Dependencies in Stochastic Loss Reserve Models

Glenn Meyers

ggmeyers@metrocast.net

Presentation to CAS Annual Meeting

November 6, 2016

Recent History

- Set up CAS Loss Reserve Database in 2011
  - Both upper and lower Schedule P triangles for hundreds of insurers.
  - Purpose was to enable "Aggressive Retrospective Testing" of stochastic loss reserve models.

Recent History

- Set up CAS Loss Reserve Database in 2011
  - Both upper and lower Schedule P triangles for hundreds of insurers.
  - Purpose was to enable "Aggressive Retrospective Testing" of stochastic loss reserve models.
- Published the monograph "Stochastic Loss Reserving Using Bayesian MCMC Models" in 2015
  - Used retrospective testing to identify shortcomings in two currently popular stochastic loss reserve models.
  - Proposed new models to address these shortcomings.

Unfinished Business

- Original purpose for creating the database - quantify dependencies
  - Joint project with the Australian Institute of Actuaries.
  - Project did not succeed!
  - Lesson learned - Pointless to quantify dependencies until we had a good univariate (i.e. single-line) model.
  - That was the purpose of the monograph!

Unfinished Business

- Original purpose for creating the database - quantify dependencies
  - Joint project with the Australian Institute of Actuaries.
  - Project did not succeed!
  - Lesson learned - Pointless to quantify dependencies until we had a good univariate (i.e. single-line) model.
  - That was the purpose of the monograph!
- Australian Objective - Risk margin for total loss reserve liability
  - Dependencies matter!

Unfinished Business

- Original purpose for creating the database - quantify dependencies
  - Joint project with the Australian Institute of Actuaries.
  - Project did not succeed!
  - Lesson learned - Pointless to quantify dependencies until we had a good univariate (i.e. single-line) model.
  - That was the purpose of the monograph!
- Australian Objective - Risk margin for total loss reserve liability
  - Dependencies matter!
- I will address the risk margin issue on Wednesday where I present my paper "A Cost of Capital Risk Margin Formula for Non-Life Insurance Liabilities."

Outline of Presentation

- Update Changing Settlement Rate (CSR) model for paid loss triangles
  - Reason - Risk margins deal with discounted loss reserves.

Outline of Presentation

- Update Changing Settlement Rate (CSR) model for paid loss triangles
  - Reason - Risk margins deal with discounted loss reserves.
- Propose a model to deal with dependencies between CSR models by line.
- "Compare" this model with one that assumes independence between CSR models by line.
- By "compare" I mean test to see if the differences between the models are significantly significant.

The Changing Settlement Rate (CSR) Model

- The monograph makes the case that the Mack (chain ladder) model and the bootstrap ODP are biased upward on the CAS Loss Reserve Database data.
- CSR is an attempt to correct this bias.
- Modification of CSR model in the monograph
  - Monograph version assumes constant change in settlement rate.
  - New version allows settlement rate to change.
- Notation
  - w = Accident Year
  - d = Development Year
  - X_{wd} = Cumulative Paid Loss

The Changing Settlement Rate (CSR) Model

- *logelr* ∼uniform(-1.5,0.5)
- $\alpha_1 = 0, \alpha_w \sim$ normal$(0,\sqrt{10})$ for $w = 2, \ldots, 10$.
- $\beta_{10} = 0, \beta_d \sim$ uniform(-5,5) for $d = 1, \ldots, 9$.
- $\sigma_d^2 \sim \sum_{i=d}^{10} a_i$ for $d = 1, \ldots, 10$, where $a_i \sim$ uniform(0,1)
- $\mu_{w,d} =$log$(Premium_w)+logelr + \alpha_w + \beta_d \cdot speedup_w$
- $X_{w,d} \sim$ lognormal$(\mu_{w,d}, \sigma_d)$

