CLRS 2008

What Color is Your Copula: The Language of Uncertainty

Terminology Surrounding Loss Reserve Variability/Ranges

Daniel Murphy, FCAS, MAAA
Trinostics

September 18, 2008

Agenda

- Why Analyze Reserve Ranges?
- General Terminology
- Popular Stochastic Methods
- Aggregation

Agenda

- **Why Analyze Reserve Ranges?**
- General Terminology
- Popular Stochastic Methods
- Aggregation

Why analyze uncertainty of claim liabilities?

- **Rating Agencies**
  - Increasing focus on economic capital, potential for reserves to vary from their stated values
- **Solvency II**
  - **Solvency Capital Requirement (SCR)** – capital to absorb significant unforeseen losses and give reasonable assurance to policyholders (0.5% probability of ruin over a one year timeframe)
- **Market value**
  - Value of an asset in the market reflects the uncertainty of its future cash flows
- **Actuarial Standards of Practice**
  - Encouraged

For ASB, *Uncertainty* is becoming more abstract, broad

<table>
<tr><th>ASOP 36 (eff. 10-15-00)</th><th>ASOP 43 (eff. 9-1-07)</th></tr>
<tr><td>▸ Written SOP – narrow</td><td>▸ Estimates of unpaid claims – broad</td></tr>
<tr><td>▸ Stated reserves “make a reasonable provision” if within actuary’s “range of reasonable estimates” = {estimates | method and assumptions are reasonable in actuary’s judgment}

▸ In determining range, actuary should consider implications of uncertainty; sources include
  ▸ Changes in: operations, environment, data, trends, actuarial patterns, types of claims, frequency or severity
  ▸ Erratic development data
  ▸ Random chance

PRACTICAL</td><td>▸ Actuary should consider – but is not required to measure – the uncertainty of the estimate

▸ If measuring uncertainty, actuary should choose appropriate methods, models, and assumptions

▸ Types and sources of uncertainty may include
  — Model risk
  — Parameter risk
  — Process risk

MATHEMATICAL</td></tr>
</table>

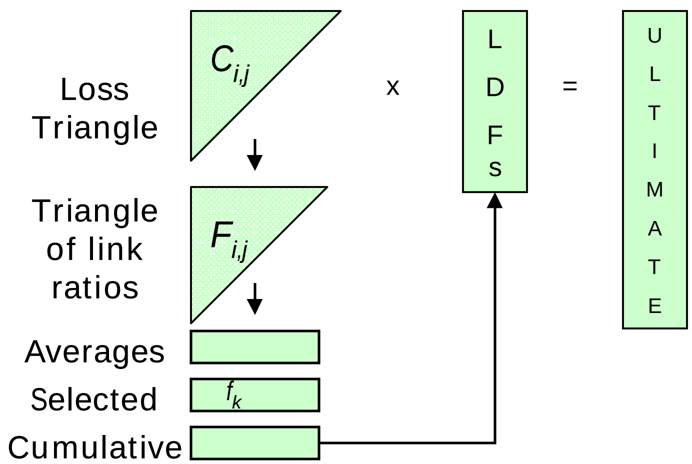
Agenda

- Why Analyze Reserve Ranges?
- **General Terminology**
  - Methods, Models, and Assumptions
  - The "Risks"
    - — Model Risk
    - — Parameter Risk
    - — Process Risk
- Popular Stochastic Methods
- Aggregation

Methods vs. Models, and Assumptions of each

Method	Model
‣ Mathematical algorithm for estimating unpaid claim amount	‣ Mathematical description of the unpaid claim phenomenon
‣ Parameters are selected ‣ *Judged* appropriate	‣ Parameters are estimated ‣ Can be tested
‣ Method assumed appropriate ‣ *Judged* appropriate	‣ Model assumed appropriate ‣ Can be tested

Chain Ladder Method



Chain Ladder Model*

Generalization of Murphy, Mack model:

$$C_{i,k+1} = f_k C_{i,k} + \sigma_k \varepsilon_{i,k} C_{i,k}^{\alpha_k}$$

$\varepsilon_{i,k}$ ~ independent Standard Normal rv's

i corresponds to accident year (row)

k corresponds to development age (column)

* Majidi, Bardis, Murphy, CAS *E-Forum*, Fall 2008

Sources and quantifiability of uncertainty in unpaid claim estimate from a modeler's perspective

- Model risk
  - Perhaps paid loss development method is inappropriate
    - Does AY '06 contain different state?
  - A.k.a, "Bias"
  - Quantifiability of model risk requires new data, broader model
- Parameter risk
  - Selected LDF is itself subject to the whim of the data
    - May not coincide with "true" LDF
  - Final estimate will change with variable parameters
  - Quantifiability of impact of potential parameter variability depends on model
- Process risk
  - Even if selection luckily coincides with "true" LDF, still expect final Age-2 amount to be different from expected amount $66,360
  - A.k.a, Residual or "unexplained" risk
    - Variability of age-2 losses not explained by paid losses at age 1
  - Quantifiability depends on model

ABC Insurance Company

Paid Losses

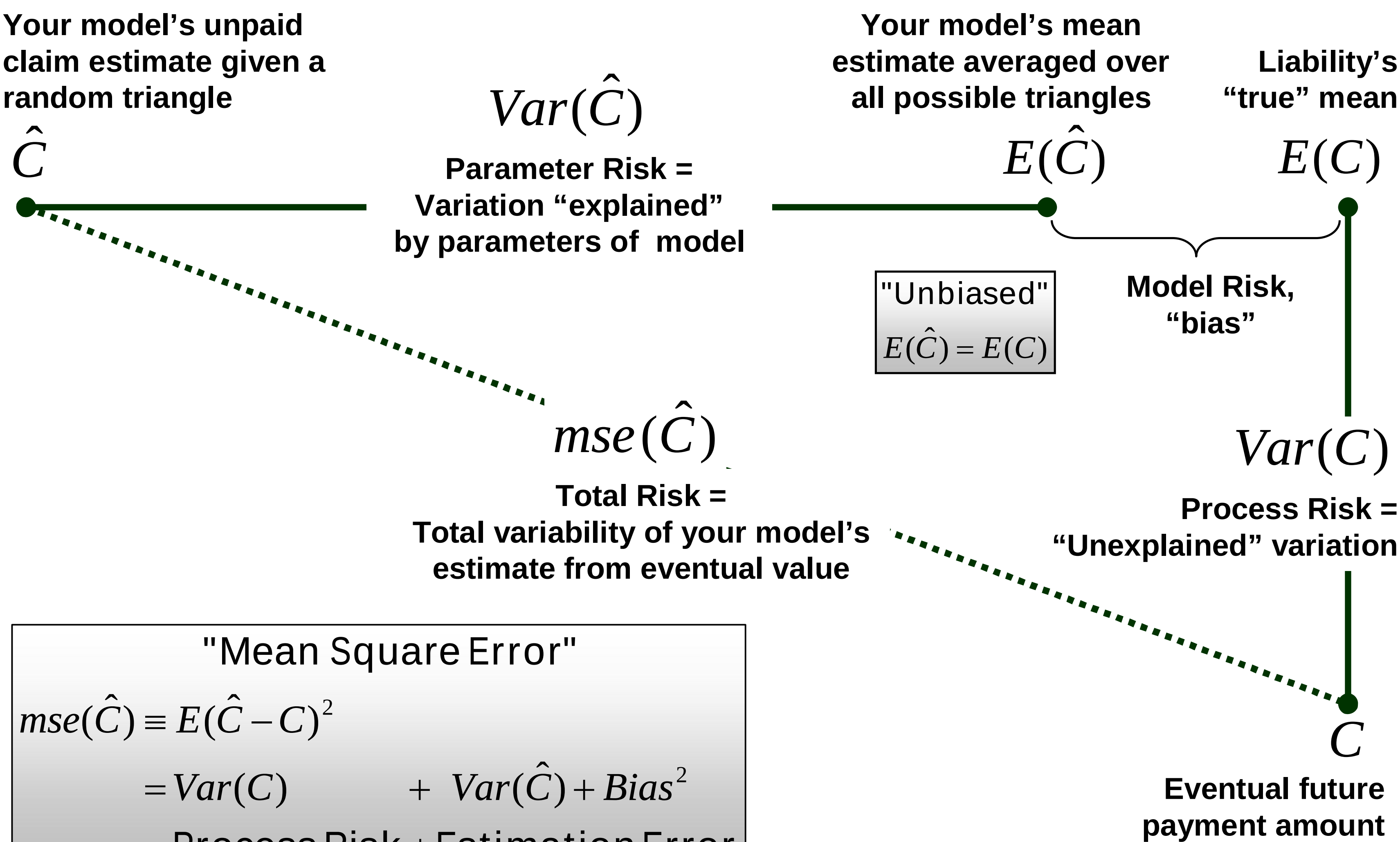
AY\Age	1	2
1999	10,238	24,654
2000	5,508	16,235
2001	7,374	20,620
2002	6,153	19,182
2003	7,253	25,066
2004	10,855	38,520
2005	10,313	34,341
2006	16,411	42,228
2007	21,234	

