Full Credibility for GLM's and GLMM's

José Garrido

Concordia University

ARC 2014, University of California at Santa Barbara
July 13-16, 2014

Joint project with Oscar Alberto Quijano Xacur and Jian Tang

Warning

This talk contains explicit mathematical content.

Actuarially inspired guidance is advised (Rated AIG).

Are GLM and GLMM estimators always credible?

Since Mowbray (1914, PCAS) up to de Vylder (1985, IME) actuaries have worried about the reliability of statistical premium estimates obtained from the scarce data observed in some risk classes of a segmented portfolio of policies.

An extensive actuarial literature was produced over those 70 years to study this problem.

As actuaries have been adopting pricing models based on GLM's and GLMM's, the credibility question seems to have disappeared.

Here is a partial list of the few papers on the topic: Nelder and Verrall (1997, AB), Schmitter (2004, AB), Antonio and Beirlant (2007, IME), and G. and Zhou (2009, AB).

$\bar{Y}$

$\bar{Y}$ M

$$z\bar{Y} + (1 - z)M$$

$$z\bar{Y} + (1 - z)M$$

- The distribution of the loss depends on some random parameter Θ.

$$z\bar{Y} + (1 - z)M$$

- The distribution of the loss depends on some random parameter Θ.
- The optimal z is found by minimizing

$$\mathbb{E}\Big\{\big[\mu(\Theta) - (a\bar{Y} + b)\big]^2\Big\},$$

where $\mu(\Theta) = \mathbb{E}(Y \mid \Theta)$.

$$z\bar{Y} + (1 - z)M$$

- The distribution of the loss depends on some random parameter Θ.
- The optimal z is found by minimizing

$$\mathbb{E}\Big\{\big[\mu(\Theta) - (a\bar{Y} + b)\big]^2\Big\},$$

where $\mu(\Theta) = \mathbb{E}(Y \mid \Theta)$.

- z depends only on $\mathbb{E}\big[\mu(\Theta)\big]$, $\mathbb{V}\big[\sigma^2(\Theta)\big]$ and $\mathbb{V}\big[\mu(\Theta)\big]$, where $\sigma^2(\Theta) = \mathbb{V}\big[Y \mid \Theta\big]$, and $M = \mathbb{E}\big[\mu(\Theta)\big]$.

$$z\bar{Y} + (1 - z)M$$

- The distribution of the loss depends on some random parameter Θ.
- The optimal z is found by minimizing

$$\mathbb{E}\Big\{\big[\mu(\Theta) - (a\bar{Y} + b)\big]^2\Big\},$$

  where $\mu(\Theta) = \mathbb{E}(Y \mid \Theta)$.
- z depends only on $\mathbb{E}\big[\mu(\Theta)\big]$, $\mathbb{V}\big[\sigma^2(\Theta)\big]$ and $\mathbb{V}\big[\mu(\Theta)\big]$, where $\sigma^2(\Theta) = \mathbb{V}\big[Y \mid \Theta\big]$, and $M = \mathbb{E}\big[\mu(\Theta)\big]$.
- It is distribution free.

Jewell's prior

For (conditional) loss distributions, given $\Theta = \theta$, from the exponential (dispersion) family:

$$p(y, \theta, \lambda) = a(y, \lambda) \exp\Big\{\lambda\big[\theta y - \kappa(\theta)\big]\Big\}.$$

For (conditional) loss distributions, given $\Theta = \theta$, from the exponential (dispersion) family:

$$p(y, \theta, \lambda) = a(y, \lambda) \exp\Big\{\lambda\big[\theta y - \kappa(\theta)\big]\Big\}.$$

Consider the following prior distribution for Θ:

$$u_{n_0, x_0}(\theta) = A(n_0, x_0) \exp\Big\{n_0\big[\theta x_0 - \kappa(\theta)\big]\Big\},$$

then:

For (conditional) loss distributions, given $\Theta = \theta$, from the exponential (dispersion) family:

$$p(y, \theta, \lambda) = a(y, \lambda) \exp \Big\{ \lambda \big[\theta y - \kappa(\theta) \big] \Big\}.$$

Consider the following prior distribution for Θ:

$$u_{n_0, x_0}(\theta) = A(n_0, x_0) \exp \Big\{ n_0 \big[\theta x_0 - \kappa(\theta) \big] \Big\},$$

then:

- $u_{n_0, x_0}(\theta)$ is a conjugate prior.