Features of the CSR Model

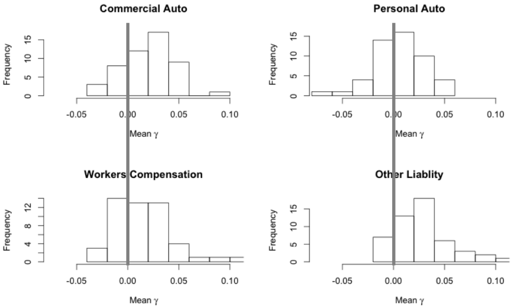
- $\mu_{w,d} = \log(Premium_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
  - The α_w parameter allows the expected loss ratio to change by accident year.
  - The $\beta_d \cdot speedup_w$ product (or interaction) allows the loss development factors to change by accident year.
- $speedup_1 = 1$
- $speedup_w = speedup_{w-1} \cdot (1 - \gamma - (w-2)) \cdot \delta)$
- Speedup Rate $= \gamma - (w-2) \cdot \delta$.
  - $\gamma \sim$ normal(0,0.05), $\delta \sim$ normal(0,0.01)
  - If positive, claim settlement speeds up.
  - If negative, claim settlement slows down
  - The δ parameter allows the speedup rate to change over time.

Posterior Sample of Size 10,000 with Bayesian MCMC

- For each parameter set in the sample get
  - $\{\alpha_w\}_{w=2}^{10}, \{\beta_d\}_{d=1}^{9}, \{\sigma_d\}_{d=1}^{10}, logelr, \gamma, \delta$
- Calculate $\mu_{w,10}$
- Simulate $X_{w,10}$ ~lognormal$(\mu_{w,10}, \sigma_{10})$
- Calculate $\sum_{w=1}^{10} X_{w,10}$

Result is a sample of 10,000 outcomes from the predictive distribution of total losses.

Posterior Means of γ Over All Insurers

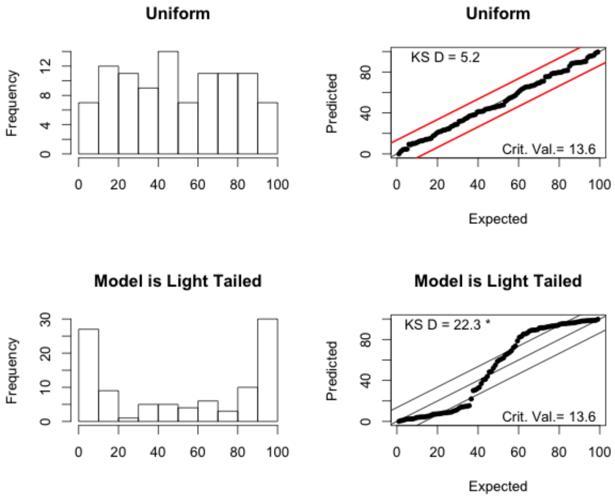


Generally, claim settlement is speeding up.

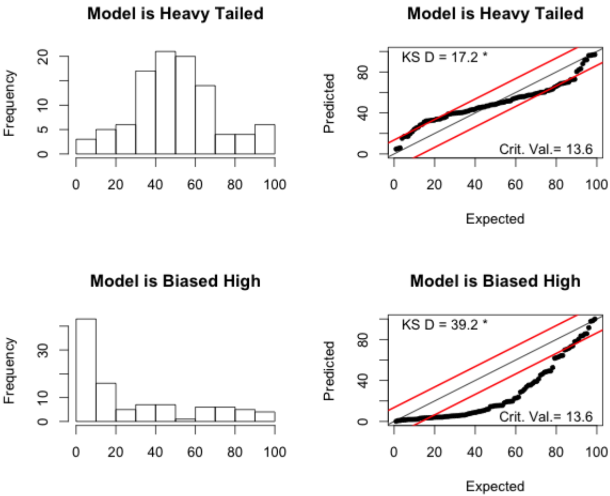
Criteria for Testing Stochastic Models

- Using the predictive distributions, find the percentiles of the outcome data for several loss triangles.
- The percentiles should be uniformly distributed.
  - Histograms
  - PP Plots and the Kolmogorov-Smirnov Test
    - Plot Expected vs Predicted Percentiles
    - KS Critical Values - 19.2 for N = 50 or 9.6 for N = 200