Variance of Age-2 Losses: 91,785,515

AY\Age	LDF
1999	2.408
2000	2.948
2001	2.796
2002	3.118
2003	3.456
2004	3.549
2005	3.330
2006	2.573
Simple	3.022
Weighted	2.980
5-Yr Wtd	3.125
3-Yr SA	3.151
Selected	3.125

2007 Estimated Age-2 Loss 66,360

Total Risk is statistical equivalent of Pythagorean Theorem



"Mean Square Error"

$$mse(\hat{C}) \equiv E(\hat{C} - C)^2$$

$$= Var(C) \quad + \quad Var(\hat{C}) + Bias^2$$

$$= \text{Process Risk} + \text{Estimation Error}$$

A Risk by any other name …

- The word **Risk** can be ambiguous
  - Layperson uses **risk** broadly
  - For a "quant," **risk** usually refers to "variance" or "standard deviation"
  - Other terms
    - — Value at Risk (VaR) is a quantile (e.g., the 75th percentile)
    - — Tail Value at Risk (TVar) is the expected value of tail losses
- Most stochastic methods estimate risk of ultimate loss first, then back into risk of outstanding loss
  - *std(OS) = std(Ult - Paid) = std(Ult)* because paid loss is a scalar
- **Coefficient of Variation**, or CV, is a popular measure of relative risk
  - $cv(X) = \frac{std(X)}{mean(X)} = \frac{\sigma_X}{\mu_X}$
- **Scalability**
  - Often, *cv(X)* determined from one stochastic method is applied to the mean μ_Y of another method – or a carried reserve – to impute the standard deviation of the other method/reserve:
    - — If reasonable to assume $cv(Y) = cv(X)$ $\left(\text{ie, } \frac{\sigma_Y}{\mu_Y} = \frac{\sigma_X}{\mu_X} \right)$, then $\sigma_Y = \frac{\sigma_X}{\mu_X} \mu_Y$
    - — Justification: lognormals with same shape parameter σ have same cv

Agenda

- Why Analyze Reserve Ranges?
- General Terminology
- **Popular Stochastic Methods**
  - **Mack/Murphy**
  - Monte Carlo Simulation
  - Bootstrapping
- Aggregation

Mack/Murphy Method: Overview

- The Mack/Murphy method derives formulas for standard error of the chain ladder unpaid claim estimate
  - Method only uses data in the triangle
  - Tail variability beyond the triangle can be incorporated in various ways
- There are formulas for parameter risk, process risk, and total risk
- This is an analytic calculation, analogous to finding the standard deviation of a random sample
- Given the central estimate of the unpaid claim liability and this method's standard error,
  - One can fit almost any two-parameter probability distribution to model the *distribution* of unpaid claims
  - Mack recommends normal or lognormal, Murphy suggests student-t
  - Benefit: confidence levels, VaR's, TVars, etc.

But first, the M/M model and Mack's formula

- Mack and Murphy start with the same three simple assumptions

  (CL1) $E(C_{i,k+1} \mid \text{the triangle}) = C_{ik} f_k$

  (CL2) $\boldsymbol{Var}(\boldsymbol{C}_{i,k+1} \mid \text{the triangle}) = \boldsymbol{C}_{ik}\sigma_k^2$ for unknown parameters σ_k^2

  (CL3) accident years are independent

- Mack derives the closed-form formula

$$m\hat{s}e(\hat{C}_{iI}) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

  where

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1} \sum_{i=1}^{I-k} C_{ik} \left(\frac{C_{i,k+1}}{C_{ik}} - \hat{f}_k \right)^2,$$

  the "*f*-hats" are the weighted average link ratios, *I* = # AYs, *k* denotes age, and the "*C*-hats" are the chain ladder estimates of future loss for accident yr *i*.

- Mack's formula is a thing of beauty!
- Murphy's formula is not closed-form but recursive, with an extra term in the parameter risk formula

M/M Method Example: Point Estimate

ABC Insurance Company
Chain Ladder Projection of Paid Losses

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	10,238	24,654	38,025	46,550	52,842	58,722	65,227	67,604	69,559
2000	5,508	16,235	25,586	32,863	38,111	42,315	45,171	47,666	49,045
2001	7,374	20,620	34,220	43,438	50,898	55,475	58,367	60,943	62,706
2002	6,153	19,182	31,005	40,424	46,949	50,942	54,931	57,354	59,014
2003	7,253	25,066	40,134	51,063	58,376	64,144	69,166	72,218	74,307
2004	10,855	38,520	62,348	82,710	95,382	104,806	113,011	117,998	121,411
2005	10,313	34,341	51,110	65,632	75,688	83,166	89,677	93,634	96,343
2006	16,411	42,228	66,770	85,743	98,879	108,649	117,155	122,324	125,863
2007	21,234	66,360	104,927	134,743	155,386	170,740	184,106	192,230	197,791
sum below diagonal	0	66,360	171,697	286,118	425,335	531,505	628,046	716,701	786,478
								Total O/S =	**373,845**
	5 year								
RTRs	3.125	1.581	1.284	1.153	1.099	1.078	1.044	1.029	
LDFs	9.315	2.981	1.885	1.468	1.273	1.158	1.074	1.029	1.000

Original data triangle

Estimated future values based on weighted average RTR factors

Total variance is estimated recursively ala Murphy

ABC Insurance Company
Total Risk
Based on Chain Ladder Projection of Paid Losses

Variance ($\$^2$)

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999									
2000									1,902,220
2001								2,088,335	4,923,348
2002							1,580,389	3,648,369	6,344,119
2003						1,788,756	4,196,447	7,208,241	11,127,982
2004					1,125,201	4,718,811	9,582,414	15,800,096	24,311,856
2005				3,419,436	5,389,037	8,988,368	13,431,950	18,448,005	24,759,471
2006			4,769,700	12,580,548	17,909,193	25,149,906	33,550,992	42,230,778	52,757,030
2007		61,452,982	161,965,332	275,457,646	368,485,450	451,544,103	533,388,343	593,026,502	645,188,188
sum below diagonal		61,452,982	169,601,444	303,662,011	413,691,192	539,824,418	692,048,696	870,481,277	1,141,875,645

mse = $ 33,792

	2	3	4	5	6	7	8	9 = Ult
σ_k	*45.201*	*9.558*	*7.402*	*3.133*	*4.838*	*4.838*	*4.838*	*4.838*
σ_β	*0.200*	*0.023*	*0.015*	*0.007*	*0.011*	*0.012*	*0.015*	*0.019*

- Parameter variance, process variance, and their sum = total variance are calculated separately using Murphy recursive formulas
- Mack's closed-form formula gives ultimate, not intermediate, values
- σ_k we saw in Mack's formula, σ_β = standard deviation of VW RTR