For (conditional) loss distributions, given $\Theta = \theta$, from the exponential (dispersion) family:

$$p(y, \theta, \lambda) = a(y, \lambda) \exp\Big\{\lambda\big[\theta y - \kappa(\theta)\big]\Big\}.$$

Consider the following prior distribution for Θ:

$$u_{n_0, x_0}(\theta) = A(n_0, x_0) \exp\Big\{n_0\big[\theta x_0 - \kappa(\theta)\big]\Big\},$$

then:

- $u_{n_0, x_0}(\theta)$ is a conjugate prior.
- and the mean of the posterior distribution is

$$\mathbb{E}\big[\mu(\Theta) \mid Y_1, \ldots, Y_n\big] = \frac{n_0}{n_0 + n\lambda} x_0 + \frac{n\lambda}{n_0 + n\lambda} \bar{Y}.$$

Jewell's prior

For (conditional) loss distributions, given $\Theta = \theta$, from the exponential (dispersion) family:

$$p(y, \theta, \lambda) = a(y, \lambda) \exp\Big\{ \lambda \big[\theta y - \kappa(\theta) \big] \Big\}.$$

Consider the following prior distribution for Θ:

$$u_{n_0, x_0}(\theta) = A(n_0, x_0) \exp\Big\{ n_0 \big[\theta x_0 - \kappa(\theta) \big] \Big\},$$

then:

- $u_{n_0, x_0}(\theta)$ is a conjugate prior.
- and the mean of the posterior distribution is

$$\mathbb{E}\big[\mu(\Theta) \mid Y_1, \ldots, Y_n \big] = \frac{n_0}{n_0 + n\lambda} x_0 + \frac{n\lambda}{n_0 + n\lambda} \bar{Y}.$$

- Bühlmann's formula is exact in this case, i.e. the best linear estimator is also the best overall Bayesian estimator.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.

.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.
- Linear credibility estimation for $\vec{\beta}(\Theta_j)$ of the form

$$\vec{B}_j = \mathbf{Z}_j \hat{\vec{\beta}}_j + (\mathbf{I} - \mathbf{Z}_j)\vec{b}$$

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.
- Linear credibility estimation for $\vec{\beta}(\Theta_j)$ of the form

$$\vec{B}_j = \mathbf{Z}_j \hat{\vec{\beta}}_j + (\mathbf{I} - \mathbf{Z}_j)\vec{b}$$

- $\hat{\vec{\beta}}_j$ is the estimator of $\vec{\beta}(\Theta_j)$, $\mathbf{Z}_j = \mathrm{diag}(z_{j1}, \ldots, z_{jk})$, $\vec{b} = \mathbb{E}\left[\vec{\beta}(\Theta_j)\right]$.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.
- Linear credibility estimation for $\vec{\beta}(\Theta_j)$ of the form

$$\vec{B}_j = \mathbf{Z}_j \hat{\vec{\beta}}_j + (\mathbf{I} - \mathbf{Z}_j)\vec{b}$$

- $\hat{\vec{\beta}}_j$ is the estimator of $\vec{\beta}(\Theta_j)$, $\mathbf{Z}_j = \mathrm{diag}(z_{j1}, \ldots, z_{jk})$, $\vec{b} = \mathbb{E}\left[\vec{\beta}(\Theta_j)\right]$.
- The optimal $\mathbf{Z}_j$ is found by minimizing $\mathbb{E}\left\{[\vec{\beta}(\Theta_j) - \vec{B}_j]^T \boldsymbol{\Sigma}[\vec{\beta}(\Theta_j) - \vec{B}_j]\right\}$, for some positive definite matrix $\boldsymbol{\Sigma}$.

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \ldots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.
- Linear credibility estimation for $\vec{\beta}(\Theta_j)$ of the form

$$\vec{B}_j = \mathbf{Z}_j\hat{\vec{\beta}}_j + (\mathbf{I} - \mathbf{Z}_j)\vec{b}$$

- $\hat{\vec{\beta}}_j$ is the estimator of $\vec{\beta}(\Theta_j)$, $\mathbf{Z}_j = \text{diag}(z_{j1}, \ldots, z_{jk})$, $\vec{b} = \mathbb{E}\big[\vec{\beta}(\Theta_j)\big]$.
- The optimal $\mathbf{Z}_j$ is found by minimizing $\mathbb{E}\big\{[\vec{\beta}(\Theta_j) - \vec{B}_j]^T \boldsymbol{\Sigma}[\vec{\beta}(\Theta_j) - \vec{B}_j]\big\}$, for some positive definite matrix $\boldsymbol{\Sigma}$.
- Credibility is linear for $\hat{\vec{\beta}}_j$ and the means $\vec{\mu}_j = \mathbb{E}(\vec{Y}_j)$.