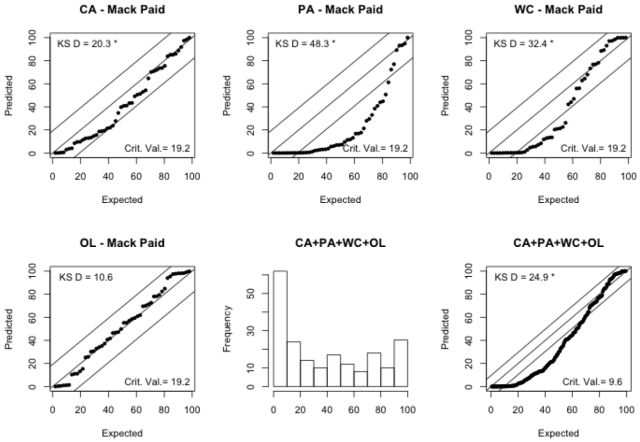
Illustrative Tests of Uniformity



Illustrative Tests of Uniformity

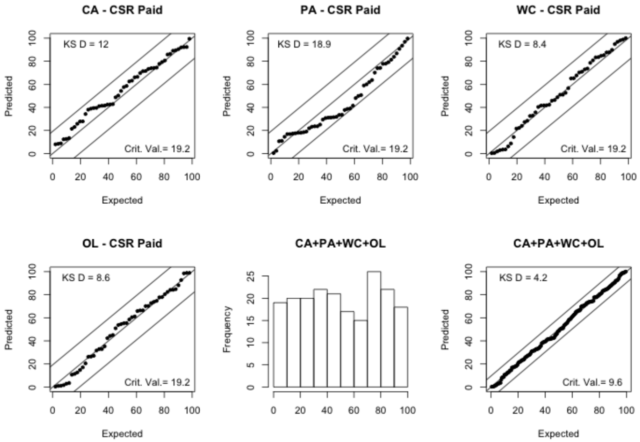


Mack Model on Paid Data



Conclusion - Mack model is biased upward.

CSR on Paid Data



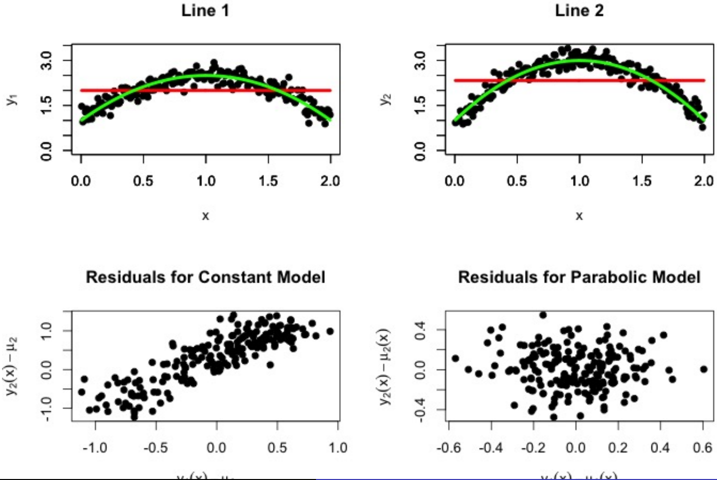
Conclusion - Validates within KS Boundaries.

Meaning of the Successful Validation

- Recall the lesson learned from the dependency project with the Australian Institute of Actuaries.
- Pointless to quantify dependencies until we had a good univariate (i.e. single-line) model.
- We have a univariate model that is suitable for the study of dependencies.

