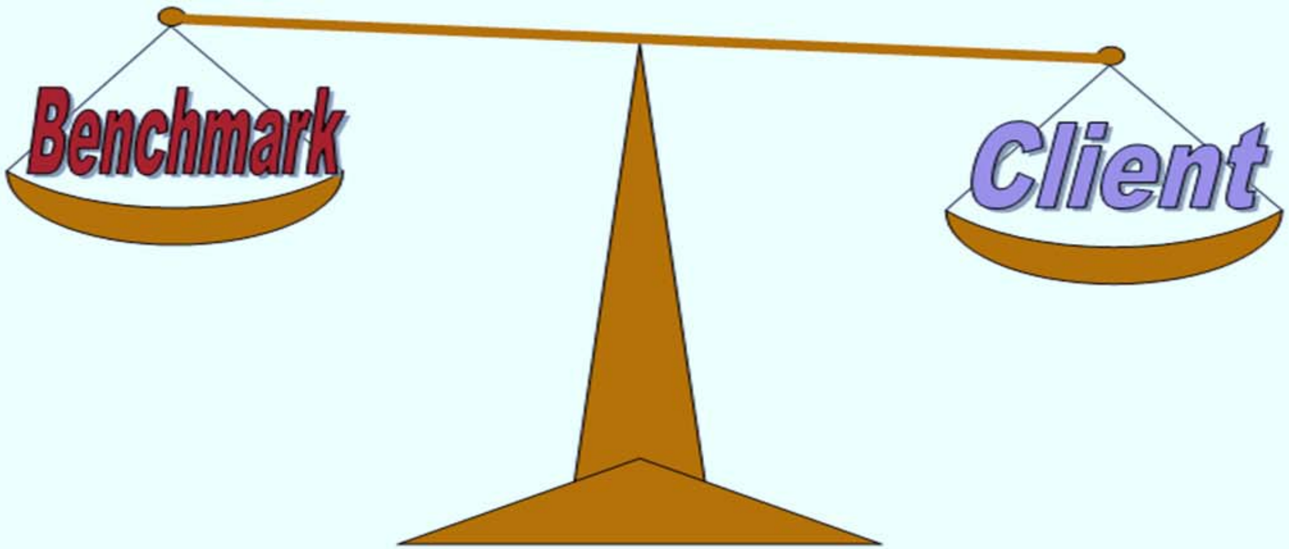
Antitrust Notice

Munich RE

The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.



INTRODUCTION TO BAYESIAN LOSS DEVELOPMENT

CAS Seminar on Reinsurance – June 6-7, 2016

Dave Clark
Munich Reinsurance America, Inc

Munich RE

Agenda

Munich RE

1. Loss Development in Reinsurance Pricing
2. Bayesian Theory and Mathematics (optional)
3. Practical Implementation

Loss Development Blending

Reinsurance pricing problem:

We have a loss development triangle from our client:

- May be sparse, not fully credible
- No tail beyond latest age in triangle

We have “benchmark” pattern from other sources:

- ISO / RAA / Reserving / Peer Companies
- Uncertain estimation and relevance for this client

Loss Development Blending

(numbers for illustration only)

Munich RE

Single Benchmark Example								
	12	24	36	48	60	72	84	96
1990	73	262	469	528	536	591	604	606
1991	148	346	391	502	522	514	567	
1992	99	198	219	394	408	430		
1993	118	255	352	412	581			
1994	275	415	645	803				
1995	261	446	637					
1996	130	471						
1997	148							
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
1990	3.589	1.790	1.126	1.015	1.103	1.022	1.003	
1991	2.338	1.130	1.284	1.040	0.985	1.103		
1992	2.000	1.106	1.799	1.036	1.054			
1993	2.161	1.380	1.170	1.410				
1994	1.509	1.554	1.245					
1995	1.709	1.428						
1996	3.623							
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	
Avg ATA	2.168	1.412	1.271	1.115	1.047	1.060	1.003	

Bayesian Philosophy

“The Bayesian paradigm offers a formal mechanism for incorporating into one's analysis information not contained in the available data.”

- Zhang, Dukic & Guszcza (2012)

“…the prior probability distributions in Bayesian inference provide a powerful mechanism for incorporating information from previous studies, and for controlling confounding.”

