Proxies

Glenn Meyers, FCAS, MAAA, Ph.D.
Chief Actuary, ISO Innovative Analytics
Presented at the ASTIN Colloquium
June 4, 2009

Objective
Estimate Loss Liabilities with Limited Data

- "The term 'proxy' is used to denote simplified methods for the valuation of technical provisions that are applied when there is only insufficient data to apply a reliable statistical actuarial method, or when there is insufficient actuarial expertise available to the insurer."
  - CEIOPS – Groupe Consultatif Coordination Group on Proxies, "Draft Interim report including testing proposals for proxies under QIS 4, " page 23, November 2007.

Concepts Deemed Appropriate

- Benchmark Data
  - Derived from "industry"
- Credibility
  - Recognize that benchmark data comes from a heterogeneous collection of insurers
  - Depending on the volume of data, the final estimate of the liability should fall between a raw estimate based on the insurer's own data and the industry average.

High Level Outline of Method

- Given "benchmark" scenarios for expected loss
  - Initially assume they are equally likely.
- Given the data, use Bayes'Theorem to assign posterior probabilities to each benchmark scenario.

$$\Pr\{\text{Scenario}|\text{Data}\} \propto \text{Likelihood}\{\text{Data}|\text{Scenario}\} \times \Pr\{\text{Scenario}\}$$

- Calculate a "statistic of interest" from the posterior probability weighted mixture of scenarios.

Insurer #1 – Very Small

Amounts in Thousands

Insurer	AY	Premium	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
1	1986	3,874	478	587	398	114	111	78	22	4	1	
	1987	4,755	702	500	458	353	150	93		12		
	1988	3,960	799	566	447	57	162		15			
	1989	3,094	802	605	308	67	22	1	1			
	1990	2,694	690	538	119	78	118	130				
	1991	1,958	526	277	289	50	17					
	1992	1,548	488	318	78	13						
	1993	1,710	540	367	114							
	1994	1,742	518	405								
	1995	1,740	564									

Statistics of interest

Let X = sum of outstanding loss from the triangle of paid losses.

1. Expected value of X
2. Standard deviation of X
3. 99th percentile of X

An Illustration of the Method

- Example based on US Commercial Auto data
- Approach can be adopted to fit other jurisdictions

The Benchmark Scenarios – Part 1

Distribution of (conditional) Expected Loss

- Parameters determining the expected loss
  - {*ELR*} = Expected Loss Ratio by accident year
  - {*Dev*} = Paid loss incremental development factors

$$\mathrm{E}\left[Loss_{AY,Lag}\right] = \mathrm{Premium}_{AY} \times ELR_{AY} \times Dev_{Lag}$$

  - Benchmark models include 5000 scenarios.
  - Prior distribution assigns equal probability to each scenario

First 10 of 5000 Scenarios

elr1	elr2	elr3	elr4	elr5	elr6	elr7	elr8	elr9	elr10
0.64238	0.51981	0.51953	0.63213	0.67768	0.67399	0.86061	0.68820	1.02861	1.06770
0.55981	0.60917	0.70393	0.95286	0.58794	0.82253	0.68061	0.44846	0.70584	0.58142
0.74129	0.80828	0.79765	0.84994	1.04280	0.87760	0.65143	1.02369	0.76180	0.70327
0.86351	0.66209	0.80791	0.77055	0.82531	0.85447	0.86690	0.56974	0.79379	0.85215
0.49830	0.55237	0.77189	0.77043	0.70321	0.69503	0.77223	0.80198	0.91289	0.98609
0.41553	0.57165	0.47195	0.59233	0.69693	0.43060	0.54119	0.75860	0.65576	0.72832
0.63947	0.62849	0.78387	0.84099	0.73037	0.65107	0.76595	1.05825	1.18335	1.07327
0.72711	0.61534	0.81855	0.69928	0.67066	0.72449	0.58961	0.70398	0.87130	0.86999
0.87818	0.82552	0.62798	0.70623	0.69164	0.69997	0.69734	0.73586	0.86176	0.61927
0.59965	0.59008	0.66362	0.62165	0.60918	0.67158	0.67992	0.74368	0.90779	0.86630

