The Bayesian Revolution in Stochastic Loss Reserving

Glenn Meyers

ggmeyers@metrocast.net

Presentation to Bayesian Meetup Group
London, England

November 23, 2017

Outline of Presentation

1. Outline of Bayesian Methodology and Software
2. Stochastic Loss Reserve Models
   - Correlated Chain Ladder (CCL) Model for Incurred Losses
   - Changing Settlement Rate (CSR) Model for Paid Losses
   - CSR ∪ CSR Model for both Paid and Incurred Losses
3. Dependencies

Bayesian MCMC

- Input - Process and prior distributions of a model
- Output - A large sample (say 10,000) from the posterior distribution
- The output is the limiting distribution of a special Markov Chain.
- I recommend using specialized software to do Bayesian MCMC. The software converges very rapidly to the posterior distribution.

Bayesian MCMC

- Input - Process and prior distributions of a model
- Output - A large sample (say 10,000) from the posterior distribution
- The output is the limiting distribution of a special Markov Chain.
- I recommend using specialized software to do Bayesian MCMC. The software converges very rapidly to the posterior distribution.
- With the introduction of MCMC in 1990, the complexity of practical Bayesian models has increased by several orders of magnitude.

Evolution of MCMC Software

- WinBUGS (Original — now discontinued)
- OpenBUGS (Continuation of WinBUGS)
  - Designed mainly for the Windows operating system.
- JAGS — **J**ust **A**nother **G**ibbs **S**ampler
  - Originated by Martyn Plummer.
  - Runs on multiple operating systems.
  - Callable from R ("runjags" package.)
- Stan (in honor of Stanislaw Ulam)
  - Stan team led by Andrew Gelman at Columbia University.
  - Runs on multiple operation systems.
  - Callable from R ("rstan" package) and other languages, e.g. Python and Matlab.

The CAS Loss Reserve Database

- The National Association of Insurance Commissioners (NAIC) gave the University of Wisconsin free access to their database for academic studies.
- Jed Frees and I were involved in a joint project with the Australian Institute of Actuaries to study dependencies between lines of business. One of Jed's graduate students at the time, Peng Shi, assembled the database from the NAIC Schedule P data.
- I (with the help of the CAS) requested to make the data available on the CAS web site. It was granted by Terri Vaughan, the then president of the NAIC and CAS member. The link to the database is below.

http://www.casact.org/research/index.cfm?fa=loss_reserves_data

Schedule P Data for Incurred Losses

Table: Illustrative Insurer Net Written Premium

AY	1	2	3	4	5	6	7	8	9	10
Premium	5812	4908	5454	5165	5214	5230	4992	5466	5226	4962

Table: Illustrative Insurer Incurred Losses Net of Reinsurance

AY \ Lag	1	2	3	4	5	6	7	8	9	10	Source
1988	1722	3830	3603	3835	3873	3895	3918	3918	3917	3917	1997
1989	1581	2192	2528	2533	2528	2530	2534	2541	2538	2532	1998
1990	1834	3009	3488	4000	4105	4087	4112	4170	4271	4279	1999
1991	2305	3473	3713	4018	4295	4334	4343	4340	4342	4341	2000
1992	1832	2625	3086	3493	3521	3563	3542	3541	3541	3587	2001
1993	2289	3160	3154	3204	3190	3206	3351	3289	3267	3268	2002
1994	2881	4254	4841	5176	5551	5689	5683	5688	5684	5684	2003
1995	2489	2956	3382	3755	4148	4123	4126	4127	4128	4128	2004
1996	2541	3307	3789	3973	4031	4157	4143	4142	4144	4144	2005
1997	2203	2934	3608	3977	4040	4121	4147	4155	4183	4181	2006

Insurer 353 - Commercial Auto

Schedule P Data for Paid Losses

Table: Illustrative Insurer Net Written Premium

AY	1	2	3	4	5	6	7	8	9	10
Premium	5812	4908	5454	5165	5214	5230	4992	5466	5226	4962

Table: Illustrative Insurer Paid Losses Net of Reinsurance

AY \ Lag	1	2	3	4	5	6	7	8	9	10	Source
1988	952	1529	2813	3647	3724	3832	3899	3907	3911	3912	1997
1989	849	1564	2202	2432	2468	2487	2513	2526	2531	2527	1998
1990	983	2211	2830	3832	4039	4065	4102	4155	4268	4274	1999
1991	1657	2685	3169	3600	3900	4320	4332	4338	4341	4341	2000
1992	932	1940	2626	3332	3368	3491	3531	3540	3540	3583	2001
1993	1162	2402	2799	2996	3034	3042	3230	3238	3241	3268	2002
1994	1478	2980	3945	4714	5462	5680	5682	5683	5684	5684	2003
1995	1240	2080	2607	3080	3678	2004	4117	4125	4128	4128	1997
1996	1326	2412	3367	3843	3965	4127	4133	4141	4142	4144	2005
1997	1413	2683	3173	3674	3805	4005	4020	4095	4132	4139	2006

Insurer 353 - Commercial Auto

Made use of the database in my monograph

The url for the monograph
http://www.casact.org/pubs/monographs/index.cfm?fa=meyers-monograph01

- Analyzed 50 insurers in each of four lines of insurance.
- Backtested the Mack model with the R "ChainLadder" package on incurred data and found that it understated the variability of the lower triangle holdout data.
- Backtested the Mack and ODP Bootstrap model with the R "ChainLadder" package on paid data and found that it tended to overstate the losses in the lower triangle holdout data.
- Proposed new models that improve the performance on the lower triangle holdout data.

