A Practical Approach for Estimating the Unpaid Claim Liability Distribution

Casualty Loss Reserve Seminar – September 2020

By John A. Dawson, FCAS, MAAA
Ashurst Beacon LLC

Ashurst Beacon LLC
Actuarial Consulting

Introduction

- Origin of the approach stemmed from reviewing loss reserve studies while working at Ernst & Young
- My objective was to develop an approach for estimating the unpaid claim liability that could be incorporated easily into most traditional loss reserve analyses without introducing complex statistical techniques.
- It is not intended to be a replacement for other techniques for estimating the UCL
  - e.g. Mack method, Bootstrapping, ICRFS, Simulation
- I hope you'll find the approach useful

"All models are approximations. Essentially, all models are wrong, but some are useful."

George Edward Pelham Box

Single Line/Segment of Business

Unpaid Claim Liability ("UCL"):

- Amount that will be paid after the valuation date to fulfil all claim obligations incurred under the policies issued
- Is the sum of payments that will be paid on all individual claims (C_i) incurred but unpaid as of the valuation date
  - i.e. UCL = $\sum C_i$ where C_i is a random variable representing the unpaid on an individual claim.
- Will not be known with certainty until all claims are closed and their settlement values known. Depending on the line of business it could be many years before this occurs.
- The value of C_i is a random variable from some unknown distribution

Single Line/Segment of Business

Let us assume for the moment that:

- C_i are independent random variables
- The number of unpaid claims (n) is known
- The true mean of the individual claim distribution (μ_c) is known
- The Coefficient of Variation (C_V) of the individual claim distribution $C_V(C_i)$ is known.

Given these assumptions, we know that the true mean of the UCL distribution (μ_L) is μ_c*n

Single Line of Business

The Central Limit Theorem says that:

- UCL distribution is approximately Normal
- Mean of the UCL distribution (μ_L) = $n * \mu_c$
- The variance of the UCL distribution (σ_L^2) = $n * \sigma^2_c$

Given the above, the C_V of the UCL distribution = $\sigma_L / \mu_L = \sqrt{(n * \sigma^2_c)}/(n * \mu_c) = C_V(C_i)/\sqrt{n}$

This C_V of the UCL distribution is **<u>conditional</u> on knowing the (1) number of unpaid claims, (2) the true mean of the UCL distribution and (3) the C_V of the individual claim size distribution**

We will denote this <u>conditional</u> coefficient of variation as $C_V(L|\mu_L)$

Ashurst Beacon LLC
Actuarial Consulting

C_V of Conditional UCL Distribution

Suppose a line/segment has 1,500 unpaid claims and the C_V of the claim distribution was chosen to be 3.000

Then $C_V(L|\mu_L) = 3.000 / \sqrt{1{,}500} = 0.077$

Should one infer from this that the UCL distribution for this line/segment is a Normal Distribution with a standard deviation equal to 7.7% of its mean?

The answer is "yes", **but only if**:

- we know with certainty the true mean value of the UCL distribution (i.e. μ_L) and the number of unpaid claims (n)

