Chain ladder with random effects

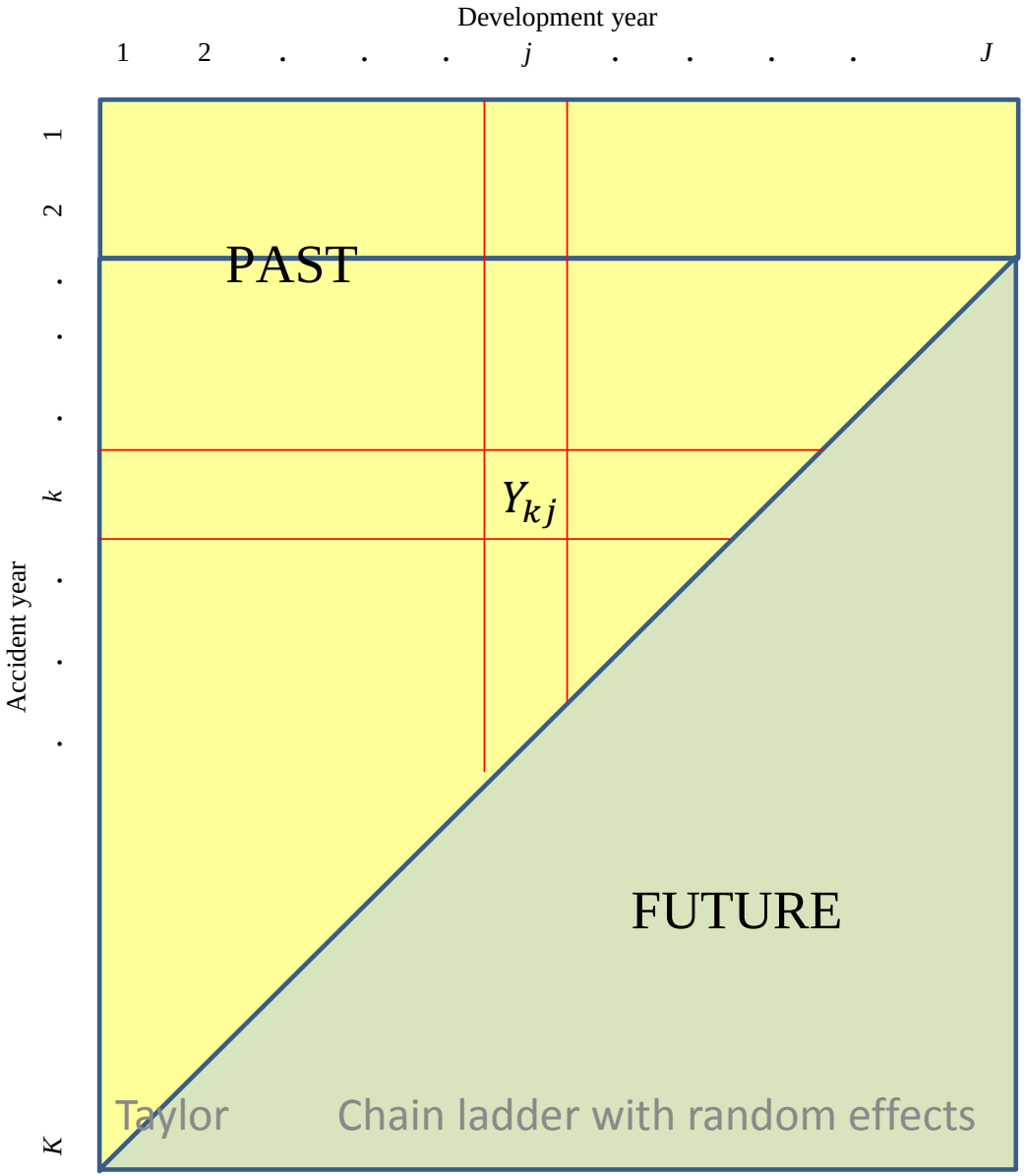
Greg Taylor
Taylor Fry Consulting Actuaries
Sydney Australia

