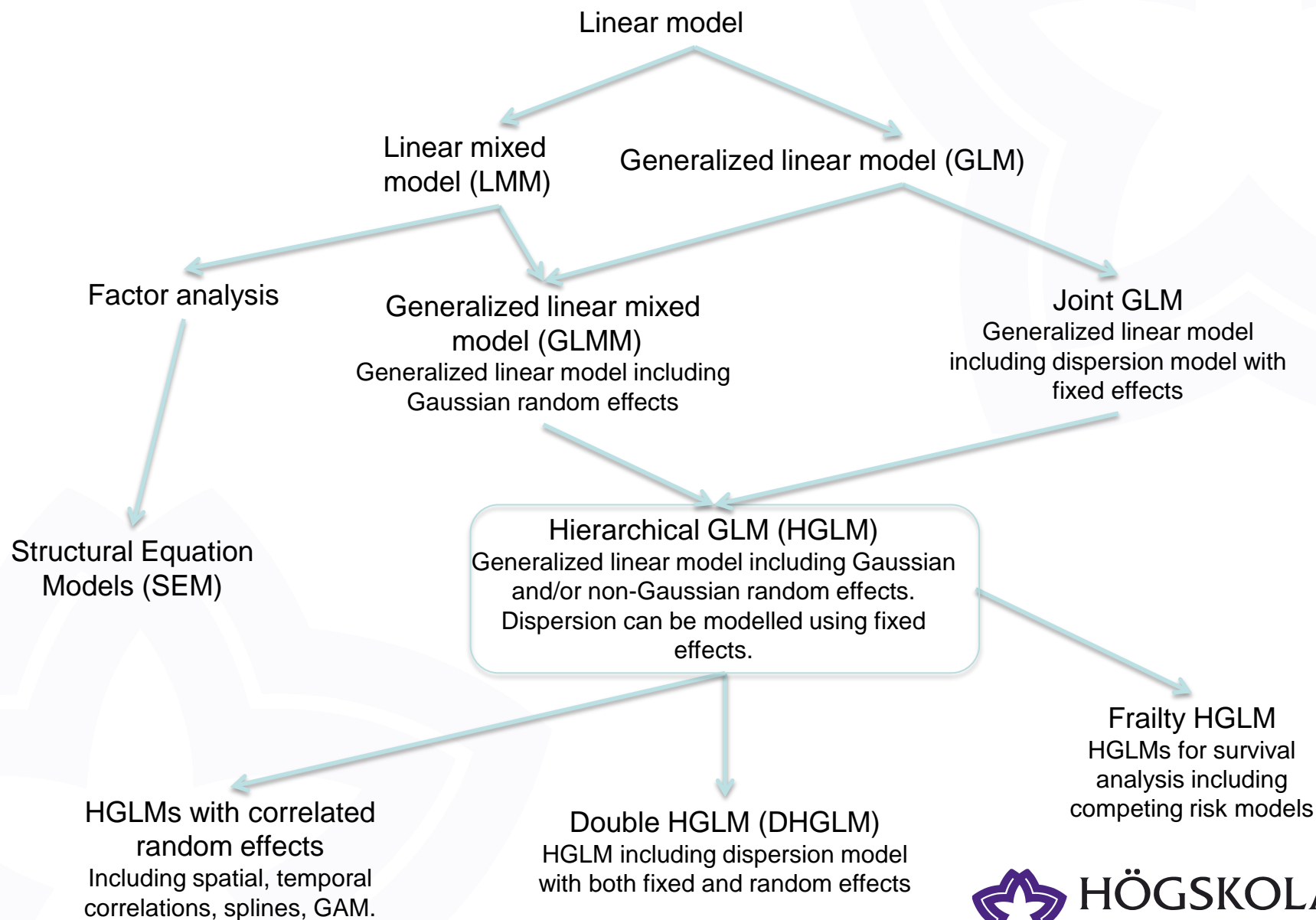


Youngjo Lee
Lars Rönnegård
Maengseok Noh

 CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

 HÖGSKOLAN
DALARNA

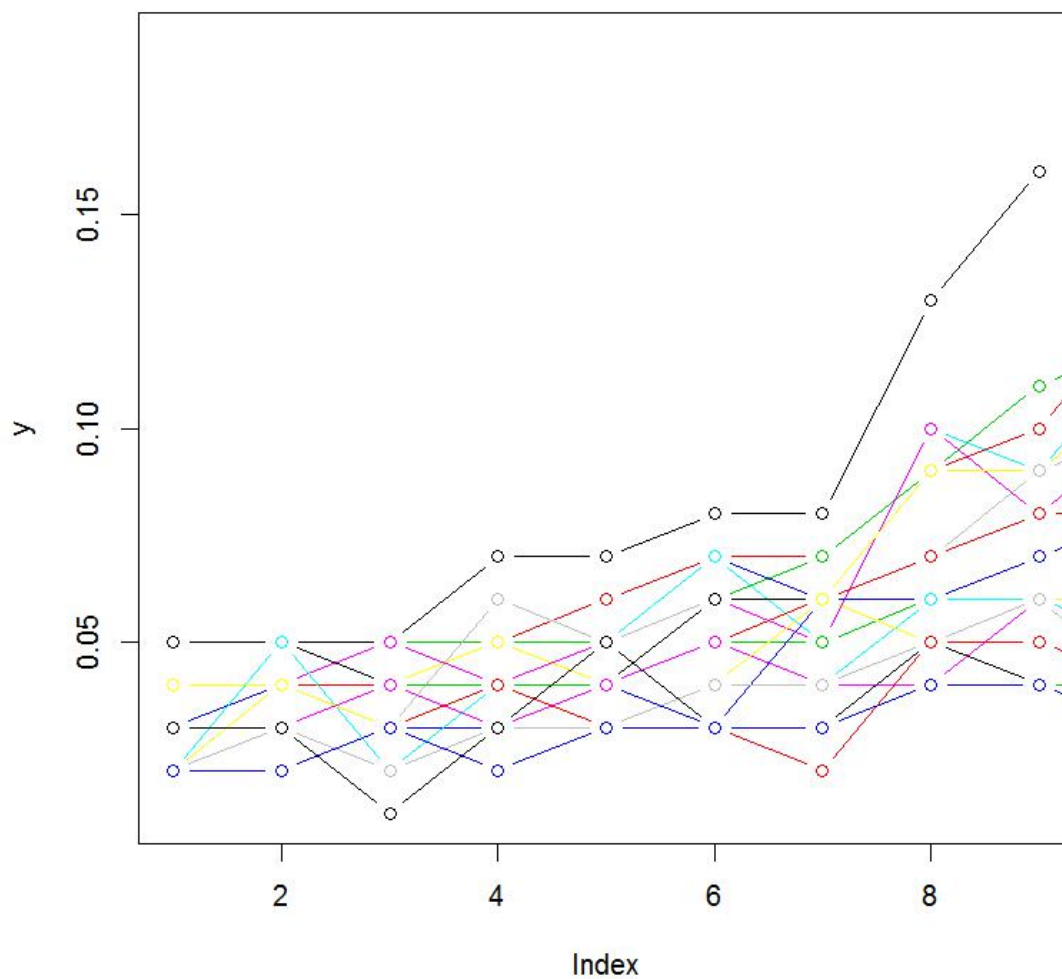




Crack growth data

Hudak et al. (1978) crack lengths measured on a compact tension steel.

- 21 metallic specimens, crack lengths recorded every 104 cycles
- y = increment of crack length
- crack0 = covariate for the mean part of the model
- cycle = covariate for the dispersion part of the model




```
## GLM ##
```

```
res_glm <- glm(y ~ crack0, family= Gamma(link=log), data=data_crack_growth)
```

```
## GLM ##  
res_glm <- glm(y ~ crack0, family= Gamma(link=log), data=data_crack_growth)  
  
library(dhglm)  
## HGLM I ##  
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
                           LinPred = y ~ crack0 + (1|specimen), RandDist = "inverse-gamma")  
model_phi <- DHGLMMODELING(Model="dispersion")  
res_hglm1 <- dhglmfit(RespDist="gamma", DataMain=data_crack_growth,  
                    MeanModel=model_mu, DispersionModel=model_phi)
```



```
## GLM ##
```

```
res_glm <- glm(y ~ crack0, family= Gamma(link=log), data=data_crack_growth)
```

```
library(dhglm)
```

```
## HGLM I ##
```

```
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
                          LinPred = y ~ crack0 + (1|specimen), RandDist = "inverse-gamma")
```

```
model_phi <- DHGLMMODELING(Model="dispersion")
```

```
