Deloitte.

Loss Reserving Using Claim-Level Data

James Guszcza, FCAS, MAAA
Jan Lommele, FCAS, MAAA
CAS Annual Meeting
San Francisco
November, 2006

Audit. Tax. Consulting. Financial Advisory.

Topics

Deloitte.

Overview

Why do reserving at the claim level?

2 Basic Motivations (and a Fringe Benefit)

1. Better reserve estimates for a changing book of business
   - Do triangles "summarize away" patterns that could aid prediction?
   - Could adding predictive variables help?

2. More accurate estimate of reserve variability
   - Insight of bootstrapping: empirical distribution can often serve as a proxy for the true distribution.
   - But loss triangles summarize away this distributional information.

- Fringe Benefit: once a suitable claim-level data warehouse is built, it can also be used for various claim analytics projects
  - Claim Severity / duration modeling
  - Soft fraud modeling
  - These types of projects have high ROI

Motivation #1: Better Reserve Estimates

- Key idea: use predictive variables to supplement loss development patterns
  - Most reserving approaches analyze summarized loss/claim triangles.
  - Does not allow the use of covariates to predict ultimate losses (other than time-indicators).

- Actuaries use predictive variables to construct rating plans & underwriting models.

- Why not loss reserving too?

Why Use Predictive Variables?

- Suppose a company's book of business has been deteriorating for the past few years.
- This decline might not be reflected in a summarized loss development triangle.
- However: The resulting change in distributions of certain predictive variables might allow us to refine our ultimate loss estimates.
- Predictive variables can be "leading indicators" of changes in development patterns.

Examples of Predictive Variables

- Policy metrics:
  - limits written, classification, jurisdiction, distribution by product type, rating territory…
  - age of policy, agent characteristics, bill paying history, financial information on the policyholder, motor vehicle records

- Operational metrics:
  - case reserve philosophy, new rating strategies…

- Changes in policy processing:
  - Role of agent vs. underwriter, new system

- Financial Metrics:
  - Rate Changes

Examples of Predictive Variables

- Qualitative metrics:
  - Mix of Preferred/standard/non-standard, schedule rating, premium discounts, renewal credits, multiple policy credits

- Claim metrics:
  - Date of accident, time of day, report lag, coverage…
  - nature of injury (BI, PD, medical vs. indemnity etc)
  - type of injury (e.g. back strain, broken bone, damaged vehicle, house fire, car theft…)
  - cause of loss/injury, diagnosis and treatment codes, financial information on the claimant, attorney retained by claimant or not

- Claimant metrics:
  - Age, gender, marital status, medical history (if available)

Aside: note that these variables are also useful for claim duration modeling and soft fraud analysis.

The Advantages of More Data Points

- Typical reserving projects use claim data summarized to the year/quarter level.
- Probably an artifact of the era of pencil-and-paper statistics.
- In certain cases important patterns might be "summarized away".
- In the computer age, why restrict ourselves?
- More data points ➔ less chance of over-fitting the model.
- More data points ➔ easier to use modern out-of-sample validation techniques to evaluate models.

Motivation #2: Reserve Variability

- Bootstrapping is a modern, simulation-based technique for estimating the variance of a complex estimator.

- Most current discussions: we can re-sample ("bootstrap") the *residuals* of a stochastic reserving model.
  - Add the residuals back to the data ➔ gives you a pseudo-dataset
  - Resampling the residuals 1000 times gives you 1000 pseudo datasets
  - Compute the reserves on each pseudo dataset
  - This constitutes a distribution of reserves

- If we perform our reserving analysis on claim-level data, we can bootstrap the *observations* rather than the *residuals* of a statistical model.
  - A much simpler procedure

Further Claim Analytics

- A claim-level database built for loss reserving can also be used for high-ROI claims analytics projects. For example:

- Soft Fraud
  - Many policyholders unconsciously think it OK to inflate claims.
  - Chance to "get back" money they've sunk into insurance over the years.
  - Natural application of predictive modeling: models to predict which claimants are displaying patterns of "opportunism", "abuse" or "inflation".

- Claim Duration / Severity Management Models
  - Use to optimally allocate resources to claims of different expected durations.
  - Promotes a consistent, rational claims management process.
  - Models can be used to reduce average claim duration; diminish the frequency of long-term claims.