Astin Colloquium, The Hague
21-24 May 2013

Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Framework and notation



- Claims trapezoid
  - K rows
  - J columns ($J \leq K$)
  - Entries Y_{kj}
  - Triangle if $K = J$
- Nature of Y_{kj} left unspecified
- Define cumulative row sums

$$X_{kj} = \sum_{i=1}^{j} Y_{ki}$$

Overview

- Fixed effects models
  - **Families of chain ladder models**
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Families of chain ladder models

- Two distinct families

EDF Mack models

- Accident periods stochastically independent
- For each k, the X_{kj} form a Markov chain
- For each k, j,
  - $E[X_{k,j+1}|X_{kj}] = f_j X_{kj}$
  - Distribution of $Y_{k,j+1}/X_{kj}$ drawn from **exponential dispersion family (EDF)** with dispersion parameter $\phi_{kj}(X_{kj})$

EDF cross-classified (ANOVA) models

- Y_{kj} stochastically independent
- For each k, j,
  - Distribution of Y_{kj} drawn from EDF with dispersion parameter ϕ_{kj}
  - $E[Y_{kj}] = \alpha_k \beta_j$
  - $\sum_{j=1}^{J} \beta_j = 1$

Age-to-age factor (column effect)

Row effect

Column effect

Chain ladder algorithm

- Define **estimators** of Mack age-to-age factors

$$\hat{f}_j = \frac{\sum_{k=1}^{K-j} X_{k,j+1}}{\sum_{k=1}^{K-j} X_{kj}}$$

- Define **forecasts** of future claims experience

$$\hat{X}_{kj} = \underbrace{X_{k,K-k+1}}_{\text{Most recent diagonal}} \hat{f}_{K-k+1}\, \hat{f}_{K-k+2} \cdots \hat{f}_{j-1}$$

Overview

- Fixed effects models
  - Families of chain ladder models
  - **Maximum likelihood estimators**
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Maximum likelihood estimators

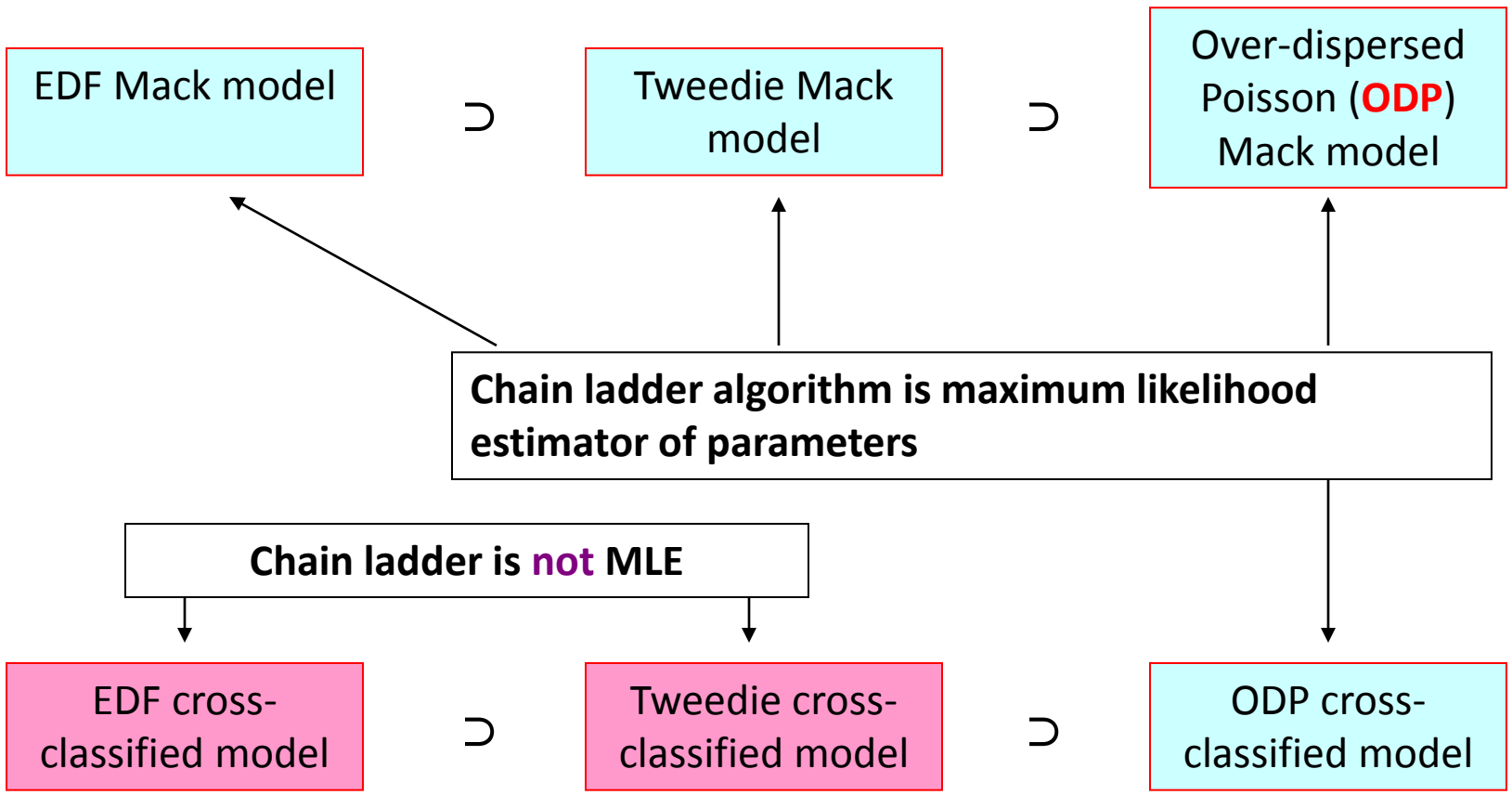
- Estimators from chain ladder algorithm generate MLE forecasts for some models subject to certain restrictions on dispersion parameters:
  - **EDF Mack model**