- Dong & Chan (2013)

Bayesian Philosophy

Munich RE

Bayes' Theorem:

$$\pi(\theta|X) = \frac{f(X|\theta) \cdot \pi(\theta)}{\int f(X|\theta) \cdot \pi(\theta)\, d\theta}$$

This formula has three components:

$\pi(\theta)$ A distribution representing "prior" knowledge of the parameters θ

$f(X|\theta)$ A likelihood function representing the probability of observing the actual data X given a certain parameter set.

$\pi(\theta|X)$ The "posterior" probability of the parameters, revised based on the data

Bayesian Philosophy

Munich RE

Good News: Bayesian model allows us to incorporate expert judgment or prior knowledge in a coherent way.

Bad News: Bayesian model requires us to set up explicit statistical distributions to incorporate this prior knowledge.

If θ is a parameter vector (e.g. a set of ten age-to-age factors),

Then $\pi(\theta)$ is a ten-dimensional probability density function.

And this requires evaluating a ten-dimensional integral:

$$\int f(X|\theta) \cdot \pi(\theta)\, d\theta$$

Bayesian Philosophy

Munich RE

Tools for Evaluating the Mathematics:

1) Conjugate Families
2) Linear Approximation to Bayes Formula => Bühlmann-Straub
3) Numerical Approximation of the Formula
    a) Quadrature integration (old method)
    b) Simulation via MCMC (the new favorite)

Conjugate family has advantage of simple calculation and interpretability.

Conjugate Prior

Munich RE

When the prior distribution $\pi(\theta)$ and likelihood $f(X|\theta)$ are chosen such that the posterior distribution $\pi(\theta|X)$ has the same distribution form as the prior, then we have a *conjugate* relationship.

Common examples from the Exponential Family are:

$\pi(\theta)$ => $f(X|\theta)$

Gamma => Poisson

Beta => Binomial

For the loss development pattern problem, we need a multivariate conjugate relationship.

Dirichlet => Multinomial

Dirichlet Distribution

Munich RE

The Dirichlet distribution is a multivariate version of the beta distribution.

