A cookbook, not a panacea, for claims reserving

Rezervovacia kuchárka (GLM ako aperitív, hlavné menu po bayesovsky, GEE ako digestív)

Michal Pešta

Charles University in Prague

Faculty of Mathematics and Physics

Actuarial Seminar, Prague

12 October 2012

Overview

- Chain ladder models
- GLM in stochastic reserving
- Nonparametric smoothing models in reserving
- Bayesian reserving models
- GEE in stochastic reserving

- Support:

  Czech Science Foundation project "DYME Dynamic Models in Economics" No. P402/12/G097

Terminology

- $X_{i,j} \ldots$ claim amounts in development year j with accident year i
- $X_{i,j}$ stands for the incremental claims in accident year i made in accounting year $i+j$
- $n \ldots$ current year – corresponds to the most recent accident year and development period
- Our data history consists of right-angled isosceles triangles $X_{i,j}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, n+1-i$

Run-off (incremental) triangle

Accident year i	Development year j				
	1	2	$\cdots$	$n-1$	n
1	$X_{1,1}$	$X_{1,2}$	$\cdots$	$X_{1,n-1}$	$X_{1,n}$
2	$X_{2,1}$	$X_{2,2}$	$\cdots$	$X_{2,n-1}$	
			$\ddots$		
$\vdots$	$\vdots$	$\vdots$	$X_{i,n+1-i}$		
$n-1$	$X_{n-1,1}$	$X_{n-1,2}$			
n	$X_{n,1}$				

Notation

- $C_{i,j}$ ... cumulative payments in origin year i after j development periods

$$C_{i,j} = \sum_{k=1}^{j} X_{i,k}$$

- $C_{i,j}$ ... a random variable of which we have an observation if $i + j \leq n + 1$
- Aim is to estimate the ultimate claims amount $C_{i,n}$ and the outstanding claims reserve

$$R_i = C_{i,n} - C_{i,n+1-i}, \quad i = 2, \ldots, n$$

by completing the triangle into a square

Run-off (cumulative) triangle

Accident year i	Development year j				
	1	2	$\cdots$	$n-1$	n
1	$C_{1,1}$	$C_{1,2}$	$\cdots$	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$	$\cdots$	$C_{2,n-1}$	
			$\ddots$		
$\vdots$	$\vdots$	$\vdots$	$C_{i,n+1-i}$		
$n-1$	$C_{n-1,1}$	$C_{n-1,2}$			
n	$C_{n,1}$				

Chain ladder

[1] $\mathsf{E}[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = f_j C_{i,j}$

[2] $\mathsf{Var}[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = \sigma_j^2 C_{i,j}^{\alpha},\ \alpha \in \mathbb{R}$

[3] Accident years $[C_{i,1},\ldots,C_{i,n}]$ are independent vectors

Development factors (link ratios) f_j

$$\widehat{f}_j^{(n)} = \frac{\sum_{i=1}^{n-j} C_{i,j}^{1-\alpha} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}^{2-\alpha}}, \quad 1 \leq j \leq n-1$$

$\widehat{f}_n^{(n)} \equiv 1$ (assuming no tail)

Mack or linear regression

- $\alpha = 0 \ldots$ linear regression (no intercept, homoscedastic) for $[C_{\bullet,j}, C_{\bullet,j+1}]$ satisfies CL
- $\alpha = 1 \ldots$ Mack (1993), but also the Aitken (no intercept, heteroscedastic) regression model with weights $C_{i,j}^{-1}$
- smoothing (and extrapolation) of development factors possible

Ultimates and reserves

- Ultimate claims amounts $C_{i,n}$ are estimated by

$$\widehat{C}_{i,n} = C_{i,n+1-i} \times \widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1}$$

- Reserves R_i are, thus, estimated by

$$\widehat{R}_i = \widehat{C}_{i,n} - C_{i,n+1-i} = C_{i,n+1-i}\left(\widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1} - 1\right)$$