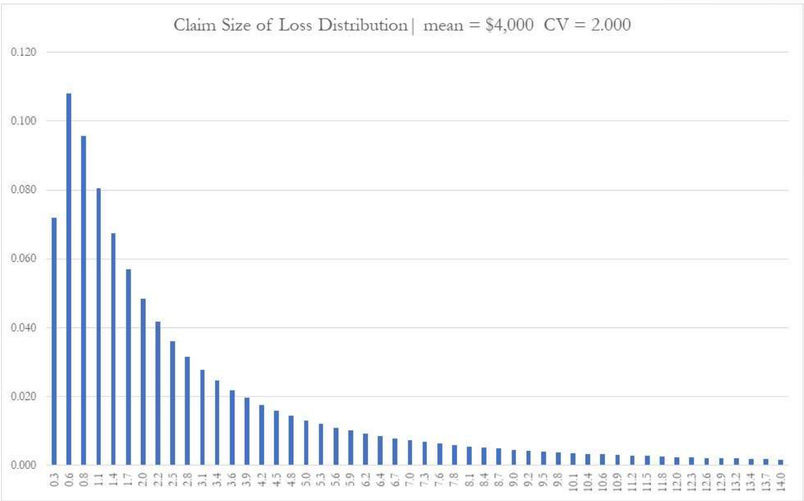
Example

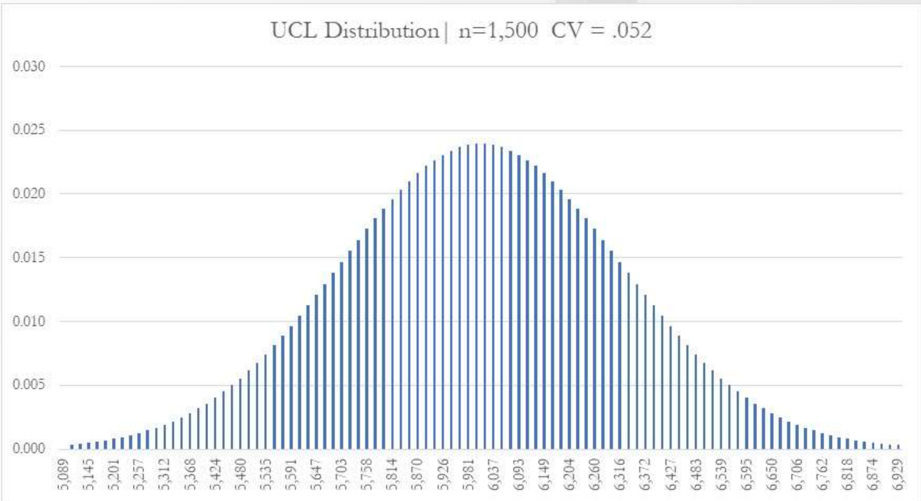
Table 1 $ in thousands

	(1)	(2)	(3)	(4)	(5)
		No, Unpaid	UCL		Process
Line	C_V(C)	Claims (n)	Mean	$C_V(L\|\mu)$	Std. Error
A	1.000	800	$2,000	0.035	$71
B	2.000	1,500	$6,000	0.052	$310
C	3.000	1,500	$15,000	0.077	$1,162
D	4.000	1,000	$15,000	0.126	$1,897
E	5.000	300	$10,000	0.289	$2,887
Notes:					
Col (1) = selected					
Cols (2) & (3) are assumed given					
Col (4) = Col (1) /SQRT(Col (2)) = Process volatility					
Col (5) = Col (3) * Col (4)					

- The C_V's shown in Col (4) are a measure of the overall uncertainty of the UCL distribution **only if** we know the mean of the UCL distribution with certainty
- They can be thought of as measures of the **process volatility** inherent in the line of business
- Their value is proportional to C_V(C) and inversely proportional to the number of unpaid claims (n)
- This is the essence of insurance – the volatility of aggregate loss on many claims is lower than the volatility on any individual claim

