Hierarchical Generalized Linear Models in Practice

Roger Payne

VSN International, 5 The Waterhouse, Waterhouse Street, Hemel Hempstead, UK and Rothamsted Research, Harpenden, UK

email: Roger.Payne@vsni.co.uk

RSS Statistical Computing Section, 3rd April 2008
HGLMs – introduction

• Hierarchical generalized linear models
  • extend generalized linear models to >1 source of error
  • include generalized linear mixed models as a special case
    • but the additional random terms are not constrained to follow a Normal distribution, nor to have an identity link
  • allow for modelling of the dispersion of the error terms
    • extending quasi-likelihood methods of Nelder & Pregibon (1987)
  • have an efficient fitting algorithm
    • no numerical integration is required
    • examples available in GenStat for Windows 9th Edition onwards

• Hierarchical generalized nonlinear models
  • include nonlinear parameters in the HGLM fixed model
    • in GenStat for Windows 10th Edition
HGLMs – definition

- expected value
  \[ E(y) = \mu \]
- link
  \[ \eta = g(\mu) \]
- distribution – Normal, Binomial, Poisson or Gamma (from exponential family)
- but linear predictor
  \[ \eta = X \beta + \sum Z_i \nu_i \]
  now contains additional vectors of random effects \( \nu_i \) with Normal, beta, gamma or inverse gamma distributions and with their own link functions
  - Normal-identity gives a GLMM but HGLM algorithms use much improved Laplace approximations in their use of adjusted profile likelihood
- inference by \textit{h-likelihood}

..
H-likelihood (LNP §4.5)

- is an extended likelihood \( L(\theta, \nu; y, \nu) \equiv L(\theta; y) \times L(\theta, \nu; \nu|y) \)
  where the random parameters \( \nu \) are on the canonical scale
  - i.e. \( L(\theta_1, \nu_{\theta_1}; y, \nu) / L(\theta_2, \nu_{\theta_2}; y, \nu) = L(\theta_1; y) / L(\theta_2; y) \)
  - where \( \nu_{\theta_1} \) and \( \nu_{\theta_2} \) are estimates of \( \nu \) for \( \theta \) at \( \theta_1 \) and \( \theta_2 \)
  - can treat like an ordinary log-likelihood
  - estimates and information matrices from joint maximization of h-likelihood are identical to those from maximizing the marginal likelihood
  - requires profile likelihood of \( \theta \) from the extended likelihood to be proportional to the marginal likelihood \( L(\theta) \)

- minor extension: canonical scale may not be available for all parameters – e.g. (say) for \( \nu \) but not for \( \phi \)
  - then make joint inferences for \( (\theta, \nu) \), but estimation of \( \phi \) needs marginal likelihood, approximated by adjusted profile likelihood

- if no canonical scale
  - use extended likelihood to estimate random parameters
  - use adjusted profile likelihood to estimate fixed parameters (but enhanced Laplace approximations available)
E.g. Normal-Normal HGLM

- \( y = X \beta + Z \nu + \epsilon \)
  - \( \epsilon \) follows multivariate Normal\((0, \Sigma)\), \( \nu \) follows multivariate Normal\((0, D)\)
  - \( \Sigma \) and \( D \) parameterized by variance-component parameters \( \tau = (\sigma^2, \sigma_v^2) \)

- so (LNP Example 6.1) we have linear predictor
  - \( \eta = g(\mu) = X \beta + Z \nu \)
  - where \( g() \) is the identity function, and \( u = \nu \)

- and GLMs
  - \( y|u \) follows a GLM distribution with
    - \( E(y|u) = \mu \)
    - \( \text{var}(y|u) = \phi V(\mu) \) where \( \phi = \sigma^2, V(\mu) = 1 \)
    - \( u \) is distributed as Normal\((0, \lambda)\), with \( \lambda = \sigma_v^2 \)

- constraint is \( E(u) = 0 \)

- extended likelihood (LNP §5.4) is
  - \( \log L(\theta, \nu; y, \nu) = \log f(y, \nu) = \log f(y|\nu) + \log f(\nu) \)
    - \( = \frac{1}{2} \log |2\pi \Sigma| - \frac{1}{2}(y-X\beta-Z\nu)^t \Sigma^{-1}(y-X\beta-Z\nu) - \frac{1}{2} \log |2\pi D| - \nu^t D^{-1} \nu \)
E.g. Normal-Normal HGLM

- extended likelihood (LNP §5.4)
  - \( \log L(\theta, \nu; y, \nu) = \log f(y, \nu) = \log f(y|\nu) + \log f(\nu) \)
  - \( = \frac{1}{2} \log |2\pi \Sigma| - \frac{1}{2}(y-X\beta-Z\nu)^t \Sigma^{-1}(y-X\beta-Z\nu) - \frac{1}{2} \log |2\pi D| - \nu^t D^{-1} \nu \)
  - note: \( \tau = (\sigma^2, \sigma^2) \) appears only in \( \Sigma \) and \( D \)

- Fisher information
  - \( I(\nu^\wedge) = Z^t \Sigma^{-1} Z + D^{-1} \)
  - depends on \( \tau \) but not on \( \beta \)

- so scale \( \nu \) is canonical for \( \beta \) (but not for \( \tau \), which must be estimated by a marginal likelihood)

- i.e. the extended likelihood is an \( h \)-likelihood
  - joint estimation is possible for \( \beta \) and \( \nu \)
  - but dispersion parameters \( \tau \) estimated by adjusted profile likelihood

...
E.g. Normal-Normal HGLM

- extended likelihood (LNP §5.4)
  - \[ \log L(\theta,\nu; y,\nu) = \log f(y,\nu) = \log f(y|\nu) + \log f(\nu) \]
    \[ = \frac{1}{2}\log|2\pi\Sigma| - \frac{1}{2}(y-X\beta-Z\nu)^t\Sigma^{-1}(y-X\beta-Z\nu) - \frac{1}{2}\log|2\pi D| - \frac{1}{2}\nu^tD^{-1}\nu \]

- differentiate with respect to \( \beta \) and \( \nu \)
  - \[ \frac{\partial L(\beta,\tau,\nu; y,\tau,\nu)}{\partial \beta} = X^t\Sigma^{-1}(y - X\beta - Z^t\nu) \]
  - \[ \frac{\partial L(\beta,\tau,\nu; y,\tau,\nu)}{\partial \nu} = Z^t\Sigma^{-1}(y - X^t\beta - Z^t\nu) \]

- normal equations
  \[
  \begin{pmatrix}
  X^t\Sigma^{-1}X & X^t\Sigma^{-1}Z \\
  Z^t\Sigma^{-1}X & Z^t\Sigma^{-1}Z + D^{-1}
  \end{pmatrix}
  \begin{pmatrix}
  \hat{\beta} \\
  \hat{\nu}
  \end{pmatrix}
  =
  \begin{pmatrix}
  X^t\Sigma^{-1}y \\
  Z^t\Sigma^{-1}y
  \end{pmatrix}
  \]