Deloitte.

Modeling Approach

Sample Model Design

Audit.Tax.Consulting.Financial Advisory.

Philosophy

- Provide an illustration of how reserving might be done at the claim or policy level.

- Simple starting point: consider a generalization of the chain-ladder.
    - Just one possible model
    - Connects this emerging topic with traditional practice
    - Simpler, more appropriate models along these lines are possible

- Analysis is suggestive rather than definitive
    - No consideration of superimposed inflation
    - No consideration of calendar year effects
    - Model risk not considered here
    - etc…

“Chain Ladder” Model Design

- Basic idea: predict L_{j+1} using covariates along with L_j.

- Build 9 successive GLM models
  - Regress L_{24} on L_{12}; L_{36} on L_{24} … etc
  - Notation: L_j = paid losses at time *j*.
  - The models incorporate covariates $\{X_i\}$ in addition to L_j.
  - 1 record per claim

- Each GLM analogous to a link ratio.

- The $L_j \rightarrow L_{j+1}$ model is applied to either
  - *Actual* values @ *j*
  - *Predicted* values from the $L_{j-1} \rightarrow L_j$ model

Model Design

- Idea: model each claim's loss development from period
  $L_j \rightarrow L_{j+1}$

$$\frac{L_{j+1}}{L_j} = f(\alpha + \beta_1 X_1 + \ldots + \beta_n X_n)$$

- …as a function of a linear combination of several covariates.

$$\frac{L_{j+1}}{L_j} = f(\alpha) = LinkRatio$$

- Claim-level generalization of the chain-ladder idea:
  - Consider case where there are no covariates

- No statistical assumptions made as yet.

Bringing GLM to Bear

- Assume *f*() is a linear function of covariates $\{X_i\}$
- Adopt over-dispersed Poisson GLM:
  - Log link function
  - Pragmatic choice – often used in reserving
  - Variance of L_{j+1} is proportional to mean
  - Use L_{j+1} as the target and treat $\log(L_j)$ as the offset term
  - Or simply use the ratio (L_{j+1}/L_j) as the target and use L_j as a weight.

$$\frac{L_{j+1}}{L_j} = e^{\alpha+\beta_1 X_1+\ldots+\beta_n X_n} + \varepsilon$$

$$L_{j+1} = e^{\log(L_j)+\alpha+\beta_1 X_1+\ldots+\beta_n X_n} + \varepsilon$$

Deloitte.

Case Study

Simulated Claim Data

Comparison of the proposed method with the chain-ladder

Simulation Study

- We can illustrate the technique by applying it to simulated data.

- By construction:
    1. Policies are identical except for one characteristic:
       Good credit vs. bad credit
    2. Bad credit policies develop more slowly.
        - Build a steeper development pattern into the simulation.
    3. A greater proportion of bad credit policies have been written in more recent years.

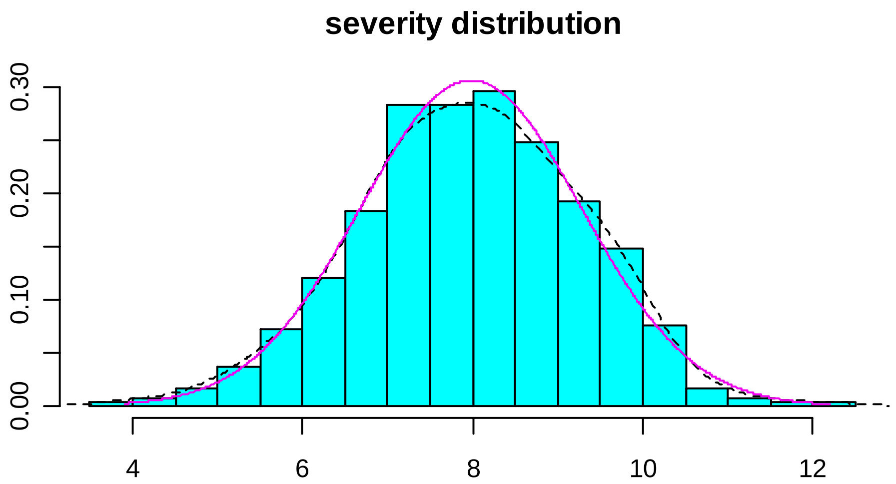
- The chain ladder will not reflect the changing mix of business in a timely way.