Instead of a yes/no probability of p or $(1-p)$,

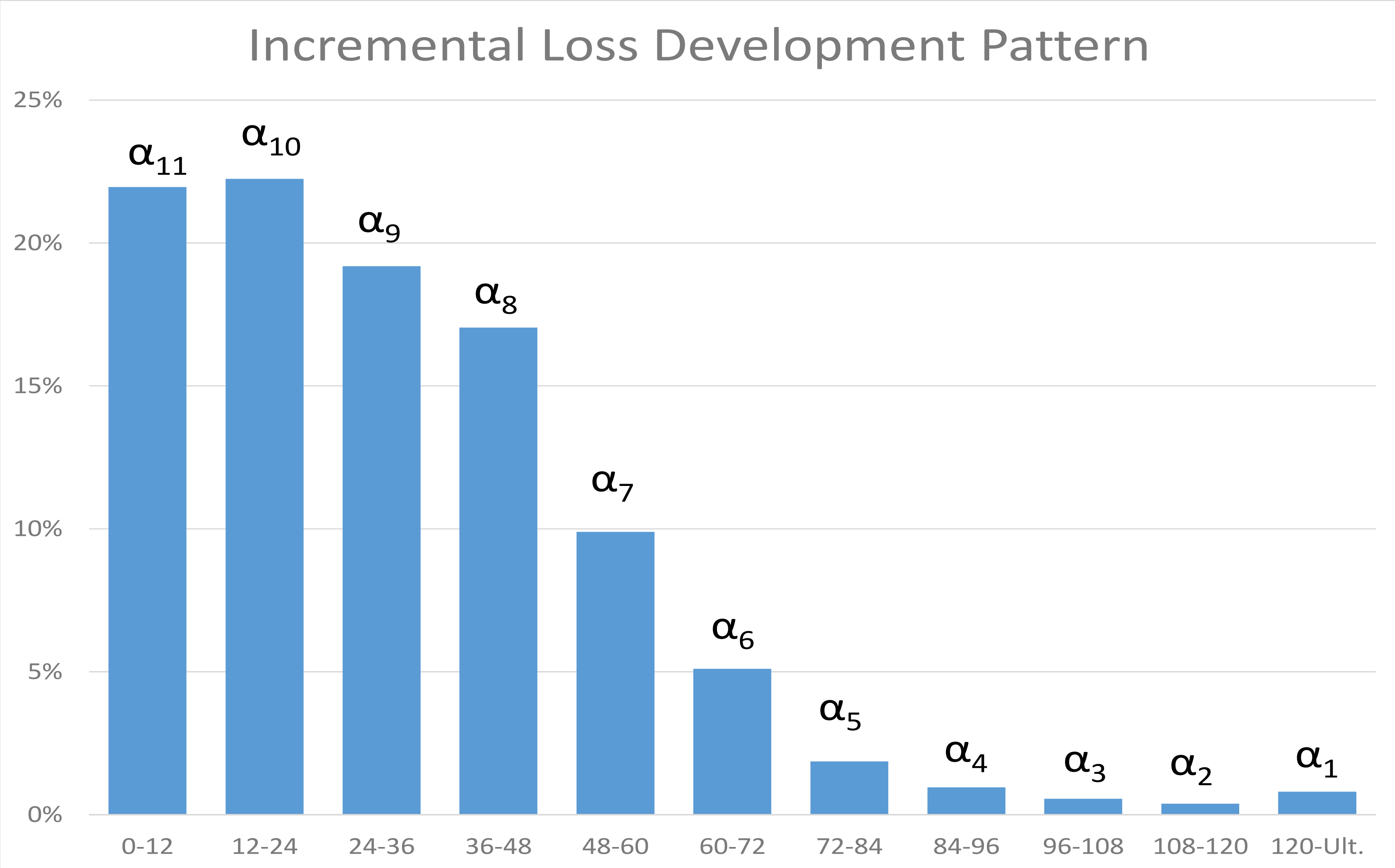
we have a vector of probabilities $\{p_1, p_2, p_3, \cdots, p_k\}$.

$$\pi(\boldsymbol{p}) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_k)}{\Gamma(\alpha_1) \cdot \Gamma(\alpha_2) \cdots \Gamma(\alpha_k)} \cdot \prod_{j=1}^{k} p_1^{\alpha_1 - 1} \cdot p_2^{\alpha_2 - 1} \cdots p_k^{\alpha_k - 1}$$

$$E(p_i) = \frac{\alpha_i}{\sum \alpha_j} \qquad 1 = \sum p_j$$

We can view this as a simulation of “k” gamma random variables, with a common scale parameter, which are then turned into percentages.

Dirichlet Distribution



Dirichlet Distribution

The Dirichlet distribution gives us a prior pattern that matches the percents paid (or reported) in each incremental period. The increments are proportional to the α_j parameters.

Given a new *observed* pattern from the client data, the "posterior" distribution simply adds a value to update the Dirichlet parameters.

$$\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k$$

becomes

$$(\alpha_1 + n_1), (\alpha_2 + n_2), (\alpha_3 + n_3), \cdots, (\alpha_k + n_k)$$

However, this assumes we have a complete pattern from the client.

We do not; we have a series of incomplete patterns.

Generalized Dirichlet Distribution

Munich RE

The Generalized Dirichlet distribution (Wong 1998) solves this challenge of incomplete data for us.

The Generalized form has twice as many parameters, but accommodates the incomplete data. Conveniently, this is also a conjugate form – meaning that the posterior distribution is again Generalized Dirichlet, with adjusted parameters.

The α's are incremental losses, the β's are cumulative losses.

$$\{\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k, \beta_1, \beta_2, \beta_3, \cdots, \beta_k\}$$

This has the remarkable interpretation that $\left\{1 + \frac{\alpha_j}{\beta_j}\right\}_{j=1}^{k}$ is the sequence of age-to-age factors.

Credibility Blending Formula

In addition to the uncertainty in the "prior" benchmark pattern, we need a measure of the process variance in the client triangle.