...
Augmented mean model

- normal equations
  \[
  \begin{pmatrix}
  X^t \Sigma^{-1} X \\
  Z^t \Sigma^{-1} X \\
  
  \end{pmatrix}
  \begin{pmatrix}
  X^t \Sigma^{-1} Z \\
  Z^t \Sigma^{-1} Z + D^{-1} \\
  
  \end{pmatrix}
  \begin{pmatrix}
  \beta \\
  \nu \\
  
  \end{pmatrix}
  =
  \begin{pmatrix}
  X^t \Sigma^{-1} y \\
  Z^t \Sigma^{-1} y \\
  
  \end{pmatrix}
  \]
- same equations are given by an *augmented* mean model

\[
\begin{pmatrix}
  y \\
  \psi_M \\
  
  \end{pmatrix}
  =
  \begin{pmatrix}
  X \\
  0 \\
  
  \end{pmatrix}
  \begin{pmatrix}
  \beta \\
  \nu \\
  
  \end{pmatrix}
  + e^*,
\]
where the error term $e^*$ is normal with mean zero and variance matrix

\[
\Sigma_a \equiv \begin{pmatrix}
  \Sigma & 0 \\
  0 & D \\
  
  \end{pmatrix},
\]

- $\psi_M \equiv E(u) = 0$
- write as $y_a = T\delta + \varepsilon^*$
- normal equations are $(T^t \Sigma_a^{-1} T)^{-1} \delta^* = T^t \Sigma_a^{-1} y$ (as above)
- can fit as an ordinary weighted regression
- see LNP §5.3.5

..
Dispersion model (LNP §5.4.2)

- marginal likelihood $L(\beta, \tau) =$
  - $|2\pi \Sigma|^{-\frac{1}{2}} \int \exp\{-\frac{1}{2}(y-X\beta-Zv)^t \Sigma^{-1}(y-X\beta-Zv)\} \times |2\pi D|^{-\frac{1}{2}} \exp\{-\frac{1}{2}v^t D^{-1}v\} dv$
  - $|2\pi \Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(y-X\beta-Zv)^t \Sigma^{-1}(y-X\beta-Zv)\}$
    $\times |2\pi D|^{-\frac{1}{2}} \exp\{-\frac{1}{2}v^\beta_t D^{-1}v^\beta_t\} \times \int \exp\{-\frac{1}{2}(v-v^\beta_t)^t I(v^\beta_t)(v-v^\beta_t)\} dv$
  - $|2\pi \Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(y-X\beta-Zv)^t \Sigma^{-1}(y-X\beta-Zv)\}$
    $\times |2\pi D|^{-\frac{1}{2}} \exp\{-\frac{1}{2}v^\beta_t D^{-1}v^\beta_t\} \times |I(v^\beta_t)/(2\pi)|^{-\frac{1}{2}}$

- marginal log-likelihood =
  - $h(\beta, \tau, \nu^\beta_t) - \frac{1}{2} \log\{|I(v^\beta_t)/(2\pi)|\}$
  - and can show that $I(v^\beta_t) = Z^t \Sigma^{-1} Z + D^{-1}$

- profile likelihood of variance components is
  - $h(\beta^\tau_t, \tau, \nu^\tau_t) - \frac{1}{2} \log\{|I(v^\tau_t)/(2\pi)|\}$
  - with $I(v^\tau_t) = Z^t \Sigma^{-1} Z + D^{-1}$

- with REML adjustment this becomes
  - $h(\beta^\tau_t, \tau, \nu^\tau_t) - \frac{1}{2} \log\{|I(v^\tau_t)/(2\pi)|\} - \frac{1}{2} \log\{|X^t V^{-1} X/(2\pi)|\}$

- fit by Extended Quasi Likelihood i.e. by gamma GLM
### Fitting algorithm

- interconnected Normal and Gamma GLMs (§5.4.4)

<table>
<thead>
<tr>
<th>Component</th>
<th>$\beta$ (fixed)</th>
<th>$\sigma^2$ (fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>$y$</td>
<td>$d^*$</td>
</tr>
<tr>
<td>Mean</td>
<td>$\mu$</td>
<td>$\sigma^2$</td>
</tr>
<tr>
<td>Variance</td>
<td>$\sigma^2$</td>
<td>$2(\sigma^2)^2$</td>
</tr>
<tr>
<td>Link</td>
<td>$\eta = \mu$</td>
<td>$\xi = h(\sigma^2)$</td>
</tr>
<tr>
<td>Linear Pred.</td>
<td>$X\beta + Z\nu$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Dev. Comp.</td>
<td>$d$</td>
<td>$\text{gamma}(d^*, \sigma^2)$</td>
</tr>
<tr>
<td>Prior Weight</td>
<td>$1/\sigma^2$</td>
<td>$(1-q)/2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Component</th>
<th>$\nu$ (random)</th>
<th>$\sigma^2_{\nu}$ (fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>$\psi_M$</td>
<td>$d_M^*$</td>
</tr>
<tr>
<td>Mean</td>
<td>$u$</td>
<td>$\sigma^2_{\nu}$</td>
</tr>
<tr>
<td>Variance</td>
<td>$\sigma^2_{\nu}$</td>
<td>$2(\sigma^2_{\nu})^2$</td>
</tr>
<tr>
<td>Link</td>
<td>$\eta_M = g_M(u)$</td>
<td>$\xi_M = h_M(\sigma^2_{\nu})$</td>
</tr>
<tr>
<td>Linear Pred.</td>
<td>$\nu$</td>
<td>$\gamma_M$</td>
</tr>
<tr>
<td>Deviance</td>
<td>$d_M$</td>
<td>$\text{gamma}(d_M^*, \sigma^2_{\nu})$</td>
</tr>
<tr>
<td>Prior Weight</td>
<td>$1/\sigma^2_{\nu}$</td>
<td>$(1-q_M)/2$</td>
</tr>
</tbody>
</table>

\[d_i = (y_i - X_i \hat{\beta} - Z_i \hat{\nu})^2,\]
\[d_{Mi} = \hat{\nu}_i^2,\]
\[\text{gamma}(d^*, \phi) = 2\{-\log(d^*/\phi) + (d^* - \phi)/\phi\}\]
and
\[(q, q_M)\] are leverages, given by the diagonal elements of
\[T(T^T\Sigma_a^{-1}T)^{-1}T^T\Sigma_a^{-1}.\]
HGLMs in GenStat

• procedures (Payne, Lee, Nelder & Noh 2008)
  • HGFIXEDMODEL – defines the fixed model for an HGLM or DHGLM
  • HGRANDOMMODEL – defines the random model for an HGLM
  • HGDRANDOMMODEL – adds random terms into the dispersion models of an HGLM, so that the whole model becomes a DHGLM
  • HGNONLINEAR – defines nonlinear parameters for the fixed model
  • HGANALYSE – fits a hierarchical generalized linear model (HGLM) or a double hierarchical generalized linear model (DHGLM)
  • HGDISPLAY – displays results from an HGLM or DHGLM
  • HGPLOT – produces model-checking plots for an HGLM or DHGLM
  • HGPREDICT – forms predictions from an HGLM or DHGLM analysis
  • HGKEEP – saves information from an HGLM or DHGLM analysis
  • HGGGRAPH – plots predictions from an HGLM or DHGLM analysis
  • HGWALD – gives Wald tests for fixed terms that can be dropped