$$Var\left[Y_{k,j+1}|X_{kj}\right] = \sigma_{j+1}^2 X_{kj}$$

  - **EDF cross-classified model**

$$\phi_{kj} = \phi$$

Chain ladder algorithm as MLE



Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - **Bayes, linear Bayes and MAP estimators in the EDF**
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Bayes estimators in the EDF

- EDF log-likelihood

Location parameter

$$\ell(y|\theta,\phi) = [y\theta - \kappa(\theta)]/a(\phi) \ + \lambda(y,\theta)$$

Dispersion parameter

- Natural conjugate prior log-likelihood

Location parameter

$$\ell(\theta; m, \psi) = \frac{[m\theta - \kappa(\theta)]}{\psi} + \lambda^*(\psi)$$

Dispersion parameter

- Posterior log-likelihood

$$\ell(\theta|y; m, \phi, \psi) = \frac{\left[\left(\boldsymbol{y\phi^{-1} + m\psi^{-1}}\right)\left(\boldsymbol{\phi^{-1} + \psi^{-1}}\right)^{\boldsymbol{-1}}\theta - \kappa(\theta)\right]}{(\phi^{-1} + \psi^{-1})^{-1}}$$

NOTE: linear in prior location parameter m and observation y

Bayes estimators in the EDF (cont.)

- Posterior log-likelihood

$$\ell(\theta|y;m,\phi,\psi) = \frac{[(y\phi^{-1}+m\psi^{-1})(\phi^{-1}+\psi^{-1})^{-1}\theta - \kappa(\theta)]}{(\phi^{-1}+\psi^{-1})^{-1}}$$

- Bayes estimator

$$E[\mu(\theta)|y;m,\phi,\psi] = (y\phi^{-1}+m\psi^{-1})(\phi^{-1}+\psi^{-1})^{-1}$$
$$= \boldsymbol{zy + (1-z)m}$$

Mean value of observation

where

$$z = \phi^{-1}/(\phi^{-1}+\psi^{-1}) = 1/(1+\phi/\psi)$$

- Credibility result
  - Linear combination of prior mean and observation

Linear Bayes and MAP estimators in the EDF

- **Bayes estimator** is linear in observation
- Hence same as **linear Bayes estimator**
- **Maximum a posteriori (MAP) estimator** is mode of posterior distribution of observation
  - In the EDF case may be shown to be equal to the Bayes estimator