Hachemeister's credibility for regression models

- We assume that there are different risk classes. A regression is carried out for each class separately.
- For each class the response variable is a vector $\vec{Y}_j = (Y_{j1}, \dots, Y_{jn})$.
- Each class has its own random risk parameter Θ_j, such that

$$\mathbb{E}(\vec{Y}_j \mid \Theta_j) = \mathbf{X}\vec{\beta}(\Theta_j)$$

- The Θ_js are identically distributed.
- Linear credibility estimation for $\vec{\beta}(\Theta_j)$ of the form

$$\vec{B}_j = \mathbf{Z}_j \hat{\vec{\beta}}_j + (\mathbf{I} - \mathbf{Z}_j)\vec{b}$$

- $\hat{\vec{\beta}}_j$ is the estimator of $\vec{\beta}(\Theta_j)$, $\mathbf{Z}_j = \mathrm{diag}(z_{j1}, \dots, z_{jk})$, $\vec{b} = \mathbb{E}\big[\vec{\beta}(\Theta_j)\big]$.
- The optimal $\mathbf{Z}_j$ is found by minimizing $\mathbb{E}\big\{[\vec{\beta}(\Theta_j) - \vec{B}_j]^T \boldsymbol{\Sigma} [\vec{\beta}(\Theta_j) - \vec{B}_j]\big\}$, for some positive definite matrix $\boldsymbol{\Sigma}$.
- Credibility is linear for $\hat{\vec{\beta}}_j$ and the means $\vec{\mu}_j = \mathbb{E}(\vec{Y}_j)$.
- Distribution free.

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

Linear credibility formula for the $\vec{\beta}$'s:

$$\vec{B} = \mathbf{Z}\hat{\vec{\beta}} + (\mathbf{I} - \mathbf{Z})\vec{b}.$$

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

Linear credibility formula for the $\vec{\beta}$'s:

$$\vec{B} = \mathbf{Z}\hat{\vec{\beta}} + (\mathbf{I} - \mathbf{Z})\vec{b}.$$

- The optimal $\mathbf{Z}$ is found in such a way that the distance from $\vec{B}$ to $\vec{\beta}(\Theta)$ is minimum.

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

Linear credibility formula for the $\vec{\beta}$'s:

$$\vec{B} = \mathbf{Z}\hat{\vec{\beta}} + (\mathbf{I} - \mathbf{Z})\vec{b}.$$

- The optimal $\mathbf{Z}$ is found in such a way that the distance from $\vec{B}$ to $\vec{\beta}(\Theta)$ is minimum.
- The definition of distance comes from an inner product $\langle X, Y \rangle = \mathbb{E}(\mathbf{X}\boldsymbol{\Sigma}\vec{Y})$ for some positive definite matrix $\boldsymbol{\Sigma}$.

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

Linear credibility formula for the $\vec{\beta}$'s:

$$\vec{B} = \mathbf{Z}\hat{\vec{\beta}} + (\mathbf{I} - \mathbf{Z})\vec{b}.$$

- The optimal $\mathbf{Z}$ is found in such a way that the distance from $\vec{B}$ to $\vec{\beta}(\Theta)$ is minimum.
- The definition of distance comes from an inner product $\langle X, Y \rangle = \mathbb{E}(\mathbf{X}\boldsymbol{\Sigma}\vec{Y})$ for some positive definite matrix $\boldsymbol{\Sigma}$.
- Credibility is linear for the $\vec{\beta}$'s but not for the mean.

$$\mathbb{E}(\vec{Y} \mid \Theta) = f(\vec{\beta}(\Theta))$$

Linear credibility formula for the $\vec{\beta}$'s:

$$\vec{B} = \mathbf{Z}\hat{\vec{\beta}} + (\mathbf{I} - \mathbf{Z})\vec{b}.$$

- The optimal $\mathbf{Z}$ is found in such a way that the distance from $\vec{B}$ to $\vec{\beta}(\Theta)$ is minimum.
- The definition of distance comes from an inner product $\langle X, Y \rangle = \mathbb{E}(\mathbf{X}\boldsymbol{\Sigma}\vec{Y})$ for some positive definite matrix $\boldsymbol{\Sigma}$.
- Credibility is linear for the $\vec{\beta}$'s but not for the mean.
- Distribution free.