res_hglm1 <- dhglmfit(RespDist="gamma", DataMain=data_crack_growth,  
                    MeanModel=model_mu, DispersionModel=model_phi)
```

```
## HGLM II ##
```

```
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
                          LinPred = y ~ crack0 + (1|specimen), RandDist="inverse-gamma")
```

```
model_phi <- DHGLMMODELING(Model = "dispersion", Link = "log",  
                          LinPred = phi ~ cycle)
```

```
res_hglm2 <- dhglmfit(RespDist = "gamma", DataMain = data_crack_growth,  
                    MeanModel = model_mu, DispersionModel = model_phi)
```

```
## GLM ##
```

```
res_glm <- glm(y ~ crack0, family= Gamma(link=log), data=data_crack_growth)
```

```
library(dhglm)
```

```
## HGLM I ##
```

```
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
  LinPred = y ~ crack0 + (1|specimen), RandDist = "inverse-gamma")
```

```
model_phi <- DHGLMMODELING(Model="dispersion")
```

```
res_hglm1 <- dhglmfit(RespDist="gamma", DataMain=data_crack_growth,  
  MeanModel=model_mu, DispersionModel=model_phi)
```

```
## HGLM II ##
```

```
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
  LinPred = y ~ crack0 + (1|specimen), RandDist="inverse-gamma")
```

```
model_phi <- DHGLMMODELING(Model = "dispersion", Link = "log",  
  LinPred = phi ~ cycle)
```

```
res_hglm2 <- dhglmfit(RespDist = "gamma", DataMain = data_crack_growth,  
  MeanModel = model_mu, DispersionModel = model_phi)
```

```
## DHGLM I ##
```