• menus
  • Stats | Regression Analysis | Mixed Models | Hierarchical Generalized Linear Models
  • cover the standard situations, but not dispersion modelling nor DHGLMs
  ..
GenStat HGLM examples
Example – chocolate cakes

- LNP §5.5
- breaking angle of chocolate cakes
- split plot: Replicate/Batch/Cake
- treatment factors: Recipe (whole-plot factor, between Batches), Temperature (sub-plot factor, within Batches)
- analyse as a Normal-Normal HGLM to compare with familiar REML
HGLM menu – for cakes

- find menu in Mixed models section of Regression Analysis on Stats menu
Output: Normal-Normal HGLM

Hierarchical generalized linear model

Response variate: Angle

Mean model

Fixed terms: Recipe*Temperature
Distribution: normal
Link: identity
Random terms: Replicate/Batch
Distribution: normal
Link: identity
Dispersion: free

Dispersion model

Distribution: gamma
Link: logarithm

Estimates from the mean model

<table>
<thead>
<tr>
<th></th>
<th>estimate</th>
<th>s.e.</th>
<th>t(252)</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>29.133</td>
<td>2.038</td>
<td>14.294</td>
</tr>
<tr>
<td>Recipe 2</td>
<td>-2.267</td>
<td>1.796</td>
<td>-1.262</td>
</tr>
<tr>
<td>Recipe 3</td>
<td>-1.200</td>
<td>1.796</td>
<td>-0.668</td>
</tr>
<tr>
<td>Temperature 185</td>
<td>2.400</td>
<td>1.652</td>
<td>1.453</td>
</tr>
<tr>
<td>Temperature 195</td>
<td>1.667</td>
<td>1.652</td>
<td>1.009</td>
</tr>
<tr>
<td>Temperature 205</td>
<td>4.400</td>
<td>1.652</td>
<td>2.663</td>
</tr>
<tr>
<td>Temperature 215</td>
<td>9.533</td>
<td>1.652</td>
<td>5.771</td>
</tr>
<tr>
<td>Temperature 225</td>
<td>5.933</td>
<td>1.652</td>
<td>3.591</td>
</tr>
<tr>
<td>Recipe 2.Temperature 185</td>
<td>0.133</td>
<td>2.336</td>
<td>0.057</td>
</tr>
<tr>
<td>Recipe 2.Temperature 195</td>
<td>3.200</td>
<td>2.336</td>
<td>1.370</td>
</tr>
<tr>
<td>Recipe 2.Temperature 205</td>
<td>0.867</td>
<td>2.336</td>
<td>0.371</td>
</tr>
<tr>
<td>Recipe 2.Temperature 215</td>
<td>-1.933</td>
<td>2.336</td>
<td>-0.827</td>
</tr>
</tbody>
</table>

← mean model

← dispersion model here just fits variance components

← estimates of parameters in the mean model

..
Output: Normal-Normal HGLM

Estimated parameters in mean model (continued) fixed terms only by default

Parameters in dispersion models (logged variance components)

Assess random & dispersion models by $-2 \times P_{\beta,v}(h)$

Fixed model by $-2 \times P_v(h)$

For DIC use $-2 \times (h/v)$

H-likelihood of mean model is $-2 \times (h)$

EQD's are approximations to profile likelihoods ..

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0190</td>
<td>0.0945</td>
<td>31.93</td>
<td>20.47</td>
</tr>
<tr>
<td>lambda Replicate</td>
<td>3.641</td>
<td>0.390</td>
<td>9.35</td>
<td>38.12</td>
</tr>
<tr>
<td>lambda Replicate.Batch</td>
<td>1.315</td>
<td>0.365</td>
<td>3.61</td>
<td>3.723</td>
</tr>
</tbody>
</table>

Message: s.e.s are based on dispersion parameter with value 1.

Likelihood statistics:

- $-2 \times h(y|v)$
  - $1535.137$
- $-2 \times h$
  - $1787.389$
- $-2 \times P_v(h)$
  - $1639.073$
- $-2 \times P_{\beta,v}(h)$
  - $1595.346$
- $-2 \times EQD(y|v)$
  - $1535.137$
- $-2 \times EQD$
  - $1787.389$
- $-2 \times P_v(EQD)$
  - $1639.073$
- $-2 \times P_{\beta,v}(EQD)$
  - $1595.346$

Fixed parameters in mean model: 18
Random parameters in mean model: 60
Fixed dispersion parameters: 3
Random dispersion parameters: 0
Further output and model checking

- click on *Further Output* button in *HGLM* menu
- click on *Model Checking* button in *HGLM Further Output* menu to obtain *HGLM Model Checking* menu
Residual plots for cakes
Residual plots for batches
Residual plots for replicates
Residual plots for dispersion model
Compare to REML

Hierarchical generalized linear model

Estimates from the dispersion model

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0190</td>
<td>0.0945</td>
<td>31.93</td>
<td>20.47</td>
</tr>
<tr>
<td>lambda Replicate</td>
<td>3.641</td>
<td>0.390</td>
<td>9.35</td>
<td>38.12</td>
</tr>
<tr>
<td>lambda Replicate.Batch</td>
<td>1.315</td>
<td>0.365</td>
<td>3.61</td>
<td>3.723</td>
</tr>
</tbody>
</table>
Assess fixed model using \(-2 \, p_v(h)\)

```
" assess the fixed model - drop terms one at a time "
VARIATE [NVALUES=3; IPRINT=extra] Change;
EXTRA='Change in -2 \sim \{times\} P_{-\nu(h)}'
TEXT [VALUES='Temperature','Recipe','Recipe.Temperature'] Drop
HGDISPLAY [PRINT=likelihood]
HGKEEP LIKELIHOOD=L1

HGFIXEDMODEL Recipe+Temperature
HGANALYSE [PRINT=likelihood] Angle
HGKEEP LIKELIHOOD=L2

HGFIXEDMODEL Temperature
HGANALYSE [PRINT=likelihood] Angle
HGKEEP LIKELIHOOD=L3

HGFIXEDMODEL
HGANALYSE [PRINT=likelihood] Angle
HGKEEP LIKELIHOOD=L4
PRINT Drop,Change; FIELD=24

" compare with REML Wald statistics "
VDISPLAY [PRINT=wald]
```
Assess fixed model using $-2 \ p_v(h)$

Tests for fixed effects

Sequentially adding terms to fixed model

<table>
<thead>
<tr>
<th>Fixed term</th>
<th>Wald statistic</th>
<th>n.d.f.</th>
<th>F statistic</th>
<th>d.d.f.</th>
<th>F pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipe</td>
<td>3.16</td>
<td>2</td>
<td>1.58</td>
<td>28.0</td>
<td>0.224</td>
</tr>
<tr>
<td>Temperature</td>
<td>102.60</td>
<td>5</td>
<td>20.52</td>
<td>210.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Recipe.Temperature</td>
<td>10.06</td>
<td>10</td>
<td>1.01</td>
<td>210.0</td>
<td>0.439</td>
</tr>
</tbody>
</table>

Dropping individual terms from full fixed model

<table>
<thead>
<tr>
<th>Fixed term</th>
<th>Wald statistic</th>
<th>n.d.f.</th>
<th>F statistic</th>
<th>d.d.f.</th>
<th>F pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipe.Temperature</td>
<td>10.06</td>
<td>10</td>
<td>1.01</td>
<td>210.0</td>
<td>0.439</td>
</tr>
</tbody>
</table>

Message: denominator degrees of freedom for approximate F-tests are calculated using algebraic derivatives ignoring fixed/boundary/singular variance parameters.