Line of Business - B

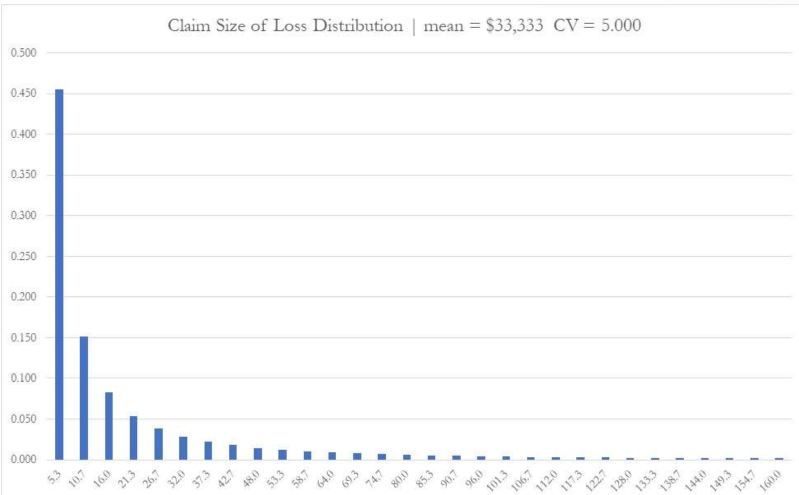




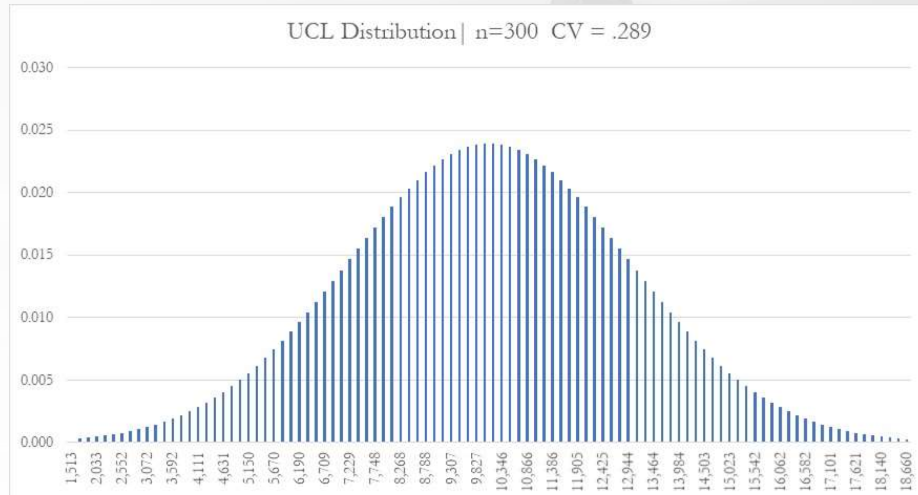
Claim distribution with a mean of $4,000 and CV = 2.0

With 1,500 unpaid claims, the UCL distribution has a mean of $6.0M and a standard deviation of $0.310M ($C_V$ = 0.052).

Line of Business - E



Claim distribution with a mean of $33,333 and CV = 5.0



With 300 unpaid claims, the UCL distribution has a mean of $10.0M and a standard deviation of $2.877M ($C_V$ = 0.289).

Ashurst Beacon LLC
Actuarial Consulting

C_V of Conditional UCL Distribution

Table 2: Coefficient of Variation of Conditional UCL Distribution								
		Number of Unpaid Claims						
CV(Ci)		100	500	1,000	2,500	5,000	10,000	50,000
0.500		0.050	0.022	0.016	0.010	0.007	0.005	0.002
1.000		0.100	0.045	0.032	0.020	0.014	0.010	0.004
1.500		0.150	0.067	0.047	0.030	0.021	0.015	0.007
2.000		0.200	0.089	0.063	0.040	0.028	0.020	0.009
2.500		0.250	0.112	0.079	0.050	0.035	0.025	0.011
3.000		0.300	0.134	0.095	0.060	0.042	0.030	0.013
3.500		0.350	0.157	0.111	0.070	0.049	0.035	0.016
4.000		0.400	0.179	0.126	0.080	0.057	0.040	0.018
4.500		0.450	0.201	0.142	0.090	0.064	0.045	0.020
5.000		0.500	0.224	0.158	0.100	0.071	0.050	0.022

Key Assumptions – Number of Unpaid Claims

- Often estimated in the loss reserve analysis.
- Could use the number of open claims - conservative approach since the lower the number the higher the C_V of the UCL distribution.
- Assumed no uncertainty in the number of unpaid claims
  - chosen to simplify the calculations but approach could be amended to incorporate uncertainty