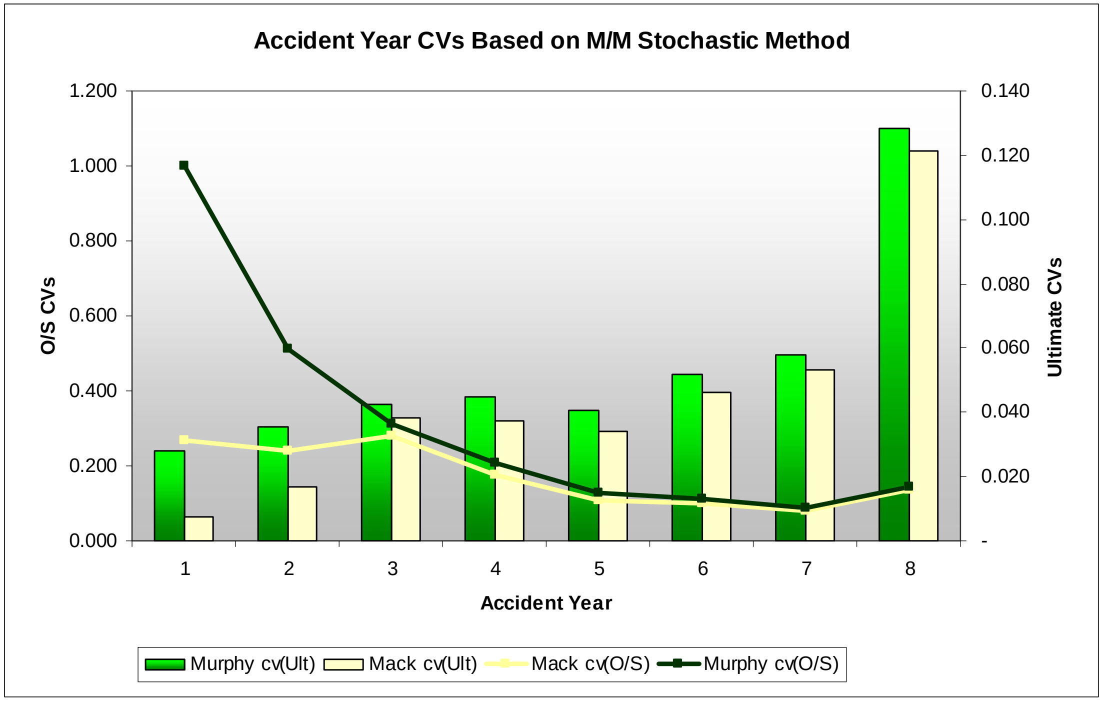
M/M Method Example:
Numerical summary that actuaries love to see

ABC Insurance Company
M/M Stochastic Analysis based on Chain Ladder Projection of Paid Losses

			Murphy s.e. Formulation			cv of O/S Loss			
AY (i)	**Ultimate_i**	**O/S_i**	**Proc risk**	**Parm risk**	**Total risk**	Process	Parameter	Total	Tot cv of Ultimate
1999	69,559	-	-	-	-				
2000	49,045	1,379	1,056	887	1,379	0.766	0.643	1.000	0.028
2001	62,706	4,338	1,695	1,432	2,219	0.391	0.330	0.511	0.035
2002	59,014	8,071	2,020	1,505	2,519	0.250	0.186	0.312	0.043
2003	74,307	15,931	2,640	2,039	3,336	0.166	0.128	0.209	0.045
2004	121,411	38,701	3,564	3,407	4,931	0.092	0.088	0.127	0.041
2005	96,343	45,233	4,014	2,940	4,976	0.089	0.065	0.110	0.052
2006	125,863	83,635	5,896	4,242	7,263	0.070	0.051	0.087	0.058
2007	197,791	176,557	20,977	14,323	25,401	0.119	0.081	0.144	0.128
Total:	786,478	**373,845**	22,774	24,964	**33,792**	0.061	0.067	**0.090**	0.043

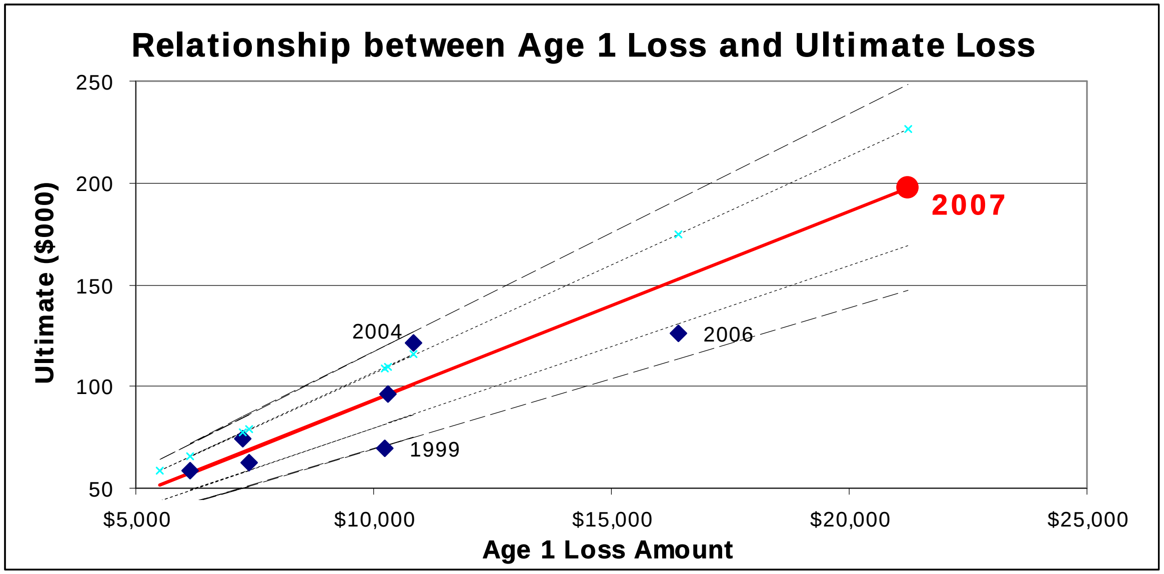
- Here, “risk” = standard deviation (“standard error”)
  - Total Risk is the square root of the variances in the ultimate column on previous slide
- Total cv of O/S and of Ultimate use same standard error in the numerator; only denominators differ

M/M Method Example: Graphical summary displays "smile" of cv(O/S) and "blow up" of cv(Ultimate)



- The further an accident year from ultimate resolution, the more relative uncertainty in its estimated ultimate value
- Same cannot be said for cv metric for outstanding loss
  - Mature AY cv's of O/S are larger due to smaller amounts in denominator
- Slide illustrates slight difference between Mack and Murphy
  - Formulation of parameter risk
  - Treatment of limited data in tail

Familiar regression graph illustrates M/M theory



- Graph illustrates actuary's selected linear relationship

  $y = 9.315\,x$

- Statistical theory cautions against extrapolations beyond experience interval

- The "true" linear relationship will almost certainly be different from our selection
  - Dotted line = ±2 parameter risks (standard errors)
- In addition, whatever the linear relationship "truly" is, actual results will deviate from that mean value
  - Dashed line = ±2 total risks (s.e.'s)
- Recall: ±2 standard deviations enclose a 98% confidence interval (standard normal)

M/M VaR estimates of total outstanding can vary substantially depending on assumed probability distribution

ABC Insurance Company
M/M Stochastic Analysis Based on Paid Loss

			Murphy s.e. Formulation		
AY (i)	**$Ultimate_i$**	**O/S_i**	**Proc risk**	**Parm risk**	**Total risk**
Total:	786,478	**373,845**	22,774	24,964	**33,792**

From slide 16

VaR %-ile (p)	.994 ('A'*)	.999 ('AAA'*)
Norminv(p, 373845, 33792)	$458,735	$478,270
Loginv(p, 12.828, 0.090)	$467,025	$492,025
difference	~$8,300	~$13,800

- Even assuming the more skewed distribution, the Mack/Murphy method has been criticized for understating tail risk (GIRO working party, July 2007)

* S&P confidence levels under a one year time horizon

Mack/Murphy method: wrapup

- Advantages
  - Strictly analytical method, no simulation required
  - Instantaneously fast
  - Well-known
- Disadvantages
  - Not robust to outliers
  - Probably understates cv, tail variability
    - See 2007 GIRO report
  - Overparameterizes the data
  - Does not necessarily model situation when actuary selects factors other than weighted or simple average

Agenda

- Why Analyze Reserve Ranges?
- General Terminology
- **Popular Stochastic Methods**
  - Mack/Murphy
  - **Monte Carlo Simulation**
  - Bootstrapping
- Aggregation

Monte Carlo Simulation takes a Methodizer's approach

- A simulation of the distribution of unpaid claims generally follows these steps:
  - Choose a deterministic method (a **process**) to generate ultimate loss outcomes
    - — Model inputs as random variables
  - Randomly generate input values
  - Calculate ultimate outcome, unpaid claim value (save result!)
  - Repeat many many times
  - Empirical distribution estimates the theoretical distribution of unpaid claims
    - — Reflects variability of assumed process (**process risk**)
- Chosen method determines types of inputs to simulate
  - Chain ladder, BF: link ratios (or link_ratio–1) usually simulated as normal or lognormal random variables
  - Pure premium, BF: loss ratios usually simulated as lognormals
- Parameters (e.g., μ and σ) of simulated distributions must be selected before random draws can occur
  - μ, σ estimated from the data, selected from benchmarks, judgment
  - How to reflect risk that selected μ, σ might not equal "true" value (**parameter risk**)?
    - — Various approaches exist in the literature (e.g., Kreps, *PCAS* 1997; see also Hodes, Feldblum, and Blumsohn, *PCAS* 1999)
    - — Beyond scope of this presentation