First - Beware of Artificial Correlations

- Red model - Constant — Green model - Parabolic



Dependencies - A Recent Development

- "Predicting Multivariate Insurance Loss Payments Under a Bayesian Copula Framework"
  - by Yanwei (Wayne) Zhang - FCAS and Vanja Dukic
  - Awarded the 2014 ARIA Prize by CAS

The General Idea Behind Zhang/Dukic

Given Bayesian MCMC models:

- $X_1 \sim$ Bayesian MCMC Model 1
- $X_2 \sim$ Bayesian MCMC Model 2, then:
- Fit the joint (X_1, X_2) with a joint Bayesian MCMC model.

The General Idea Behind Zhang/Dukic

Given Bayesian MCMC models:

- $X_1 \sim$ Bayesian MCMC Model 1
- $X_2 \sim$ Bayesian MCMC Model 2, then:
- Fit the joint (X_1, X_2) with a joint Bayesian MCMC model.
  - The marginal distributional model is of the same parametric form as the original models.
  - However the parameters of the univariate and marginal models may differ.

The General Idea Behind Zhang/Dukic

Given Bayesian MCMC models:

- $X_1 \sim$ Bayesian MCMC Model 1
- $X_2 \sim$ Bayesian MCMC Model 2, then:
- Fit the joint (X_1, X_2) with a joint Bayesian MCMC model.
  - The marginal distributional model is of the same parametric form as the original models.
  - However the parameters of the univariate and marginal models may differ.
- Marginal and univariate parameters were significantly different when I applied their approach with the CSR model.

The General Idea Behind Zhang/Dukic

Given Bayesian MCMC models:

- $X_1 \sim$ Bayesian MCMC Model 1
- $X_2 \sim$ Bayesian MCMC Model 2, then:
- Fit the joint (X_1, X_2) with a joint Bayesian MCMC model.
  - The marginal distributional model is of the same parametric form as the original models.
  - However the parameters of the univariate and marginal models may differ.
- Marginal and univariate parameters were significantly different when I applied their approach with the CSR model.
- I obtained better agreement between the marginal and univariate parameters with the model that Zhang/Dukic used in their paper.

Two Steps to Fitting a Bivariate Model That Preserves Univariate Fits

Joint Lognormal Distribution

$$\begin{pmatrix} \log\left(X_{wd}^{1}\right) \\ \log\left(X_{wd}^{2}\right) \end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix} \mu_{wd}^{1} \\ \mu_{wd}^{2} \end{pmatrix}, \begin{pmatrix} \left(\sigma_d^1\right)^2 & \rho \cdot \sigma_d^1 \cdot \sigma_d^2 \\ \rho \cdot \sigma_d^1 \cdot \sigma_d^2 & \left(\sigma_d^2\right)^2 \end{pmatrix}\right)$$

Two Steps to Fitting a Bivariate Model That Preserves Univariate Fits

Joint Lognormal Distribution

$$\begin{pmatrix} \log\left(X_{wd}^1\right) \\ \log\left(X_{wd}^2\right) \end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix} \mu_{wd}^1 \\ \mu_{wd}^2 \end{pmatrix}, \begin{pmatrix} \left(\sigma_d^1\right)^2 & \rho \cdot \sigma_d^1 \cdot \sigma_d^2 \\ \rho \cdot \sigma_d^1 \cdot \sigma_d^2 & \left(\sigma_d^2\right)^2 \end{pmatrix}\right)$$

1. Use Bayesian MCMC to get a sample of 10,000 μ_{wd}s and σ_ds for each line 1 and 2 (= CA, PA, WC and OL).

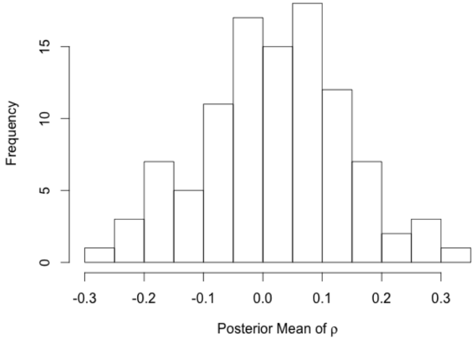
Two Steps to Fitting a Bivariate Model That Preserves Univariate Fits