dev1	dev2	dev3	dev4	dev5	dev6	dev7	dev8	dev9	dev10
0.26065	0.27554	0.19793	0.13306	0.06310	0.03381	0.01743	0.00622	0.00616	0.00610
0.18221	0.27828	0.22338	0.14106	0.07270	0.05640	0.02265	0.00788	0.00777	0.00767
0.19116	0.24405	0.22792	0.14219	0.08522	0.04176	0.02254	0.01541	0.01505	0.01470
0.20829	0.25630	0.21802	0.15399	0.07876	0.04478	0.02079	0.00642	0.00635	0.00629
0.21311	0.27703	0.22048	0.12817	0.07705	0.02758	0.03724	0.00651	0.00645	0.00639
0.24295	0.23668	0.21317	0.13694	0.08393	0.05420	0.01981	0.00414	0.00411	0.00408
0.14528	0.21675	0.21128	0.15596	0.10748	0.06433	0.04220	0.01958	0.01889	0.01825
0.19997	0.25765	0.20300	0.13895	0.09342	0.02764	0.05433	0.00847	0.00835	0.00823
0.17924	0.23408	0.22115	0.15868	0.08765	0.04008	0.01691	0.02150	0.02072	0.02000
0.24456	0.27650	0.21135	0.13281	0.06064	0.03184	0.01715	0.00849	0.00838	0.00827

The Benchmark Scenarios – Part 2

Distribution of Outcomes Given the Expected Loss

- Collective Risk Model – For each settlement lag
  1. Select random claim count, N_{Lag}.
  2. Select random claim sizes, $Z_{AY,Lag,i}$, for i=1,… ,N_{Lag}.
  3. Set $$X_{AY,Lag} = \sum_{i=1}^{N_{Lag}} Z_{AY,Lag,i}$$

The Benchmark Scenarios – Part 2

Claim Severity Distribution

- Pareto distribution $F_{Lag}(z) = 1 - \left(\frac{\theta}{z+\theta}\right)^{\alpha}$

- Average claim severity increases with settlement lag

Settlement Lag	Expected Claim Severity	Coefficient of Variation	θ (000)	α
1	5	3.00	4.5244	1.8978
2	20	3.50	6.4195	1.2099
3	75	2.00	39.1784	1.3639
4	100	1.75	62.9849	1.4565
5	125	1.60	84.1527	1.4715
6	150	1.55	77.6372	1.2481
7	175	1.50	68.9063	1.0566
8	185	1.45	78.6878	1.0837
9	195	1.40	91.6444	1.1261
10	200	1.35	117.6357	1.2623

Likelihood{Data|Scenario}

- Use insurer premium and benchmark scenarios to get expected losses by *AY* and *Lag*

$$\text{E}\left[Loss_{AY,Lag}\right] = \text{Premium}_{AY} \times ELR_{AY} \times Dev_{Lag}$$

- Divide expected loss by expected severity to get expected count

$$\lambda_{AY,Lag} = \text{E}\left[Loss_{AY,Lag}\right] / \text{E}\left[Z_{Lag}\right]$$

- Likelihood calculated by collective risk model – Issues in flux
- Likelihood reflects correlation
- Ways to get likelihood
  - FFT (Accurate but slow)
  - Overdispersed Negative Binomial (Approximate and fast)
  - Later versions could use the Tweedie likelihood

Examples

- Real data – American Commercial Auto
  - Obtained from NAIC Schedule P
- Examples start with small insurer and increase in size.