Monte Carlo Simulation of the Loss Development Method: All losses at age k use the same simulated RTR at age k

ABC Insurance Company
Chain Ladder Simulation of Paid Losses

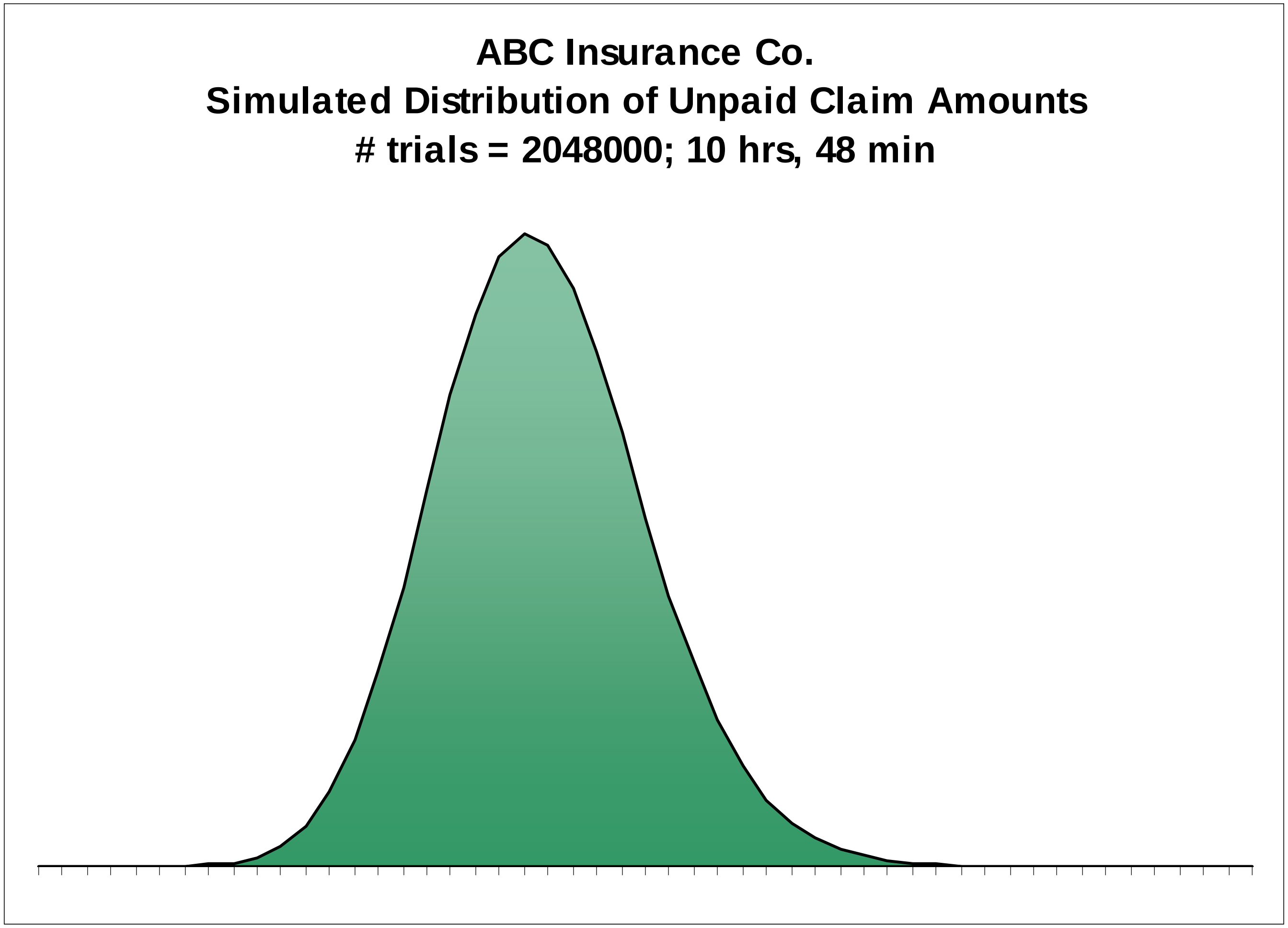
AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	10,238	24,654	38,025	46,550	52,842	58,722	65,227	67,604	69,559
2000	5,508	16,235	25,586	32,863	38,111	42,315	45,171	47,666	48,651
2001	7,374	20,620	34,220	43,438	50,898	55,475	58,367	60,119	61,000
2002	6,153	19,182	31,005	40,424	46,949	50,942	54,371	55,998	58,579
2003	7,253	25,066	40,134	51,063	58,376	65,348	70,832	73,891	76,454
2004	10,855	38,520	62,348	82,710	95,125	105,683	111,751	116,266	119,504
2005	10,313	34,341	51,110	65,000	75,571	82,604	89,805	93,425	101,151
2006	16,411	42,228	67,339	87,368	100,911	110,832	118,540	122,824	127,805
2007	21,234	68,053	105,940	135,776	156,966	171,817	183,325	192,057	195,364
sum below diagonal		68,053	173,279	288,144	428,573	536,284	628,624	714,579	788,509
									375,876 =Est'd O/S
	5 year								
Selected	3.125	1.581	1.284	1.153	1.099	1.078	1.044	1.029	
LDFs	9.315	2.981	1.885	1.468	1.273	1.158	1.074	1.029	1.000

Steps to simulate RTRs

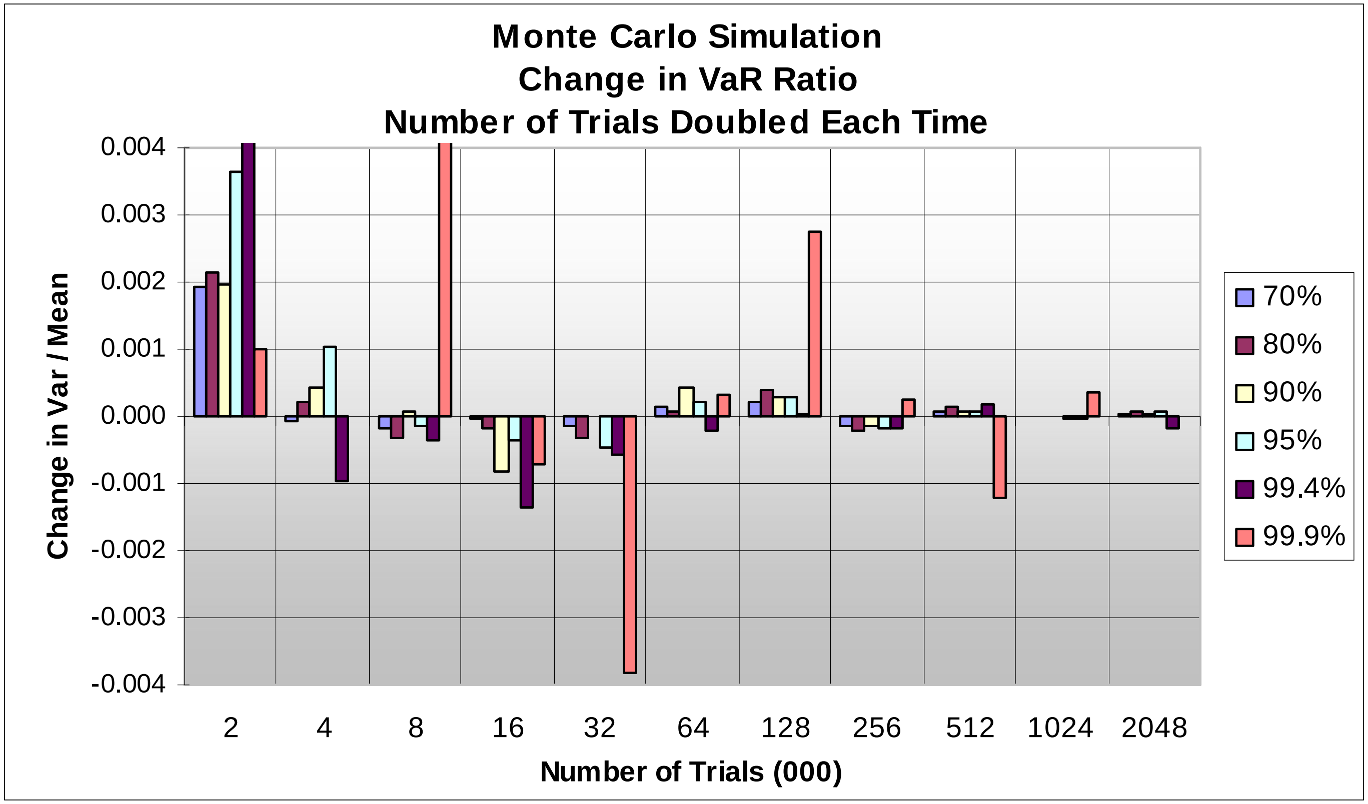
	1	2	3	4	5	6	7	8	
RTR-1	2.125	0.581	0.284	0.153	0.099	0.078	0.044	0.029	
Parameter risk cv's from M/M formula; will assume appropriate for selected 1-2 RTR:									
cv_β	0.079	0.039	0.054	0.044	0.113	0.156	0.330	0.643	
σ^2	0.006	0.00151	0.00293	0.00195	0.01262	0.02411	0.10336	0.34622	$=\ln(1+cv^2)$
μ	0.751	-0.543	-1.260	-1.877	-2.321	-2.559	-3.172	-3.716	$=\ln(\text{mean})-\sigma^2/2$
Sim'd RTR	3.205	1.595	1.272	1.150	1.119	1.067	1.030	1.021	=loginv(rand(),μ,σ)+1