Generalized Linear Models

- a flexible generalization of ordinary linear regression
- formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression and Poisson regression

GLM: 3 elements

1. random component: outcome of the dependent variables Y from the exponential family, i.e.,

$$f_Y(y;\theta,\phi) = \exp\{[y\theta - b(\theta)]/a(\phi) + c(y,\phi)\}$$

   where θ is canonical parameter, ϕ is dispersion parameter and $\mathrm{E}Y_i = \mu_i$

2. systematic component: linear predictor (mean structure)

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$$

3. link: function g

$$\eta_i = g(\mu_i)$$

Exponential family

- include many of the most common distributions, including the normal, exponential, gamma, chi-squared, beta, Dirichlet, Bernoulli, categorical, Poisson, Wishart, Inverse Wishart and many others
- a number of common distributions are exponential families only when certain parameters are considered fixed and known, e.g., binomial (with fixed number of trials), multinomial (with fixed number of trials), and negative binomial (with fixed number of failures)
- common distributions that are not exponential families are Student's t, most mixture distributions, and even the family of uniform distributions with unknown bounds

Canonical link – sufficient statistic

- exponential family

$$\mathsf{E}Y = \mu = b'(\theta), \quad \mathsf{Var}Y = b''(\theta)a(\phi) \equiv V(\mu)\tilde{a}(\phi)$$

- distribution $\longleftrightarrow$ link function (sufficient statistic $\longleftrightarrow$ canonical link)
  - normal ... identity: $\mu_i = \mathbf{X}_{i,\bullet}\boldsymbol{\beta}$
  - Gamma (exponential) ... inverse (reciprocal): $\mu_i^{-1} = \mathbf{X}_{i,\bullet}\boldsymbol{\beta}$
  - Poisson ... logarithm: $\log(\mu_i) = \mathbf{X}_{i,\bullet}\boldsymbol{\beta}$
  - Binomial (multinomial) ... logit: $\log\left(\frac{\mu_i}{1-\mu_i}\right) = \mathbf{X}_{i,\bullet}\boldsymbol{\beta}$
  - inverse Gaussian ... reciprocal squared: $\mu_i^{-2} = \mathbf{X}_{i,\bullet}\boldsymbol{\beta}$

Link functions

- logit $\eta = \log\{\mu/(1-\mu)\}$
- probit $\eta = \Phi^{-1}(\mu)$
- complementary log-log $\eta = \log\{-\log(1-\mu)\}$
- power family of links

$$\eta = \begin{cases} (\mu^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \log\mu, & \lambda = 0; \end{cases} \quad \text{or} \quad \eta = \begin{cases} \mu^\lambda, & \lambda \neq 0, \\ \log\mu, & \lambda = 0. \end{cases}$$

Estimation

- estimation of the parameters via maximum likelihood, quasi-likelihood or Bayesian techniques

$N(\mu, \sigma^2)$

- support $(-\infty, +\infty)$
- dispersion parameter $\phi = \sigma^2$
- cumulant function $b(\theta) = \theta^2/2$
- $c(y, \phi) = -\frac{1}{2}\left(\frac{y^2}{\phi} + \log(2\pi\phi)\right)$
- $\mu(\theta) = \mathsf{E}_\theta Y = \theta$
- canonical link $\theta(\mu)$: identity
- variance function $V(\mu) = 1$

$Po(\mu)$

- support $\{0, 1, 2, \ldots\}$
- dispersion parameter $\phi = 1$
- cumulant function $b(\theta) = \exp\{\theta\}$
- $c(y, \phi) = -\log y!$
- $\mu(\theta) = \mathsf{E}_\theta Y = \exp\{\theta\}$
- canonical link $\theta(\mu)$: log
- variance function $V(\mu) = \mu$

