Biennial Convention 2007
Adventures in Risk
23-26 September 2007 • Christchurch, New Zealand
Institute of Actuaries of Australia

Recent stochastic developments of the chain ladder

Greg Taylor

Overview

- Chain ladder
  - – Theoretical basis?
  - – Theoretical basis for extension to Bornhuetter-Ferguson?
  - – Extension to allow for diversification benefit

Overview

- No new material here
- All material drawn from the literature
  - Generally widely known among academic actuaries
  - Not so well known among practitioners

Overview

- **Chain ladder**
  - **Theoretical basis?**
  - Theoretical basis for extension to Bornhuetter-Ferguson?
  - Extension to allow for diversification benefit

Chain ladder – theoretical justification

- Notation
  - i = accident period
  - j = development period
  - C_{ij} = claims experience in (i,j) cell
    - Can be counts, claim payments, incurred costs, anything
  - $S_{ij} = \sum_{k=1}^{j} C_{ik}$ = cumulative claims experience

Chain ladder – theoretical justification (cont’d)

- Chain ladder based on age-to-age factors $f_j = S_{i,j+1} / S_{ij}$
- Strongly heuristic device
- **BUT** does it have a theoretical basis?
  - If so, when?
  - Are there occasions when it is **not** theoretically justified?

Chain ladder – theoretical justification (cont'd)

- Original justification given by Hachemeister & Stanard (1975)
- They assumed that
  - C_{ij} ~ Poisson($\alpha_i\beta_j$) for parameters α_i, β_j
  - All C_{ij} are stochastically independent
- Then showed that standard chain ladder algorithm yields the **maximum likelihood predictor** of future C_{ij}

Chain ladder – theoretical justification (cont'd)

- Hachemeister & Stanard's result quoted in my 1986 book (Taylor, 1986)
- Nonetheless languished for many years
- Eventually re-discovered by Renshaw & Verrall (1998)
- Extended by England & Verrall (2002)

Chain ladder – theoretical justification (cont'd)

- Extended by England & Verrall (2002)
- They work with **over-dispersed Poisson (ODP)** distribution
  - Also called **quasi-Poisson**
  - C ~ ODP(μ,φ) means that

$$C/\varphi \sim \text{Poisson}(\mu/\varphi)$$

N.B. $E[C] = \mu$, $Var[C] = \varphi\mu$, $CoV[C] = (\varphi/\mu)^{½}$

Chain ladder – theoretical justification (cont'd)

Hachemeister & Stanard

- Assumed that
  - C_{ij} ~ Poisson($\alpha_i\beta_j$) for parameters α_i, β_j
  - All C_{ij} are stochastically independent

Chain ladder – theoretical justification (cont'd)

Hachemeister & Stanard

- Assumed that
  - C_{ij} ~ Poisson($\alpha_i\beta_j$) for parameters α_i, β_j
  - All C_{ij} are stochastically independent

England & Verrall

- Assumed that
  - C_{ij} ~ ODP($\alpha_i\beta_j$,φ) for parameters α_i, β_j , φ
  - All C_{ij} are stochastically independent

Chain ladder – theoretical justification (cont'd)

Hachemeister & Stanard

- Assumed that
  - $C_{ij} \sim \text{Poisson}(\alpha_i \beta_j)$ for parameters α_i, β_j
  - All C_{ij} are stochastically independent

England & Verrall

- Assumed that
  - $C_{ij} \sim \text{ODP}(\alpha_i \beta_j, \varphi)$ for parameters α_i, β_j , φ
  - All C_{ij} are stochastically independent

In each case standard chain ladder algorithm yields the **maximum likelihood predictor** of future C_{ij}

Cases of unjustified chain ladder

- Hertig (1985) assumes that

$$S_{i,j+1} / S_{ij} \sim \text{logN}(\mu_j,\sigma^2_j)$$

  which implies that

$$E[S_{ij}] = \alpha_i\beta_j \text{ (as before)}$$

- This model is often referred to as the **stochastic chain ladder**
- Hertig derives an estimator of future S_{ij} as a function of quantities ln $(S_{i,j+1} / S_{ij})$
  - c.f. $S_{i,j+1} / S_{ij}$ (unlogged) for standard chain ladder
  - The estimator is ML