This is incorporated via a "dispersion" or variance/mean parameter ϕ. Equivalent to the factor used in GLM or Bootstrapping for chain ladder.

$$Credibility\ K = \frac{Expected\ Process\ Variance}{Variance\ of\ Hypothetical\ Means}$$

In the N/(N+K) formula, the "K" acts as ballast. It can be interpreted as counts or dollars depending upon the application.

Estimate of Process Variance/Mean Parameter

The selection of the scale parameter can be taken from a collective risk model.

Table M for Workers' Compensation is one example.

This parameter can also vary by layer and line of business as needed.

Theoretical "Table M" (for illustration)				
Gamma Shape Parameter	Insurance Charge at Entry=1	Expected Loss Group	Aggregate Loss Size (example)	Implied Variance/Mean
0.5	0.484	48	360,000	720,000
1	0.368	37	1,000,000	1,000,000
1.5	0.308	31	2,000,000	1,333,333
2	0.271	27	3,750,000	1,875,000

Credibility Blending Formula

Munich RE

Bayesian Combinations:

General Dirichlet:

$$ATA_d = \frac{\alpha_{k-d} + \beta_{k-d}}{\beta_{k-d}}$$

Chainladder:

$$ATA_d = \frac{\sum_{t=1}^{k-d} C_{t,d+1}}{\sum_{t=1}^{k-d} C_{t,d}}$$

Blended:

$$ATA_d = \frac{\phi \cdot (\alpha_{k-d} + \beta_{k-d}) + \sum_{t=1}^{k-d} C_{t,d+1}}{\phi \cdot \beta_{k-d} + \sum_{t=1}^{k-d} C_{t,d}}$$

Credibility Blending Formula

Munich RE

The credibility blending becomes a simple dollar-weighted average.

If you can calculate an age-to-age factor, then you can do a Bayesian model!

		Example of Blending Client and Benchmark Patterns						
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
ATA from Triangle								
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	-
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	-
ATA	**2.168**	**1.412**	**1.271**	**1.115**	**1.047**	**1.060**	**1.003**	
Benchmark Pattern								
Col. 1	1,419	2,027	2,546	2,933	3,383	3,633	3,717	3,042
Col. 2	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000
ATA	**2.819**	**1.973**	**1.571**	**1.364**	**1.182**	**1.101**	**1.076**	**1.315**
Blended Pattern								
Col. 1	2,523	3,949	4,622	4,769	4,849	4,738	4,321	3,042
Col. 2	6,393	6,713	6,639	6,047	5,535	5,171	4,606	4,000
ATA	**2.534**	**1.700**	**1.436**	**1.268**	**1.141**	**1.091**	**1.066**	**1.315**

Credibility Blending – Value of Conjugate Priors

Munich RE

"Conjugate priors… have the desirable feature that prior information can be viewed as 'fictitious sample information' in that it is combined with the sample in exactly the same way that additional sample information would be combined.

"The only difference is that the prior information is 'observed' in the mind of the researcher, not in the real world."

- Bayesian Econometric Methods; Koop, Poirier & Tobias

Creation of an Informative Prior

Munich RE

How do we create the prior distribution $\pi(\theta)$?

1) Empirical Bayes - estimate from collection of available patterns
2) Elicit ranges from expert users: e.g., select slow/medium/fast patterns
3) Reverse engineer – what prior is implied by the credibility percents that have been applied by users on actual accounts?

Different form for pricing and reserving actuaries?

Sources of Variance in the Prior

1) Market Heterogeneity

"…the market experience is not fully relevant to a particular client. This is usually captured by the spread, or heterogeneity, of the client risk premiums around the standard market rate."

