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 $\phi_{kj}$
  - $E[Y_{kj}] = \alpha_k \beta_j$
  - $\sum_{j=1}^{J} \beta_j = 1$

**Age-to-age factor** (column effect)
**Row effect**
**Column effect**
Chain ladder algorithm

- Define **estimators** of Mack age-to-age factors
  \[
  \hat{f}_j = \frac{\sum_{k=1}^{K-j} X_{k,j+1}}{\sum_{k=1}^{K-j} X_{k,j}}
  \]

- Define **forecasts** of future claims experience
  \[
  \hat{X}_{kj} = X_{k,K-k+1} \hat{f}_{K-k+1} \hat{f}_{K-k+2} \ldots \hat{f}_{j-1}
  \]
  Most recent diagonal
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
    \[
    \text{Var}[Y_{k,j+1} | X_{kj}] = \sigma_{j+1}^2 X_{kj}
    \]
  – EDF cross-classified model
    \[
    \phi_{kj} = \phi
    \]
Chain ladder algorithm as MLE

- EDF Mack model
- Tweedie Mack model
- Over-dispersed Poisson (ODP) Mack model

Chain ladder algorithm is maximum likelihood estimator of parameters

Chain ladder is not MLE

- EDF cross-classified model
- Tweedie cross-classified model
- ODP cross-classified model
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**
  \[ \ell(y|\theta, \phi) = \frac{y\theta - \kappa(\theta)}{a(\phi)} + \lambda(y, \theta) \]

- **Natural conjugate prior log-likelihood**
  \[ \ell(\theta; m, \psi) = \frac{m\theta - \kappa(\theta)}{\psi} + \lambda^*(\psi) \]

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

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

- **Credibility result**
  - Linear combination of prior mean and observation

Mean value of observation

Taylor    Chain ladder with random effects
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

Chain ladder

EDF Mack model

Fixed effects model

Random effects model

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

England, Verrall & Wüthrich (2012)

Random effects

Row effects

Fixed effects

Column effects

Fixed effects

Random effects

Wüthrich (2007)
Wüthrich & Merz (2008)

ODP
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 \( \phi_{kj} \)

EDF cross-classified (ANOVA) models
• \( Y_{kj} \) stochastically independent
• For each \( k, j \),
  - Distribution of \( Y_{kj} \) drawn from EDF with dispersion parameter \( \phi_{kj} \)
  - \( E[Y_{kj}] = \alpha_k \beta_j \)
  - \( \sum_{j=1}^J \beta_j = 1 \)

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

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

Taylor (2013)

- More general variance structure

\[ \text{Var}[Y_{k,j+1}|X_{kj}] = \text{function of } X_{kj} \]
Differences from previous models – EDF cross-classified


• \(Y_{kj}/a_kb_j \sim EDF\) with location parameter \(\theta_k\)
• \(\theta_k\) a random parameter with natural conjugate distribution and location parameter 1
• Only row effects are randomised
• Tweedie sub-family
  – Location parameter = mean
  \[E[Y_{kj}] = (\theta_k a_k)b_j \quad E[\theta_k] = 1\]

Taylor (2013)

• \(Y_{kj} \sim EDF\) with location parameter \(a_k\beta_j\)
• \(a_k, \beta_j\) 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

Prior mean

Credibility: depends on dispersion factors as usual

**EDF cross-classified model**

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

  Row sum

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

  Column sum

- These equations are implicit
- Can be solved by iteration

Taylor

Chain ladder with random effects
Bayes EDF cross-classified model

• Generic result

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

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

• 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

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

• 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 \( \alpha_k, \beta_j \)

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

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

– Similarly for \( \bar{Y}_k^{(\beta)}, z_k^{(\beta)} \)
Bayes ODP cross-classified model – further simplification

\[ \bar{Y}_k^{(\alpha)} = \sum_{R(k)} Y_{kj} \phi_{kj}^{-1} \bigg/ \sum_{R(k)} \hat{\beta}_j \phi_{kj}^{-1} \]

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

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

\[ \bar{Y}_k^{(\alpha)} = \sum_{R(k)} Y_{kj} \bigg/ \sum_{R(k)} \hat{\beta}_j \]

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

Taylor Chain ladder with random effects 25
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

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

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

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

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

Uninformative priors

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

$$Z_k^{(\alpha)}, Z_j^{(\beta)} \to 1$$

Total reliance on data

Chain ladder estimators for ODP cross-classified model

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

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

Taylor Chain ladder with random effects

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**

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

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

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

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

**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:

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

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

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

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

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)

References (2)