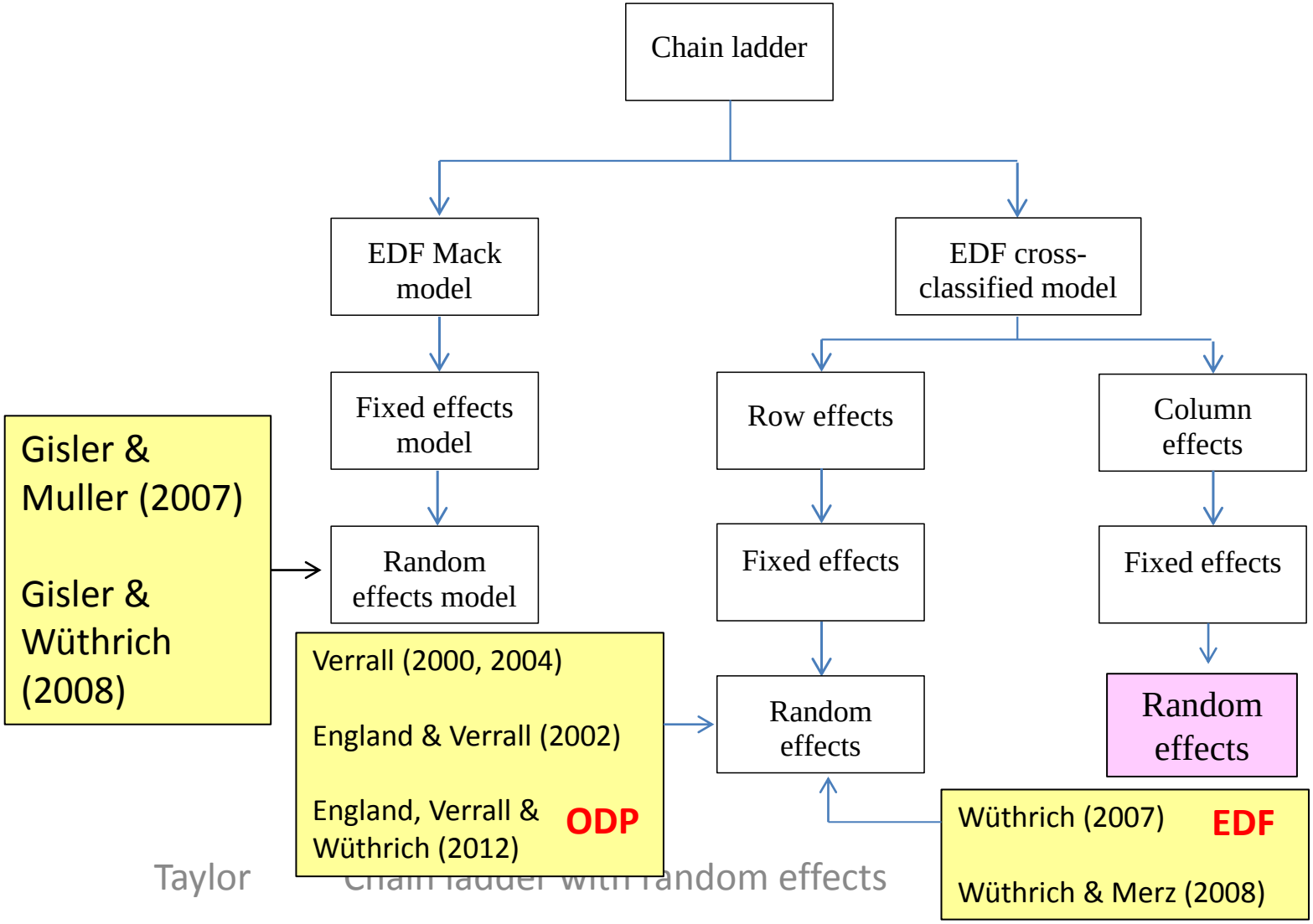
Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - **Families of chain ladder models**
  - Bayes estimators for chain ladder
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Families of random effects chain ladder models

- Previously considered models were **fixed effects models**
  - All parameters were fixed (i.e. non-stochastic) quantities
- Any fixed effects model can be converted to a **random effects model** by randomising its parameters
  - Parameters subject to a prior distribution