2) Estimation Uncertainty

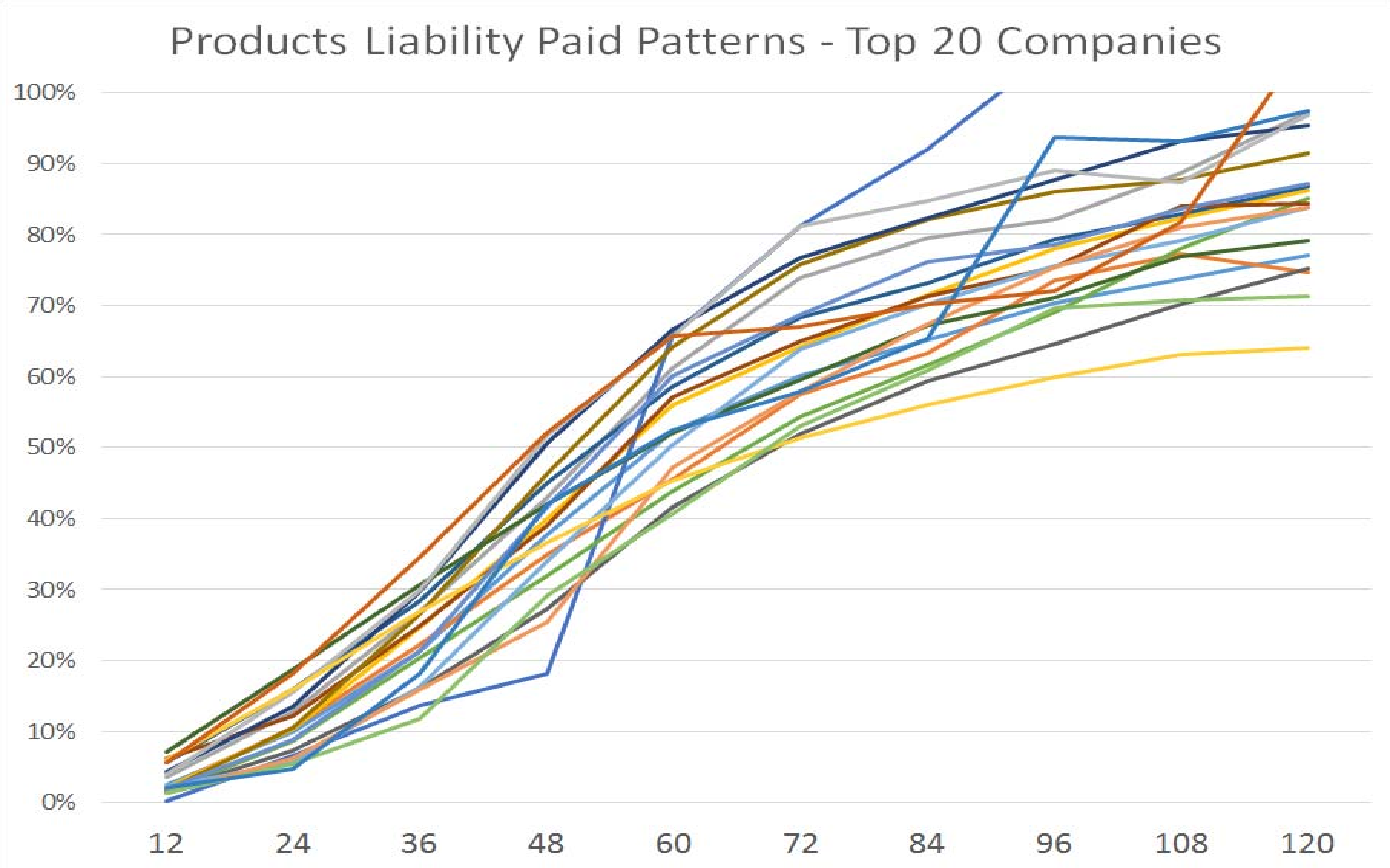
"…although the market rate is typically computed from a larger data set than that of a client, it, too, is based on a loss database of limited size and is therefore affected by the same type of uncertainty."

- Parodi & Bonche

"Uncertainty-Based Credibility and its Applications" *Variance* 2010

Estimate from collection of available patterns

Products Liability



Extending the Model

Munich RE

We can extend this model further by including mixtures of prior distributions.

Perhaps we know that companies are naturally grouped into Fast, Medium, or Slow payment patterns. But we do not know to which group our client belongs.

		Cumulative Loss Development Factors						
	12	24	36	48	60	72	84	96
Fast	14.014	4.930	2.607	1.759	1.406	1.263	1.191	1.155
Medium	21.950	7.787	3.946	2.512	1.842	1.558	1.415	1.315
Slow	49.240	15.860	7.407	4.163	2.706	2.057	1.750	1.567

Extending the Model

Munich RE

We assign initial weights to each of the groupings (perhaps 33%/33%/33%) and then apply Bayes' theorem to update the weights.