$\Gamma(\mu, \nu)$

- support $(0, +\infty)$
- Var$Y = \mu^2\nu$
- dispersion parameter $\phi = \nu^{-1}$
- cumulant function $b(\theta) = -\log\{-\theta\}$
- $c(y, \phi) = \nu\log(\nu y) - \log y - \log\Gamma(\nu)$
- $\mu(\theta) = \mathsf{E}_\theta Y = -1/\theta$
- canonical link $\theta(\mu)$: reciprocal
- variance function $V(\mu) = \mu^2$

Mack's model as GLM

- reformulate Mack's model as a model of ratios

$$\mathrm{E}\left[\frac{C_{i,j+1}}{C_{i,j}}\right] = f_j \quad \text{and} \quad \mathrm{Var}\left[\frac{C_{i,j+1}}{C_{i,j}}\middle| C_{i,1},\ldots,C_{i,j}\right] = \frac{\sigma_j^2}{C_{i,j}}$$

- conditional weighted normal GLM

$$\frac{C_{i,j+1}}{C_{i,j}} \sim \mathbb{N}\left(f_j, \frac{\sigma_j^2}{C_{i,j}}\right)$$

- Mack's model was not derived/designed as a GLM, but a conditional weighted normal GLM gives the same estimates

GLM for triangles

- independent incremental claims $X_{ij},\ i+j \leq n+1$
  - overdispersed Poisson distributed X_{ij}

$$\mathsf{E}[X_{ij}] = \mu_{ij} \quad \text{and} \quad \mathsf{Var}[X_{ij}] = \phi\mu_{ij}$$

  - Gamma distributed X_{ij}

$$\mathsf{E}[X_{ij}] = \mu_{ij} \quad \text{and} \quad \mathsf{Var}[X_{ij}] = \phi\mu_{ij}^2$$

- logarithmic link function

$$\log(\mu_{ij}) = \gamma + \alpha_i + \beta_j, \quad \alpha_1 = \beta_1 = 0$$

GLM for triangles II

- overdispersed Poisson with log link provides asymptotically same parameter estimates, predicted values and prediction errors
- possible extensions:
  - Hoerl curve

$$\log(\mu_{ij}) = \gamma + \alpha_i + \beta_j \log(j) + \delta_j j$$

  - smoother (semiparametric)

$$\log(\mu_{ij}) = \gamma + \alpha_i + s_1(\log(j)) + s_2(j)$$

Estimation in triangles

- ML (maximum likelihood) … likelihood

$$L(\boldsymbol{\theta}, \phi; \mathbf{X}) = \prod_{i=1}^{n} \prod_{j=1}^{n+1-i} f(X_{ij}; \theta_{ij}, \phi)$$

- maximize log-likelihood w.r.t. parameters of μ, which is an argument of θ, i.e., $\theta(\boldsymbol{\mu}(\alpha, \beta))$

! there is no overdispersed Poisson distribution (only if thinking of negative binomial)

- QML (quasi-maximum likelihood)… quasi-likelihood for ODP

$$\log Q(\boldsymbol{\mu}; \mathbf{X}) = \sum_{i=1}^{n} \sum_{j=1}^{n+1-i} \phi(X_{ij} \log \mu_{ij} - \mu_{ij}) + const$$

Generalized additive models

- GAM …extension of GLM, with the linear predictor being replaced by a non-parametric smoother

$$\eta_{ij} = \sum_{k=1}^{p} s_k(X_{ij})$$

- $s(x)$ represents a non-parametric smoother on x, which may be chosen from several different types of smoother, such as locally weighted regression smoothers (loess), cubic smoothing splines and kernel smoothers

GAM

- Ex: smoothing (trade-off between smoothness and fit) for univariate ($p = 1$) cubic spline with normal distribution

$$\min \left\{ \sum_{i,j} [X_{ij} - s(X_{ij})]^2 + \lambda \int [s''(t)]^2 \mathrm{d}t \right\}$$

Bayesian approach

- problem of instability in the proportion of ultimate claims paid in the early development years, causing a method such as the CL to produce unsatisfactory results when applied mechanically
- to stabilize the results using an external initial estimate of ultimate claims