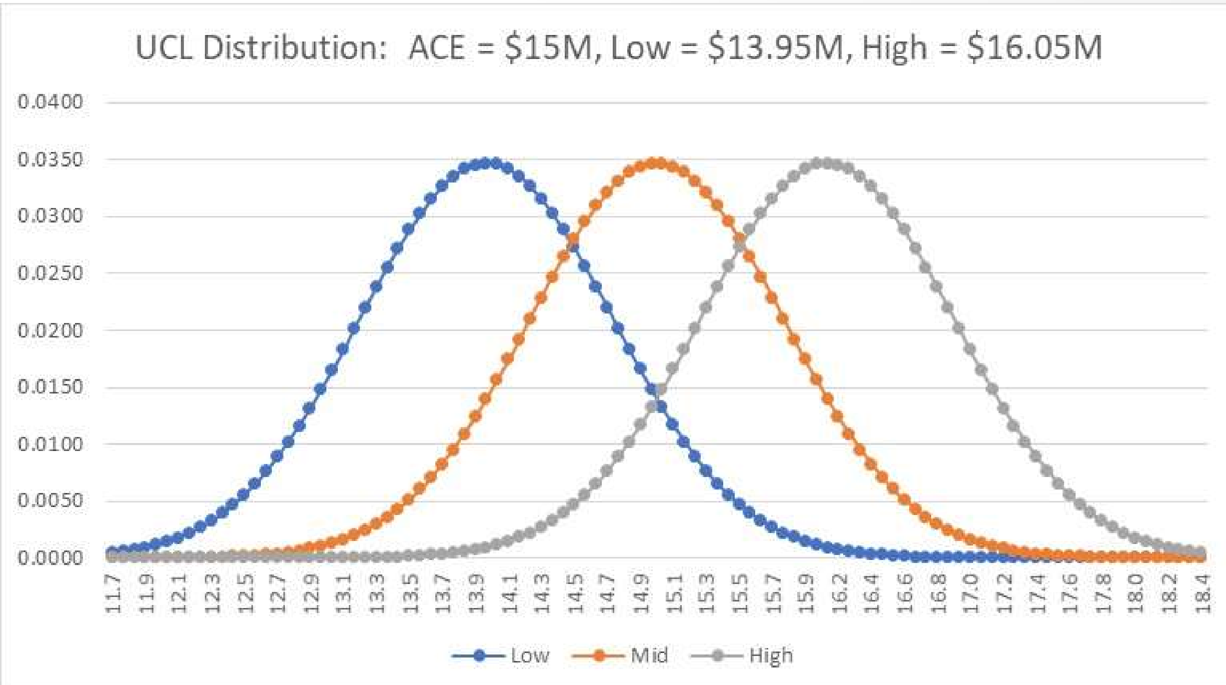
Key Assumptions – C_V of Claim Size Distribution

- Use the mean and standard deviation of open case reserves as a proxy
  - Case reserve are an estimate of amount remaining to be paid
  - Clustering of case reserves and tendency for larger claims to emerge later will likely cause C_V of open claims to be lower than C_V of unpaid amounts
- Review unpaid amounts on a block of closed claims, e.g. settlement value less paid to date as of 12/31/17 on claims closed as of 12/31/19
- Establish industry benchmark C_V's of each major line/segment. Clearly some lines/segments of business have a much higher dispersion of individual claim values than other lines.
- We will proceed under the assumption that reasonable estimates for $C_V(C_i)$ are available

Key Assumption – Mean of the UCL Distribution

- So far, we have assumed that the we know the true mean of the UCL distribution
- Clearly, we don't know for sure what the true mean is
- But we have a good estimate of its value – the Actuarial Central Estimate ("ACE")
- Most reserve evaluations acknowledge the uncertainty in estimating the mean of the UCL distribution by presenting a range of ACE values
- We will assume:
  - that the range appropriately reflects the uncertainty in estimating the mean of the UCL distribution
  - All values inside the range are reasonable alternative estimates of the ACE and are equally likely (i.e. uniformly distributed)

Accounting for Uncertainty in the Mean



Every point inside the range has a UCL distribution associated with it

If the true mean is the "low" end of the range, the UCL distribution is the blue line

If the true mean is the "high" end of the range, the UCL distribution is the grey line

If the true mean is the "mid" ACE estimate, the UCL distribution is the orange line

Accounting for Uncertainty in the ACE

The question now is how to account for the uncertainty in the estimate of the ACE in the calculation of the UCL distribution

- An empirical approach could be used by giving weight to the distribution at each point inside the range.