Insurer #1 – Very Small

Insurer	AY	Premium	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
1	1986	3,874	478	587	398	114	111	78	22	4	1	
	1987	4,755	702	500	458	353	150	93		12		
	1988	3,960	799	566	447	57	162		15			
	1989	3,094	802	605	308	67	22	1	1			
	1990	2,694	690	538	119	78	118	130				
	1991	1,958	526	277	289	50	17					
	1992	1,548	488	318	78	13						
	1993	1,710	540	367	114							
	1994	1,742	518	405								
	1995	1,740	564									

Amounts in Thousands

Insurer #2 – Medium

Insurer	AY	Premium	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
2	1986	29,478	4,452	7,003	5,944	4,822	2,362	2,564	1,261	14	36	64
	1987	63,180	5,564	6,635	7,942	4,284	3,239	2,300	899	251	560	
	1988	65,904	6,034	9,246	9,348	6,131	3,000	1,768	699	670		
	1989	61,693	6,359	12,282	10,682	12,594	2,401	1,999	1,145			
	1990	62,975	7,773	9,479	13,255	6,557	3,900	1,625				
	1991	59,433	5,705	6,583	7,537	8,179	1,827					
	1992	48,175	5,045	6,762	10,313	4,463						
	1993	43,479	4,061	5,972	4,973							
	1994	34,748	3,203	4,245								
	1995	29,300	2,907									

Amounts in Thousands

Insurer #3 – Somewhat Larger

Insurer	AY	Premium	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
3	1986	48,844	5,886	6,128	4,739	5,442	1,741	1,183	678	158		303
	1987	52,622	5,632	7,487	6,630	5,195	2,465	1,149	459	372	184	
	1988	53,507	6,419	4,697	10,567	5,715	2,280	2,031	191	296		
	1989	56,949	7,300	8,939	9,495	6,966	3,960	787	1,060			
	1990	58,611	8,249	11,302	9,038	5,687	3,452	1,483				
	1991	61,692	8,074	9,454	7,913	3,455	3,485					
	1992	66,755	8,747	10,542	11,235	4,356						
	1993	81,119	10,258	15,376	11,697							
	1994	98,632	15,540	23,594								
	1995	91,311	14,289									