Literature on random effects chain ladder models



Families of chain ladder models

- Two distinct families

EDF Mack models

- Accident periods stochastically independent
- For each k, the X_{kj} form a Markov chain
- For each k, j,
  - $E[X_{k,j+1}|X_{kj}] = f_j X_{kj}$
  - Distribution of $Y_{k,j+1}/X_{kj}$ drawn from **exponential dispersion family (EDF)** with dispersion parameter $\phi_{kj}(X_{kj})$

Distribution of $g_j = f_j - 1$ from EDF, conjugate to distribution of $Y_{k,j+1}/X_{kj}$

EDF cross-classified (ANOVA) models

- Y_{kj} stochastically independent
- For each k, j,
  - Distribution of Y_{kj} drawn from EDF with dispersion parameter ϕ_{kj}
  - $E[Y_{kj}] = \alpha_k \beta_j$
  - ~~$\sum_{j=1}^{J} \beta_j = 1$~~

Distributions of $(\boldsymbol{\kappa}')^{-1}(\boldsymbol{\alpha_k})$, $(\boldsymbol{\kappa}')^{-1}(\boldsymbol{\beta_j})$ from EDF, conjugate to distribution of Y_{kj}

In each model the prior distributions of distinct parameters are stochastically independent

Differences from previous models – EDF Mack

Gisler & Muller (2007), Gisler & Wüthrich (2008)

- Used variance structure of fixed effects Mack model

$$Var\left[Y_{k,j+1}|X_{kj}\right] = \sigma_j^2 X_{kj}$$

Taylor (2013)

- More general variance structure

$$Var\left[Y_{k,j+1}|X_{kj}\right] = \textit{function of } X_{kj}$$

Differences from previous models – EDF cross-classified

Wüthrich (2007), Wüthrich & Merz (2008)

- $Y_{kj}/a_k b_j \sim EDF$ with location parameter θ_k
- θ_k a random parameter with natural conjugate distribution and location parameter 1
- Only row effects are randomised
- Tweedie sub-family
  - Location parameter = mean

$E[Y_{kj}] = (\theta_k a_k) b_j \qquad E[\theta_k] = 1$

Taylor (2013)

- $Y_{kj} \sim EDF$ with location parameter $a_k \beta_j$
- a_k, β_j random parameters, $(\kappa')^{-1}(\alpha_k)$, $(\kappa')^{-1}(\beta_j)$ each with natural conjugate distribution
- Both row and column parameters randomised
- Different distribution from Wüthrich & Merz

Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - **Bayes estimators for chain ladder**
  - Uninformative priors
  - Bornhuetter-Ferguson estimators

Bayes estimation for random effects chain ladder models

EDF Mack model

- Generic result

$$E[\mu(\theta)|y; m, \phi, \psi] = zy + (1-z)m$$

- Bayes estimator for chain ladder (recall that f_j is chain ladder age-to-age factor)

$$\tilde{f}_j = z_j \hat{f}_j + (1 - z_j)\, m_j$$

"Observation" = conventional chain ladder estimator

Credibility: depends on dispersion factors as usual

Prior mean

EDF cross-classified model

- Generic result

$$\hat{\alpha}_k = \sum_{\mathcal{R}(k)} z_{kj}^{(\alpha)} \left[Y_{kj}/\hat{\beta}_j\right] + \left[1 - \sum_{\mathcal{R}(k)} z_{kj}^{(\alpha)}\right] a_k$$

Row sum

$$\hat{\beta}_j = \sum_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)} \left[Y_{kj}/\hat{\alpha}_k\right] + \left[1 - \sum_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)}\right] b_j$$

Column sum

- These equations are implicit
- Can be solved by iteration

Abstract location parameters of α_k, β_j

Bayes EDF cross-classified model

- Generic result

$$\hat{\alpha}_k = \sum_{\mathcal{R}(k)} z_{kj}^{(\alpha)} \left[Y_{kj}/\hat{\beta}_j\right] + \left[1 - \sum_{\mathcal{R}(k)} z_{kj}^{(\alpha)}\right] a_k$$

$$\hat{\beta}_j = \sum_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)} \left[Y_{kj}/\hat{\alpha}_k\right] + \left[1 - \sum_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)}\right] b_j$$