- Values in first subsequent diagonal use Sim'd RTRs from box
- Subsequent cells have that formula buried within

Simulated distribution mellows over time



Finding minimum number of MC trials has no neat solution



- Compared with a pdf graph, a convergence graph gives a better picture of how closely statistics of interest settle down
  - E.g., if you want to estimate the 99.9th percentile to within 0.1% of the mean, you should run at least 1 million iterations
  - For the 70th percentile, just 4,000 iterations might be sufficient

Monte Carlo simulation: wrapup

- Advantages
  - Well-known in many sciences
  - Extremely flexible
    - — Technique can model highly complex processes
- Disadvantages
  - Parameters describing the process inputs must be selected ahead of time
  - Slow to execute
    - — Quantity of random deviates increases with complexity
    - — Number of trials increases with complexity
  - Without extra steps, only measures process risk
    - — Ask your consultant if/how parameter risk is incorporated

Agenda

- Why Analyze Reserve Ranges?
- General Terminology
- **Popular Stochastic Methods**
  - Mack/Murphy
  - Monte Carlo Simulation
  - **Bootstrapping**
- Aggregation

Bootstrapping is a “modern” simulation technique

- Whereas Monte Carlo simulates parametric *inputs* (e.g., LDFs) to a complex process, Bootstrapping simulates the *data*
- If the data’s distribution is “known,” then
  - Estimate parameters of the distribution
  - Sample *from that distribution*
  - Calculate your desired output
  - Repeat
  - Make inferences (cv, VaR) from output’s empirical distribution
  - Called **Parametric Bootstrapping**
- If cannot assume data’s distribution is of a specific type
  - Sample *from the data itself* (with replacement; “resample”)
  - Continue as above
  - Called **Nonparametric Bootstrapping**
- When the process is modelled, can **bootstrap the residuals**
  - Resample the residuals
  - Recast history with synthetic data
  - Continue as above

Simple example of non-parametric bootstrapping of residuals: estimate the distribution of AY 2008 paid as of 12 months

- Same data as before, now with premium
- Based on a selected 12-month paid loss ratio of 22%, a **deterministic estimate** of AY 2008's paid amount at 12/31/08 is about $21.6 million
- To get a range around that point estimate, we could try good old fashioned Monte Carlo simulation:
  - Model the partial L/Rs as lognormals
  - Fit μ and σ to the 8 L/R data points
  - Randomly draw from the fitted L/R distribution
  - Calculate the indicated 12-month paid loss (the desired "output")
  - Repeat many times
  - Builds the empirical distribution of AY 2008 paid @ 12 months
- How would bootstrapping differ from this approach?

AY \ Age	Earned Premium A	Paid @ 12 mo B	Partial L/R B/A
1999	61,428	10,238	16.7%
2000	16,524	5,508	33.3%
2001	43,297	7,374	17.0%
2002	24,016	6,153	25.6%
2003	21,100	7,253	34.4%
2004	139,052	10,855	7.8%
2005	59,535	10,313	17.3%
2006	67,864	16,411	24.2%
2007	128,447	21,234	16.5%
Average			21.4%
C. Selected			**22.0%**
		deterministic	
2008	98,156	**21,594**	

With Bootstrapping we create synthetic historical data and select as we normally would in our deterministic analysis

Usual Deterministic Estimate				Based on Selected		Simulated Deterministic Estimate			
AY \ Age	Earned Premium	Paid @ 12 mo	Partial L/R	Expected Paid @ 12 mo	Residuals	"Resampled" Residuals index	"Resampled" Residuals Residuals	"pseudo" Paid @ 12 mo	Partial L/R
	A	B	B/A	D=AC	E=B-D	=random	F	G=D+F	H=G/A
1999	61,428	10,238	16.7%	13,514	(3,276)	4	870	14,384	23.4%
2000	16,524	5,508	33.3%	3,635	1,873	4	870	4,505	27.3%
2001	43,297	7,374	17.0%	9,525	(2,151)	3	(2,151)	7,374	17.0%
2002	24,016	6,153	25.6%	5,283	870	8	1,481	6,764	28.2%
2003	21,100	7,253	34.4%	4,642	2,611	4	870	5,511	26.1%
2004	139,052	10,855	7.8%	30,591	(19,736)	1	(3,276)	27,315	19.6%
2005	59,535	10,313	17.3%	13,098	(2,785)	5	2,611	15,709	26.4%
2006	67,864	16,411	24.2%	14,930	1,481	5	2,611	17,541	25.8%
2007	128,447	21,234	16.5%	28,258	(7,024)	6	(19,736)	8,522	6.6%
Average			21.4%						22.3%
C. Selected			**22.0%**						**23.0%**
		deterministic						simulated	
2008	98,156	**21,594**						**22,576**	

Legend
actual data
calculations
synthetic data
selection

- Given selected L/R, calculate the residuals ("noise") that we assume could have happened at any time
- Resample residuals to create synthetic historical data
- Given that data, select a L/R and calculate the paid amount
- Repeat to generate a set of actuarial central estimates for paid @ 12 months
  - Seen that way, Bootstrapping measures **parameter risk**

Nonparametric bootstrapping applied to Chain Ladder model also bootstraps the residuals

- From standard chain ladder approach, back into expected cumulative

ABC Insurance Company
Bootstrapping Paid Loss Development Method

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	10,238	24,654	38,025	46,550	52,842	58,722	65,227	67,604	69,559
2000	5,508	16,235	25,586	32,863	38,111	42,315	45,171	47,666	49,045
2001	7,374	20,620	34,220	43,438	50,898	55,475	58,367		62,706
2002	6,153	19,182	31,005	40,424	46,949	50,942			59,014
2003	7,253	25,066	40,134	51,063	58,376				74,307
2004	10,855	38,520	62,348	82,710					121,411
2005	10,313	34,341	51,110						96,343
2006	16,411	42,228							125,863
2007	21,234								197,791
sum below diagonal									786,478
									373,845 =Est'd O/S

	5 year								
Selected	3.125	1.581	1.284	1.153	1.099	1.078	1.044	1.029	tail
LDFs	9.315	2.981	1.885	1.468	1.273	1.158	1.074	1.029	1.000

Expected cumulative paid amounts based on selected link ratios

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	*7,468*	*23,337*	*36,901*	*47,386*	*54,646*	*60,046*	*64,747*	*67,604*	69,559
2000	*5,265*	*16,455*	*26,018*	*33,411*	*38,530*	*42,337*	*45,651*	47,666	
2001	*6,732*	*21,038*	*33,265*	*42,717*	*49,262*	*54,130*	58,367		
2002	*6,335*	*19,799*	*31,306*	*40,202*	*46,362*	50,942			
2003	*7,977*	*24,930*	*39,420*	*50,621*	58,376				
2004	*13,034*	*40,734*	*64,408*	82,710					
2005	*10,343*	*32,324*	51,110						
2006	*13,512*	42,228							
2007	21,234								