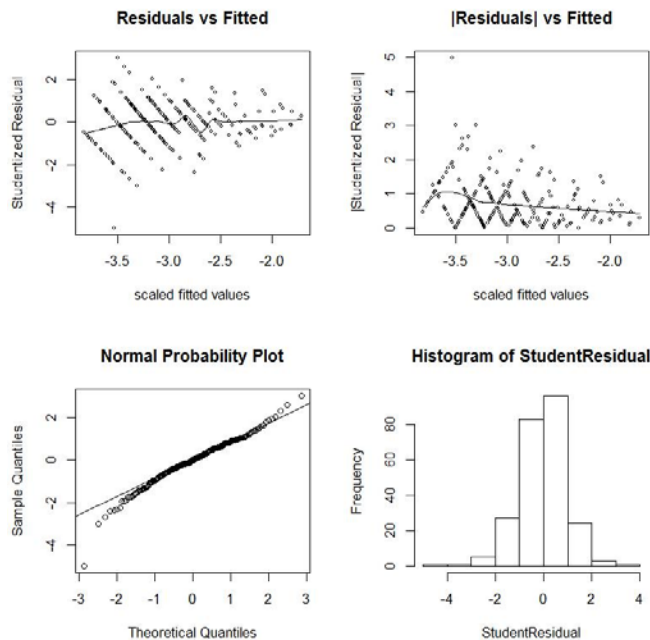
```
model_mu <- DHGLMMODELING(Model="mean", Link="log",  
  LinPred = y ~ crack0 + (1|specimen), RandDist="inverse-gamma")
```

```
model_phi <- DHGLMMODELING(Model="dispersion", Link="log",  
  LinPred = phi ~ cycle + (1|specimen), RandDist="gaussian")
```

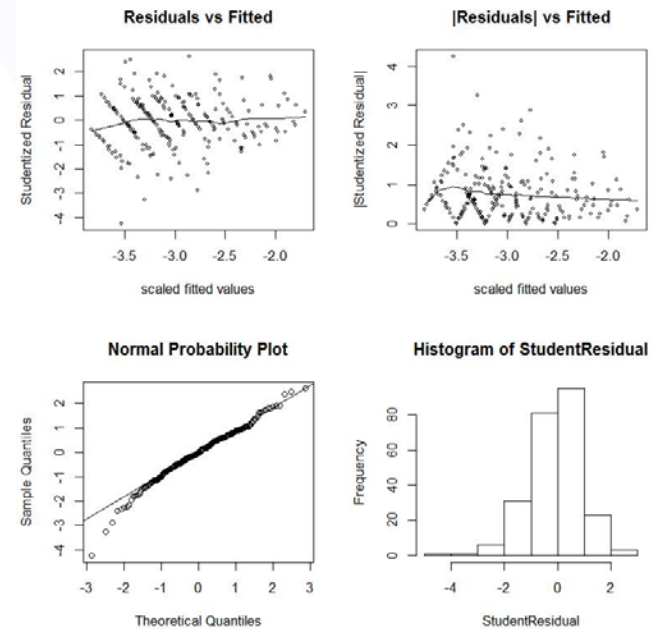
```
res_dhglm1 <- dhglmfit(RespDist = "gamma", DataMain = data_crack_growth,  
  MeanModel = model_mu, DispersionModel = model_phi)
```