		Bayesian Updating of Probabilities			
		Difference	Relative	Original	Revised
	LogLikelihood	in LL	Likelihood	Weights	Weights
	A	B=A-Max(A)	C=exp(B)	D	E=C*D/Avg(C)
Slow	-4.61	-0.77	0.464	33.33%	20.41%
Baseline	-4.06	-0.21	0.810	33.33%	35.61%
Fast	-3.84	0.00	1.000	33.33%	43.98%
			0.758	100.00%	100.00%

This allows us to adjust our "tail" based on which group is closest to our client's data.

Conclusions

- Credibility in Loss Development pattern selection has benefits
  - Stability in estimation – therefore can break data into small homogeneous pieces
  - Consistency in pricing
  - Even very sparse data from a client can update the benchmark
- The Bayesian framework can be implemented practically for pricing
- The Bayesian framework can be extended to include benchmarks for every uncertain parameter

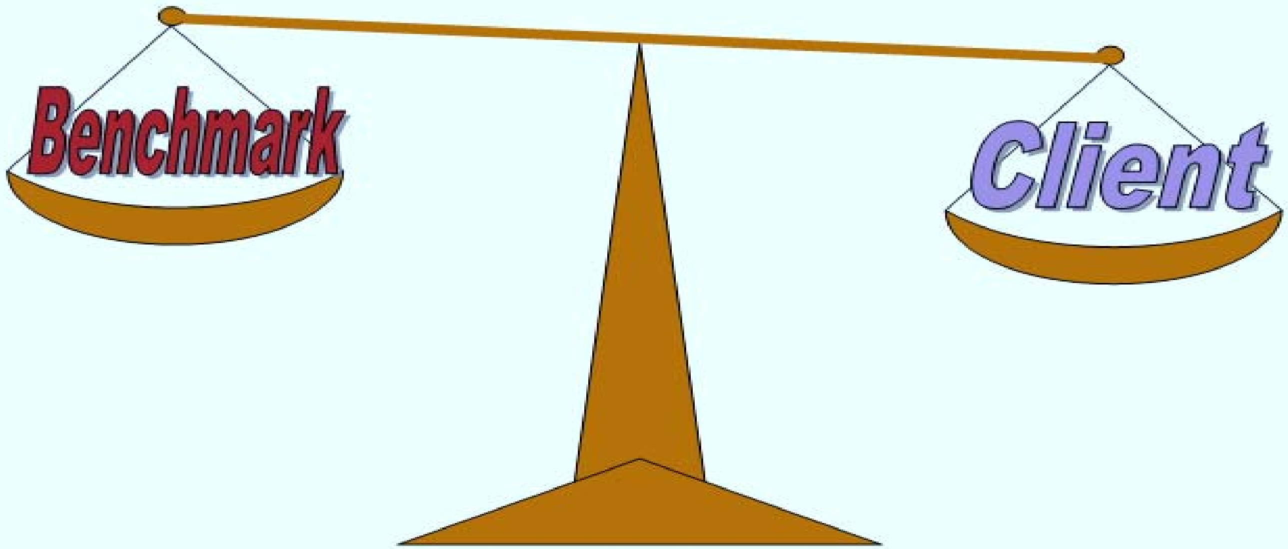
Selected References

Bolstad, William M., “Introduction to Bayesian Statistics,” Wiley, 2007.

Mildenhall Stephen J., “A Multivariate Bayesian Claim Count Development Model with Closed Form Posterior and Predictive Distributions,” *CAS Forum* Winter 2006, 451-493.

Wong, Tzu-Tsung, “Generalized Dirichlet Distribution in Bayesian Analysis,” Applied Mathematics and Computation 97 (1998), 165-181.

Zhang, Yanwei, Vanja Dukic, James Guszcza, “A Bayesian Non-Linear Model for Forecasting Insurance Loss Payments,” *Journal of the Royal Statistical Society* 2012, Vol. 175, Issue 2, 637-656



Dave Clark
daveclark@munichreamerica.com
https://twitter.com/DaveclarkR

Munich RE