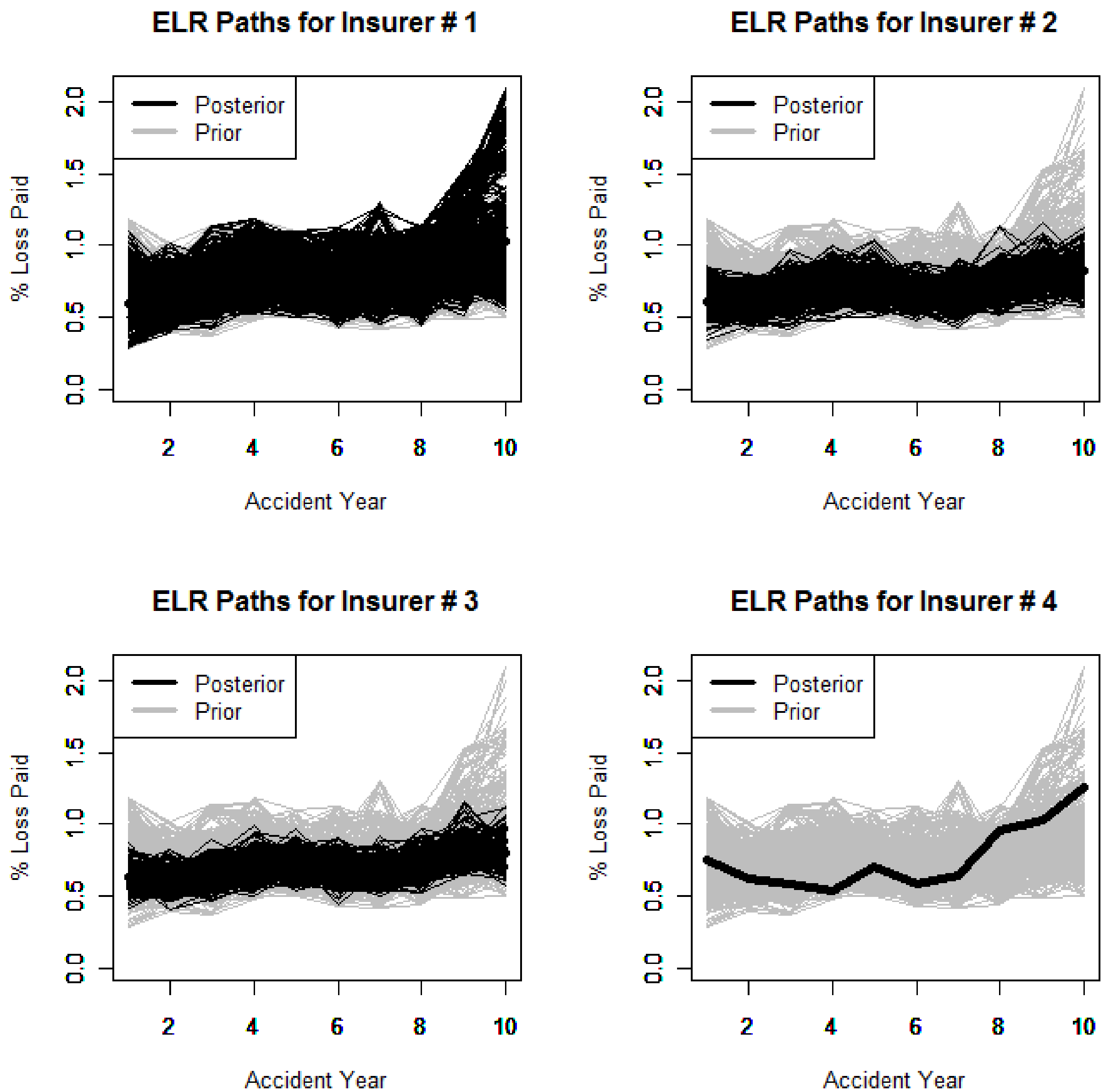
Amounts in Thousands

Insurer # 4 – Large

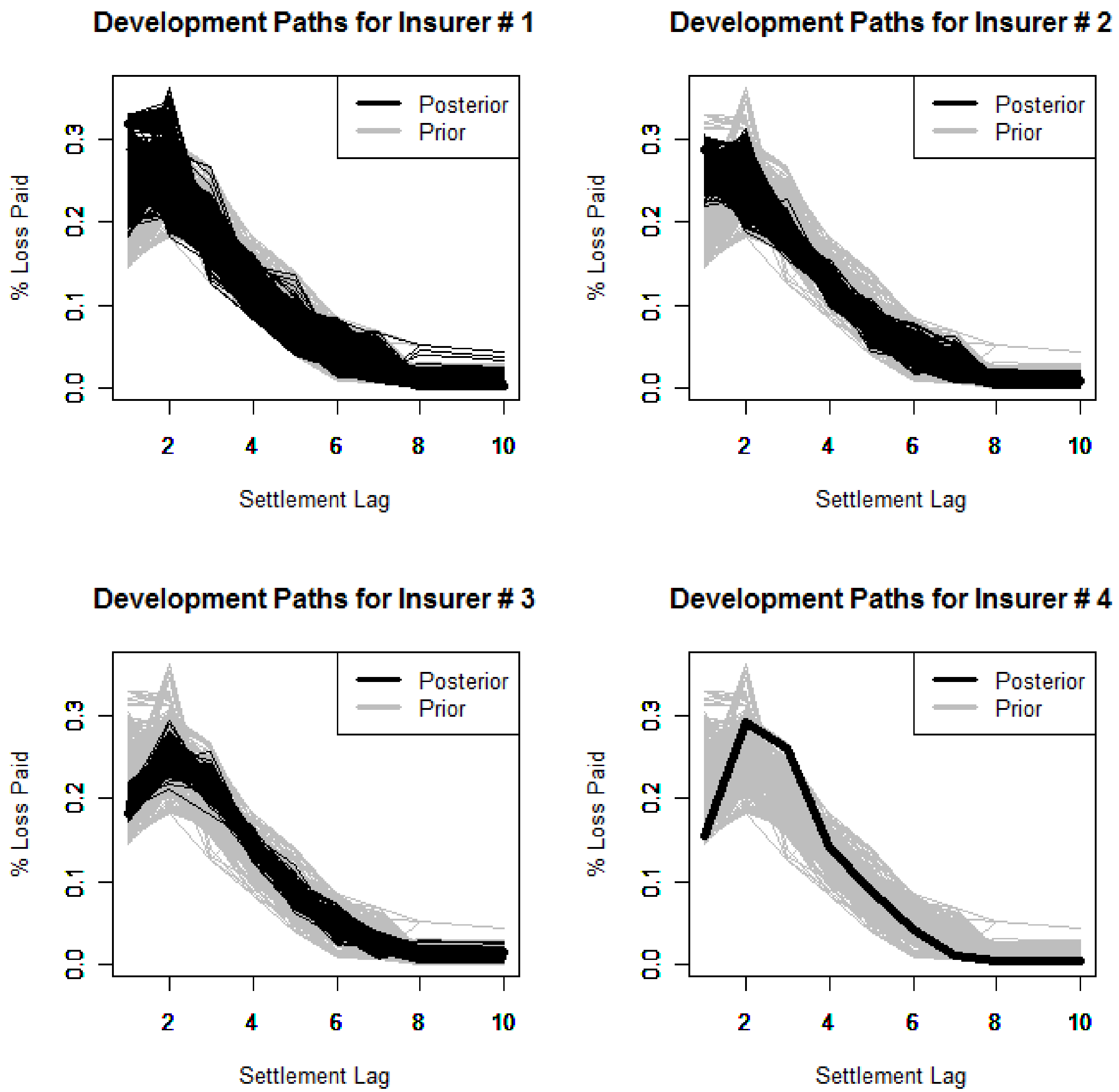
Insurer	AY	Premium	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
4	1986	154,769	22,459	28,908	19,921	21,410	10,174	2,989	4,612			
	1987	166,929	19,945	14,711	29,697	24,723	9,032	6,564				
	1988	132,704	10,036	42,385	20,782	3,156	16,903		143	173		
	1989	212,198	21,380	40,015	37,410	12,050	5,553	6,711	1,087			
	1990	218,364	22,505	52,592	31,367	24,107	18,357	9,619				
	1991	235,788	26,633	47,039	33,072	24,141	11,471					
	1992	253,306	19,292	64,884	64,489	32,434						
	1993	210,428	28,867	57,014	48,262							
	1994	222,016	37,892	59,734								
	1995	212,360	33,500									