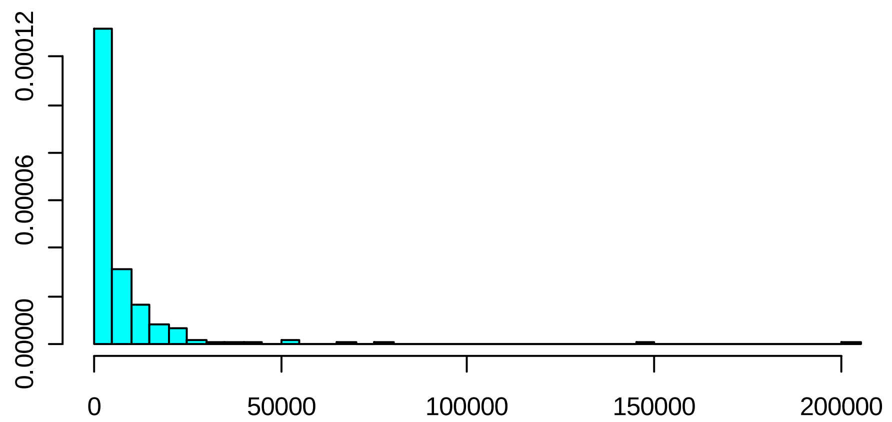
Simulated Data

- 500 claims each year * 10 years
- Each record represents 1 claim
  - IBNR not treated in this example.
- Each record has multiple loss evaluations
  - @ 12, 24, …,120 months
- "Losses @ *j* months" means:
  - *j* months from the beginning of the accident year
- Assume losses are fully developed @120 months.

Loss Distribution

- Each of the 5000 claims at 12 months was drawn from a lognormal distribution with parameters
  - $\mu=8; \sigma=1.3$
- 500 claims per year
- Idealized assumption: policies that generated the claims are identical except for their credit scores.