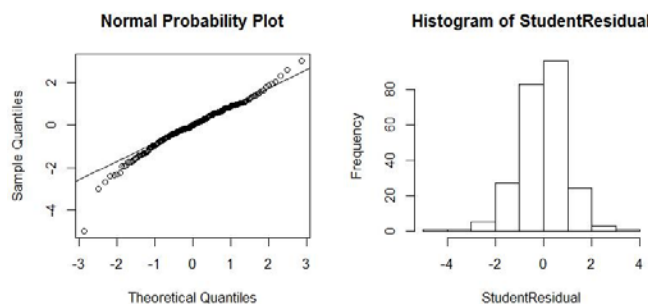
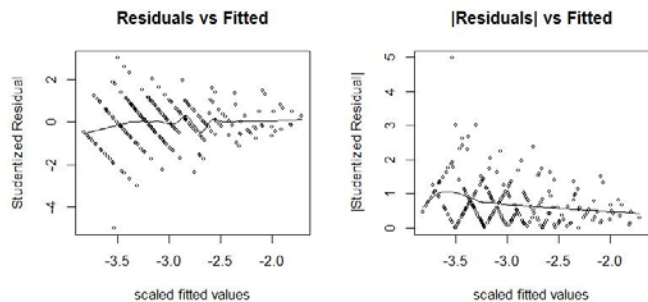
HGLM I



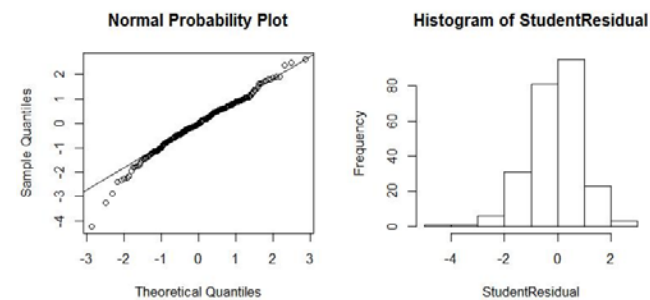
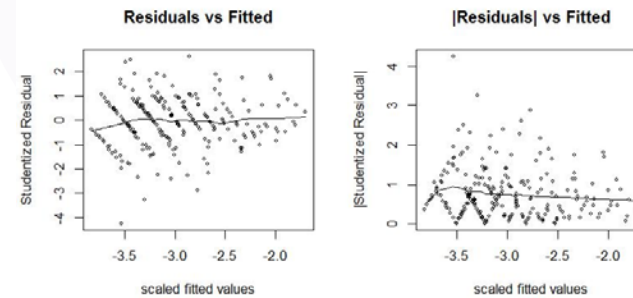
HGLM II



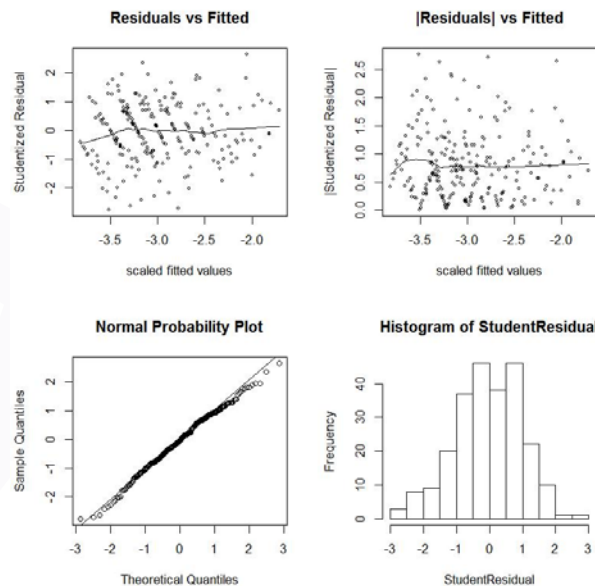
HGLM I



HGLM II



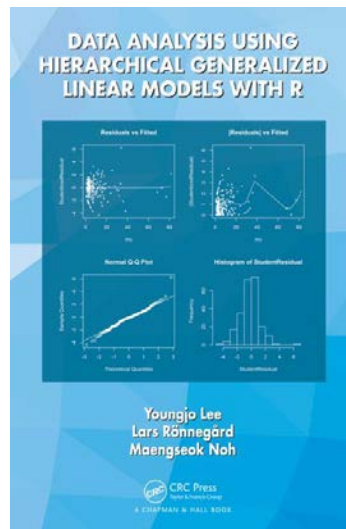
DHGLM I



LRN@DU.SE

Lee, Rönnegård, Noh

DATA ANALYSIS USING HIERARCHICAL GENERALIZED LINEAR MODELS WITH R



HÖGSKOLAN
DALARNA