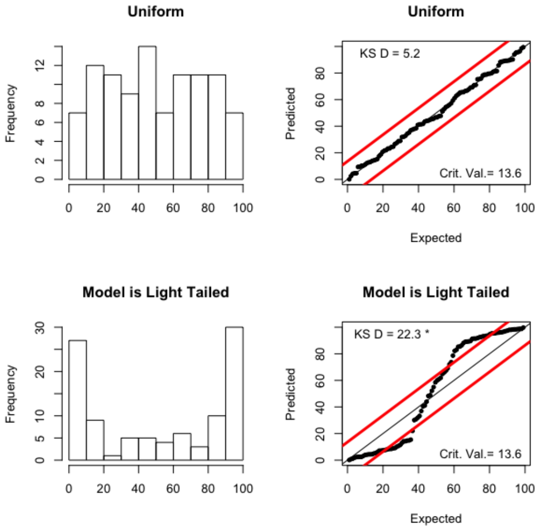
Meta-Testing Percentiles of Outcomes with the Lower Triangle Holdout Data.

- Use the model to calculate the percentile of the predictive distribution of the actual outcome.
- We should expect the percentiles of the outcomes to be uniformly distributed.
- Uniformity is testable with model fits and outcomes of several insurers.

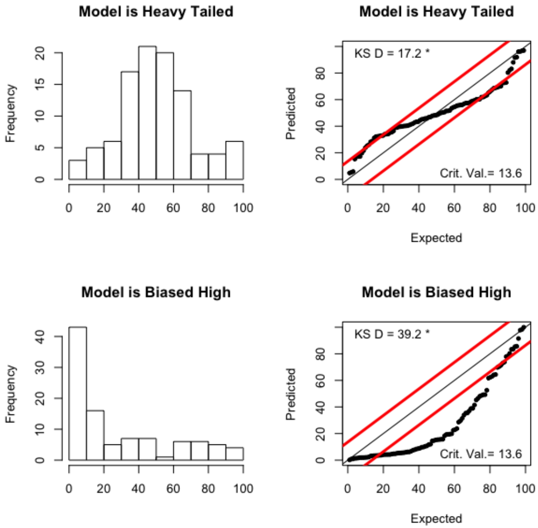
Meta-Testing Percentiles of Outcomes with the Lower Triangle Holdout Data.

- Use the model to calculate the percentile of the predictive distribution of the actual outcome.
- We should expect the percentiles of the outcomes to be uniformly distributed.
- Uniformity is testable with model fits and outcomes of several insurers.
- PP Plots - Plot the sorted values of a uniformly distributed set of numbers, Expected, against the sorted percentiles of the outcomes predicted by the model, Predicted.
  - We expect the plot to lie along a 45^o line.
- Kolmogorov-Smirnov test puts bounds around how far the difference between the predicted and expected can be.

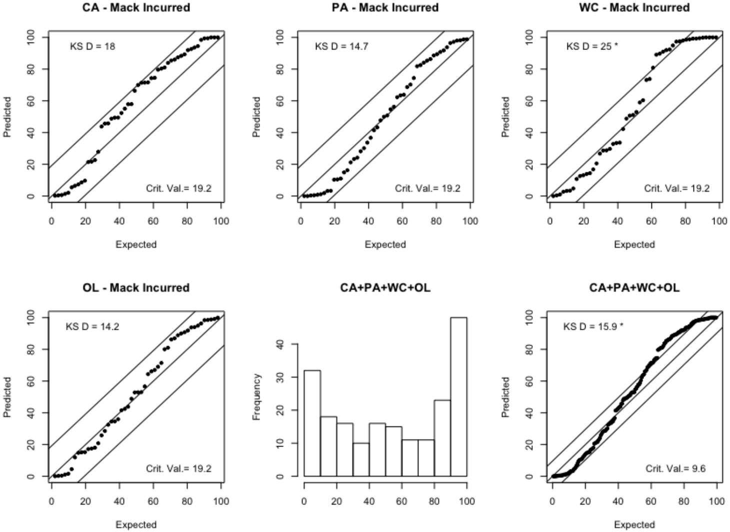
PP Plot Characteristics



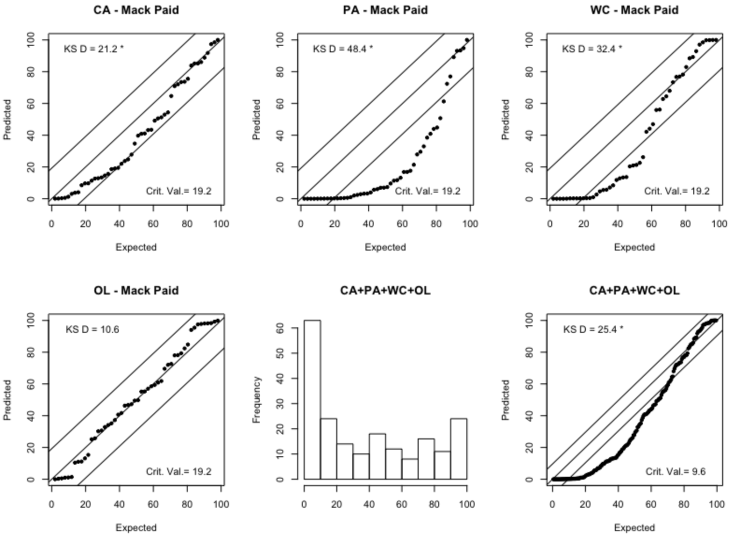
PP Plot Characteristics



Mack Model on Incurred Triangles



Mack Model on Paid Triangles



Stochastic Loss Reserving Examples

- My initial paper on the subject of stochastic loss reserving was "Estimating Predictive Distributions for Loss Reserve Models" *Variance* 2007

  http://www.variancejournal.org/issues/?fa=article&abstrID=6417
- Two features of that paper
  - It was *very primitive* Bayesian — Not MCMC.
  - It included (what I call) "aggressive backtesting."

Stochastic Loss Reserving Examples

- My initial paper on the subject of stochastic loss reserving was "Estimating Predictive Distributions for Loss Reserve Models" *Variance* 2007

  http://www.variancejournal.org/issues/?fa=article&abstrID=6417
- Two features of that paper
  - It was *very primitive* Bayesian — Not MCMC.
  - It included (what I call) "aggressive backtesting."
- Like many others (e.g. Scollnik (2001), De Alba (2002) and Verrall (2007)) - I saw Bayesian MCMC as a promising tool for stochastic loss reserving.
- But it was the availability of good data for backtesting that drew me into the game.