Amounts in Thousands

ELR Paths with 99.9% Posterior Probability



Dev Paths with 99.9% Posterior Probability



Historical Background

- Two risk theory books
  - Pentikäinen “bundles”
  - Bühlmann hyper parameters

Statistics of Interest

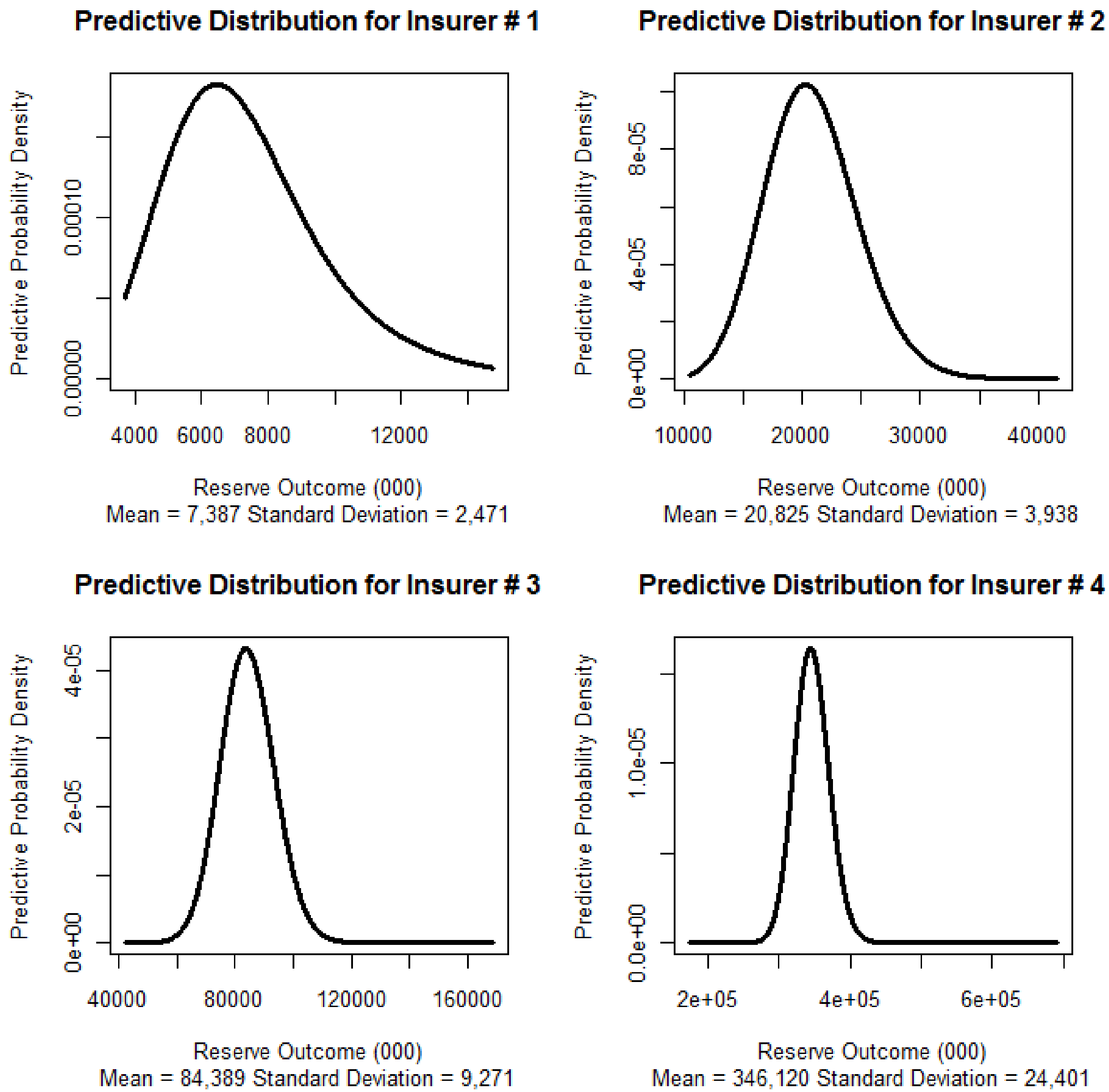
$$\text{Outcome} \quad X = \sum_{AY=2}^{10} \sum_{12-AY}^{10} X_{AY,Lag}$$

- Use FFT for each Scenario
  - FFT-1 Compound Frequency/Severity formulas
  - FFT-2 Convolute by multiplying FFT-1s
  - Invert FFT-2 to get $F(x|Scenario)$

$$F(x) = \sum_{Scenario} F(x \mid Scenario) \cdot \Pr(Scenario \mid Data)$$

- Mean and standard deviation of X
  - Calculate directly from $F(x)$

Distribution of Outcomes



Desirable Characteristics of Proxies

- Method “requires” little data
  - The only requirement is premium
    - In this case, the predictive distribution of outcomes is an equally weighted mixture of the scenarios
  - Any incremental paid loss data point will reweight the scenarios
- As an insurer has more data, the predictive distribution of outcomes becomes tighter.

Obtaining Relevant Scenarios

- American – NAIC Schedule P lines
  - Private Passenger Auto, Homeowners, Homeowners, Workers'Compensation, Medical Malpractice, … … .
- Pick 50 largest insurers
  - Calculate MLE estimates of {ELR} and {Dev}
  - Find 100 other {ELR} and {Dev} parameter sets that are "close" to MLE
    - Likelihood is close to maximum likelihood
    - Use the Gibbs sampler

Historical Background

- Bühlmann-Straub Credibility
  - Idea is to get information about prior distribution by examining statistics from other similar insurers.

References

"Proxies" *Actuarial Review*, November 2008
http://www.casact.org/newsletter/index.cfm?fa=viewart&id=5677

"Stochastic Loss Reserving with the Collective Risk Model" *CAS E-Forum*, Fall 2008
http://www.casact.org/pubs/forum/08fforum/11Meyers.pdf

"Estimating Predictive Distributions for Loss Reserve Models" *Variance* 2007, Issue 2
http://www.variancejournal.org/issues/?fa=article&abstrID=6417