Bornhuetter-Ferguson

- reminder from CL: outstanding claims

$$R_i = C_{i,n+1-i}(f_{n+1-i} \times \ldots \times f_{n-1} - 1)$$

Assumptions

(1) $\mathsf{E}[C_{i,j+k}|C_{i,1}, \ldots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i$ and
$\mathsf{E}[C_{i,1}] = \beta_1 \mu_i$
$\beta_j > 0,\ \mu_i > 0,\ \beta_n = 1$
$1 \leq i \leq n,\ 1 \leq j \leq n,\ 1 \leq k \leq n - j$

(2) accident years $[C_{i,1}, \ldots, C_{i,n}],\ 1 \leq i \leq n$ are independent

- estimate $\widehat{C}_{i,n}^{BF}$ for ultimates
- Bayesian approach to CL

Bornhuetter-Ferguson Method

- a very robust method since it does not consider outliers in the observations

Implied Assumptions

(1) $\mathsf{E}[C_{i,j}] = \beta_j \mu_i$
$\beta_j > 0,\ \mu_i > 0,\ \beta_n = 1$
$1 \leq i \leq n,\ 1 \leq j \leq n$

(2) accident years $[C_{i,1}, \ldots, C_{i,n}],\ 1 \leq i \leq n$ are independent

- "implied"assumptions are weaker than the original BF assumptions (and, hence, not equivalent)

BF Estimator

- BF estimator of ultimate from the latest

$$\widehat{C}_{i,n}^{BF} = C_{i,n-i+1} + (1 - \widehat{\beta}_{n-i+1})\widehat{\mu}_i$$

- comparing CL and BF model $\rightsquigarrow \prod_{k=j}^{n-1} f_k^{-1}$ plays the role of β_j and, therefore,

$$\widehat{\beta}_j = \prod_{k=j}^{n-1} \frac{1}{\widehat{f}_k}$$

- need a prior estimate for μ_i
- $\widehat{\mu}_i$ is often a plan value from a strategic business plan or the value used for premium calculations
- $\widehat{\mu}_i$ should be estimated before one has any observations (i.e., should be a pure prior estimate based on expert opinion) !

Comparison of BF and CL Estimators

- if the prior estimate of ultimates μ_i is equal to the CL estimate of ultimates, then the BF and CL estimators underline
- BF:

$$\widehat{C}_{i,n}^{BF} = C_{i,n-i+1} + (1 - \widehat{\beta}_{n-i+1})\widehat{\mu}_i$$

- CL:

$$\widehat{C}_{i,n}^{CL} = C_{i,n-i+1} + (1 - \widehat{\beta}_{n-i+1})\widehat{C}_{i,n}^{CL}$$

- BF differs from the CL in that the CL estimate of ultimate claims is replaced by an alternative estimate based on external information and expert judgement

Bayesian Models

- a prior distribution for row (ultimate) parameter
- Ex: ODP model, where an obvious candidate is

$$\mu_i \sim \text{independent } \Gamma(\gamma_i, \delta_i)$$

such that

$$\mathsf{E}\mu_i = \frac{\gamma_i}{\delta_i}$$

Predictive distribution

- posterior predictive distribution of incremental claims $X_{i,j}$ is an over-dispersed negative binomial distribution with mean

$$\left[Z_{i,j}C_{i,j-1} + (1 - Z_{i,j})\frac{\gamma_i}{\delta_i}\frac{1}{f_{j-1} \times \ldots \times f_{n-1}}\right](f_{j-1} - 1)$$

where

$$Z_{i,j} = \frac{\frac{1}{f_{j-1}\times\ldots\times f_{n-1}}}{\delta_i\phi + \frac{1}{f_{j-1}\times\ldots\times f_{n-1}}}$$

Credibility formula

- a natural trade-off between two competing estimates for $X_{i,j}$

$$C_{i,j-1} \quad \text{and} \quad \frac{\gamma_i}{\delta_i} \frac{1}{f_{j-1} \times \ldots \times f_{n-1}} = \mathsf{E}[\mu_i] \frac{1}{f_{j-1} \times \ldots \times f_{n-1}}$$