Models Covered in this Presentation

- A current version of the "Correlated Chain Ladder" (CCL) model on cumulative incurred triangle data.
  - Allowing for correlations between accident years can increase the variability of the predictive distribution.
- A current version of the "Changing Settlement Rate" (CSR) model on cumulative paid triangle data.
  - Allows for a gradual shift in the payout pattern to account for a change in the claim settlement rate.
- A model that simultaneously fits both the paid and incurred data with the CCL and CSR models.

The Correlated Chain Ladder (CCL) Model

- Let C^I_{wd} be the cumulative incurred loss for a 10 x 10 triangle for accident year w and development year d.
- Let P_w be the earned premium for accident year w.
- Let $\alpha_w \sim Normal(0, \sqrt{10})$ for $w = 2, \ldots, 10$. Set $\alpha_1 \equiv 0$.
- Let $\beta^I_d \sim Normal(0, \sqrt{10})$ for $d = 1, \ldots, 9$. Set $\beta^I_{10} \equiv 0$.
- Let $logelr \sim Normal(0, \sqrt{10})$.
- Let $\rho \sim \beta(2, 2)$ scaled to go between -1 and 1, where $\beta(.,.)$ denotes the β distribution.
- Let $a^I_i \sim Uniform(0, 1)$ for $d = 1, \ldots, 10$. Then set $(\sigma^I_d)^2 = \sum_{i=d}^{10} a^I_i$. This forces $\sigma^I_1 < \sigma^I_2 < \ldots < \sigma^I_{10}$.

Continuing with the CCL Model

- Set $\mu_{1d}^{l} = log(P_1) + logelr + \beta_d^{l}$ for $d = 1, \ldots, 10$.
- Set $\mu_{wd}^{l} =$
  $log(P_w) + logelr + \alpha_w + \beta_d^{l} + \rho \cdot (log(C_{w-1,d}^{l}) - \mu_{w-1,d}^{l})$
  for $w = 2, \ldots, 10$ and $d = 1, \ldots, 11 - w$.
- Then $C_{wd}^{l} \sim lognormal(\mu_{wd}^{l}, \sigma_d^{l})$.

Outline of the "CCL.R" Script

1. Read in data from the CAS Loss Reserve Database."
2. Call the stan function that outputs a sample of the parameters.
3. Turn the parameters into predicted outcomes by taking random draws from a lognormal distribution with log-mean $\mu^{I}_{w,10}$ and log-standard deviation $\sigma^{I}_{w,10}$ for $w = 1, \ldots, 10$.
4. Calculate statistics of interest, e.g. Overall Estimate, Standard Deviation, LOOIC, percentile of the outcome.

CCL Model Parameter Summary

	mean	sd
alpha[1]	0.0000	0.0000
alpha[2]	-0.2618	0.0155
alpha[3]	0.1109	0.0209
alpha[4]	0.2137	0.0248
alpha[5]	0.0095	0.0295
alpha[6]	-0.0584	0.0396
alpha[7]	0.4498	0.0530
alpha[8]	0.0265	0.0833
alpha[9]	0.1599	0.1456
alpha[10]	0.1700	0.2989
beta[1]	-0.5950	0.1176
beta[2]	-0.1892	0.0687
beta[3]	-0.1002	0.0469
beta[4]	-0.0268	0.0347
beta[5]	-0.0128	0.0307
beta[6]	-0.0014	0.0279
beta[7]	0.0015	0.0257
beta[8]	0.0052	0.0247
beta[9]	-0.0012	0.0228
beta[10]	0.0000	0.0000

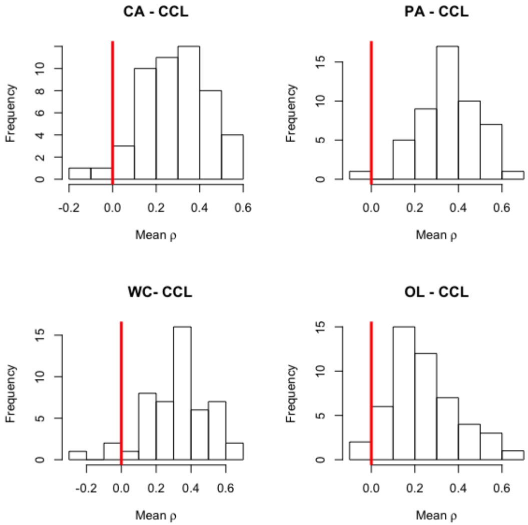
CCL Model Parameter Summary - Continued

	mean	sd
rho	0.1700	0.2068
logelr	-0.3947	0.0143
sig[1]	0.2630	0.0985
sig[2]	0.1507	0.0446
sig[3]	0.0909	0.0273
sig[4]	0.0583	0.0206
sig[5]	0.0438	0.0173
sig[6]	0.0353	0.0148
sig[7]	0.0291	0.0128
sig[8]	0.0239	0.0113
sig[9]	0.0187	0.0097
sig[10]	0.0126	0.0079