Hierarchical generalized linear model

Wald tests for dropping HGLM fixed terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Wald statistic</th>
<th>d.f.</th>
<th>approx. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipe.Temperature</td>
<td>10.06</td>
<td>10</td>
<td>0.435</td>
</tr>
</tbody>
</table>
Conjugate HGLMs

- random parameters are on the *canonical scale*
- the contribution of the random parameters to the extended likelihood (Ξ the h-likelihood) has the same form as the likelihood of the base GLM
- so it can easily be fitted together with the base GLM in the augmented mean model (same variance function, same iterative reweighting scheme etc...)
- examples
  - Normal – Normal most obvious
  - Poisson – Gamma most useful?
  - Binomial – Beta next most useful?
  - Gamma – Inverse Gamma
- algorithmically and intuitively appealing
Conjugate HGLM e.g. Poisson-Gamma

- Poisson-gamma HGLM (LNP Ex. 6.2 & 6.3): linear predictor
  - $\eta = \log(\mu) = X \beta + Z \nu$
  - where $\nu = \log(u)$

- Poisson-gamma HGLM: distributions
  - $y|u$ follows a Poisson distribution with $E(y|u) = \mu$
    - log-likelihood is $\sum \{ y \log(\mu) - \mu \}$
  - $u$ has a Gamma distribution
    - log-likelihood is $\sum \{ (\psi_M \log(u) - u - \log(\lambda)/\lambda - \log\Gamma(1/\lambda)) \}$
    - $\psi_M = E(u) = 1$
    - this is the conjugate distribution for the Poisson (so this is a conjugate HGLM)
    - note: gamma distribution for random effects has $V_M(u) = u$ and log canonical link, but standard gamma GLM has $V(\mu) = \mu^2$ and reciprocal canonical link

- kernel of h-loglikelihood (LNP §6.3) is
  - $\sum \{ y \log(\mu) - \mu \} + \sum \{ (\psi_M \log(u) - u - \log(\lambda)/\lambda - \log\Gamma(1/\lambda)) \}$
  - $\nu$ is canonical for $\beta$ (but not $\lambda$)
  - estimate $\lambda$ by profile lhd (c.f. variance components in normal-normal)
Non-conjugate HGLMs

- random parameters no longer on the canonical scale
  - use extended likelihood to estimate random parameters
  - use adjusted profile likelihood to estimate fixed parameters
  - but enhanced Laplace approximations available (Noh & Lee 06)
- augmented mean model now has a different GLM for the base GLM from the augmented units
- examples
  - Poisson – Normal          Poisson GLMM
  - Binomial – Normal         Binomial GLMM
  - Gamma – Normal            Gamma GLMM
- algorithmically more difficult, but can still be fitted within the GLM framework

..
## HGLMs

- examples of HGLMs (LNP Table 6.2)

<table>
<thead>
<tr>
<th>$y \mid u$ dist.</th>
<th>$g(\mu)^*$</th>
<th>$u$ dist.</th>
<th>$v(u)$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>id</td>
<td>Normal</td>
<td>id</td>
<td>Conjugate HGLM</td>
</tr>
<tr>
<td>Binomial</td>
<td>logit</td>
<td>Beta</td>
<td>logit</td>
<td>Conjugate HGLM</td>
</tr>
<tr>
<td>Binomial</td>
<td>logit</td>
<td>Normal</td>
<td>id</td>
<td>Binomial GLMM</td>
</tr>
<tr>
<td>Binomial</td>
<td>comp</td>
<td>Gamma</td>
<td>log</td>
<td>HGLM</td>
</tr>
<tr>
<td>Gamma</td>
<td>recip</td>
<td>Inverse-gamma</td>
<td>recip</td>
<td>Conjugate HGLM</td>
</tr>
<tr>
<td>Gamma</td>
<td>log</td>
<td>Inverse-gamma</td>
<td>recip</td>
<td>Conjugate HGLM with non-canonical link</td>
</tr>
<tr>
<td>Gamma</td>
<td>log</td>
<td>Gamma</td>
<td>log</td>
<td>HGLM</td>
</tr>
<tr>
<td>Poisson</td>
<td>log</td>
<td>Normal</td>
<td>id</td>
<td>Poisson GLMM**</td>
</tr>
<tr>
<td>Poisson</td>
<td>log</td>
<td>Gamma</td>
<td>log</td>
<td>Conjugate HGLM</td>
</tr>
</tbody>
</table>

* id = identity, recip = reciprocal, comp = complementary-log-log

** In GLMMs, we take $v = v(u) = u$
Birds in Tasmania

- HGLM
  - base GLM – Poisson distribution, Log link
  - random terms – Gamma distribution, Log link
  - i.e. Poisson-Gamma conjugate HGLM

- random terms
  - site \((Site)\)
  - treatment locations within site \((SiteTreat)\)
  - sample plots within treatment locations \((Plot)\)

- fixed terms
  - connected by habitat strips \((Treatment)\)
  - vegetation type \((Vegetation)\)
  - time of day \((AM\_vs\_PM)\)

- data set used by Steve Candy, Forestry Tasmania, at the Workshop *Extensions of Generalized Linear Models* (Nelder, Payne & Candy) before the Australasian Genstat Conference, Surfers Paradise, 30 January 2001

...
Hierarchical generalized linear model

Response variate: TotalNo

Mean model

Fixed terms: Vegetation*Treatment*AM_vs_PM
Distribution: poisson
Link: logarithm
Random terms: Site + SiteTreat + Plot
Distribution: gamma
Link: logarithm
Dispersion: fixed

Dispersion model

Distribution: gamma
Link: logarithm

Likelihood statistics

\[
\begin{align*}
-2 \times h(y|h) &= 2017.696 \\
-2 \times h &= 1867.242 \\
-2 \times P_v(h) &= 2173.327 \\
-2 \times P_{p,v}(h) &= 2200.500 \\
-2 \times EQD(y|h) &= 2018.567 \\
-2 \times EQD &= 1866.632 \\
-2 \times P_v(EQD) &= 2172.917 \\
-2 \times P_{p,v}(EQD) &= 2200.069
\end{align*}
\]