- Credibilities $z_{kj}^{(\alpha)}, z_{kj}^{(\beta)}$ complicated functions
- Much simpler results are obtained for EDF sub-families:
  - Tweedie
  - Over-dispersed Poisson

Bayes ODP cross-classified model

- Generic result

$$\hat{\alpha}_k = \sum\nolimits_{\mathcal{R}(k)} z_{kj}^{(\alpha)} \left[Y_{kj}/\hat{\beta}_j\right] + \left[1 - \sum\nolimits_{\mathcal{R}(k)} z_{kj}^{(\alpha)}\right] a_k$$

$$\hat{\beta}_j = \sum\nolimits_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)} \left[Y_{kj}/\hat{\alpha}_k\right] + \left[1 - \sum\nolimits_{\mathcal{C}(\mathrm{j})} z_{kj}^{(\beta)}\right] b_j$$

- Consider ODP (Tweedie case is similar and omitted here)

$$\hat{\alpha}_k = z_k^{(\alpha)} \bar{Y}_k^{(\alpha)} + \left(1 - z_k^{(\alpha)}\right) a_k$$

$$\hat{\beta}_j = z_j^{(\beta)} \bar{Y}_j^{(\beta)} + \left(1 - z_j^{(\beta)}\right) b_j$$

  – Classical credibility form

Bayes ODP cross-classified model (cont.)

- ODP result

$$\hat{\alpha}_k = z_k^{(\alpha)} \bar{Y}_k^{(\alpha)} + \left(1 - z_k^{(\alpha)}\right) a_k$$
$$\hat{\beta}_j = z_j^{(\beta)} \bar{Y}_j^{(\beta)} + \left(1 - z_j^{(\beta)}\right) b_j$$

- Interpretation of terms
  - a_k, b_j now prior means of α_k, β_j

$$\bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj} \phi_{kj}^{-1} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \phi_{kj}^{-1}$$

$$z_k^{(\alpha)} = 1 \Big/ \left[1 + \left[\psi_k^{(\alpha)}\right]^{-1} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \phi_{kj}^{-1} \right]$$

  - Similarly for $\bar{Y}_k^{(\beta)}, z_k^{(\beta)}$

Weighted average of terms $Y_{kj}/\hat{\beta}_j$ (which estimate α_k)

Dispersion parameter of Y_{kj}

Dispersion parameter of α_k

Bayes ODP cross-classified model – further simplification

$$\bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj}\phi_{kj}^{-1} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \phi_{kj}^{-1}$$

$$z_k^{(\alpha)} = 1 \Big/ \left[1 + \left[\psi_k^{(\alpha)}\right]^{-1} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \phi_{kj}^{-1} \right]$$

- Assume $\phi_{kj} = \phi$ (which justifies chain ladder)
- Then

$$\bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j$$

Simplest possible estimate of α_k

$$z_k^{(\alpha)} = 1 \Big/ \left[1 + \left[\phi/\psi_k^{(\alpha)}\right] \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \right]$$

Familiar credibility ratio of dispersions

Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - **Uninformative priors**
  - Bornhuetter-Ferguson estimators

Uninformative priors

- Consider the previous case of ODP cross-classified model with $\phi_{kj} = \phi$ (chain ladder)
  - Could have considered others

Informative priors

$$z_k^{(\alpha)} = 1 \Big/ \left[1 + \left[\phi / \psi_k^{(\alpha)} \right] \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \right]$$

$$\bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j$$

$$\hat{\alpha}_k = z_k^{(\alpha)} \bar{Y}_k^{(\alpha)} + \left(1 - z_k^{(\alpha)}\right) a_k$$

$$\hat{\beta}_j = z_j^{(\beta)} \bar{Y}_j^{(\beta)} + \left(1 - z_j^{(\beta)}\right) b_j$$

Uninformative priors

Let $\psi_k^{(\alpha)}, \psi_j^{(\beta)} \to \infty$

$$z_k^{(\alpha)}, z_j^{(\beta)} \to 1$$

Total reliance on data

$$\hat{\alpha}_k = \bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j$$

$$\hat{\beta}_j = \sum_{\mathcal{C}(j)} Y_{kj} \Big/ \sum_{\mathcal{C}(j)} \hat{\alpha}_k$$