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.
- If $\vec{\mu} = \mathbb{E}(\vec{Y})$ and **g** is the link function

$$\mathbf{g}(\vec{\mu}) = \begin{pmatrix} g(\mu_1) \\ \vdots \\ g(\mu_k) \end{pmatrix} = \mathbf{X}\vec{\beta}, \qquad \vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}).$$

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.
- If $\vec{\mu} = \mathbb{E}(\vec{Y})$ and $\mathbf{g}$ is the link function

$$\mathbf{g}(\vec{\mu}) = \begin{pmatrix} g(\mu_1) \\ \vdots \\ g(\mu_k) \end{pmatrix} = \mathbf{X}\vec{\beta}, \qquad \vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}).$$

Is it possible to have linear credibility for $\vec{\mu}$ under this setup?

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.
- If $\vec{\mu} = \mathbb{E}(\vec{Y})$ and $\mathbf{g}$ is the link function

$$\mathbf{g}(\vec{\mu}) = \begin{pmatrix} g(\mu_1) \\ \vdots \\ g(\mu_k) \end{pmatrix} = \mathbf{X}\vec{\beta}, \qquad \vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}).$$

Is it possible to have linear credibility for $\vec{\mu}$ under this setup?

$$\mathbf{Z}\mathbf{g}^{-1}(\mathbf{X}\hat{\vec{\beta}}) + (\mathbf{I} - \mathbf{Z})\vec{M} \;\; = \;\; \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}_{cred})$$

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.
- If $\vec{\mu} = \mathbb{E}(\vec{Y})$ and $\mathbf{g}$ is the link function

$$\mathbf{g}(\vec{\mu}) = \begin{pmatrix} g(\mu_1) \\ \vdots \\ g(\mu_k) \end{pmatrix} = \mathbf{X}\vec{\beta}, \qquad \vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}).$$

Is it possible to have linear credibility for $\vec{\mu}$ under this setup?

$$\begin{aligned} \mathbf{Z}\mathbf{g}^{-1}(\mathbf{X}\hat{\vec{\beta}}) + (\mathbf{I} - \mathbf{Z})\vec{M} &= \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}_{cred}) \\ \mathbf{g}(\mathbf{Z}\mathbf{g}^{-1}(\mathbf{X}\hat{\vec{\beta}}) + (\mathbf{I} - \mathbf{Z})\vec{M}) &= \mathbf{X}\vec{\beta}_{cred} \end{aligned}$$

- $\vec{Y}$ is a vector, with components in the exponential dispersion family.
- If $\vec{\mu} = \mathbb{E}(\vec{Y})$ and $\mathbf{g}$ is the link function

$$\mathbf{g}(\vec{\mu}) = \begin{pmatrix} g(\mu_1) \\ \vdots \\ g(\mu_k) \end{pmatrix} = \mathbf{X}\vec{\beta}, \qquad \vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}).$$

Is it possible to have linear credibility for $\vec{\mu}$ under this setup?

$$\begin{aligned} \mathbf{Z}\mathbf{g}^{-1}(\mathbf{X}\hat{\vec{\beta}}) + (\mathbf{I} - \mathbf{Z})\vec{M} &= \mathbf{g}^{-1}(\mathbf{X}\vec{\beta}_{cred}) \\ \mathbf{g}(\mathbf{Z}\mathbf{g}^{-1}(\mathbf{X}\hat{\vec{\beta}}) + (\mathbf{I} - \mathbf{Z})\vec{M}) &= \mathbf{X}\vec{\beta}_{cred} \end{aligned}$$

the left side is not in $R(\mathbf{X})$ for all $\mathbf{Z}$. So, not possible.

- Either we give up linearity or modify the GLM.

- Either we give up linearity or modify the GLM.
- GLMM's are used in credibility. Linear credibility for the mean can be obtained in this case.

- Either we give up linearity or modify the GLM.
- GLMM's are used in credibility. Linear credibility for the mean can be obtained in this case.
- Instead we propose to give up linearity without modifying the GLM:
  - Closed form formula.
  - Interpretable.

- Either we give up linearity or modify the GLM.
- GLMM's are used in credibility. Linear credibility for the mean can be obtained in this case.
- Instead we propose to give up linearity without modifying the GLM:
  - Closed form formula.
  - Interpretable.
- The problem consists in finding credibility formulas for which we can find a corresponding prior for $\vec{\beta}$.

- Either we give up linearity or modify the GLM.
- GLMM's are used in credibility. Linear credibility for the mean can be obtained in this case.
- Instead we propose to give up linearity without modifying the GLM:
    - Closed form formula.
    - Interpretable.
- The problem consists in finding credibility formulas for which we can find a corresponding prior for $\vec{\beta}$.
- Alternatively we could find a prior for $\vec{\mu}$ and find the corresponding distribution for $\vec{\beta}$ with the relation $\vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta})$.