Joint Lognormal Distribution

$$\begin{pmatrix} \log\left(X_{wd}^1\right) \\ \log\left(X_{wd}^2\right) \end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix} \mu_{wd}^1 \\ \mu_{wd}^2 \end{pmatrix}, \begin{pmatrix} \left(\sigma_d^1\right)^2 & \rho \cdot \sigma_d^1 \cdot \sigma_d^2 \\ \rho \cdot \sigma_d^1 \cdot \sigma_d^2 & \left(\sigma_d^2\right)^2 \end{pmatrix} \right)$$

1. Use Bayesian MCMC to get a sample of 10,000 $\mu_{wd}s$ and σ_ds for each line 1 and 2 (= CA, PA, WC and OL).
2. For each **parameter set** in the univariate sample for each line, use Bayesian MCMC to get a single ρ from the bivariate distribution of $\left(\log(X_{wd}^1), \log(X_{wd}^2)\right)$.

Posterior Mean of ρ for 102 Pairs of Triangles



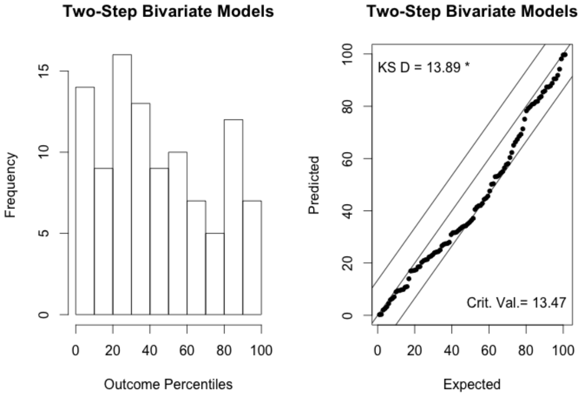
Note - $\overline{\rho}$ is fairly symmetric around 0.

Of Particular Interest - The Distribution of the Sum of Losses for Two Lines of Insurance

$$\sum_{w=1}^{10} X_{w,10}^{\{1\}} + \sum_{w=1}^{10} X_{w,10}^{\{2\}}$$

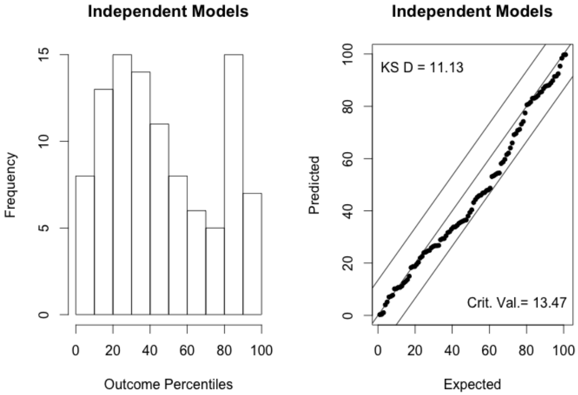
- From the 2-step bivariate model
- From the independent model formed as a random sum of losses from the univariate models

Retro Test of the Sum from the Two-Step Bivariate Model on 102 Pairs of Lines



Just outside the 95% confidence band.

Retro Test of the Sum from the Independent Model on 102 Pairs of Lines



Just inside the 95% confidence band.