Residuals are based on incremental, not cumulative, paids

Incremental Actuals

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	8,239	14,416	13,371	8,525	6,292	5,880	6,505	2,377	1,956
2000	3,508	10,727	9,351	7,277	5,248	4,204	2,856	2,495	
2001	5,373	13,246	13,600	9,218	7,460	4,577	2,892		
2002	4,151	13,029	11,823	9,419	6,525	3,993			
2003	5,250	17,813	15,068	10,929	7,313				
2004	8,851	27,665	23,828	20,362					
2005	8,308	24,028	16,769						
2006	14,405	25,817							
2007	19,227								

Incremental Expecteds

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	*5,469*	*15,870*	*13,563*	*10,485*	*7,260*	*5,399*	*4,701*	*2,857*	*1,956*
2000	*3,265*	*11,190*	*9,563*	*7,393*	*5,119*	*3,807*	*3,314*	*2,014*	
2001	*4,731*	*14,306*	*12,227*	*9,452*	*6,545*	*4,867*	*4,238*		
2002	*4,333*	*13,464*	*11,507*	*8,896*	*6,159*	*4,581*			
2003	*5,974*	*16,953*	*14,489*	*11,201*	*7,755*				
2004	*11,030*	*27,700*	*23,674*	*18,302*					
2005	*8,338*	*21,981*	*18,786*						
2006	*11,506*	*28,716*							
2007	*19,227*								

Incremental Residuals

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	2,770	(1,454)	(192)	(1,960)	(968)	481	1,804	(480.37)	
2000	243	(463)	(212)	(116)	129	397	(458)	480	
2001	642	(1,060)	1,373	(234)	915	(290)	(1,346)		
2002	(182)	(435)	316	523	366	(587)			
2003	(724)	860	579	(272)	(442)				
2004	(2,179)	(35)	154	2,060					
2005	(30)	2,047	(2,017)						
2006	2,899	(2,899)							
2007	0								

Residuals are generally ill-behaved without their massage

- Here the residuals are massaged into “scaled pearson residuals”
  - “Standardized” residuals can be scrambled among all AY’s and ages
  - We’ll skip the details … (and the massage)
- Resample with replacement

Random index

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	23	15	39	14	3	40	28	1	8
2000	6	17	29	10	14	26	43	37	
2001	5	5	21	41	19	42	42		
2002	24	44	23	20	19	20			
2003	23	19	44	23	33				
2004	4	22	16	26					
2005	39	42	16						
2006	10	5							
2007	35								

"Resampled" residuals (pretend post-massage)

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	*154*	*(35)*	*(587)*	*860*	*642*	*1,804*	*523*	*2,770*	*2,899*
2000	*(2,179)*	*(2,899)*	*(272)*	*(1,454)*	*860*	*(116)*	*(480)*	*397*	
2001	*(724)*	*(724)*	*316*	*(458)*	*(212)*	*(1,346)*	*(1,346)*		
2002	*(2,017)*	*480*	*154*	*1,373*	*(212)*	*1,373*			
2003	*154*	*(212)*	*480*	*154*	*915*				
2004	*(182)*	*579*	*2,047*	*(116)*					
2005	*(587)*	*(1,346)*	*2,047*						
2006	*(1,454)*	*(724)*							
2007	*(442)*								

Recast historical cumulative triangle, reproject ultimates

- Use resampled residuals to create synthetic incremental amounts
- Re-cumulate
- Now you have a synthetic triangle from which to pick factors and project ultimates, unpaid amount

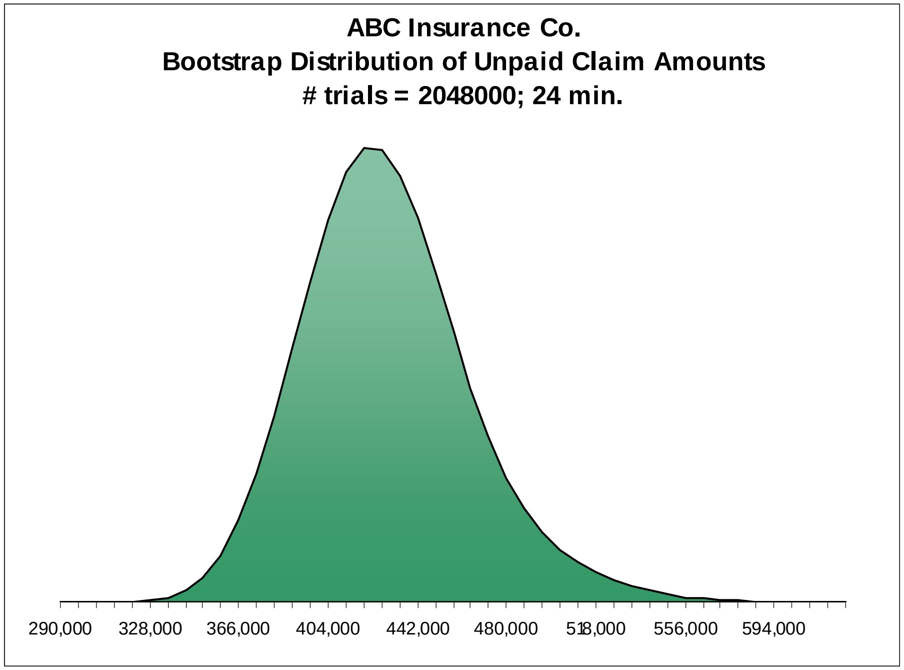
Synthetic incremental amounts

AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	*5,623*	*15,835*	*12,976*	*11,345*	*7,902*	*7,204*	*5,224*	*5,627*	*4,854*
2000	*1,086*	*8,291*	*9,291*	*5,939*	*5,979*	*3,691*	*2,834*	*2,411*	
2001	*4,007*	*13,582*	*12,543*	*8,994*	*6,332*	*3,522*	*2,892*		
2002	*2,316*	*13,944*	*11,661*	*10,269*	*5,947*	*5,954*			
2003	*6,128*	*16,741*	*14,970*	*11,355*	*8,671*				
2004	*10,848*	*28,279*	*25,721*	*18,186*					
2005	*7,751*	*20,635*	*20,833*						
2006	*10,052*	*27,991*							
2007	*18,785*								