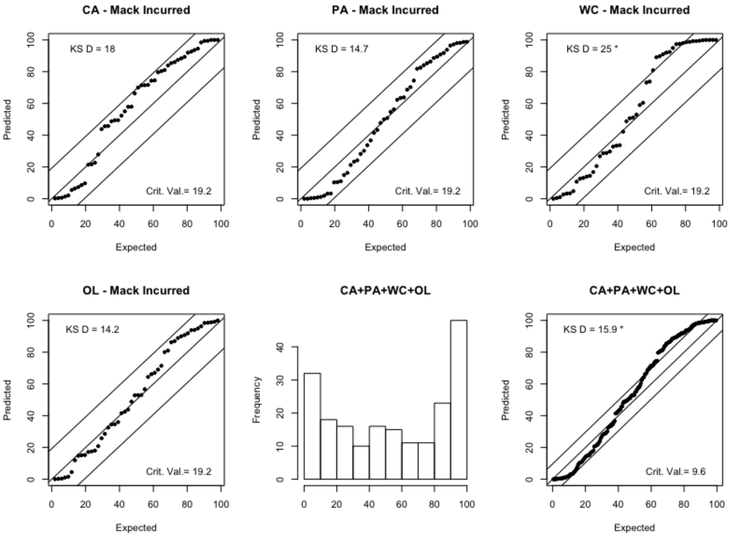
CCL Model Output

W	Premium	Estimate	S.E.	CV	Outcome	Pct
1	5812	3917	0	0	3917	
2	4908	2546	63	0.02	2532	
3	5454	4108	121	0.03	4279	
4	5165	4312	141	0.03	4341	
5	5214	3550	128	0.04	3587	
6	5230	3328	150	0.05	3268	
7	4992	5283	300	0.06	5684	
8	5466	3797	329	0.09	4128	
9	5226	4178	622	0.15	4144	
10	4962	4156	1499	0.36	4181	
Total	52429	39175	1973	0.05	40061	73.50

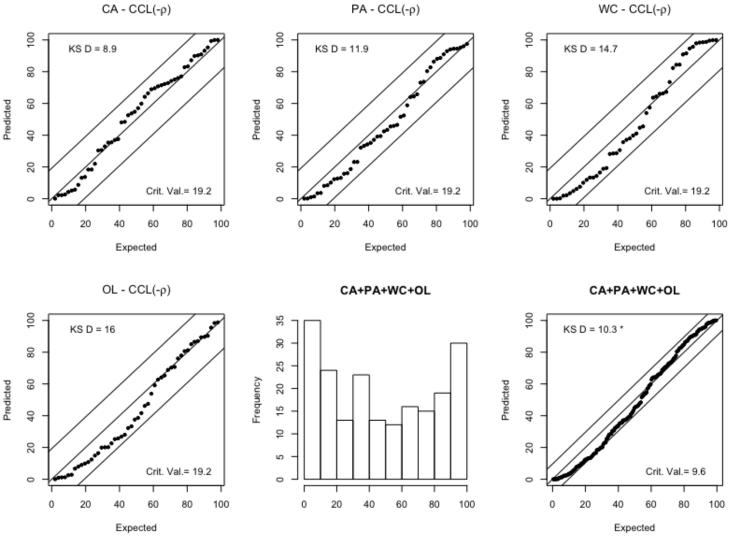
Distribution of the Mean of ρ for the CCL Model



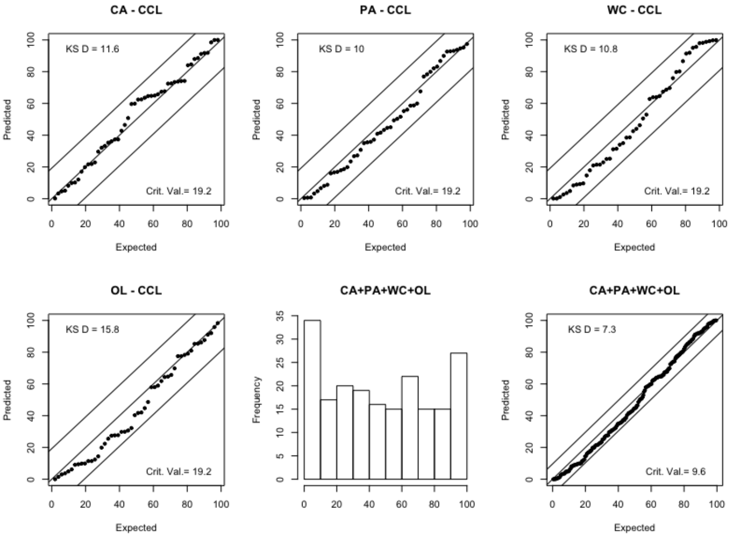
Mack Model on Incurred Triangles



CCL($-\rho$) Model on Incurred Triangles



CCL Model on Incurred Triangles



The Changing Settlement Rate (CSR) Model

- Let C^P_{wd} be the cumulative paid loss for a 10 x 10 triangle for accident year w and development year d.
- Let P_w be the earned premium for accident year w.
- Let $\alpha_w \sim Normal(0, \sqrt{10})$ for $w = 2, \ldots, 10$. Set $\alpha_1 \equiv 0$.
- Let $\beta^P_d \sim Normal(0, \sqrt{10})$ for $d = 1, \ldots, 9$. Set $\beta^P_{10} \equiv 0$.
- Let $logelr \sim Normal(0, \sqrt{10})$.
- Let $\gamma \sim Normal(0, 0.05)$.
- Let $a^P_i \sim Uniform(0, 1)$ for $d = 1, \ldots, 10$. Then set $(\sigma^P_d)^2 = \sum_{i=d}^{10} a^P_i$. This forces $\sigma^P_1 < \sigma^P_2 < \ldots < \sigma^P_{10}$.

Continuing with the CSR Model.

- Set $\mu^P_{wd} = log(P_w) + logelr + \alpha_w + \beta^p_d \cdot (1-\gamma)^{w-1}$ for $w = 1,\ldots,10$ and $d = 1,\ldots,11-w$.
- Then $C^P_{wd} \sim lognormal(\mu^P_{wd}, \sigma^P_d)$.

Continuing with the CSR Model.

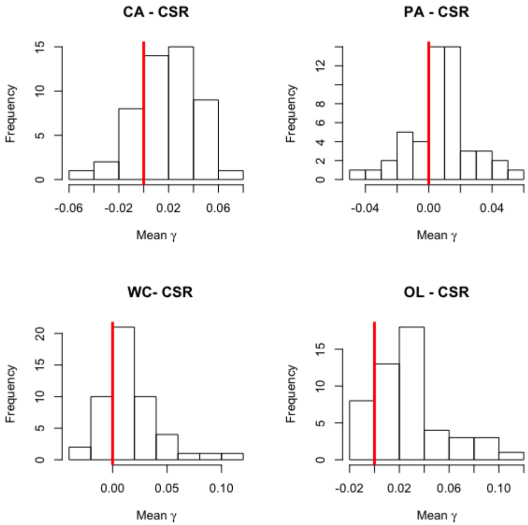
- Set $\mu_{wd}^{P} = log(P_w) + logelr + \alpha_w + \beta_d^p \cdot (1-\gamma)^{w-1}$ for $w = 1, \ldots, 10$ and $d = 1, \ldots, 11 - w$.
- Then $C_{wd}^{P} \sim lognormal(\mu_{wd}^{P}, \sigma_d^{P})$.