- As an alternative, we can derive the C_V of the overall distribution simply by using the Law of Total Variance:

  Var (Y) = E(Var[Y | X]) + Var[E(Y | X)] where X & Y are independent variables

  This is the familiar formula used in credibility theory that the total variance is equal to the expected value of the process volatility plus the variance of the hypothetical means

Ashurst Beacon LLC
Actuarial Consulting

Accounting for Uncertainty in the ACE

Setting Y as the UCL and X as the ACE of the UCL, the overall variance of the UCL distribution can be stated as:

$$Var (L) = \{ACE * C_V(L|\mu_L)\}^2 + Var[ACE(L)]$$

Dividing both sides by the ACE^2 and taking the square root we get

$$C_V(L) = \sqrt{[\{C_V(L|\mu_L)\}^2 + \{C_V(ACE(L))\}^2]}$$

This equation says that the **unconditional** coefficient of variation of the UCL distribution is simply the square root of the sum of squares of the **conditional** coefficient of variation of the UCL distribution and the coefficient of variation of the distribution of ACE estimates.

We already have the first term $C_V(L|\mu_L)$. The question now is how to calculate $C_V(ACE(L))$

Accounting for Uncertainty in the ACE

Assuming that all estimates inside the range are equally likely, then the distribution of estimates of the ACE is uniformly distributed across the range from low (A) to high (B), and the variance of the ACE distribution is

$$Var(ACE) = (B - A)^2 / 12$$

Dividing both sides by ACE^2 and taking square roots we get

$$C_V(ACE) = (B - A)/\{ACE*\sqrt{12})$$

For Line C in our example, assuming the range of estimates is $13.95M to $16.05M, with a select estimate of $15.0M (i.e. range is +/- 7.0% around the ACE), then

$$C_V(ACE) = (16.05 - 13.95)/\{15*\sqrt{12}) = 2.1 / 51.96 = .040$$

We will view the value of $C_V(ACE)$ as a measure of **parameter uncertainty**

Example continued

Table 3	$ in thousands								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		No, Unpaid	UCL	UCL Range			Process		
Line	$C_V(C)$	Claims (n)	ACE	Low	High	$C_V(L\|\mu)$	Std. Error	$C_V(ACE)$	$C_V(L)$
A	1.000	800	$2,000	$1,900	$2,100	0.035	$71	0.029	0.046
B	2.000	1,500	$6,000	$5,580	$6,420	0.052	$310	0.040	0.066
C	3.000	1,500	$15,000	$13,950	$16,050	0.077	$1,162	0.040	0.087
D	4.000	1,000	$15,000	$13,500	$16,500	0.126	$1,897	0.058	0.139
E	5.000	300	$10,000	$8,750	$11,250	0.289	$2,887	0.072	0.298

Notes:

- Col (1) = selected
- Cols (2), (3), (4) & (5) from actuarial study
- Col (6) = Col (1) /SQRT(Col (2)) = Process volatility
- Col (7) = Col (3) * Col (6)
- Col (8) = [Col (5) - Col (4)] / [Col (3) * √12]
- Col (9) = √[Col (6)^2 + Col (8)^2]

Ashurst Beacon LLC
Actuarial Consulting

Aggregating across all Lines/Segments of Business

Now that we have the $C_V(L)$ for each line/segment, the next question is how to get the results for all lines/segments in aggregate.

My suggested approach is to first calculate the process volatility for the aggregate UCL distribution, i.e. $C_V(L|\mu_L)$ using the standard statistical formula for the variance of the sum of random variables

i.e. $Var(X+Y+Z) = Var(X) + Var(Y) + Var(Z) + 2*\rho_{x,y}*\sigma_x\sigma_y + 2*\rho_{x,z}*\sigma_x\sigma_z + 2*\rho_{y,z}*\sigma_y\sigma_z$

where $\rho_{x,y}$ is the correlation coefficient of the variables X and Y.

It is my contention that there is little correlation in process volatility across most lines of business. Hence, setting ρ as zero or a low value is, in my opinion, not unreasonable.