Model Selection on Training Data

- If we fit model, f, by maximum likelihood define

$$AIC = 2 \cdot p - 2 \cdot L\left(x|\hat{\theta}\right)$$

- Where
  - p is the number of parameters.
  - $L\left(x|\hat{\theta}\right)$ is the maximum log-likelihood of the model specified by f.
- Lower AIC indicates a better fit.
  - Encourages larger log-likelihood
  - Penalizes for increasing the number of parameters

Bayesian Model Selection the WAIC Statistic

- Given an MCMC model with parameters $\{\theta_i\}_{i=1}^{10,000}$

$$WAIC = 2 \cdot \hat{p} - 2 \cdot \overline{\{L(x|\theta_i)\}_{i=1}^{10,000}}$$

- Where
  - $\hat{p}$ is the **effective** number of parameters.
  - $\hat{p}$ decreases as the prior distribution becomes more "informative" i.e. less influenced by the data.
  - $\overline{\{L(x|\theta_i)\}_{i=1}^{10,000}}$ = average log-likelihood.
- WAIC is calculated with the "loo" package in R.

The Leave One Out Information Criteria (LOOIC)

- Given an MCMC model with data vector, x, and parameter vectors $\{\theta_i\}_{i=1}^{10,000}$, define:

$$\text{LOOIC} = 2 \cdot \hat{p}_{LOOIC} - 2 \cdot \overline{\{L(x|\theta_i)\}_{i=1}^{10,000}}$$

- L denotes the log-likelihood of x.
- $\hat{p}_{LOOIC}$ is the **effective number of parameters**.

The Leave One Out Information Criteria (LOOIC)

- Given an MCMC model with data vector, x, and parameter vectors $\{\theta_i\}_{i=1}^{10{,}000}$, define:

$$\text{LOOIC} = 2 \cdot \hat{p}_{LOOIC} - 2 \cdot \overline{\{L(x|\theta_i)\}}_{i=1}^{10{,}000}$$

- L denotes the log-likelihood of x.
- $\hat{p}_{LOOIC}$ is the <u>**effective**</u> **number of parameters**.

$$\hat{p}_{LOOIC} = \overline{\{\{L(x|\theta_i)\}}_{i=1}^{10{,}000} - \overline{\left\{\sum_{j=1}^{J}\{L(x_j|x_{-j},\theta_i)\}\right\}}_{i=1}^{10{,}000}$$

- $x_{-j} = (x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_J)$.
- LOOIC is approximated with the "loo" package in R.

Choosing Between 2-Step and Independent Models

- The WAIC and LOOIC statistics indicate that the independent model is preferred

Choosing Between 2-Step and Independent Models

- The WAIC and LOOIC statistics indicate that the independent model is preferred

For all 102 pairs of lines!

Choosing Between 2-Step and Independent Models

- The WAIC and LOOIC statistics indicate that the independent model is preferred

For all 102 pairs of lines!

- Counterintuitive to many actuaries.
  - Inflation affects all claims simultaneously.
  - Underwriting cycle effects

Choosing Between 2-Step and Independent Models

- The WAIC and LOOIC statistics indicate that the independent model is preferred

For all 102 pairs of lines!

- Counterintuitive to many actuaries.
  - Inflation affects all claims simultaneously.
  - Underwriting cycle effects
- I think I owe an explanation.

The Changing Settlement Rate (CSR) Model

Dependencies in Stochastic Loss Reserve Models

Glenn Meyers

Background

CSR Model

Dependencies

Bivariate Models

Model Selection

SCC Model

Conclusions

- *logelr* $\sim$uniform(-5,0)
- $\alpha_1 = 0, \alpha_w \sim$ normal$(0, \sqrt{10})$ for $w = 2, \ldots, 10$.
- $\beta_{10} = 0, \beta_d \sim$ uniform(-5,5) for $d = 1, \ldots, 9$.
- $\sigma_d^2 \sim \sum_{i=d}^{10} a_i$ for $d = 1, \ldots, 10$, where $a_i \sim$ uniform(0,1)
- $\mu_{w,d} = \log(Premium_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
- $X_{w,d} \sim$ lognormal$(\mu_{w,d}, \sigma_d)$