- A positive γ moves the development year component closer to zero with each accident year.
- Hence a speedup of claim settlement.
- If γ is negative, the claim settlement slows down.

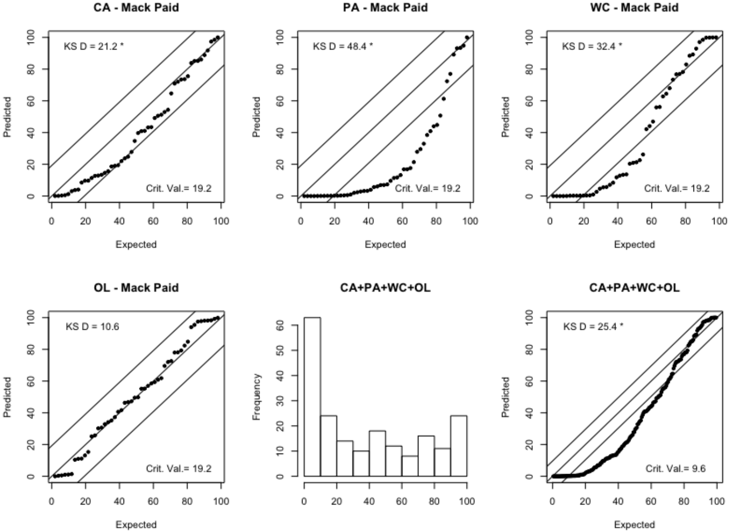
CSR Model Output

Group	W	Premium	Estimate	SE	CV	Outcome	Pct
353	1	5812	3912	0	0	3912	
353	2	4908	2562	108	0.04	2527	
353	3	5454	4132	183	0.04	4274	
353	4	5165	4284	205	0.05	4341	
353	5	5214	3512	185	0.05	3583	
353	6	5230	3320	221	0.07	3268	
353	7	4992	4964	422	0.09	5684	
353	8	5466	3318	413	0.12	4128	
353	9	5226	3761	747	0.20	4144	
353	10	4962	3798	1478	0.39	4139	
353	Total	52429	37563	2401	0.06	40000	86.66

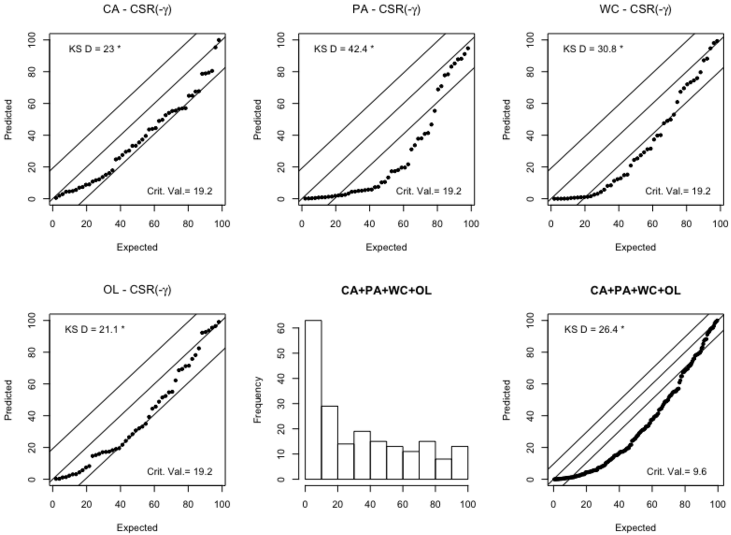
Mean γ=0.045


CA - CSR
PA - CSR
WC- CSR
OL - CSR
Frequency
Mean γ
-0.06
-0.02
0.02
0.06
-0.04
0.00
0.04
0.05
0.10
0
5
10
15
20
2
4
6
8
12

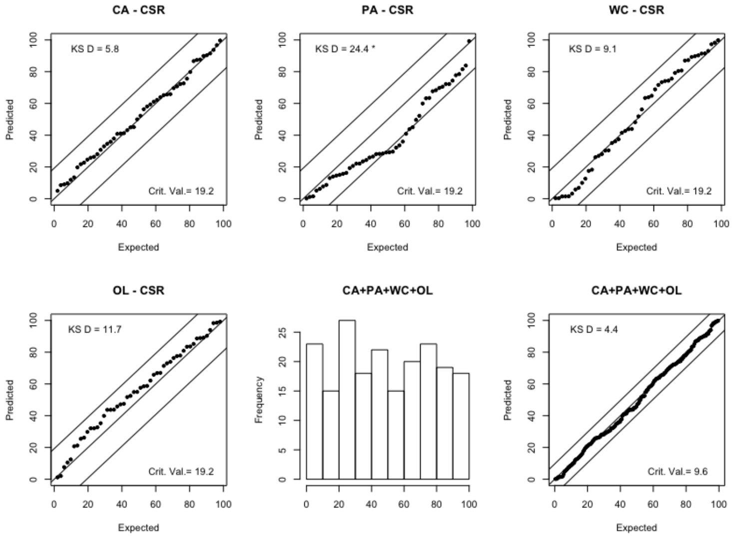
Mack Model on Paid Triangles



CSR($-\gamma$) Model on Paid Triangles



CSR Model on Paid Triangles



Some Intricacies of stan

- The model block R/stan script "StanIntro.R" reads:

```
model {
alpha ~ gamma(1,1);
beta ~ gamma(1,1);
y ~ gamma(alpha,beta);
}
```

- What these sampling statements do is increment (add to) the log-likelihood of the parameters (or data) to a hidden variable called "target."