Fixed parameters in mean model: 12
Random parameters in mean model: 132
Fixed dispersion parameters: 3
Random dispersion parameters: 0
Wald tests

Wald tests for dropping HGLM fixed terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Wald statistic</th>
<th>d.f.</th>
<th>approx. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation.Treatment.AM_vs_PM</td>
<td>2.576</td>
<td>2</td>
<td>0.276</td>
</tr>
</tbody>
</table>

12 HGKEEP LIKELIHOOD=L1; LDF=D1
13 CALCULATE T1 = CPUTIME(0)
14 PRINT T1 - TO; DECIMALS=2; HEADING='Time taken (seconds)'

Time taken (seconds)

4.58

- no evidence of a 3-factor interaction

..
Hierarchical generalized linear model

Response variate: TotalNo

Mean model

Fixed terms: Vegetation*Treatment*AM_vs_PM - (Vegetation.Treatment AM_vs_PM)
Distribution: poisson
Link: logarithm
Random terms: Site + SiteTreat + Plot
Distribution: gamma
Link: logarithm
Dispersion: fixed

Dispersion model

Distribution: gamma
Link: logarithm

Likelihood statistics

-2 × h(yn)   2019.526
-2 × h       1871.466
-2 × Pν(h)   2175.924
-2 × Pν(h)   2200.036
-2 × EQD(yn) 2020.197
-2 × EQD     1871.047
-2 × Pν(EQD) 2175.504
-2 × Pν(EQD) 2199.617

Fixed parameters in mean model  10
Random parameters in mean model 132
Fixed dispersion parameters  3
Random dispersion parameters  0
Wald tests and Change

Wald tests for dropping HGLM fixed terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Wald statistic</th>
<th>d.f.</th>
<th>approx. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation.Treatment</td>
<td>4.489</td>
<td>2</td>
<td>0.106</td>
</tr>
<tr>
<td>Vegetation.AM_vs_PM</td>
<td>4.951</td>
<td>2</td>
<td>0.084</td>
</tr>
<tr>
<td>Treatment.AM_vs_PM</td>
<td>5.471</td>
<td>1</td>
<td>0.019</td>
</tr>
</tbody>
</table>

20 HGKEEP          LIKELIHOOD=L2; LDF=D2
21 CALCULATE      Lchange,Dchange = L2$[3],D2$[1] - L1$[3],D1$[1]
22 PRINT           Lchange,Dchange; HEAD='Change','d.f.>'; DECIMALS=2,0

Change                  d.f.
2.60                     -2

23 CALCULATE          T2 = CPUTIME(0)
24 PRINT               T2 - T1; DECIMALS=2; HEADING='Time taken (seconds)'

Time taken (seconds)
4.27

• change deviance 2.60 on 2 d.f. for omitting Vegetation.Treatment.AM_vs_PM (c.f. Wald 2.58)
• next omit Vegetation.Treatment

...
Hierarchical generalized linear model

Response variate: TotalNo

Mean model

Fixed terms: Vegetation*Treatment*AM vs PM - (Vegetation.Treatment.AM vs PM) - (Vegetation.Treatment)
Distribution: poisson
Link: logarithm
Random terms: Site + SiteTreat + Plot
Distribution: gamma
Link: logarithm
Dispersion: fixed

Dispersion model

Distribution: gamma
Link: logarithm

Likelihood statistics

\[-2 \times h(y|y)\] 2018.510
\[-2 \times h\] 1881.726
\[-2 \times P_v(h)\] 2180.339
\[-2 \times P_{v,v}(h)\] 2201.810
\[-2 \times EQD(y|y)\] 2019.131
\[-2 \times EQD\] 1881.270
\[-2 \times P_v(EQD)\] 2179.883
\[-2 \times P_{v,v}(EQD)\] 2201.354

Fixed parameters in mean model 8
Random parameters in mean model 132
Fixed dispersion parameters 3
Random dispersion parameters 0
Wald tests and Change

Wald tests for dropping HGLM fixed terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Wald statistic</th>
<th>d.f.</th>
<th>approx. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation.AM_vs_PM</td>
<td>4.996</td>
<td>2</td>
<td>0.082</td>
</tr>
<tr>
<td>Treatment.AM_vs_PM</td>
<td>5.183</td>
<td>1</td>
<td>0.023</td>
</tr>
</tbody>
</table>

31 HGKEEP LIKELIHOOD=L3; LDF=D3
32 CALCULATE Lchange, Dchange = L3$[3], D3$[1] - L2$[3], D2$[1]
33 PRINT Lchange, Dchange; HEAD='Change', 'd.f.'; DECIMALS=2, 0

Change d.f.
4.42 -2

34 CALCULATE T3 = CPU_TIME(0)
35 PRINT T3 - T2; DECIMALS=2; HEADING='Time taken (seconds)'

Time taken (seconds)
4.19

• change deviance 4.42 on 2 d.f. for omitting Vegetation.Treatment (c.f. Wald 4.49)
• next omit Vegetation.AM_PM

..
Hierarchical generalized linear model

Response variate: TotalNo

Mean model

Fixed terms: Vegetation*Treatment*AM_vs_PM - (Vegetation.Treatment.AM_vs_PM) - (Vegetation.Treatment) - (Vegetation.AM_vs_PM)
Distribution: poisson
Link: logarithm
Random terms: Site + SiteTreat + Plot
Distribution: gamma
Link: logarithm
Dispersion: fixed

Dispersion model

Distribution: gamma
Link: logarithm

Likelihood statistics

-2 × h(y|y) 2024.828
-2 × h 1884.965
-2 × P_v(h) 2186.312
-2 × P_{p,v}(h) 2200.968
-2 × EQD(y|y) 2025.499
-2 × EQD 1884.525
-2 × P_v(EQD) 2184.872
-2 × P_{p,v}(EQD) 2200.517

Fixed parameters in mean model 6
Random parameters in mean model 132
Fixed dispersion parameters 3
Random dispersion parameters 0
### Wald tests for dropping HGLM fixed terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Wald statistic</th>
<th>d.f.</th>
<th>approx. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>21.71</td>
<td>2</td>
<td>0.000</td>
</tr>
<tr>
<td>Treatment.AM_vs_PM</td>
<td>5.48</td>
<td>1</td>
<td>0.019</td>
</tr>
</tbody>
</table>

- change deviance 4.97 on 2 d.f. for omitting Vegetation.AM_PM (c.f. Wald 5.00)
- now study results

```plaintext
42 HGKEEP  LIKELIHOOD=L4; LDF=D4
43 CALCULATE Lchange,Dchange = L4$[3],D4$[1] - L3$[3],D3$[1]
44 PRINT   Lchange,Dchange; HEAD='Change','d.f.'; DECIMALS=2,0

  Change       d.f.
  4.97         -2

45 CALCULATE  T4 = CPETIME(0)
46 PRINT      T4 - T3; DECIMALS=2; HEADING='Time taken (seconds)'
```