Possible exception occurs if loss is analyzed separately from expense and needs to be combined

Example continued

Table 4 $ in thousands

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		No, Unpaid	UCL	UCL Range			Process		
Line	$C_V(C)$	Claims (n)	ACE	Low	High	$C_V(L\|\mu)$	Std. Error	$C_V(ACE)$	$C_V(L)$
A	1.000	800	$2,000	$1,900	$2,100	0.035	$71	0.029	0.046
B	2.000	1,500	$6,000	$5,580	$6,420	0.052	$310	0.040	0.066
C	3.000	1,500	$15,000	$13,950	$16,050	0.077	$1,162	0.040	0.087
D	4.000	1,000	$15,000	$13,500	$16,500	0.126	$1,897	0.058	0.139
E	5.000	300	$10,000	$8,750	$11,250	0.289	$2,887	0.072	0.298
Total		**5,100**	**$48,000**	**$44,400**	**$51,600**	**0.076**	**$3,658**	**0.043**	**0.088**

Notes:

Col (1) = selected

Cols (2), (3), (4) & (5) from actuarial study

Cols (2), (3), (4) & (5) Totals from actuarial study

Col (6) = Col (1) /√(Col (2)) = Process volatility

Col (6) Total = Col (7) Total / Col (3) Total

Col (7) = Col (3) * Col (6)

Col (7) Total = √[sum of squares Col (7)]

Col (8) = [Col (5) - Col (4)] / [Col (3) * √12]

Col (8) Total = [Col (5) - Col (4)]/[Col (3) * √12]

Col (9) = $\sqrt{[\text{Col }(6)^2 + \text{Col }(8)^2]}$

Col (9) Total = $\sqrt{[\text{Col }(6)^2 + \text{Col }(8)^2]}$

Aggregating across all Lines of Business

Information for individual lines are the same as in the earlier slide

The information for the aggregate total line is calculated as follows:

- Number of unpaid claims [Col (2)], ACE [Col (3)] and Aggregate range [Cols (4) & (5)] come from the actuarial study
- Aggregate standard deviation [Col (7)] is calculated using the variance of the sum of random variables assuming lines are uncorrelated, i.e. square root of the sum of squares of Col (7)
- Aggregate process volatility [Col (6)] is simply calculated as Col (7) / Col (3) = 0.076
- Aggregate parameter uncertainty [Col (8)] is calculated using the aggregate “low” and “high” estimates, i.e. = (51.6 – 44.4)/(48.0 * √12) = 7.2 / 166.3 = 0.043
- C_V of aggregate UCL distribution Col (9) is calculated as $\sqrt{(0.076^2 + 0.043^2)} = 0.088$

Ashurst Beacon LLC
Actuarial Consulting

Aggregating across all Lines of Business

The following table illustrates the sensitivity of results to the choice of process correlation

(ρ is assumed to be the same across all lines/segments):

ρ =	0.000	0.200	0.400	0.600	0.800	1.000
$C_V(L\|\mu)$ =	0.076	0.090	0.102	0.113	0.123	0.132
$C_V(L)$ =	0.088	0.100	0.111	0.121	0.130	0.139

In this example, regardless of correlation assumption, the uncertainty associated with the aggregate UCL distribution is driven mostly by process volatility. Parameter uncertainty has a small impact on $C_V(L)$.

This finding may suggest that the range around the ACE is too small for a book of business that has so much process volatility.

Thought – is there a relationship between the level of process volatility and parameter uncertainty?

Consistency between process volatility & parameter uncertainty

Company A

		No, Unpaid	UCL	UCL Range				Process		
Line	$C_V(C)$	Claims (n)	ACE	Low	High	Range%	$C_V(L\|\mu)$	Std. Error	$C_V(ACE)$	$C_V(L)$
A	1.000	800	$2,000	$1,900	$2,100	5.0%	0.035	$71	0.029	0.046
B	2.000	1,500	$6,000	$5,580	$6,420	7.0%	0.052	$310	0.040	0.066
C	3.000	1,500	$15,000	$13,950	$16,050	7.0%	0.077	$1,162	0.040	0.087
D	4.000	1,000	$15,000	$13,500	$16,500	10.0%	0.126	$1,897	0.058	0.139
E	5.000	300	$10,000	$8,750	$11,250	12.5%	0.289	$2,887	0.072	0.298
Total		**5,100**	**$48,000**	**$44,400**	**$51,600**	7.5%	**0.076**	**$3,658**	**0.043**	**0.088**