Some Intricacies of stan

- Here is something you could do.

```
model {
alpha ~ gamma(1,1);
beta_1 ~ gamma(1,1);
beta_2 ~ gamma(1,1);
y_1 ~ gamma(alpha,beta_1);
y_2 ~ gamma(alpha,beta_2);
}
```

- The log-likelihoods of y_1 and y_2 are added successively to "target."
- What this does is fit two models to the data $\{y\_1\}$ and $\{y\_2\}$ with separate rate parameters β_1 and β_2 and common shape parameter α.

Apply this idea to the CCL and CSR Models

- Look at accident year (level) parameters
  - *logelr* - Should be the same if we have a long enough development period.
  - The paid and incurred should be close after 10 years.

$$logelr^I \sim normal(logelr^P, 0.05)$$

  - Provisionally assume that the α_w parameters are the same. Else why bother?
- The β_d, γ and ρ will differ for paid and incurred losses.

Apply this idea to the CCL and CSR Models

- Look at accident year (level) parameters
  - *logelr* - Should be the same if we have a long enough development period.
  - The paid and incurred should be close after 10 years.

$$logelr^{I} \sim normal(logelr^{P}, 0.05)$$

  - Provisionally assume that the α_w parameters are the same. Else why bother?
- The β_d, γ and ρ will differ for paid and incurred losses.
- I want to credit Ned Tyrrell, FCAS, who suggested this integrated approach to me.

IP-CCL Model Output for CA #353

W	Premium	Estimate	S.E.	CV	Outcome	Pct
1	5812	3910	0	0	3917	
2	4908	2545	43	0.02	2532	
3	5454	4126	80	0.02	4279	
4	5165	4313	94	0.02	4341	
5	5214	3546	86	0.02	3587	
6	5230	3301	101	0.03	3268	
7	4992	5189	212	0.04	5684	
8	5466	3619	231	0.06	4128	
9	5226	4023	422	0.10	4144	
10	4962	3991	790	0.20	4181	
Total	52429	38571	1214	0.03	40061	89.24

CCL Model Output for CA #353

W	Premium	Estimate	S.E.	CV	Outcome	Pct
1	5812	3917	0	0	3917	
2	4908	2546	63	0.02	2532	
3	5454	4108	121	0.03	4279	
4	5165	4312	141	0.03	4341	
5	5214	3550	128	0.04	3587	
6	5230	3328	150	0.05	3268	
7	4992	5283	300	0.06	5684	
8	5466	3797	329	0.09	4128	
9	5226	4178	622	0.15	4144	
10	4962	4156	1499	0.36	4181	
Total	52429	39175	1973	0.05	40061	73.50

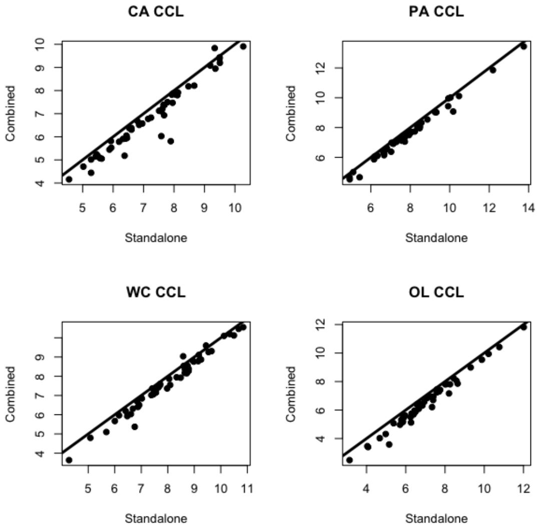
IP-CSR Model Output for CA #353

W	Premium	Estimate	SE	CV	Outcome	Pct
1	5812	3910	0	0	3912	
2	4908	2540	55	0.02	2527	
3	5454	4119	94	0.02	4274	
4	5165	4305	105	0.02	4341	
5	5214	3540	95	0.03	3583	
6	5230	3295	107	0.03	3268	
7	4992	5179	220	0.04	5684	
8	5466	3612	233	0.06	4128	
9	5226	4015	423	0.11	4144	
10	4962	3984	789	0.20	4139	
Total	52429	38500	1261	0.03	40000	88.50

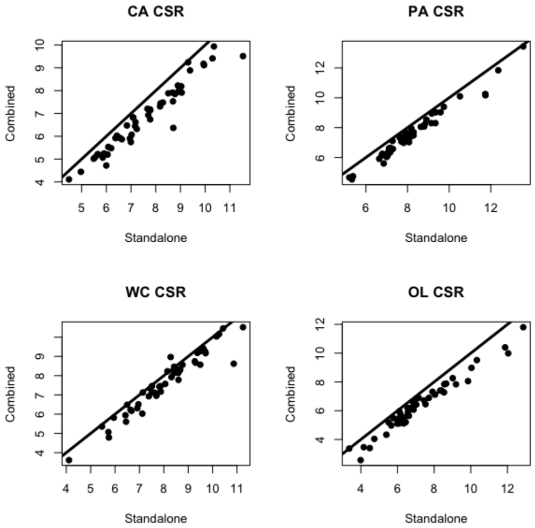
CSR Model Output for CA #353

W	Premium	Estimate	SE	CV	Outcome	Pct
1	5812	3912	0	0	3912	
2	4908	2562	108	0.04	2527	
3	5454	4132	183	0.04	4274	
4	5165	4284	205	0.05	4341	
5	5214	3512	185	0.05	3583	
6	5230	3320	221	0.07	3268	
7	4992	4964	422	0.09	5684	
8	5466	3318	413	0.12	4128	
9	5226	3761	747	0.20	4144	
10	4962	3798	1478	0.39	4139	
Total	52429	37563	2401	0.06	40000	86.66