- Bayesian model has the CL as one extreme (no prior information about the row parameters) and the BF as the other (perfect prior information about the row parameters)

Bayesian trade-off

- BF assumes that there is perfect prior information about the row parameters (does not use the data at all for one part of estimation) ... heroic assumption
- prefer to use something between BF and CL
- credibility factor $Z_{i,j}$ governs the trade-off between the prior mean and the data
- the further through the development we are, the larger $\frac{1}{f_{j-1} \times \ldots \times f_{n-1}}$ becomes, and the more weight is given to the CL ladder estimate
- choice of δ_i is governed by the prior precision of the initial estimate for ultimates ... with regard given to the over-dispersion parameter (e.g., an initial estimate from the ODP)

Cape Cod

- another Bayesian-like approach

Assumptions

(1) $\mathsf{E}[C_{i,j}] = \kappa \pi_i \beta_j$
$\kappa > 0,\ \pi_i > 0,\ \beta_j > 0,\ \beta_n = 1$
$1 \leq i \leq n,\ 1 \leq j \leq n,\ 1 \leq k \leq n - j$

(2) accident years $[C_{i,1}, \ldots, C_{i,n}], \quad 1 \leq i \leq n$ are independent

- estimate $\widehat{C}_{i,n}^{CC}$ for ultimates
- equivalent to implied assumptions of BF with $\mu_i = \kappa \pi_i$

Cape Cod Estimator

- main deficiency of DFM (CL) ... ultimate claims completely depend on the latest diagonal claims $\rightsquigarrow$ not robust (sensitive to outliers)
- moreover, in long-tailed LoBs (e.g., liability) the first observation is not representative
- one possibility is to smoothen outliers from the latest (diagonal) $\rightsquigarrow$ combine BF and CL into Benktander-Hovinen method
- another way is to make the diagonal claims more robust $\rightsquigarrow$ Cape Cod method
- CC estimator of ultimate from the latest

$$\widehat{C}_{i,n}^{CC} = C_{i,n-i+1} - \widehat{C}_{i,n-i+1}^{CC} + \prod_{j=n-i+1}^{n-1} f_j \widehat{C}_{i,n-i+1}^{CC}$$

Generalised Cape Cod

- π_i can be interpreted as the premium received for accident year i
- κ reflects the average loss ratio
- loss ratio for an accident year using the CL estimate for the ultimate claim

$$\widehat{\kappa}_i = \frac{\widehat{C}_{i,n}^{CL}}{\pi_i} = \frac{C_{i,n-i+1}}{\prod_{j=n-i+1}^{n-1} f_j \pi_i} = \frac{C_{i,n-i+1}}{\beta_{n-i+1}\pi_i}$$

- initial expected ratio may be set to the same value derived from an overall weighted ("robusted") average ratio (simple CC method)

$$\widehat{\kappa}^{CC} = \sum_{i=1}^{n} \frac{\beta_{n-i+1}\pi_i}{\sum_{k=1}^{n} \beta_{n-k+1}\pi_k} \widehat{\kappa}_i = \frac{\sum_{i=1}^{n} C_{i,n-i+1}}{\sum_{i=1}^{n} \beta_{n-i+1}\pi_i}$$

Generalised Cape Cod II

- robusted value for latest (diagonal)

$$\widehat{C}_{i,n-i+1} = \widehat{\kappa}^{CC} \pi_i \beta_{n-i+1}$$

- in the CC method, the CL iteration is applied to the robusted diagonal

$$\widehat{C}_{i,n}^{CC} = C_{i,n-i+1} + (1 - \beta_{n-i+1}) \widehat{\kappa}^{CC} \pi_i$$