Cases of unjustified chain ladder (cont'd)

- Is there any consistent relation between the assumed distribution of the C_{ij} and estimators of $E[C_{ij}]$?
- Consider **maximally efficient unbiased estimators**, i.e. having minimum variance out of all unbiased estimators
- **Lehmann-Scheffé theorem** says that these must be based on the **sufficient statistic** of the parameter set to be estimated

Cases of unjustified chain ladder (cont'd)

- **Rao-Blackwell theorem** says that these must be based on the **sufficient statistic** of the parameter set to be estimated
  - What does this mean?
- A function $t(X_1,X_2,\ldots,X_n)$ of a random sample $\{X_1,X_2,\ldots,X_n\}$ from a distribution that dpends on a parameter θ is called a **sufficient statistic** for θ if the likelihood

$$L(X_1,\ldots,X_n; t(X_1,X_2,\ldots,X_n))$$

is independent of θ

- i.e. all of the information about θ contained in the whole sample $\{X_1,\ldots,X_n\}$ Is also contained in the value $t(X_1,\ldots,X_n)$

Cases of unjustified chain ladder (cont'd)

Distribution	**Sufficient statistic for mean**
ODP	Sample mean
Gamma	Sample mean
Any member of exponential dispersion family	Sample mean
Log normal	Sample mean of logged observations
Pareto	Sample mean of logged observations

General justifiability of chain ladder

- Appears to be reasonably close to MLE for “short tailed” cell distributions
  - “short tailed” if sample mean is sufficient statistic for population mean
  - Implies that cell probability density function tail converges to zero exponentially or faster
- Will be quite different from MLE for “long tailed” cell distributions

Overview

- **Chain ladder**
  - Theoretical basis?
  - **Theoretical basis for extension to Bornhuetter-Ferguson?**
  - Extension to allow for diversification benefit

Bornhuetter-Ferguson estimation

- Typical form

Estimated ultimate incurred =

Actual incurred to date

\+

Prior estimate of ultimate incurred

X

Chain ladder estimate of future incurred proportion

Bornhuetter-Ferguson estimation

- Typical form

Estimated ultimate incurred =

Actual incurred to date

+

Prior estimate of ultimate incurred

e.g. written premium X prior loss ratio

X

Chain ladder estimate of future incurred proportion

Bornhuetter-Ferguson estimation

- Typical form

Estimated ultimate incurred =

Actual incurred to date

\+

Prior estimate of ultimate incurred

e.g. written premium X prior loss ratio

X

Chain ladder estimate of future incurred proportion

Sounds Bayesian

Bayesian formulation of chain ladder

- From England & Verrall (2002)
- Assume that

$$C_{ij} \sim ODP(\alpha_i\beta_j,\varphi) \text{ with } \sum\beta_j = 1$$

Each α_i subject to prior

$$\alpha_i \sim Gamma(\gamma_i,\delta_i)$$

$$\text{pdf proportional to } \alpha^\gamma \exp -\alpha\gamma$$

$$E[\alpha_i] = \gamma_i/\delta_i$$

Bayesian formulation of chain ladder (cont'd)

$$C_{ij} \sim ODP(\alpha_i \beta_j, \varphi)$$

$$\alpha_i \sim Gamma(\gamma_i, \delta_i)$$

- Posterior-to-data distribution of a future C_{ij} has mean

$$E[C_{ij}|data] = Z_{ij} \text{ X chain ladder estimate}$$

$$+$$

$$(1- Z_{ij}) \text{ X prior estimate}$$

where

$$Z_{ij} = 1/(1 + \varphi \delta_i f_{j:\infty})$$

with $f_{j:\infty}$ denoting the true age-j-to-ultimate development factor

Bayesian formulation of chain ladder - interpretation

$E[C_{ij}|\text{data}] = Z_{ij}$ X chain ladder estimate

+

$(1- Z_{ij})$ X prior estimate

Note that

- Case $Z_{ij} = 1$ is case of accepting unmodified chain ladder forecasts
- Case $Z_{ij} = 0$ is case of forecasting on the basis of the prior estimate
  - i.e. Bornhuetter-Ferguson
- Cases $0<Z_{ij}<1$ are intermediate
  - Blend of chain ladder and Bornhuetter-Ferguson results