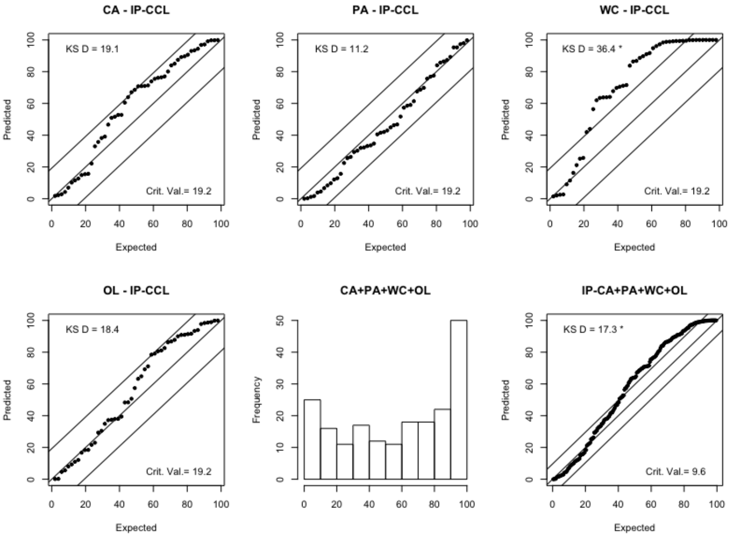
CCL log(SE) Comparisons for all 200 Triangles



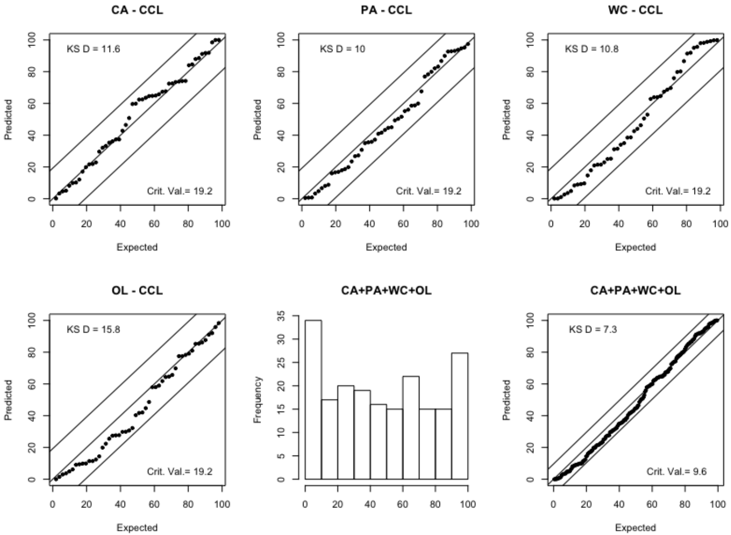
CSR log(SE) Comparisons for all 200 Triangles



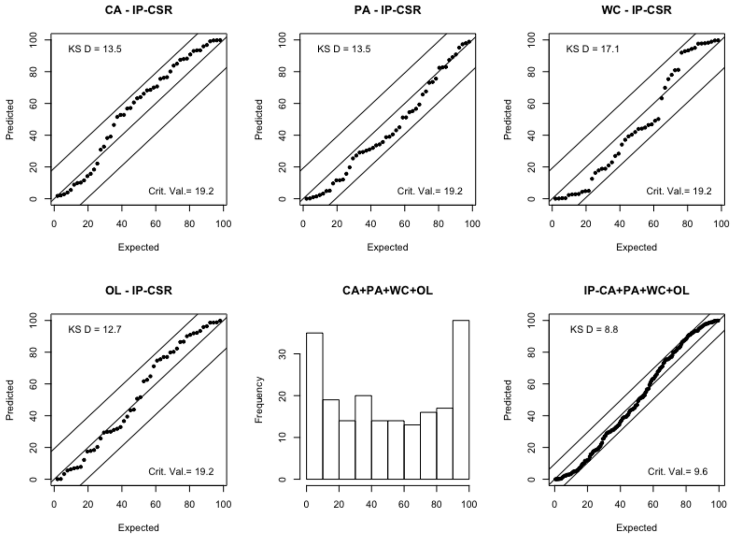
IP-CCL PP Plots



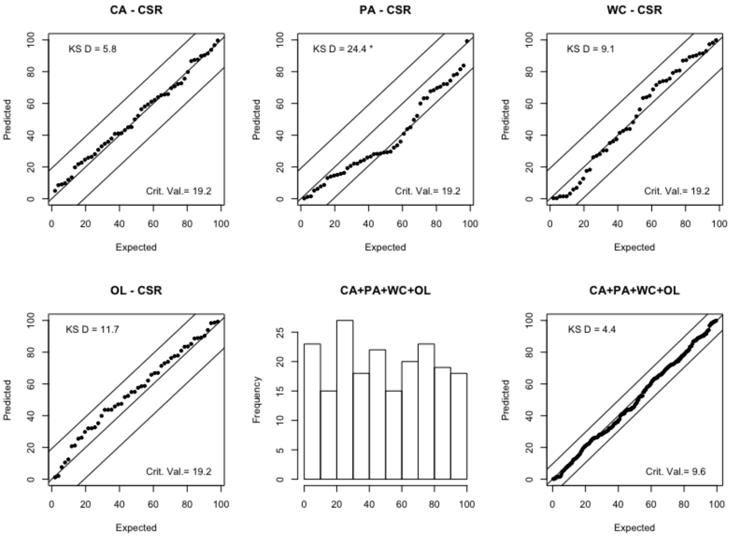
Compare with Standalone CCL PP Plots



IP-CSR PP Plots



Compare with Standalone CSR PP Plots



Comparing Combined IP and Standalone Models

- For the IP-CSR, all lines are within the KS bounds.
- For the standalone CSR, CA, WC and OL are within the KS bounds.
- For the IP-CCL, all lines except WC are within the KS bounds.
- For the standalone CCL, all lines are within the KS bounds.