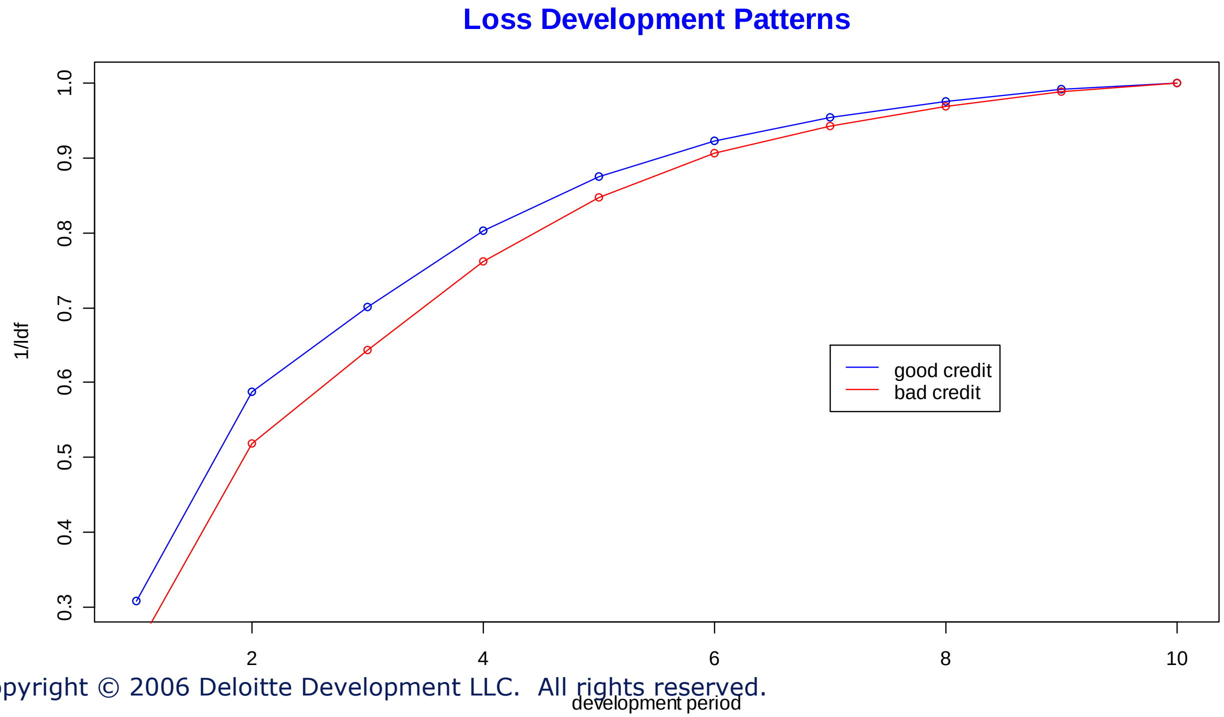




By Design

1. Build in slower average development patterns for losses from "bad credit" policies.
   - Random error around the average also used in the simulation of individual claims' development.

2. Proportion of "bad credit" policies higher in more recent accident years.



Year	%bad credit
1990	30%
1991	35%
1992	40%
1993	45%
1994	50%
1995	55%
1996	60%
1997	65%
1998	70%
1999	75%

Covariates

- In this example, credit is the only covariate.
  - Bad credit {0,1} indicator
- This suffices to illustrate the point.
  - Other covariates could easily be built into the simulation.
- Here, we know the "true" covariates
  - In real life, this is a significant source of model risk.

Approach

- We simulate data for all 10 accident years to ultimate.
- We therefore know the “true” outstanding losses by year.
- Compare the results of both our method and the chain-ladder to the “truth”.

The Data

- Summarize the 5000 simulated data points to the accident year level.
  - We know the "true" o/s losses
  - Can calculate the implied development factors.

- Losses in the blue-shaded cells will be treated as unknown
  ➔ not used in the models.

	Losses in $1000's											
	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120	ultimate	o/s
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	**0**
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123	11,223	11,223	**99**
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444	11,635	11,739	11,739	**295**
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642	11,902	12,100	12,210	12,210	**567**
1994	2,912	5,563	6,644	7,629	8,329	8,808	9,112	9,321	9,484	9,571	9,571	**763**
1995	3,724	7,167	8,573	9,850	10,763	11,393	11,796	12,070	12,282	12,397	12,397	**1,634**
1996	3,213	6,202	7,423	8,540	9,337	9,885	10,232	10,473	10,656	10,757	10,757	**2,217**
1997	3,335	6,445	7,727	8,887	9,721	10,281	10,643	10,890	11,083	11,187	11,187	**3,460**
1998	3,596	6,975	8,387	9,662	10,589	11,207	11,604	11,876	12,090	12,204	12,204	**5,229**
1999	3,327	6,481	7,817	9,018	9,889	10,483	10,860	11,123	11,323	11,432	11,432	**8,105**
												22,369
implied												
Link ratios	1.964	1.209	1.149	1.094	1.060	1.036	1.022	1.017	1.009	1.000		
LDFs	3.436	1.750	1.448	1.260	1.152	1.087	1.049	1.026	1.009	1.000		

Comparing the Results

- The chain ladder produces results that are too low
    - AY 1999 CL reserves are too low by $.6M
- Unlike the C-L, the proposed method does pick up on the shifting distribution of bad credit policies.
    - Proposed method's results very close to the "truth"

acc. year	losses @ 12/99	true ultimate	true o/s	our method	chain ladder
1990	10,970	10,970	-	-	-
1991	11,123	11,223	99	99	97
1992	11,444	11,739	295	294	289
1993	11,642	12,210	567	572	558
1994	8,808	9,571	763	765	728
1995	10,763	12,397	1,634	1,629	1,535
1996	8,540	10,757	2,217	2,205	2,097
1997	7,727	11,187	3,460	3,475	3,304
1998	6,975	12,204	5,229	5,237	4,898
1999	3,327	11,432	8,105	8,057	7,466
			22,369	22,333	20,972
				-0.16%	-6.25%

Comparing the Results

- In more detail…
- Chain-ladder LDFs, ultimate losses too low in recent accident years