Bayesian formulation of chain ladder – blending coefficient

$E[C_{ij}|data] = Z_{ij}$ X chain ladder estimate

+

$(1- Z_{ij})$ X prior estimate

- **Blending coefficient $Z_{ij} = 1/(1 + \varphi\delta_i f_{j:\infty})$**
- Functions as **credibility** of chain ladder results
- Note that Z_{ij} may be re-cast:

$$Z_{ij} = 1/(1 + \varphi/\gamma_i^{-1}E[S_{ij}])$$

where

φ = measure of dispersion of C_{ij}

$\gamma_i^{-1} = CoV^2[\alpha_i]$ = measure of dispersion of α_i

Bayesian formulation of chain ladder – blending coefficient

$$Z_{ij} = 1/(1 + \varphi/\gamma_i^{-1}E[S_{ij}])$$

where

φ = measure of dispersion of C_{ij}

$\gamma_i^{-1} = \mathrm{CoV}^2[\alpha_i]$ = measure of dispersion of α_i

φ	γ_i^{-1}	Z_{ij}
$\rightarrow 0$	finite, >0	$\rightarrow 1$
$\rightarrow \infty$	finite, >0	$\rightarrow 0$
finite, >0	$\rightarrow 0$	$\rightarrow 0$
finite, >0	$\rightarrow \infty$	$\rightarrow 1$

Bayesian formulation of chain ladder – blending coefficient

$$Z_{ij} = 1/(1 + \varphi/\gamma_i^{-1}E[S_{ij}])$$

where

φ = measure of dispersion of C_{ij}

$\gamma_i^{-1} = CoV^2[\alpha_i]$ = measure of dispersion of α_i

φ	γ_i^{-1}	Z_{ij}
$\rightarrow 0$	finite, >0	$\rightarrow 1$
$\rightarrow \infty$	finite, >0	$\rightarrow 0$
finite, >0	$\rightarrow 0$	$\rightarrow 0$
finite, >0	$\rightarrow \infty$	$\rightarrow 1$

Bornhuetter -Ferguson

Overview

- **Chain ladder**
  - Theoretical basis?
  - Theoretical basis for extension to Bornhuetter-Ferguson?
  - **Extension to allow for diversification benefit**

Chain ladder – diversification benefit

- This question requires the chain ladder to be extended to multiple classes of business with possible dependencies between them
- Recent such extensions are:
  - Braun (2004)
  - Pröhl & Schmidt (2005)
  - Merz & Wüthrich (2007)
- Mention as an aside **synchronous bootstrapping** (Taylor & McGuire, 2007)
  - Not specific to chain ladder but applicable to it

Braun (2004)

Standard chain ladder

Data C_{ij} as before

$$F_{ij} = S_{i,j+1} / S_{ij}$$

$$E[F_{ij}] = f_j$$

$$Var[F_{ij}] = \sigma_j^2 / C_{ij}$$

Braun (2004) – model formulation

Standard chain ladder

Data C_{ij} as before

$F_{ij} = S_{i,j+1} / S_{ij}$

$E[F_{ij}|S_{ij}] = f_j$

$Var[F_{ij}|S_{ij}] = \sigma_j^2 / C_{ij}$

Braun's extension

Data C_{kij} (k= class of business)

[actually, Braun considers only k=1,2]

$F_{kij} = S_{ki,j+1} / S_{kij}$

$E[F_{kij}|S_{kij}] = f_{kj}$

$Var[F_{kij}|S_{kij}] = \sigma_{kj}^2 / C_{kij}$

$Cov[F_{kij},F_{mij}|S_{kij},S_{mij}]=\rho_j/[C_{kij},C_{mij}]^{½}$

Braun (2004) - results

- Braun’s extension consists of:
  - – Extension of Mack’s earlier algorithm for estimating prediction error associated with chain ladder estimate of liability
    - Including estimation of new parameters ρ_j

Pröhl & Schmidt (2005) – model formulation

- K classes (K an arbitrary natural number)