Time taken (seconds)
4.11
Predicted means: treatment x time of day
Hierarchical generalized nonlinear models

- expected value $E(y) = \mu$
- link $\eta = g(\mu)$
- distribution – Normal, Binomial, Poisson or Gamma (from exponential family)
- linear predictor $\eta = X\beta + \sum Z_i \gamma_i$
- random effects $\gamma_i$ with either beta, Normal, gamma or inverse gamma distributions, and their own link functions
- nonlinear parameters in fixed terms in the linear predictor
  - $X\beta = \sum x_i \beta_i$
  - but now some $x_i$'s are nonlinear functions of explanatory variables and parameters that are to be estimated
- extension of generalized nonlinear models of Lane (1996)
- constraint – available only for conjugate HGLM's
### Implementation – interlinked GLMs

- fit nonlinear parameters by maximizing h-likelihood of augmented mean model

<table>
<thead>
<tr>
<th>Components</th>
<th>( \beta ) (fixed)</th>
<th>( \gamma ) (fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>( y )</td>
<td>( d^* )</td>
</tr>
<tr>
<td>Mean</td>
<td>( \mu )</td>
<td>( \phi )</td>
</tr>
<tr>
<td>Variance</td>
<td>( \phi V(\mu) )</td>
<td>( 2\phi^2 )</td>
</tr>
<tr>
<td>Link</td>
<td>( \eta = g(\mu) )</td>
<td>( \xi = h(\phi) )</td>
</tr>
<tr>
<td>Linear Pred.</td>
<td>( X\beta + Zv )</td>
<td>( G \gamma )</td>
</tr>
<tr>
<td>Dev. Comp.</td>
<td>( d )</td>
<td>gamma((d^*, \phi))</td>
</tr>
<tr>
<td>Prior Weight</td>
<td>( 1/\phi )</td>
<td>( (1 - q)/2 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Components</th>
<th>( u ) (random)</th>
<th>( \lambda ) (fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>( \psi_M )</td>
<td>( d^*_M )</td>
</tr>
<tr>
<td>Mean</td>
<td>( u )</td>
<td>( \lambda )</td>
</tr>
<tr>
<td>Variance</td>
<td>( \lambda V_M(u) )</td>
<td>( 2\lambda^2 )</td>
</tr>
<tr>
<td>Link</td>
<td>( \eta_M = g_M(u) )</td>
<td>( \xi_M = h_M(\lambda) )</td>
</tr>
<tr>
<td>Linear Pred.</td>
<td>( v )</td>
<td>( G_M \gamma_M )</td>
</tr>
<tr>
<td>Deviance</td>
<td>( d_M )</td>
<td>gamma((d^*_M, \lambda))</td>
</tr>
<tr>
<td>Prior Weight</td>
<td>( 1/\lambda )</td>
<td>( (1 - q_M)/2 )</td>
</tr>
</tbody>
</table>
Growth of Hooded Parrot nestlings

- Hooded Parrot (*Psephotus dissimilis*)
- grass parrot in Northern Territory of Australia
- nests in termite mounds
- nests also inhabited by moth larvae that feed on nestling waste

Acknowledgement: S Cooney
Australian National University, Canberra
http://www.anu.edu.au/BoZo/stuart/HPP.htm
Growth of Hooded Parrot nestlings

- investigate relationship between parrot and moth
  - beneficial, commensal or parasitic
- 41 nests located and monitored
- each brood had between 1-7 chicks
- treatments randomized to nests
  - moth larvae left or experimentally removed from nest
- weight of chicks measured over time
- growth modelled over time by logistic curve
  - \( weight = a + c / (1 + \exp\{ -b \times (age - m) \}) \)
  - model linear in \( a \) and \( c \), nonlinear in \( b \) and \( m \)
- fit as HGNLM because
  - brood is a random effect
  - treatments are applied to complete broods

...
Initial values from logistic curve

3 " fit logistic standard curve (ignoring Brood) to get initial values "
4 MODEL Weight
5 FITCURVE [CURVE=logistic; NOMESSAGE=residual] Age * Treat

Nonlinear regression analysis

Response variate: Weight
Explanatory: Age
Grouping factor: Treat, all linear parameters separate
Fitted Curve: \[ A + C/(1 + \exp(-B*(X - M))) \]

Summary of analysis

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>s.s.</th>
<th>m.s.</th>
<th>v.r.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>5</td>
<td>133643.6</td>
<td>26728.51</td>
<td>899.02</td>
</tr>
<tr>
<td>Residual</td>
<td>516</td>
<td>15341.4</td>
<td>29.73</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>521</td>
<td>148984.1</td>
<td>265.96</td>
<td></td>
</tr>
</tbody>
</table>

Percentage variance accounted for 89.6
Standard error of observations is estimated to be 5.45.

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.2647</td>
<td>0.0132</td>
</tr>
<tr>
<td>M</td>
<td>9.909</td>
<td>0.236</td>
</tr>
<tr>
<td>C Treat N</td>
<td>47.81</td>
<td></td>
</tr>
<tr>
<td>A Treat N</td>
<td>0.1966</td>
<td></td>
</tr>
<tr>
<td>C Treat Y</td>
<td>48.68</td>
<td></td>
</tr>
<tr>
<td>A Treat Y</td>
<td>0.5423</td>
<td></td>
</tr>
</tbody>
</table>

6 RKEEP ESTIMATES=Estcurve
Hierarchical generalized linear model

Response variate: Weight

Mean model

Fixed terms: ExpBM
Distribution: normal
Link: identity
Random terms: Brood
Distribution: normal
Link: identity
Dispersion: free

Dispersion model

Distribution: gamma
Link: logarithm
HGNLM, common A, B, C and M

Estimates from the mean model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(518)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.26</td>
<td>0.0172</td>
<td>14.89</td>
</tr>
<tr>
<td>M</td>
<td>10.04</td>
<td>0.2897</td>
<td>34.66</td>
</tr>
<tr>
<td>constant</td>
<td>0.19</td>
<td>1.4896</td>
<td>0.13</td>
</tr>
<tr>
<td>ExpBM</td>
<td>48.87</td>
<td>1.8512</td>
<td>26.40</td>
</tr>
</tbody>
</table>

Estimates from the dispersion model

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0239</td>
<td>0.0640</td>
<td>47.23</td>
<td>20.57</td>
</tr>
<tr>
<td>lambda Brood</td>
<td>2.155</td>
<td>0.258</td>
<td>8.36</td>
<td>8.629</td>
</tr>
</tbody>
</table>

Message: s.e.s are based on dispersion parameter with value 1.