Synthetic cumulative amounts

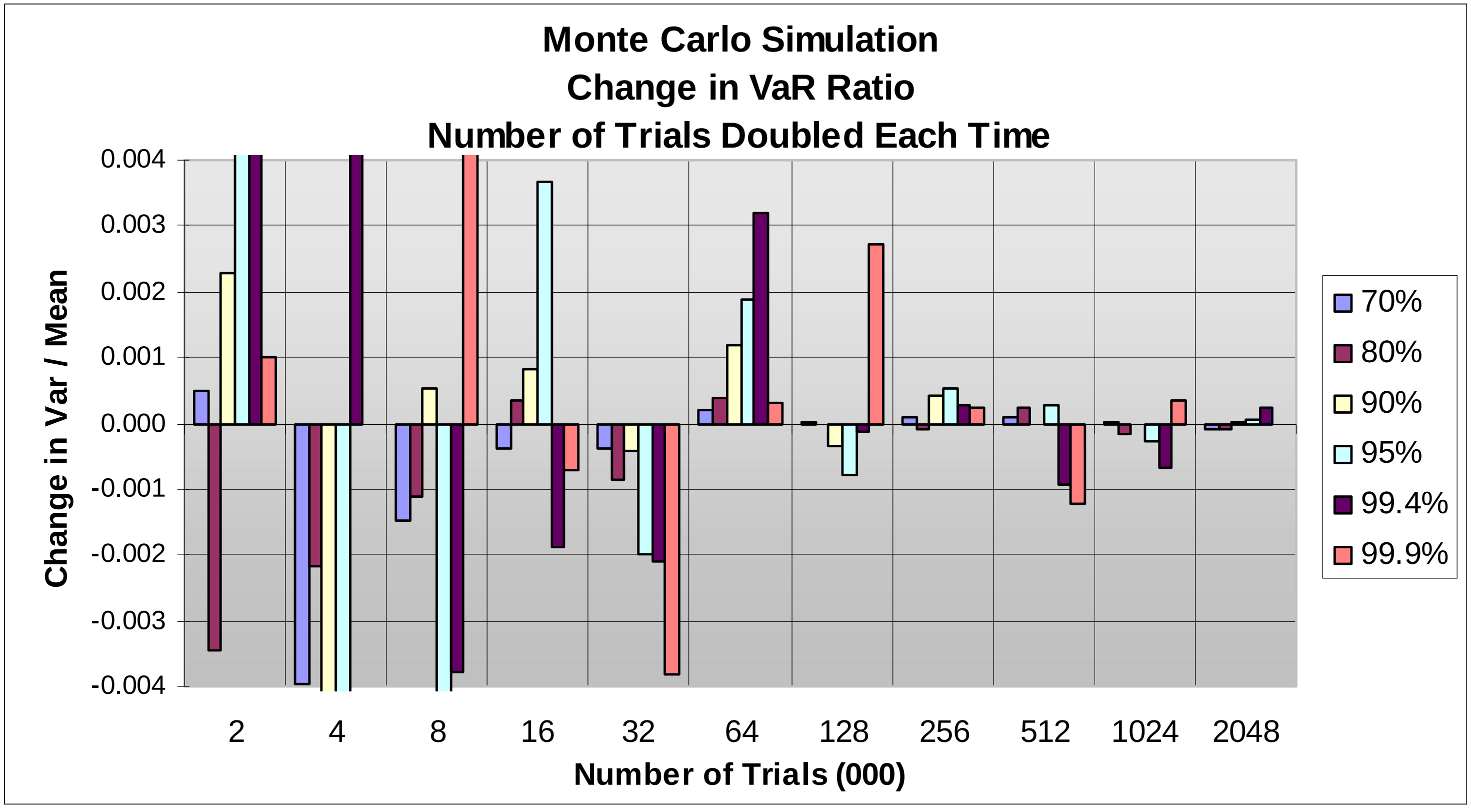
AY \ Age	1	2	3	4	5	6	7	8	9 = Ult
1999	*5,623*	*21,458*	*34,434*	*45,779*	*53,681*	*60,885*	*66,109*	*71,736*	*76,591*
2000	*1,086*	*9,377*	*18,668*	*24,607*	*30,586*	*34,277*	*37,111*	*39,522*	42,197
2001	*4,007*	*17,588*	*30,131*	*39,125*	*45,458*	*48,979*	*51,871*		59,695
2002	*2,316*	*16,261*	*27,922*	*38,191*	*44,138*	*50,091*			62,025
2003	*6,128*	*22,869*	*37,839*	*49,194*	*57,865*				80,046
2004	*10,848*	*39,126*	*64,848*	*83,033*					135,181
2005	*7,751*	*28,386*	*49,219*						104,893
2006	*10,052*	*38,043*							137,544
2007	*18,785*								264,894
sum below diagonal									886,474
									498,043
	5 year								
Selected	3.900	1.696	1.309	1.177	1.117	1.076	1.078	1.068	tail
LDFs	14.102	3.615	2.131	1.628	1.383	1.238	1.151	1.068	1.000

Bootstrap distribution appears slightly skewed



- Bootstrapping takes less time than MC to run same number of trials
- The Bootstrapping mean is about $425,000 whereas the deterministic estimate was about $375,000
  - Not uncommon
  - Usual solution is to scale the empirical output so the means coincide

Finding minimum number of MC trials has no neat solution

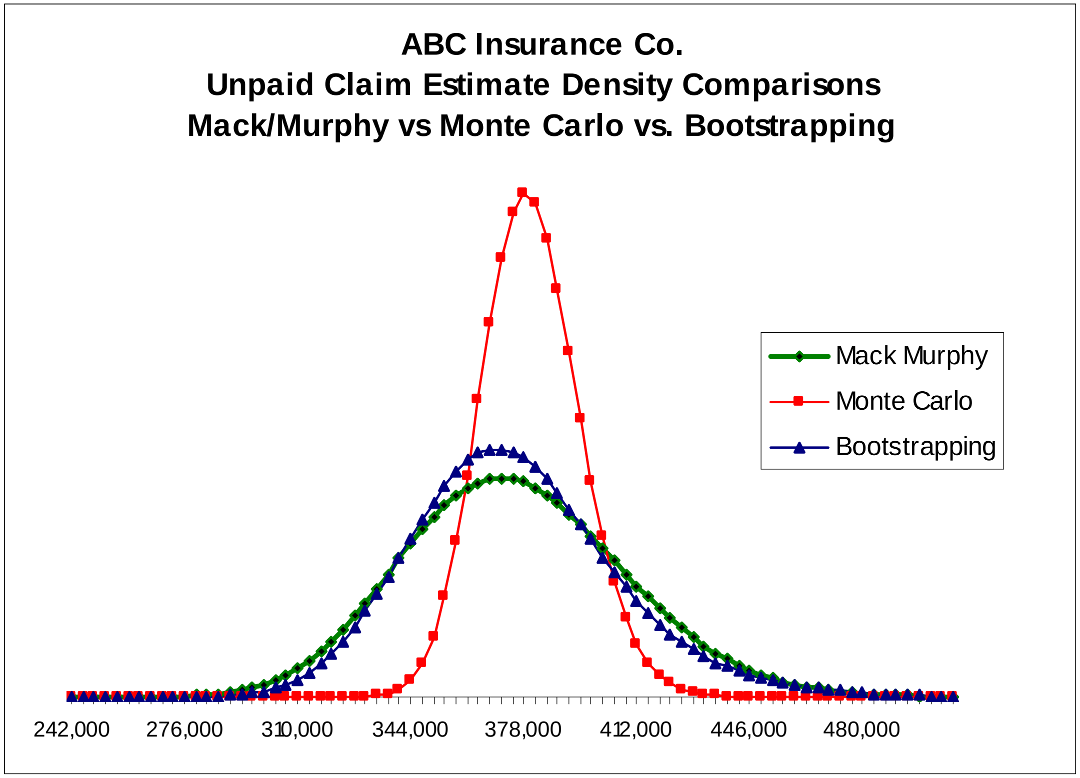


- For the same level of tolerance, Bootstrapping needed 2M trials where MC needed only 1M
- Might be wise to double the trials again to make sure the empirical data has actually stabilized; relatively quick

Bootstrapping: wrapup

- Advantages
  - The data drives the simulation
    - — No assumptions necessary for process input parameters
  - Faster than MC simulation (although more trials may be necessary)
- Disadvantages
  - Without extra steps, only measures parameter risk
    - — Ask your consultant if/how process risk is incorporated
  - Misconception: “Bootstrapping doesn’t work with incurred losses!”
    - — Some methods of massaging the residuals are better able to allow for negative development
  - Heteroscedasticity
    - — When development data is widely different by age (e.g., E&S data), separate residuals into similar resampling groups
    - — Plot residuals by age for visual cues
  - Bootstrap mean may not equal deterministic mean
    - — Scale the output
  - Bootstrapping has also been criticized by GIRO for understating tail variability

Summary of popular methods' results: MM and BS densities appear similar, MC not so



- Without parameter risk, MC clearly understates mean square error
- MM may have more spread than BS, but BS here omits process risk

Agenda

- Why Analyze Reserve Ranges?
- General Terminology
- Popular Stochastic Methods
- **Aggregation**

Here are a few familiar formulas when aggregating lines

- Mean of aggregate is the aggregate of the marginals' means
  - $E(X+Y) = E(X) + E(Y)$
- Variance of the aggregate is the aggregate of marginals' variances with an extra cross-product term
  - $V(X+Y) = V(X) + 2Cov(X,Y) + V(Y)$ (1)
  - Analogous to $(x+y)^2 = x^2 + 2xy + y^2$
- Confidence level (e.g., 75% VaR) of the aggregate is the aggregate of marginals' confidence levels less the *diversification benefit*
  - $VaR_{75\%}(X+Y) = VaR_{75\%}(X) + VaR_{75\%}(Y) - DB_{75\%}$
  - Q: When would $DB_{75\%} = 0$?
- Distribution of the aggregate is the aggregate of marginals' distributions
  - $F_{X+Y} = Copula(F_X, F_Y)$

More aggregation formulas: correlation

- Correlation scales the covariance of two lines by dividing by their standard deviations

$$\text{Correlation Coefficient}: \quad \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

  - Allows comparison of two lines of difference sizes

- Such relationships between N lines of business are encapsulated in the covariance matrix and the correlation matrix

$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & \cdots & Cov(X_1,X_N) \\ Cov(X_2,X_1) & Var(X_2) & \cdots & Cov(X_2,X_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_N,X_1) & Cov(X_N,X_2) & \cdots & Var(X_N) \end{bmatrix} \quad Corr = \begin{bmatrix} 1 & corr(X_1,X_2) & \cdots & corr(X_1,X_N) \\ corr(X_2,X_1) & 1 & \cdots & corr(X_2,X_N) \\ \vdots & \vdots & \ddots & \vdots \\ corr(X_N,X_1) & corr(X_N,X_2) & \cdots & 1 \end{bmatrix}$$