					Losses in $1000's									
	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120		C-L	truth	proposed
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	0	**0**	0
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123		11,220	97	**99**	99
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444			11,734	290	**295**	294
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642				12,200	558	**567**	572
1994	2,912	5,563	6,644	7,629	8,329	8,808					9,537	729	**763**	765
1995	3,724	7,167	8,573	9,850	10,763						12,298	1,535	**1,634**	1,629
1996	3,213	6,202	7,423	8,540							10,637	2,097	**2,217**	2,205
1997	3,335	6,445	7,727								11,031	3,304	**3,460**	3,475
1998	3,596	6,975									11,873	4,898	**5,229**	5,237
1999	3,327										10,792	7,465	**8,105**	8,057
												20,974	**22,369**	22,333
C-L	1.906	1.192	1.146	1.090	1.055	1.033	1.022	1.016	1.009	1.000				
	3.244	1.702	1.428	1.246	1.143	1.083	1.048	1.025	1.009	1.000				
truth	1.964	1.209	1.149	1.094	1.060	1.036	1.022	1.017	1.009	1.000				
	3.436	1.750	1.448	1.260	1.152	1.087	1.049	1.026	1.009	1.000				
proposed	1.954	1.208	1.152	1.093	1.059	1.036	1.023	1.017	1.009	1.000				
	3.422	1.751	1.450	1.258	1.151	1.087	1.049	1.026	1.009	1.000				

Eliminating the Covariate

- What if we run our claim-level model without including the credit predictive variable?

- The claim-level method exactly reproduces the chain ladder results.
  - The method is a proper generalization of the chain ladder.
  - Improved predictions are due to inclusion of relevant covariates.

acc. year	losses @ 12/99	true ultimate	true o/s	our method	chain ladder
1990	10,970	10,970	**-**	**-**	**-**
1991	11,123	11,223	**99**	**97**	**97**
1992	11,444	11,739	**295**	**289**	**289**
1993	11,642	12,210	**567**	**558**	**558**
1994	8,808	9,571	**763**	**728**	**728**
1995	10,763	12,397	**1,634**	**1,535**	**1,535**
1996	8,540	10,757	**2,217**	**2,097**	**2,097**
1997	7,727	11,187	**3,460**	**3,304**	**3,304**
1998	6,975	12,204	**5,229**	**4,898**	**4,898**
1999	3,327	11,432	**8,105**	**7,466**	**7,466**
			22,369	**20,972**	**20,972**
				-6.25%	-6.25%

Deloitte.

Reserve Variability

The Bootstrap

Estimating the probability distribution of ones outstanding loss estimate

Audit . Tax . Consulting . Financial Advisory.

The Bootstrap

- An estimate of outstanding losses is a point estimate.
- In statistics, we always want to complement a point estimate with a confidence interval.
  - Or better yet the distribution of the point estimate.
- In the case of loss reserving, it is hard derive formulas to calculate a confidence interval around a reserve estimate.
- Bootstrapping is a simulation-based technique for estimating the variability of any estimator.
  - No distributional assumptions needed.

The Key Idea

- Everything we know about the "true" probability distribution comes from the data.

- So let's treat the data as a *proxy* for the true distribution.

- We draw multiple samples from this proxy…
  - This is called "resampling".
  - Resampling = sampling with replacement

- And compute the statistic of interest on each of the resulting pseudo-datasets.
  - You thereby estimate the *distribution* of this statistic.

Motivating Example

- Let's look at a simple case where we all know the answer in advance.

- Pull 500 draws from the *n*(5000,100) dist.

- The sample mean ≈ 5000
  - Is a point estimate of the "true" mean μ.
  - But how sure are we of this estimate?

raw data	
statistic	**value**
#obs	500
mean	**4995.79**
sd	**98.78**
2.5%*ile*	4812.30
97.5%*ile*	5195.58

- From theory, we know that: $s.d.(\bar{X}) = \sigma / \sqrt{N} \approx {}^{100}\!\big/\!_{\sqrt{500}} \approx 4.47$

Sampling With Replacement

- Suppose we didn't know this textbook formula. How can we use resampling to estimate the s.d. of the sample mean?
  - (hopefully $\approx$ 4.47)

- Draw a data point at random from the data set.
  - Then throw it back in
- Draw a second data point.
  - Then throw it back in…
- Keep going until we've got 500 data points.
  - You might call this a "pseudo" data set.