- Either we give up linearity or modify the GLM.
- GLMM's are used in credibility. Linear credibility for the mean can be obtained in this case.
- Instead we propose to give up linearity without modifying the GLM:
    - Closed form formula.
    - Interpretable.
- The problem consists in finding credibility formulas for which we can find a corresponding prior for $\vec{\beta}$.
- Alternatively we could find a prior for $\vec{\mu}$ and find the corresponding distribution for $\vec{\beta}$ with the relation $\vec{\mu} = \mathbf{g}^{-1}(\mathbf{X}\vec{\beta})$.
- The case of 1 parameter is an equivalent problem, but for the exponential family.

- An exponential dispersion family can be parametrized as

$$p(y;\mu,\lambda) = a(y,\lambda)\exp\Big\{-\frac{\lambda}{2}d(y,\mu)\Big\},$$

where d is the unit deviance function of the family.

Conjugate priors for exponential dispersion families

- An exponential dispersion family can be parametrized as

$$p(y;\mu,\lambda) = a(y,\lambda)\exp\Big\{-\frac{\lambda}{2}d(y,\mu)\Big\},$$

  where d is the unit deviance function of the family.

- Let us consider the prior for μ defined as

$$\pi_{n_0,x_0}(\mu) = A(n_0,x_0)h(\mu)\exp\Big\{-\frac{n_0}{2}d(x_0,\mu)\Big\},$$

  for some function h.

Conjugate priors for exponential dispersion families, properties:

- For every h, we have that π_{n_0,x_0} is a conjugate prior.

Conjugate priors for exponential dispersion families, properties:

- For every h, we have that π_{n_0,x_0} is a conjugate prior.
- The credibility formula implied by π_{n_0,x_0} is linear if and only if

$$\mathbb{E}_{\pi_{n_0,x_0}}[\mu] \propto x_0.$$

Conjugate priors for exponential dispersion families, properties:

- For every h, we have that π_{n_0,x_0} is a conjugate prior.
- The credibility formula implied by π_{n_0,x_0} is linear if and only if

$$\mathbb{E}_{\pi_{n_0,x_0}}[\mu] \propto x_0.$$

- π_{n_0,x_0} corresponds to Jewell's prior when $h(\mu) = \frac{1}{V(\mu)}$, where $V(\mu)$ is the variance function of the family.

Conjugate priors for exponential dispersion families, properties:

- For every h, we have that π_{n_0,x_0} is a conjugate prior.
- The credibility formula implied by π_{n_0,x_0} is linear if and only if

$$\mathbb{E}_{\pi_{n_0,x_0}}[\mu] \propto x_0.$$

- π_{n_0,x_0} corresponds to Jewell's prior when $h(\mu) = \frac{1}{V(\mu)}$, where $V(\mu)$ is the variance function of the family.
- By choosing different h it is possible to find different credibility forms, including non-linear ones.

Conjugate priors for exponential dispersion families, properties:

- For every h, we have that π_{n_0,x_0} is a conjugate prior.
- The credibility formula implied by π_{n_0,x_0} is linear if and only if

$$\mathbb{E}_{\pi_{n_0,x_0}}[\mu] \propto x_0.$$

- π_{n_0,x_0} corresponds to Jewell's prior when $h(\mu) = \frac{1}{V(\mu)}$, where $V(\mu)$ is the variance function of the family.
- By choosing different h it is possible to find different credibility forms, including non-linear ones.
- Example: For the gamma distribution, $h(\mu) = \frac{\log(\mu)}{\mu^N}$ gives

$$\mathbb{E}_{\pi_{n_0,x_0}}[\mu] \propto x_0 \left(1 + \frac{1}{a\ln(x_0) + b}\right),$$

which implies a nonlinear credibility formula for μ.

Antonio, K. and Beirlant, J. (2007) "Actuarial statistics with generalized linear mixed models. *Insurance: Mathematics and Economics*, **40**(1), pp.58–76.

De Vylder, F.E. (1985) "Non–linear regression in credibility theory", *Insurance: Mathematics and Economics*, **4**(3), pp.163–172.

Garrido, J. and Zhou, J. (2009) "Full credibility with generalized linear and mixed models", *ASTIN Bulletin*, **39**(61).

Nelder, J.A. and Verrall, R.J. (1997) "Credibility theory and generalized linear models", *ASTIN Bulletin*, **27**(1).

Schmitter, H. (2004) "The sample size needed for the calculation of a GKM tariff", *ASTIN Bulletin*, **34**(1), 2004.

Mowbray, A.H. (1914) "How extensive a payroll exposure is necessary to give a dependable pure premium", *Proc. of the Casualty Actuarial Society*, **1**, pp.24–30.

Thank you for your attention!