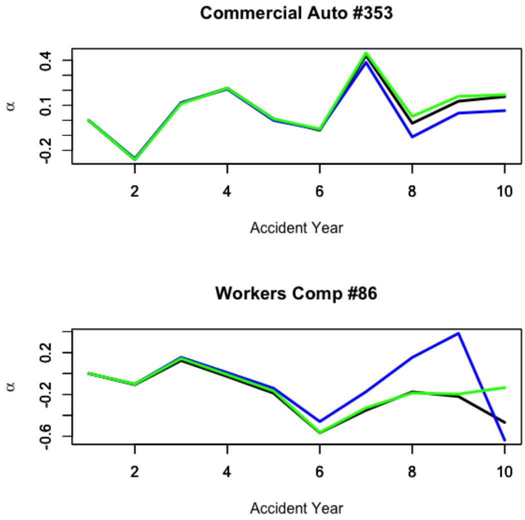
Comparing Combined IP and Standalone Models

- For the IP-CSR, all lines are within the KS bounds.
- For the standalone CSR, CA, WC and OL are within the KS bounds.
- For the IP-CCL, all lines except WC are within the KS bounds.
- For the standalone CCL, all lines are within the KS bounds.
- What is the difference between WC and the other lines?
- Run the combined model on Insurer 86 with WC.

Comparing Combined IP and Standalone Models

- For the IP-CSR, all lines are within the KS bounds.
- For the standalone CSR, CA, WC and OL are within the KS bounds.
- For the IP-CCL, all lines except WC are within the KS bounds.
- For the standalone CCL, all lines are within the KS bounds.
- What is the difference between WC and the other lines?
- Run the combined model on Insurer 86 with WC.
- Big difference between paid and incurred losses at $d = 10$ for WC. Not so for the other lines.

Common α_w parameters are not good for WC



The Quest for a Minimum Variance Estimator

- Ned Tyrrell pulled me into combined models to get more accurate, i.e. less variability in the, estimates of the loss reserve liability.
- Ned's objective appears to have been met for CA, PA and OL.
- My suggestions for using combined models

The Quest for a Minimum Variance Estimator

- Ned Tyrrell pulled me into combined models to get more accurate, i.e. less variability in the, estimates of the loss reserve liability.
- Ned's objective appears to have been met for CA, PA and OL.
- My suggestions for using combined models

1. Fit standalone CCL and CSR models.
2. Compare α_w parameters.
3. If the above parameters are sufficiently close, fit a combined CCL and CSR model.

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).

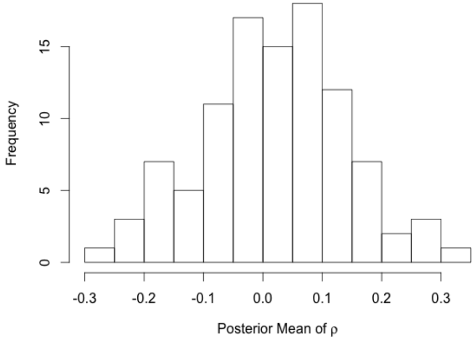
Two Steps to Fitting a Bivariate Model That Preserves Univariate Fits

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 μ_{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 $(\log(X_{wd}^1), \log(X_{wd}^2))$.

Posterior Mean of ρ for 102 Pairs of Triangles



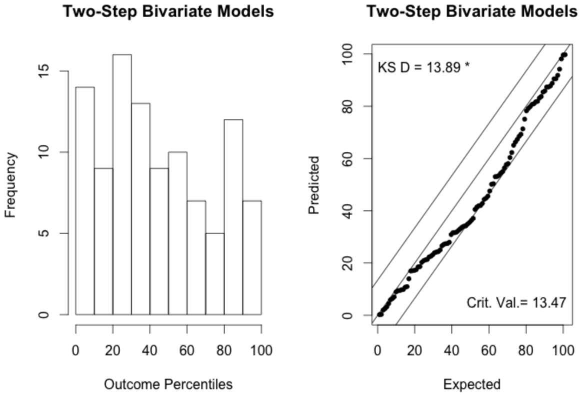
Note - $\overline{\rho}$ is fairly symmetric around 0.

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

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

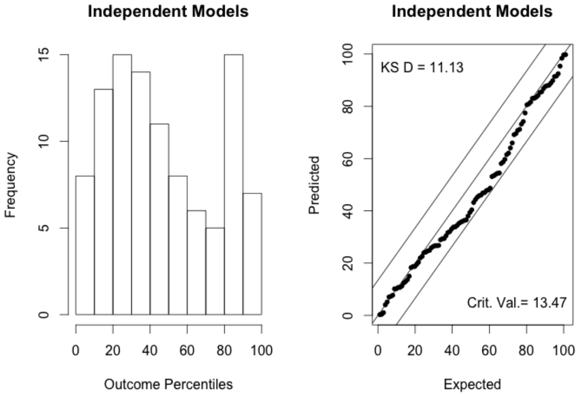
- From the 2-step bivariate model
- From the independent model formed as a random sum of losses from the univariate models

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 ***effective*** **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

Concluding Remarks

- This presentation was more focused on what Bayesian MCMC can do, and passing over many of the fine points of how to do it.

Concluding Remarks

- This presentation was more focused on what Bayesian MCMC can do, and passing over many of the fine points of how to do it.
- A common theme to all the examples was fitting a loss distribution with covariates.
- The predictive distribution is a mixture of fairly simple distributions over a large sample from the posterior distribution.

Concluding Remarks

- This presentation was more focused on what Bayesian MCMC can do, and passing over many of the fine points of how to do it.
- A common theme to all the examples was fitting a loss distribution with covariates.
- The predictive distribution is a mixture of fairly simple distributions over a large sample from the posterior distribution.
- Do not expect to master this technique quickly. It will take a lot of work.

Some References

To appear in *Variance*

- Dependencies in Stochastic Loss Reserve Models
  - R/stan scripts available in a spreadsheet
- A Cost of Capital Risk Margin Formula for Non-Life Insurance Liabilities
  - R/stan scripts available in a zip file

Some References

To appear in *Variance*

- Dependencies in Stochastic Loss Reserve Models
  - R/stan scripts available in a spreadsheet
- A Cost of Capital Risk Margin Formula for Non-Life Insurance Liabilities
  - R/stan scripts available in a zip file

Done!