- This is not merely re-sorting the data.
  - Some of the original data points will appear more than once; others won't appear at all.

Sampling With Replacement

- In fact, there is a chance of:

$$(1 - {}^{1}/_{500})^{500} \approx {}^{1}/_{e} \approx .368$$

that any one of the original data points won't appear at all if we sample with replacement 500 times.

➔ any data point is included with Prob ≈ .632

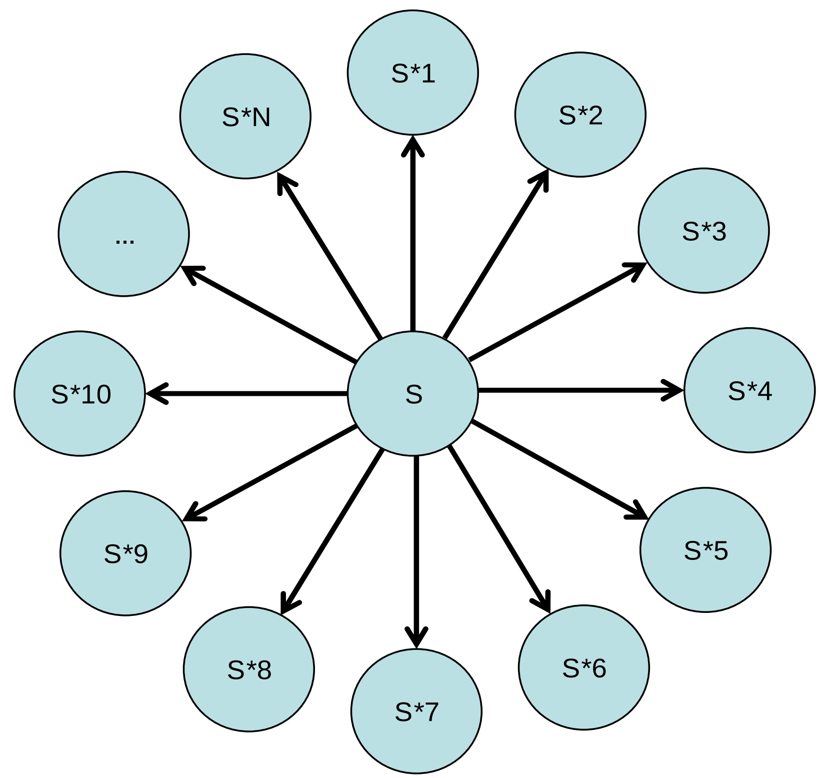
- Intuitively, we treat the original sample as the "true population in the sky".
- Each resample simulates the process of taking a sample from the "true" distribution.

Resampling

- Sample with replacement 500 data points from the original dataset S
  - Call this S^*_1

- Now do this 999 more times!
  - S^*_1, S^*_2,…, S^*_{1000}

- Compute X-bar on each of these 1000 samples.