Likelihood statistics

-2 × h(y|v)  3025.719
-2 × h     3211.537
-2 × P_v(h) 3124.117
-2 × P_{p,v}(h) 3122.544
-2 × EQD(y|v) 3025.719
-2 × EQD    3211.537
-2 × P_v(EQD) 3124.117
-2 × P_{p,v}(EQD) 3122.544
Hierarchical generalized linear model

Response variate: Weight

Mean model

Fixed terms: ExpBM + Treat
Distribution: normal
Link: identity
Random terms: Brood
Distribution: normal
Link: identity
Dispersion: free

Dispersion model

Distribution: gamma
Link: logarithm
HGNLM, common B, C and M

Estimates from the mean model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(517)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.26</td>
<td>0.0172</td>
<td>14.89</td>
</tr>
<tr>
<td>M</td>
<td>10.05</td>
<td>0.2896</td>
<td>34.70</td>
</tr>
<tr>
<td>constant</td>
<td>-0.01</td>
<td>1.5176</td>
<td>-0.01</td>
</tr>
<tr>
<td>ExpBM</td>
<td>48.82</td>
<td>1.8493</td>
<td>26.40</td>
</tr>
<tr>
<td>Treat Y</td>
<td>0.77</td>
<td>1.1578</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Estimates from the dispersion model

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0239</td>
<td>0.0640</td>
<td>47.22</td>
<td>20.57</td>
</tr>
<tr>
<td>lambda Brood</td>
<td>2.174</td>
<td>0.261</td>
<td>8.33</td>
<td>8.790</td>
</tr>
</tbody>
</table>

Message: s.e.s are based on dispersion parameter with value 1.

Likelihood statistics

-2 × h(y|v)   3025.460
-2 × h     3211.257
-2 × P_v(h) 3123.691
-2 × P_0,v(h) 3119.973
-2 × EQD(y|v) 3025.460
-2 × EQD    3211.257
-2 × P_v(EQD) 3123.691
-2 × P_0,v(EQD) 3119.973
Hierarchical generalized linear model

Response variate: Weight

Mean model

Fixed terms: ExpBM*Treat
Distribution: normal
Link: identity
Random terms: Brood
Distribution: normal
Link: identity
Dispersion: free

Dispersion model

Distribution: gamma
Link: logarithm
HGNLM, common B and M

Estimates from the mean model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(516)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.26</td>
<td>0.017</td>
<td>14.86</td>
</tr>
<tr>
<td>M</td>
<td>10.05</td>
<td>0.291</td>
<td>34.54</td>
</tr>
<tr>
<td>constant</td>
<td>0.25</td>
<td>1.529</td>
<td>0.17</td>
</tr>
<tr>
<td>ExpBM</td>
<td>48.30</td>
<td>1.897</td>
<td>25.46</td>
</tr>
<tr>
<td>Treat Y</td>
<td>-0.11</td>
<td>1.411</td>
<td>-0.08</td>
</tr>
<tr>
<td>ExpBM.Treat Y</td>
<td>1.53</td>
<td>1.379</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Estimates from the dispersion model

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0228</td>
<td>0.0641</td>
<td>47.15</td>
</tr>
<tr>
<td>lambda Brood</td>
<td>2.183</td>
<td>0.261</td>
<td>8.37</td>
</tr>
</tbody>
</table>

Message: s.e.s are based on dispersion parameter with value 1.

Likelihood statistics

-2 × h(y|y)       3023.883
-2 × h          3210.079
-2 × P_v(h)     3122.471
-2 × P_p,v(h)   3116.258
-2 × EQD(y|y)   3023.883
-2 × EQD        3210.079
-2 × P_v(EQD)   3122.471
-2 × P_p,v(EQD) 3116.258
Hierarchical generalized linear model

Response variate: Weight

**Mean model**

Fixed terms: ExpBM*Treat
Distribution: normal
Link: identity
Random terms: Brood
Distribution: normal
Link: identity
Dispersion: free

**Dispersion model**

Distribution: gamma
Link: logarithm
HGNLM, common M

Estimates from the mean model

<table>
<thead>
<tr>
<th></th>
<th>estimate</th>
<th>s.e.</th>
<th>t(515)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>0.266</td>
<td>0.021</td>
<td>12.57</td>
</tr>
<tr>
<td>B2</td>
<td>0.237</td>
<td>0.024</td>
<td>9.99</td>
</tr>
<tr>
<td>M</td>
<td>10.035</td>
<td>0.289</td>
<td>34.78</td>
</tr>
<tr>
<td>constant</td>
<td>0.706</td>
<td>1.562</td>
<td>0.45</td>
</tr>
<tr>
<td>ExpBM</td>
<td>47.464</td>
<td>2.010</td>
<td>23.61</td>
</tr>
<tr>
<td>Treat Y</td>
<td>-1.555</td>
<td>2.074</td>
<td>-0.75</td>
</tr>
<tr>
<td>ExpBM.Treat Y</td>
<td>4.124</td>
<td>3.068</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Estimates from the dispersion model

Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>estimate</th>
<th>s.e.</th>
<th>t(*)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>phi</td>
<td>3.0230</td>
<td>0.0642</td>
<td>47.11</td>
<td>20.55</td>
</tr>
<tr>
<td>lambda Brood</td>
<td>2.176</td>
<td>0.261</td>
<td>8.33</td>
<td>8.807</td>
</tr>
</tbody>
</table>

Message: s.e.s are based on dispersion parameter with value 1.

Likelihood statistics

\[-2 \times h(y|y)\] 3023.039
\[-2 \times h\] 3208.880
\[-2 \times P_v(h)\] 3121.326
\[-2 \times P_{v,v}(h)\] 3115.093
\[-2 \times EQD(y|y)\] 3023.039
\[-2 \times EQD\] 3208.880
\[-2 \times P_v(EQD)\] 3121.326
\[-2 \times P_{v,v}(EQD)\] 3115.093
Hierarchical generalized linear model

Response variate: Weight

Mean model

Fixed terms: ExpBM*Treat
Distribution: normal
Link: identity
Random terms: Brood
Distribution: normal
Link: identity
Dispersion: free

Dispersion model

Distribution: gamma
Link: logarithm
### HGNLM, different A, B, C and M

#### Estimates from the mean model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate $\hat{\beta}$</th>
<th>s.e.</th>
<th>t(514)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>0.263</td>
<td>0.022</td>
<td>12.05</td>
</tr>
<tr>
<td>B2</td>
<td>0.243</td>
<td>0.028</td>
<td>8.66</td>
</tr>
<tr>
<td>M1</td>
<td>9.958</td>
<td>0.344</td>
<td>28.97</td>
</tr>
<tr>
<td>M2</td>
<td>10.243</td>
<td>0.528</td>
<td>19.40</td>
</tr>
<tr>
<td>constant</td>
<td>0.415</td>
<td>1.740</td>
<td>0.24</td>
</tr>
<tr>
<td>ExpBM</td>
<td>47.764</td>
<td>2.177</td>
<td>21.94</td>
</tr>
<tr>
<td>Treat Y</td>
<td>-0.385</td>
<td>3.304</td>
<td>-0.12</td>
</tr>
<tr>
<td>ExpBM.Treat Y</td>
<td>2.900</td>
<td>4.070</td>
<td>0.71</td>
</tr>
</tbody>
</table>