  - These matrices should always be positive-semidefinite (you can take their “square root”)
    - — Can find the “square root” of Σ using *Cholesky decomposition*
    - — Used to simulate correlated random variables

A simple example of the "correlation matrix" approach

- Suppose monoline ABC Insurance Co. writes in two states, X and Y
- Ran the Mack / Murphy method on state X
- The reserves for Y are half those for X
- Assume cv for Y is same as for X
- Correlation based on judgment

	A	B	C	D	E	F
1		**Outstanding Loss**				
2		**Mean**	**CV**	**S.E.**	**mse**	**75%VaR**
3						
4	**Line X**	**373,845**	**0.090**	**33,792**	1,141,875,645	395,684
5	**Line Y**	186,923	0.090	16,896	285,468,911	197,842
6						
7	**Correlation**	80%		**Cov->**	456,750,258	
8						
9	**X+Y**	560,768	0.086	48,382	2,340,845,072	592,106
10						
11	Calculations:	E4=D4^2				
12		E5=D5^2				
13		E7=B7*D4*D5				
14		E9=E4+E5+2*E7				
15		F9 based on lognormal distribution with mean=B9, stdev=D9				
16	Notes:	Assume B5=B4/2				
17		Assume C5=C4. So D5=B5*C5				
18		Back into standard error of sum: = sqrt(E9)				
19		and cv of sum = D9/B9				

- As a standard practice an actuary might assume that a lognormal distribution applies to the aggregated lines
  - Fit to the calculated mean (560768) and standard error (44702)
  - Read 75% VaR off the indicated lognormal distribution
- The diversification benefit at the 75% confidence level would be about $4million

Above correlation approach violates basic mathematics

- In the preceding example we assumed that the individual lines X and Y had lognormally distributed unpaid claim amounts
- We also assumed that the unpaid claim amounts of the combined portfolio X+Y were also lognormally distributed
  - Unfortunately, the sum of two lognormals is not a lognormal
  - Loss of precision remains to be seen
- We seek a more flexible way of combining two marginal distributions into a joint distribution when
  - We have reasonable knowledge of the marginals
  - We have some idea of the strength of the dependency between the lines
  - We want to make as few additional assumptions as necessary
- Enter: the Copula!

The word *copula* comes from the discipline of Logic

- Mirriam-Webster's Dictionary defines *copula* as "the connecting link between the subject and predicate of a proposition"
- Sklar's Theorem (1959) (abbreviated):
  For every joint distribution $F_{X,Y}$ there exists a function C that breaks down $F_{X,Y}$ into its marginals F_X and F_Y:
  $$F_{X,Y}(x,y)=C(F_X(x),F_Y(y))$$
- In practice, copulas are applied in reverse, i.e.,
  - Starting with the marginals, pick a copula and form the joint distribution
- Copula you choose is unrelated to distributional form of the marginals
  - Each is simply a "marginal aggregation machine" with unique characteristics
  - Characteristic of importance to actuaries is the strength of interdependence in regions – especially tails, especially upper tails – of the combined lines

Copulas provide a convenient way to aggregate the distributions of several lines

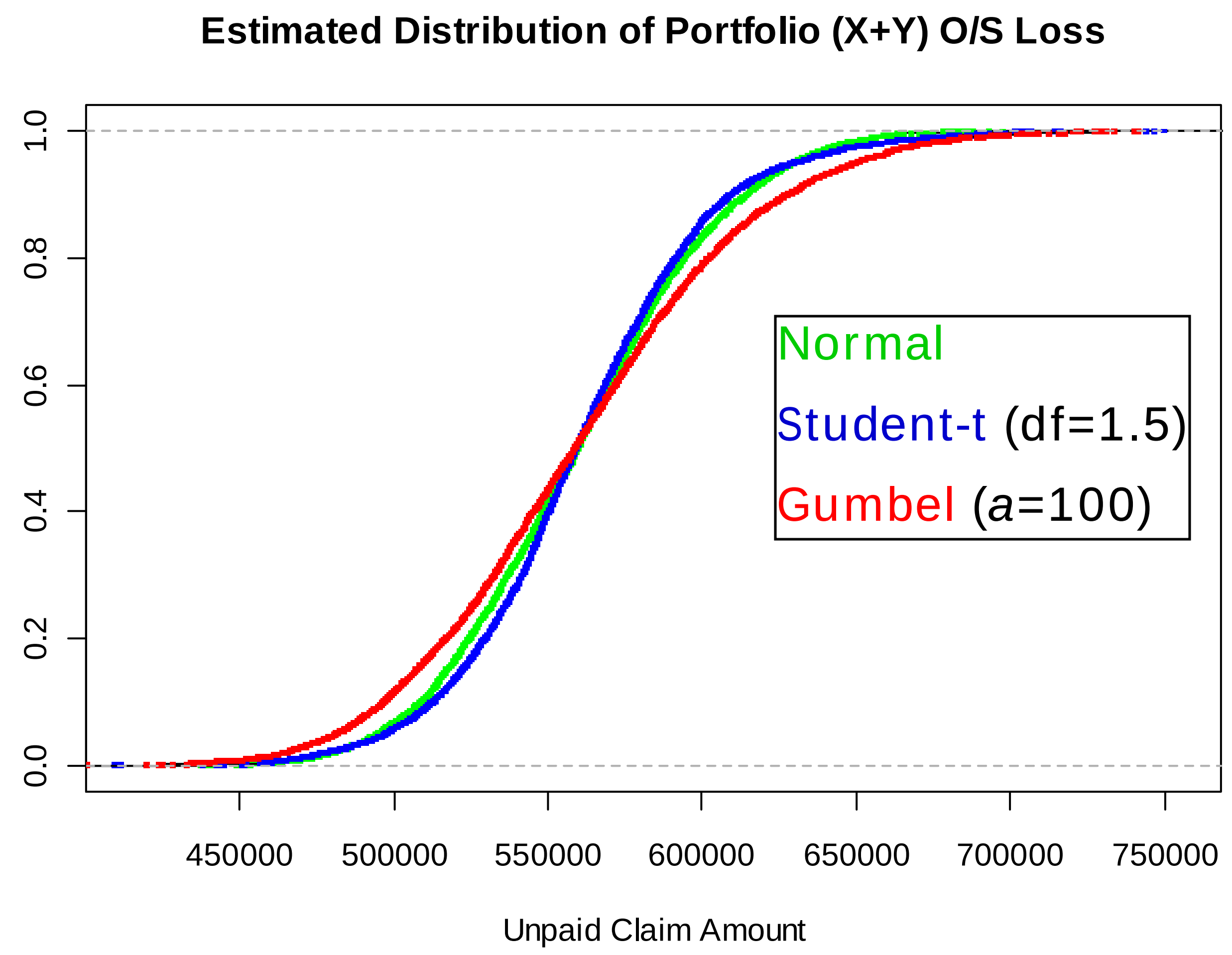
- Three popular copulas in actuarial use today are the Normal, the Student-t, and the Gumbel
- Normal copula
  - Average correlation is the sole input parameter
  - Combined lines will have no tail dependency
- Student-t
  - Uses correlation and degrees of freedom (df) as input parameters
  - The fewer df, the greater the tail dependency
- Gumbel
  - Related to extreme value theory for multivariates
  - Like normal, takes one parameter, a
    - $a = 1$ implies independence, increasing values imply greater upper tail dependence
  - Combined lines will have independent lower tails

Aggregation VaR's: correlation matrix vs. copulas

Method of Combination			Tail Percentiles 75.0%	99.0%	99.9%
Assume combined line is lognormally distributed, parameters from correlation matrix			592,106	682,622	729,038
Normal Copula	ρ=.5		589,818	673,151	715,217
Student-t	ρ=.5	df=25	589,692	673,674	718,007
		df=2	588,048	678,252	728,509
Gumbel		a=1.5	588,419	680,511	730,112
		a=20	593,472	689,223	737,617
		a=100	593,633	688,578	738,001

- Correlation-matrix values calculated analytically
- Copula values calculated from simulated empirical distribution using R
  - 1 million trials took about 5 seconds
- In this example, correlation-matrix VaR's are bracketed by the Student-t and the Gumbel

Disparity of cumulative distributions revealed when using copulas' most exaggerated parameters



Questions?

??????????????????????????????????

??????????????????????????????

??????????????????????????

??????????????????????

??????????????????

??????????????

??????????

??????

??