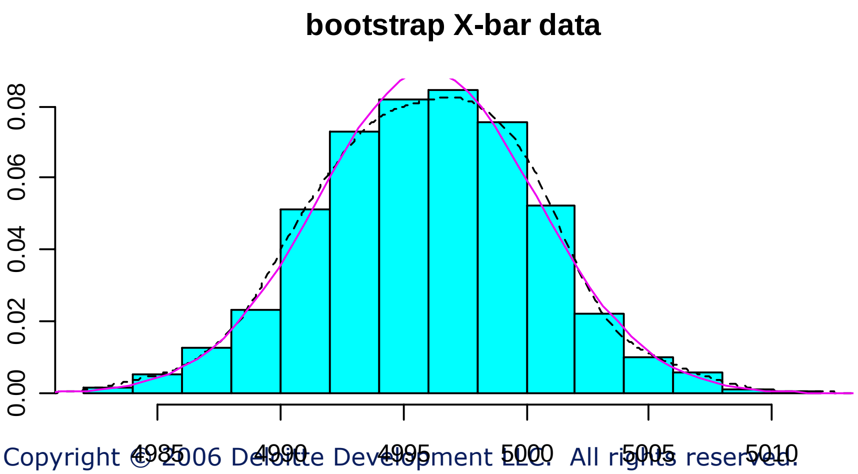
Results

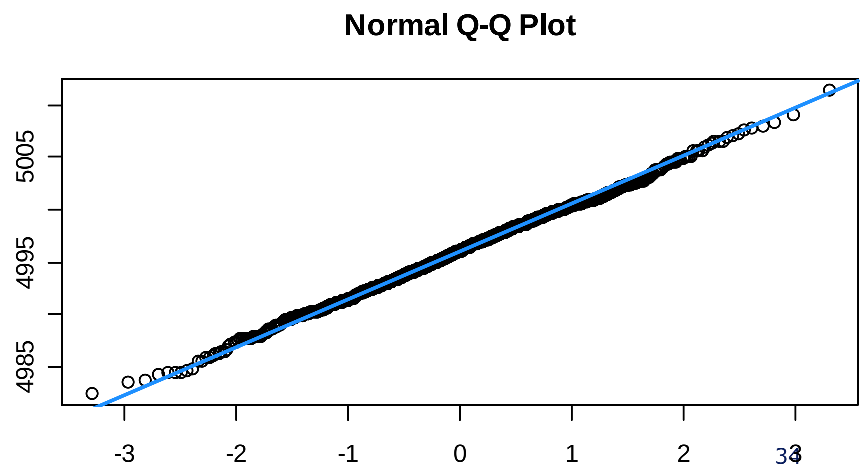
- From theory we know that
  X-bar ~ *n*(5000, 4.47)

- Bootstrapping estimates this pretty well!

- And we get an estimate of the whole distribution, not just a confidence interval.

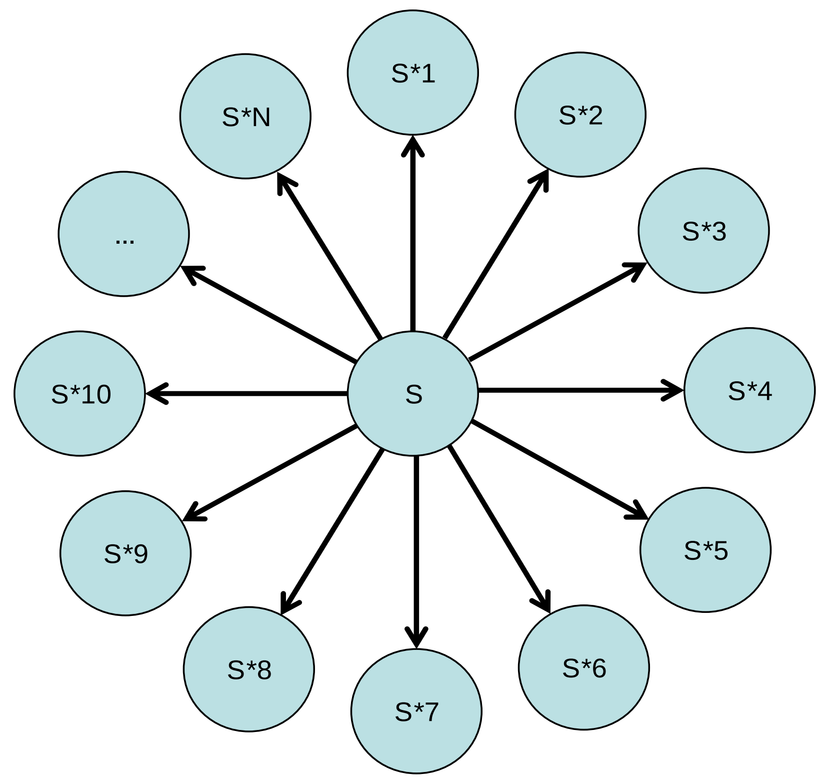
raw data		X-bar	
statistic	**value**	**theory**	**bootstrap**
#obs	500	1,000	1,000
mean	**4995.79**	**5000.00**	**4995.98**
sd	**98.78**	**4.47**	**4.43**
2.5%*ile*	4705.08	4991.23	4987.60
97.5%*ile*	5259.27	5008.77	5004.82