- modification of a BF type with modified a prior $\widehat{\kappa}^{CC} \pi_i$
- generalised CC with decay factor $0 \leq d \leq 1 \rightsquigarrow$ constant $\widehat{\kappa}^{CC}$ is replaced with $[\widehat{\kappa}_1^{CC}, \ldots, \widehat{\kappa}_n^{CC}]$ for different accident years

$$\widehat{\kappa}_i^{CC} = d \widehat{\kappa}^{CC} + (1 - d) \widehat{\kappa}_i$$

- factor of $0 \ldots$ no smoothing (CL); factor of $1 \ldots$ constant initial ratio (Simple CC)

Generalized estimating equations

- classical approaches to claims reserving problem are based on the limiting assumption that the claims in different years are independent variables
- dependencies in the development years $\rightsquigarrow$ classical techniques provide incorrect prediction
- no distributional assumptions (avoiding distribution misspecification)

GEE

- run-off triangles as one of the most typical type of actuarial data comprise correlated longitudinal data (or generally clustered data), where an accident year corresponds to a subject
- claims within "subject" should be considered as correlated by nature
- incremental claims for accident year $i \in \{1, \ldots, n\}$ create a $(n-i+1) \times 1$ vector $\mathbf{X}_i = [X_{i,1}, \ldots, X_{i,n-i+1}]^\top$ and define their expectations

$$\mathsf{E}\mathbf{X}_i = \boldsymbol{\mu}_i = [\mu_{i,1}, \ldots, \mu_{i,n+1-i}]^\top$$

Link function and linear predictor

- accident year i and development year j influence the expectation of claim amount via so-called link function g in the following manner:

$$\mu_{i,j} = g^{-1}(\mathbf{z}_{i,j}^\top \boldsymbol{\theta}),$$

where g^{-1} is referred to as the inverse of scalar link function g and $\mathbf{z}_{i,j}$ is a $p \times 1$ vector of (fictional) covariates that arranges the impact of accident and development year on the claim amount through model parameters $\boldsymbol{\theta} \in \mathbb{R}^{p\times 1}$

Example of link

- e.g., the Hoerl curve with the logarithmic link function can be "coded" by design matrix

$$\mathbf{z}_{i,j} = [1, \delta_{1,i}, \ldots, \delta_{n,i}, 1 \times \delta_{1,j}, \ldots, n \times \delta_{n,j}, \\ \delta_{1,j} \times \log 1, \ldots, \delta_{n,j} \times \log n]^\top$$

  and parameters of interest

$$\boldsymbol{\theta} = [\gamma, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n, \lambda_1, \ldots, \lambda_n]^\top,$$

  where $\delta_{i,j}$ corresponds to the Kronecker's delta
- afterwards

$$\log(\mu_{i,j}) = \gamma + \alpha_i + j\beta_j + \lambda_j \log j$$

Variance

- Besides link function g and linear predictor $\mathbf{z}_{i,j}^\top \boldsymbol{\theta}$ (i.e., mean structure μ_i), one needs to specify the variance of claim amounts
- suppose that the variance of incremental claims can be expressed as
  a known function of of their expectations

$$\text{Var} X_{i,j} = \phi h(\mu_{i,j}),$$

where $\phi > 0$ is a scale (dispersion) parameter

Working correlation matrix

- in the GEE framework, it is not necessary to know the whole distribution of the response (e.g., a distribution of the incremental claims) like in the GLM setup
- sufficient to specify the variance of $X_{i,j}$ and the working correlation matrix

$$\mathbf{R}_i(\boldsymbol{\vartheta}) \in \mathbb{R}^{(n-i+1)\times(n-i+1)}$$

for incremental claims in each accident year

Working correlation matrix II

- correlation matrix differs from accident year to accident year
- however, each correlation matrix depends only on the $s \times 1$ vector of unknown parameters ϑ, which is the same for all the accident years
- consequently, the working covariance matrix of the incremental claims is

$$\mathrm{Cov}\mathbf{X}_i = \phi \mathbf{A}_i^{1/2} \mathbf{R}_i(\vartheta) \mathbf{A}_i^{1/2},$$

  where $\mathbf{A}_i$ is an $(n-i+1) \times (n-i+1)$ diagonal matrix with $h(\mu_{i,j})$ as the jth diagonal element
- the name "working" comes from the fact that it is not expected to be correctly specified