Chain ladder estimators for ODP cross-classified model

MAP = MLE

Overview

- Fixed effects models
  - Families of chain ladder models
  - Maximum likelihood estimators
- Random effects models
  - Bayes, linear Bayes and MAP estimators in the EDF
  - Families of chain ladder models
  - Bayes estimators for chain ladder
  - Uninformative priors
  - **Bornhuetter-Ferguson estimators**

Bornhuetter-Ferguson estimators

- Still consider ODP cross-classified model with $\phi_{kj} = \phi$ (chain ladder)
  - But results hold for EDF cross-classified model

General results

$$z_k^{(\alpha)} = 1 \Big/ \left[1 + \left[\phi / \psi_k^{(\alpha)} \right] \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j \right]$$

$$\bar{Y}_k^{(\alpha)} = \sum_{\mathcal{R}(k)} Y_{kj} \Big/ \sum_{\mathcal{R}(k)} \hat{\beta}_j$$

$$\hat{\alpha}_k = z_k^{(\alpha)} \bar{Y}_k^{(\alpha)} + \left(1 - z_k^{(\alpha)}\right) a_k$$

$$\hat{\beta}_j = z_j^{(\beta)} \bar{Y}_j^{(\beta)} + \left(1 - z_j^{(\beta)}\right) b_j$$

Particular results

Forecast

$$\hat{Y}_{kj} = \hat{\alpha}_k \hat{\beta}_j = \left\{ z_k^{(\alpha)} \bar{Y}_k^{(\alpha)} + \left(1 - z_k^{(\alpha)}\right) a_k \right\} \hat{\beta}_j$$

Special cases:

$\boldsymbol{\psi_k^{(\alpha)} \to \infty}$ **(uninformative prior)** $\Rightarrow z_k^{(\alpha)} = 1$

$\hat{Y}_{kj} = \bar{Y}_k^{(\alpha)} \hat{\beta}$ (simple chain ladder forecast)

$\boldsymbol{\psi_k^{(\alpha)} \to 0}$ **($\boldsymbol{\alpha_k = a_k}$ certainly)** $\Rightarrow z_k^{(\alpha)} = 0$

$\hat{Y}_{kj} = a_k \hat{\beta}$ (Bornhuetter-Ferguson forecast)

Random effects forecast is intermediate

Conclusions

- General results
  - Fixed parameters in chain ladder models replaced by random parameters
  - Gives credibility estimator which reduces to conventional chain ladder estimator in the case of uninformative priors
- EDF Mack model
  - Results previously obtained by Gisler & Wüthrich (2008) for specific variance structure
  - Variance structure generalised here
- EDF cross-classified model
  - Previous literature introduced randomisation of row effects
  - Randomisation of column effects introduced here
  - Also different prior distribution
    - Giving choice of prior

References (1)

- England P D & Verrall R J (2002). Stochastic claims reserving in general insurance. **British Actuarial Journal**, 8(iii), 443-518.
- England P D, Verrall R J & Wüthrich M V (2012). Bayesian over-dispersed Poisson model and the Bornhuetter & Ferguson claims reserving method. **Annals of Actuarial Science**, 6(2), 258-281.
- Gisler A & Müller P (2007). Credibility for additive and multiplicative models. Paper presented to the 37th ASTIN Colloquium, Orlando FL, USA. See http://www.actuaries.org/ASTIN/Colloquia/Orlando/Presentations/Gisler2.pdf
- Gisler A & Wüthrich M V (2008). Credibility for the chain ladder reserving method. **Astin Bulletin**, 38(2), 565-597.

References (2)

- Verrall RJ (2000). An investigation into stochastic claims reserving models and the chain-ladder technique. **Insurance: mathematics and economics**, 26(1), 91-99.
- Verrall R J (2004). A Bayesian generalised linear model for the Bornhuetter-Ferguson method of claims reserving. **North American Actuarial Journal**, 8(3), 67-89.
- Wüthrich M V (2007). Using a Bayesian approach for claims reserving. **Variance**, 1(2), 292-301.
- Wüthrich M V & Merz M (2008). **Stochastic claims reserving methods in insurance**, John Wiley & Sons Ltd, Chichester UK.