$F_{kij} = S_{ki,j+1} / S_{kij}$, k=1,…,K as before

– Best to use matrix notation in multivariate situation

$S_{ij} = [S_{1ij},\ldots,S_{Kij}]^T$

$\Delta_{ij} = \text{diag}\,[S_{1ij},\ldots,S_{Kij}]$

$F_{ij} = [F_{1ij},\ldots,F_{Kij}]^T$

$\mathcal{G}_j = \{S_{kih}\text{: } h=1,\ldots,j, \text{ all } k \text{ and } i\}$

Pröhl & Schmidt (2005) – model formulation (cont'd)

$\mathbf{S}_{ij} = [S_{1ij},\ldots, S_{Kij}]^T$

$\boldsymbol{\Delta}_{ij} = \text{diag}\,[S_{1ij},\ldots, S_{Kij}]^T$

$\mathbf{F}_{ij} = [F_{1ij},\ldots, F_{Kij}]^T$

$\mathcal{G}_j = \{S_{kih}: h=1,\ldots,j, \text{ all } k \text{ and } i\}$

Assume that

$E[\mathbf{F}_{ij} \mid \mathcal{G}_j] = \mathbf{f}_j$

$Cov[\mathbf{F}_{hj},\mathbf{F}_{ij} \mid \mathcal{G}_j] = \boldsymbol{\Delta}_{ij}^{-½}\boldsymbol{\Sigma}_j\,\boldsymbol{\Delta}_{ij}^{-½}$ if $h=i$

$= 0$ if $h \neq i$

Pröhl & Schmidt (2005) – model formulation (cont’d)

$S_{ij} = [S_{1ij},\ldots, S_{Kij}]^T$

$\Delta_{ij} = \text{diag}\ [S_{1ij},\ldots, S_{Kij}]^T$

$F_{ij} = [F_{1ij},\ldots, F_{Kij}]^T$

$\mathcal{G}_j = \{S_{kih}: h=1,\ldots,j, \text{ all } k \text{ and } i\}$

Assume that

$E[F_{ij} \mid \mathcal{G}_j] = f_j$

$Cov[F_{hj},F_{ij} \mid \mathcal{G}_j] = \Delta_{ij}^{-½}\Sigma_j \Delta_{ij}^{-½}$ if h=i

$= 0$ if h≠i

Case K=2

$$\Sigma_j = \begin{pmatrix} \sigma_{1j}^2 & \rho_j \\ \rho_j & \sigma_{2j}^2 \end{pmatrix}$$

Same as Braun

Pröhl & Schmidt (2005) – results

- Pröhl & Schmidt extend the multivariate chain ladder (MVCL) to an arbitrary number of classes
- However they:
  - Do not calculate an estimate of the associated uncertainty
  - Nor suggest estimators for covariances between classes

Merz & Wüthrich (2007)

- Adopt the Pröhl-Schmidt model
- Develop an estimator for the MVCL **mean square error of prediction**
  - Multivariate version of Mack’s MSEP algorithm
- Formulate estimates of the Pröhl-Schmidt covariance matrix Σ_j
- Result reduces to:
  - Braun for K=2
  - Mack for K=1
- Heavy going computationally
  - More convenient just to bootstrap?

References (1)

- Braun C (2004). The prediction error of the chain ladder method applied to correlated run-off triangles. **ASTIN Bulletin**, 34, 399-423.
- England P D & Verrall R J (2002). Claims reserving in general insurance. **British Actuarial Journal,** 8, 443-518.
- Hachemeister C A & Stanard J N (1975). IBNR claims count estimation with static lag functions. Paper presented to the **XIIth ASTIN Colloquium**, Portimão, Portugal.
- Hertig J (1985). A statistical approach to the IBNR-reserves in marine reinsurance. **ASTIN Bulletin**, 15, 171-183.

References (2)

- Merz M & Wüthrich M (2007). Prediction error of the multivariate chain ladder reserving method. **Annals of Actuarial Science** (to appear).
- Pröhl C & Schmidt K D (2005). Multivariate chain-ladder. **Dresdner Schriften zur Versicherungsmathematik.**
- Renshaw A E & Verrall R J (1998). A stochastic model underlying the chain-ladder technique. **British Actuarial Journal,** 4, 903-923.
- Taylor G C (1986). **Claim reserving in non-life insurance.** North-Holland, Amsterdam.
- Taylor G & McGuire G (2007). A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. **North American Actuarial Journal** (to appear).