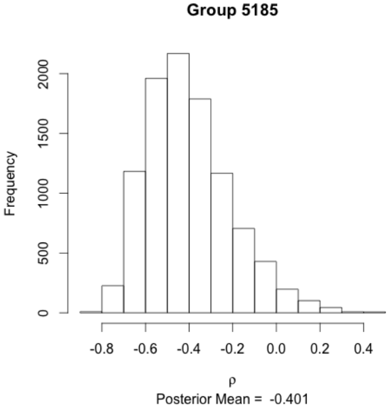
The Stochastic Cape Cod (SCC) Model

- *logelr* ∼uniform(-5,0)
- ~~$\alpha_1 = 0, \alpha_w \sim$ normal$(0,\sqrt{10})$ for $w = 2,\ldots,10$.~~
- $\beta_{10} = 0, \beta_d \sim$ uniform(-5,5) for $d = 1,\ldots,9$.
- $\sigma_d^2 \sim \sum_{i=d}^{10} a_i$ for $d = 1,\ldots,10$, where $a_i \sim$ uniform(0,1)
- $\mu_{w,d} = \log(Premium_w) + logelr$ ~~$+ \alpha_w + \beta_d \cdot speedup_w$~~
- $X_{w,d} \sim$ lognormal$(\mu_{w,d}, \sigma_d)$

The Stochastic Cape Cod (SCC) Model

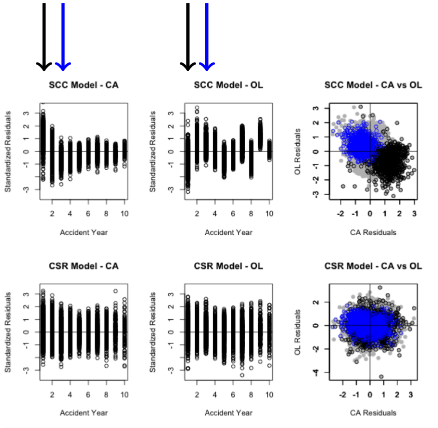
- Simpler than the CSR model
- Resembles an industry standard
  - Bornhuetter Ferguson with a constant ELR
  - Source Dave Clark and Jessica Leong in the references
- 2-Step SCC model is preferred for some insurers
- Look at a sample of standardized residual plots
- Insurer 5185 for CA and OL favors 2-Step
  - Picked as an illustration

Posterior Distribution of ρ for Insurer 5185



Note the negative posterior mean ρ.

Standardized Residual Plots for Insurer 5185



AY 1 borders are black

AY 3 borders are blue

In general, SCC residuals tend to find their own corner. If many are in the NW-SE corner, we see a negative mean ρ.

Implications of Independence

- Cost of capital risk margins should have a "diversification" credit. As an example, the EU Solvency II adds risk margins by line of business, implicitly denying a diversification credit.
- With a properly validated MCMC stochastic loss reserve model, one can get 10,000 stochastic scenarios of the future and calculate a cost of capital risk margin, and reflect diversification.

Implications of Independence

- Cost of capital risk margins should have a "diversification" credit. As an example, the EU Solvency II adds risk margins by line of business, implicitly denying a diversification credit.
- With a properly validated MCMC stochastic loss reserve model, one can get 10,000 stochastic scenarios of the future and calculate a cost of capital risk margin, and reflect diversification.
- Will address this issue on Wednesday.

A Proposed "Law" for Dependency Modeling

- Using the 2-Step procedure, we can fit bivariate distributions.
- We can compare the 2-Step model to a model that assumes independence.

A Proposed "Law" for Dependency Modeling

- Using the 2-Step procedure, we can fit bivariate distributions.
- We can compare the 2-Step model to a model that assumes independence.

The Law

- If your dependent bivariate model is "better" than the independent model, you should look for something that is missing from your model.

A Proposed "Law" for Dependency Modeling

- Using the 2-Step procedure, we can fit bivariate distributions.
- We can compare the 2-Step model to a model that assumes independence.

The Law

- If your dependent bivariate model is "better" than the independent model, you should look for something that is missing from your model.

Done!

Link to the "accepted" version of the paper