Bootstrapping Reserves

- S = our database of 5000 claims
- Sample with replacement all policies in S
  - Call this S^*_1
  - Same size as S
- Now do this 499 more times!
  - S^*_1, S^*_2,…, S^*_{500}
- Estimate o/s reserves on each sample
  - Get a distribution of reserve estimates



Sampling With Replacement

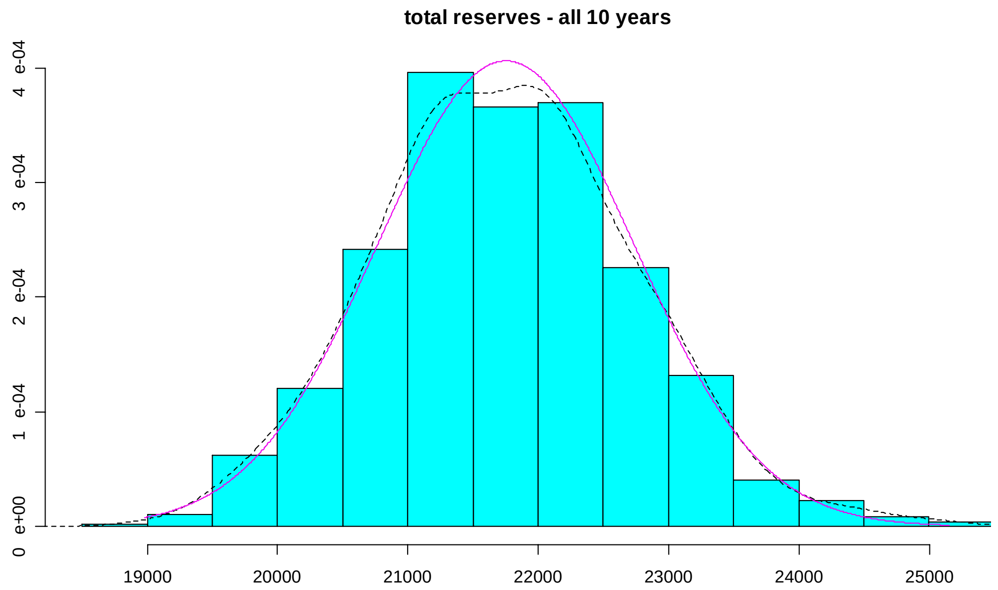
- Compute our reserve estimate on each S^*_k
  - These 500 reserve estimates constitute an estimate of the distribution of outstanding losses
  - Note: this involves running 9*500 GLM models!

- Notice that we did this by resampling our original dataset S of *claims*.
  - We resample “cases”, not residuals

- Nota bene: this method differs from methods which bootstrap the *residuals* of a model.
  - These methods rely on the assumption that your model is correct.
  - Residuals must be iid
  - If not, you have to be careful to resample only within certain specified parts of the triangle… gets messy
  - Resampling claim-level records is done at a stage prior to building the model

Distribution of Outstanding Losses

- The simulated dist of outstanding losses appears ≈ normal.
  - Mean: $21.751M
  - Median: $21.746M
  - σ : $0.982M
  - σ/μ ≈ 4.5%



- 95% confidence interval:
  (19.8M, 23.7M)

- This is a context-dependent result… it depends on:
  - The line of business
  - Number of data points …

Interpretation

- This result suggests:
  - With 95% confidence, the total o/s losses will be within +/- 9% of our estimate.
  - Assumes model is correctly specified.

- Too good to be true?
  - Yes: doesn’t include model risk.
    - In the real world we don’t know the ‘true’ model
  - Bootstrapping methodology can be refined.

- Advantage:
  - We are really using our 1000s of data points.
  - We’re not throwing away heterogeneity info.

Closing Thoughts

- The loss development triangle is not necessarily a sufficient statistic for estimating o/s losses.
  - Variability information is summarized away
  - Allows us to bootstrap observations, not just residuals or link ratios.
  - Information about differential loss development patterns might be suppressed.

- The ± 9% variability result is only suggestive
  - The data is just a primitive simulation.
  - Variability results will be context-dependent.

- Simulating $^{1}/_{10}$ the # of claims with 10 times the average size:
  - yields a similar reserve estimate…
  - …but a (-15%, +30%) 95% confidence interval...