Choice of working correlation matrix

- the simplest case is to assume uncorrelated incremental claims, i.e.,

$$\mathbf{R}_i(\boldsymbol{\vartheta}) = \mathbf{I}_{n-i+1} = \{\delta_{j,k}\}_{j,k=1}^{n-i+1,n-i+1}$$

- opposite extreme case is an unstructured correlation matrix

$$\mathbf{R}_i(\boldsymbol{\vartheta}) = \{\vartheta_{j,k}\}_{j,k=1}^{n-i+1,n-i+1}$$

such that $\vartheta_{j,j} = 1$ for $j = 1, \ldots, n-1+1$ and $\mathbf{R}_i(\boldsymbol{\vartheta})$ is positive definite

Choice of working correlation matrix II

- somewhere in between, there lies an exchangeable correlation structure

$$\mathbf{R}_i(\boldsymbol{\vartheta}) = \{\delta_{j,k} + (1 - \delta_{j,k})\vartheta\}_{j,k=1}^{n-i+1,n-i+1}, \quad \boldsymbol{\vartheta} = [\vartheta, \ldots, \vartheta]^\top$$

- an m-dependent

$$\mathbf{R}_i(\boldsymbol{\vartheta}) = \{r_{j,k}\}_{j,k=1}^{n-i+1,n-i+1}, \quad r_{j,k} = \begin{cases} 1, & j = k, \\ \vartheta_{|j-k|}, & 0 < |j-k| \leq m, \\ 0, & |j-k| > m \end{cases}$$

- an autoregresive AR(1) correlation structure

$$\mathbf{R}_i(\boldsymbol{\vartheta}) = \{\vartheta^{|j-k|}\}_{j,k=1}^{n-i+1,n-i+1}, \quad \boldsymbol{\vartheta} = [\vartheta, \ldots, \vartheta]^\top.$$

Quasi-likelihood

- parameter estimation in the GEE framework is performed in a way that the theoretical quasi-likelihood

$$Q(x;\mu) = \int \frac{x-\mu}{h(\mu)} \mathrm{d}\mu$$

  is used instead of the true log-likelihood function
- quasi-likelihood estimate in GEE setup is the solution of the score-like equation system

$$\sum_{i=1}^{n} \left[\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\theta}}\right]^\top \phi^{-1} \mathbf{A}_i^{-1/2} \mathbf{R}_i^{-1}(\boldsymbol{\vartheta}) \mathbf{A}_i^{-1/2} (\mathbf{X}_i - \boldsymbol{\mu}_i) = \mathbf{0} \in \mathbb{R}^p,$$

  where $[\partial \boldsymbol{\mu}_i / \partial \boldsymbol{\theta}]$ is a $(n-i+1) \times p$ matrix of partial derivatives of μ_i with respect to the unknown parameters θ

Further work

- claims generating process: incremental paid claims $X_{i,j}$ to be the sum of $N_{i,j}$ (independent) claims of amount $Y_{i,j}^{k}$, $k = 1, \ldots, N_{i,j}$
- Wright's model
- Tweedie compound distribution

Conclusions

- CL
- GLM
- GAM
- BF
- Bayesian framework
- CC
- GEE

References

- England, P. D. and Verrall, R. J. (2002)
  Stochastic claims reserving in general insurance (with discussion).
  British Actuarial Journal, 8 (3).
- Hudecova, S. and Pesta, M. (2012)
  Generalized estimating equations in claims reserving.
  In Komarek, A. and Nagy, S., editors, Proceedings of the 27th International Workshop on Statistical Modelling, Vol. 2, p. 555–560, Prague.
- Wüthrich, M., Merz, M. (2008)
  Stochastic claims reserving methods in insurance.
  Wiley finance series. John Wiley & Sons.

Thank you !

pesta@karlin.mff.cuni.cz