#### Estimates from the dispersion model

#### Estimates of parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate $\hat{\phi}$</th>
<th>s.e.</th>
<th>t(254)</th>
<th>antilog of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>3.0245</td>
<td>0.0642</td>
<td>47.08</td>
<td>20.58</td>
</tr>
<tr>
<td>$\lambda_{Broid}$</td>
<td>2.179</td>
<td>0.261</td>
<td>8.34</td>
<td>8.836</td>
</tr>
</tbody>
</table>

*Message: s.e.s are based on dispersion parameter with value 1.*

#### Likelihood statistics

<table>
<thead>
<tr>
<th>Likelihood Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-2 \times h(y</td>
<td>x)$</td>
</tr>
<tr>
<td>$-2 \times h$</td>
<td>3208.781</td>
</tr>
<tr>
<td>$-2 \times P_v(h)$</td>
<td>3121.155</td>
</tr>
<tr>
<td>$-2 \times P_{\nu}(h)$</td>
<td>3114.936</td>
</tr>
<tr>
<td>$-2 \times EQD(y</td>
<td>x)$</td>
</tr>
<tr>
<td>$-2 \times P_{EQD}$</td>
<td>3208.781</td>
</tr>
<tr>
<td>$-2 \times P_{\nu}(EQD)$</td>
<td>3121.155</td>
</tr>
<tr>
<td>$-2 \times P_{\nu,\nu}(EQD)$</td>
<td>3114.936</td>
</tr>
</tbody>
</table>
Likelihood statistics

```
69   " compare likelihoods "
70   PRINT   L,LA,LAC,LABC,LABCM; DECIMALS=2; FIELD=10

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>LA</th>
<th>LAC</th>
<th>LABC</th>
<th>LABCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 * h(y</td>
<td>v)</td>
<td>3025.72</td>
<td>3025.46</td>
<td>3023.88</td>
<td>3023.04</td>
</tr>
<tr>
<td>-2 * h</td>
<td>3211.54</td>
<td>3211.26</td>
<td>3210.08</td>
<td>3208.88</td>
<td>3208.78</td>
</tr>
<tr>
<td>-2 * P_v(h)</td>
<td>3124.12</td>
<td>3123.69</td>
<td>3122.47</td>
<td>3121.33</td>
<td>3121.15</td>
</tr>
<tr>
<td>-2 * P_beta,v(h)</td>
<td>3122.54</td>
<td>3119.97</td>
<td>3116.26</td>
<td>3115.09</td>
<td>3114.94</td>
</tr>
<tr>
<td>-2 * EQD(y</td>
<td>v)</td>
<td>3025.72</td>
<td>3025.46</td>
<td>3023.88</td>
<td>3023.04</td>
</tr>
<tr>
<td>-2 * EQD</td>
<td>3211.54</td>
<td>3211.26</td>
<td>3210.08</td>
<td>3208.88</td>
<td>3208.78</td>
</tr>
<tr>
<td>-2 * P_v(EQD)</td>
<td>3124.12</td>
<td>3123.69</td>
<td>3122.47</td>
<td>3121.33</td>
<td>3121.15</td>
</tr>
<tr>
<td>-2 * P_beta,v(EQD)</td>
<td>3122.54</td>
<td>3119.97</td>
<td>3116.26</td>
<td>3115.09</td>
<td>3114.94</td>
</tr>
</tbody>
</table>
```

- no evidence of any treatment effects

..
Standard curve (ignoring brood)

Nonlinear regression analysis

Accumulated analysis of variance

<table>
<thead>
<tr>
<th>Change</th>
<th>d.f.</th>
<th>s.s.</th>
<th>m.s.</th>
<th>v.r.</th>
<th>F pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Age</td>
<td>3</td>
<td>133539.09</td>
<td>44513.03</td>
<td>1498.57</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>+ Treat</td>
<td>1</td>
<td>93.51</td>
<td>93.51</td>
<td>3.15</td>
<td>0.077</td>
</tr>
<tr>
<td>+ Age.Treat</td>
<td>1</td>
<td>9.93</td>
<td>9.93</td>
<td>0.33</td>
<td>0.563</td>
</tr>
<tr>
<td>+ Separate nonlinear</td>
<td>2</td>
<td>73.31</td>
<td>36.66</td>
<td>1.23</td>
<td>0.292</td>
</tr>
<tr>
<td>Residual</td>
<td>514</td>
<td>15267.69</td>
<td>29.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>521</td>
<td>148983.54</td>
<td>285.96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- an inappropriate analysis for people without GenStat & HGNLMs
- suggestion of a treatment effect
GNLM with Brood as a fixed term

81 " compare to generalized nonlinear model (treating Brood as fixed) "
82 MODEL Weight
83 TERMS ExpBM * Treat + Brood
84 EXPRESSION [VALUE=\( \frac{1}{1+\exp(-B*(Age-M))} \)] CalcExpBM
85 RCYCLE B,M; INITIAL=Estcurve$['B','M']
86 FIT [PRINT=\*, NOMESS=alias; CONSTANT=estimate; CALC=CalcExpBM] ExpBM
87 ADD [PRINT=\*, NOMESS=alias] Treat
88 ADD [PRINT=\*, NOMESS=alias] Brood
89 ADD [PRINT=\*, NOMESS=alias] ExpBM.Treat
90 RDISPLAY [PRINT=accumulated; PPROBABILITY=\text{yes}]

Nonlinear regression analysis

Accumulated analysis of variance

<table>
<thead>
<tr>
<th>Change</th>
<th>d.f.</th>
<th>s.s.</th>
<th>m.s.</th>
<th>v.r.</th>
<th>F pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ ExpBM</td>
<td>3</td>
<td>133539.09</td>
<td>44513.03</td>
<td>2157.96</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>+ Treat</td>
<td>1</td>
<td>93.94</td>
<td>93.94</td>
<td>4.55</td>
<td>0.033</td>
</tr>
<tr>
<td>+ Brood</td>
<td>37</td>
<td>5437.11</td>
<td>146.95</td>
<td>7.12</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>+ ExpBM.Treat</td>
<td>1</td>
<td>32.42</td>
<td>32.42</td>
<td>1.57</td>
<td>0.211</td>
</tr>
<tr>
<td>Residual</td>
<td>479</td>
<td>9880.98</td>
<td>20.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>521</td>
<td>148983.54</td>
<td>285.96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

91 CALCULATE T1 = CPUTIME(0)
92 PRINT T1 - TO; DECIMALS=2; HEADING='Time taken (seconds)'

Time taken (seconds)
18.52

- an alternative inappropriate analysis
- significant (but misleading) treatment effects
Conclusions

• the HGLM menus & procedures provide very useful extensions to the standard Generalized linear Models
• represent the current state of the ongoing research by Lee, Nelder et al. on extensions to generalized linear models
• GenStat is providing a flexible and convenient framework for the collaboration – to try out, and then distribute, our ideas
• the methodology is described in the book
• there are many extensions (+ corrections) since then
  • including HGNLMs, in GenStat for Windows 10th Edition
  • Wald tests and plots of predicted means to come in the 11th Edition
• for further information, see vsni.co.uk or email roger.payne@vsni.co.uk