Company B

		No, Unpaid	UCL	UCL Range				Process		
Line	$C_V(C)$	Claims (n)	ACE	Low	High	Range%	$C_V(L\|\mu)$	Std. Error	$C_V(ACE)$	$C_V(L)$
A	1.000	8,000	$20,000	$19,000	$21,000	5.0%	0.011	$224	0.029	0.031
B	2.000	15,000	$60,000	$55,800	$64,200	7.0%	0.016	$980	0.040	0.044
C	3.000	15,000	$150,000	$139,500	$160,500	7.0%	0.024	$3,674	0.040	0.047
D	4.000	10,000	$150,000	$135,000	$165,000	10.0%	0.040	$6,000	0.058	0.070
E	5.000	3,000	$100,000	$87,500	$112,500	12.5%	0.091	$9,129	0.072	0.116
Total		**51,000**	**$480,000**	**$444,000**	**$516,000**	7.5%	**0.024**	**$11,569**	**0.043**	**0.050**

Consistency between process volatility & parameter uncertainty

Company B is 10 times the size of Company A - all other assumptions are the same

Aggregate range around the ACE for both companies is +/- 7.5% - consequently, the parameter uncertainty for both companies is the same (4.3%)

Question:

- Shouldn't the parameter uncertainty associated with Company A be higher than that for Company B since Company A's process volatility is much higher?
- Does the fact that they are the same raise questions that Company A's range is too narrow?

Increased process volatility should normally cause greater parameter uncertainty as identifying trends in the data and selecting key assumptions will be that much more difficult

Approach may be useful in identifying inconsistencies in disclosed ranges

Summary of Key Assumptions

- Number of unpaid claims, ACE & Range come from the actual study
- Number of unpaid claims is fixed
- $C_V(C_i)$ is chosen based on a review of individual claim size distribution
- Process correlation across lines/segments is zero (alternatives could be chosen)
- Actuarial analysis has range of ACE values for each line/segment and in the aggregate
- The range reflects the uncertainty in estimating the expected value of the UCL distribution and all values inside the range are equally likely (i.e. uniformly distributed)
- The aggregate range accounts for parameter correlation across the lines/segments
- UCL distribution is approximately Normal

Risk of Material Adverse Deviation

The NAIC instructions for the SAO ask that the Appointed Actuary ("AA"):

- Identify significant risks and uncertainties associated with loss reserves
- Assess whether those risks could result in a material adverse deviation ("MAD")

The word "significant" refers to the risks and uncertainties, not the possibility of a MAD occurring

The instructions do NOT ask the AA :

- to assess what is the risk (probability) of a material adverse deviation ("RMAD")
- to assess whether the risk (probability) of material adverse deviation is significant

To answer these questions requires an estimate of the probability of a material adverse deviation ("PMAD") and a definition of what is "significant"

Ashurst Beacon LLC
Actuarial Consulting

Risk of Material Adverse Deviation

Given the value $C_V(L)$, the Recorded Reserve ("R") and a Materiality Standard (M), the value of PMAD can easily be computed.

In our example, assuming the recorded reserve is $46M and M = $5.5M, then

PMAD = P(L>46.0+5.5) = P(Z>(51.5-48.0)/(.088*48.0)) = P(Z>0.71) = 0.203

The question that arises is whether this 20% probability is, by itself, sufficient to warrant an affirmative RMAD disclosure in the SAO.

The RMAD disclosure could indicate that a major contributor to the 20% PMAD is line E. This line has significant process volatility (28.9%) due to the nature of the claims and the low number of unpaid claims

Thought – should the PMAD and the threshold used for defining "significant" be a required disclosure in the SAO?

Concluding Remarks

- Approach is an alternative method for estimating the UCL distribution
- Can easily be incorporated into a typical reserve analysis
- Gives insight into the levels of process volatility and parameter uncertainty and their contribution to the overall level of uncertainty in the UCL distribution
- Identifies those lines that are contributing most to the overall uncertainty
- Can aid in the assessment of RMAD in the SAO

QUESTIONS/DISCUSSION

Ashurst Beacon LLC
Actuarial Consulting

Ashurst Beacon LLC
Actuarial Consulting